MA/MSc MT-01

# Vardhaman Mahaveer Open University, Kota

## Advanced Algebra

## Course Development Committee

**Chairman**
**Prof. (Dr.) Naresh Dadhich**
Vice-Chancellor
**Vardhaman Mahaveer Open University, Kota**

## Co-ordinator/Convener and Members

**Subject Convener**
**Prof. D.S. Chauhan**
Department of Mathematics
University of Rajasthan, Jaipur

**Co-ordinator**
**Dr. Anuradha Sharma**
Assistant Professor
Department of Botany, V.M.O.U., Kota

### Members :

1. **Prof. V.P. Saxena**
   Ex Vice-Chancellor
   Jiwaji University,
   Gwalior (MP)
2. **Prof. S.C. Rajvanshi**
   Deptt. of Mathematics
   Institute of Eng. & Tech.
   Bhaddal, Ropar (Punjab)
3. **Prof. P.K. Banerjee**
   Emeritus Fellow (UGC)
   Deptt. of Mathematics
   J.N.V. University, Jodhpur
4. **Prof. S.P. Goyal**
   Emeritus Scientist (CSIR)
   Deptt. of Mathematics
   University of Rajasthan, Jaipur
5. **Dr. A.K. Mathur**
   Associate Prof. (Retired)
   Deptt. of Mathematics
   University of Rajasthan, Jaipur
6. **Dr. K.N. Singh**
   Associate Prof. (Retired)
   Deptt. of Mathematics
   University of Rajasthan, Jaipur
7. **Dr. Paresh Vyas**
   Assistant Professor
   Deptt. of Mathematics
   University of Rajasthan, Jaipur
8. **Dr. Vimlesh Soni**
   Lecturer
   Deptt. of Mathematics
   Govt. PG College, Kota (Raj.)
9. **Dr. K.K. Mishra**
   Lecturer
   Deptt. of Mathematics
   M.S.J. Collage, Bharatpur (Raj.)
10. **Dr. K.S. Shekhawat**
    Lecturer, Deptt. of Mathematics
    Govt. Shri Kalyan College, Sikar (Raj.)

## Editing and Course Writing

**Editor**
**Prof. D.S. Chauhan**
Deptt. of Mathematics
University of Rajasthan, Jaipur

**Writers**

1. **Dr. K.N. Singh**
   Associate Prof. (Retired)
   Deptt. of Mathematics
   University of Rajasthan, Jaipur
2. **Dr. Vimlesh Soni**
   Lecturer
   Deptt. of Mathematics
   Govt. PG College, Kota (Raj.)
3. **Dr. A.K. Goyal**
   Lecturer
   Deptt. of Mathematics
   M.S.J. College, Bharatpur
4. **Dr. Keshav Sharma**
   Lecturer
   Deptt. of Mathematics
   R.R. College, Alwar

## Academic and Administrative Management

**Prof. (Dr.) Naresh Dadhich**
Vice-Chancellor
Vardhaman Mahveer Open University, Kota

**Prof. M.K. Ghadoliya**
Director (Academic)
Vardhaman Mahveer Open University, Kota

**Mr. Yogendra Goyal**
Incharge
Material Production and Distribution Department

## Course Material Production

**Mr. Yogendra Goyal**
Assistant Production Officer
Vardhaman Mahaveer Open University, Kota

# Vardhaman Mahaveer Open University, Kota

## Advanced Algebra

# PREFACE

*The Present book entitled "**Advanced Algebra**" has been designed so as to cover the unit-wise syllabus of Mathematics-First paper for M.A./M.Sc. (Previous) students of Vardhaman Mahaveer Open University, Kota. It can also be used for competitive examinations. The basic principles and theory have been explained in a simple, concise and lucid manner. Adequate number of illustrative examples and exercises have also been included to enable the students to grasp the subject easily. The units have been written by various experts in the field. The unit writers have consulted various standard books on the subject and they are thankful to the authors of these reference books.*

# UNIT 1 : Direct Products of Groups

## Structure of the Unit



## 1.0 Objectives

This unit introduces a construction process, called direct product of groups, which produces new larger groups. In this unit we shall study the external direct product of groups and the internal direct product.

## 1.1 Introduction

We know what defines a group, several examples of groups, and their subgroups. In this unit we shall see how we may build new groups from old known ones. This process of constructing new groups, which is the simplest of many others, is called a direct product.

## 1.2 External direct product

**Definition :**

Let $G_1$ and $G_2$ be any two groups. The Cartesian product of $G_1$ and $G_2$ is given by

$$G_1 \times G_2 = \{(x_1, x_2) \mid x_1 \in G_1, x_2 \in G_2\}.$$

It is a group for the binary composition defined as

$$(x_1, x_2)\,(y_1, y_2) = (x_1\, y_1,\ x_2 y_2)$$

For $(x_1, x_2), (y_1, y_2) \in G_1 \times G_2$

This group, denoted as $G_1 \times G_2$ is called the external direct product or simply direct product of $G_1$ and $G_2$.

We shall usually write all abstract groups multiplicatively unless stated otherwise.

Let us verify that $G_1 \times G_2$ forms a group for the binary composition defined above.

**1. Closure property :** Since $x_1 \in G_1, y_1 \in G_1$ therefore $x_1 y_1 \in G_1$ as $G_1$ is a group. Similarly $x_2\ y_2 \in G_2$.

Therefore $(x_1, x_2)\ (y_1, y_2) = (x_1 y_1\ ,\ x_2 y_2) \in G_1 \times G_2$

for $(x_1, x_2) \in G_1 \times G_2\ ,\ (y_1, y_2) \in G_1 \times G_2$.

**2. Associativity :** Let $(x_1\ x_2)$, $(y_1, y_2)$ and $(z_1, z_2)$ be any three members of $(G_1 \times G_2)$. Then

$$\begin{aligned}
\left[(x_1, x_2)(y_1, y_2)\right](z_1, z_2) &= (x_1 y_1, x_2 y_2)(z_1, z_2) \text{ (definition of binary composition)} \\
&= \left[(x_1 y_1) z_1, (x_2 y_2) z_2\right] \\
&= \left[x_1 (y_1 z_1), x_2 (y_2 z_2)\right] \quad \text{(Associativity in } G_1 \text{ and } G_2\text{)} \\
&= (x_1, x_2)(y_1 z_1,\ y_2 z_2) \\
&= (x_1, x_2)\left[(y_1, y_2)(z_1, z_2)\right]
\end{aligned}$$

**3. Existence of identity element :** Let $e_1$ be the identity element of $G_1$ and $e_2$ be the identity element of $G_2$. Let $(x_1, x_2)$ be any arbitrary element of $G_1 \times G_2$. Then

$$\begin{aligned}
(x_1, x_2)\ (e_1, e_2) &= (x_1\ e_1, x_2\ e_2) \\
&= (x_1, x_2) \quad \text{(definition of identity elements in } G_1 \text{ and } G_2\text{)}
\end{aligned}$$

Similarly, $(e_1, e_2)\ (x_1, x_2) = (x_1, x_2)$

Thus $(e_1, e_2)$ is the identity element in $G_1 \times G_2$.

**4. Existence of inverse element :** Let $x_1 \in G_1$ then $x_1^{-1} \in G_1$ as $G_1$ is a group. Similarly for $x_2 \in G_2\ , \exists\, x_2^{-1} \in G_2$. Thus for $(x_1, x_2) \in G_1 \times G_2\ ,\ \exists \left(x_1^{-1}, x_2^{-1}\right) \in G_1 \times G_2$ such that

$$(x_1, x_2)\left(x_1^{-1}, x_2^{-1}\right) = \left(x_1\, x_1^{-1},\ x_2\, x_2^{-1}\right) = (e_1,\ e_2)$$

and similarly

$$\left(x_1^{-1}, x_2^{-1}\right)(x_1, x_2) = (e_1, e_2)$$

Thus every element in $G_1 \times G_2$ has inverse in $G_1 \times G_2$.

Hence $G_1 \times G_2$ is a group.

**Order of the group $G_1 \times G_2$ :** Order of the group $G_1 \times G_2$ is the number of ordered pairs in the cartesian product of $G_1 \times G_2$. It is the product of the number of elements in $G_1$ to that of $G_2$.

Thus

$$\begin{aligned}
o(G_1 \times G_2) &= |\ G_1 \times G_2\ | = |G_1|\ |G_2| \\
&= o(G_1)\ o(G_2)
\end{aligned}$$

where $|\ G\ |$ denotes the number of elements is $G$.

***Ex.1.*** *Consider two groups* $(Z_2, +_2)$ *and* $(Z_3, +_3)$*, where* $Z_2 = \{0, 1\}$ *and* $Z_3 = \{0, 1, 2\}$. *Then* $Z_2 \times Z_3$ *will be a group of order 6. Thus*

$$Z_2 \times Z_3 = \{(0, 0)\ (0, 1), (0, 2), (1, 0), (1,1), (1, 2)\}$$

Hence binary composition $Z_2 \times Z_3$ is defined as follows :

$$(x_1, x_2)\ (y_1\ , y_2) = (x_1 +_2 y_1, x_2 +_3 y_2)$$

for $(x_1, x_2)\ (y_1\ , y_2) = \in Z_2 \times Z_3$

and $x_1, y_1 \in Z_2,\ x_2, y_2 \in Z_3$ .

*e.g.* $(1, 1)\ (1, 2) = (1 +_2 1, 1 +_3 2)$

$= (0, 0)$ etc.

Here (0, 0) will be the identity element.

***Ex.2.*** *Let* $(Z, +)$ *be the additive group of integers and* $(C_0, \bullet)$ *be the multiplicative group of non-zero complex numbers. Then* $Z \times C_0$ *is a group under the binary composition defined as follows :*

$$(x_1\ , x_2)\ (y_1\ , y_2) = (x_1 + y_1, x_2 \cdot y_2)$$

Since 0 is identity in $Z$ and 1 is identity in $C_0$, therefore (0, 1) is the identity in $Z \times C_0$. For each $(x, y) \in Z \times C_0$, $(-x, y^{-1})$ is inverse in $Z \times C_0$.

***Ex.3.*** *Let* $(Z, +)$ *be the group of integers, then* $G = Z \times Z$ *is a group under the composition defined as follows :*

$$(x_1\ , y_2)\ (x_2\ , y_2) = (x_1 + x_2, y_1 + y_2),\ \forall\ (x_1\ , y_1), (x_2\ , y_2) \in G\ .$$

## 1.3 Theorems on external direct product

**Theorem 1.** *Let* $G_i$ $(1 \le i \le n)$ *be n groups and G is the external direct product of these groups. Let* $e_i$ *be the identity of the group* $G_i$*, for each* $i (1 \le i \le n)$*. Then*

*(i) For each i,* $H_i = \{(e_1, e_2, \ldots, e_{i-1}, x_i, e_{i+1}, \ldots, e_n \mid x_i \in G_i)\}$ *is a normal subgroup of G*

.

*(ii)* $H_i$ *is isomorphic to* $G_i$ *i.e.* $H_i \cong G_i,\ \forall\ i$

*(iii) Each* $g \in G$ *can be written uniquely as product of elements from* $H_1, H_2, \ldots. Hn.$

**Proof :** Since $G$ is external direct product of $G_1, G_2, \ldots, G_n$

therefore $G = G_1 \times G_2 \times \ldots \times G_n$

is a group for component wise multiplication,

and $G = \{(g_1, g_2, \ldots, g_n) \mid g_i \in G_i\}$.

***(i)*** Given that

$$H_i = \{(e_1, e_2, \ldots e_{i-1}, x_i, e_{i+1}, \ldots e_n) \mid x_i \in G_i\}.$$

By the definitions of $G$ and $H_i$, it is obvious that $H_i$ is a non-empty subset of $G$ . Let $a_i$, $b_i \in G_i$, then

$a = (e_1, e_2, \ldots e_{i-1}, a_i, e_{i+1}, \ldots e_n)$ and $b = (e_1, a_2, \ldots e_{i-1}, b_i, e_{i+1}, \ldots e_n) \in H_i$.

Now, $ab^{-1} = (e_1, e_2, \ldots e_{i-1}, a_i, e_{i+1}, \ldots e_n)(e_1, e_2, e_{i-1}, b_i, e_{i+1}, \ldots e_n)^{-1}$

$= (e_1, e_2, \ldots e_{i-1}, a_i, e_{i+1}, \ldots e_n)(e_1, e_2, e_{i-1}, b_i^{-1}, e_{i+1}, \ldots e_n)$

(by the definition of inverse in $G$)

$= (e_1, e_2, \ldots e_{i-1}, a_i b_i^{-1}, e_{i+1}, \ldots e_n)$

(by the definition of binary composition in $G$)

Thus $ab^{-1} \in H_i$, whenever $a, b \in H_i$ ($a_i b_i^{-1} \in G_i$ as $G_i$ is a group)

so $H_i$ is a subgroup of $G$.

Now, let $g = (g_1, g_2, \ldots g_n) \in G$, then

$$gag^{-1} = (g_1, g_2 \ldots g_n)(e_1, e_2, \ldots e_{i-1}, a_i, e_{i+1}, \ldots e_n)(g_1, g_2 \ldots g_n)^{-1}$$

$$= (g_1, g_2 \ldots g_n)(e_1, e_2, e_{i-1}, a_i, e_{i+1}, \ldots e_n)\left(g_1^{-1}, g_2^{-1}, \ldots g_n^{-1}\right)$$

$$= \left(g_1 e_1 g_1^{-1}, g_2 e_2 g_2^{-1}, \ldots g_i a_i g_i^{-1}, g_n e_n g_n^{-1}\right)$$

$$= \left(e_1, e_2 \ldots, g_i a_i g_i^{-1}, \ldots e_n\right) \in H_i \qquad (\because \quad g_i, a_i \in G_i)$$

Thus $gag^{-1} \in H_i$, whenever $g \in G$ and $a \in H_i$.

Hence $H_i$ is a normal subgroup of $G$, $\forall\, i$

***(ii)*** Let us define a mapping $f: G_i \to H_i$

by $f(g_i) = (e_1, e_2, \ldots e_{i-1}, g_i, e_{i+1}, \ldots, e_n)$, $\forall\, g_i \in G_i$. Obviously $f$ is well-defined.

For any $a_i, b_i \in G_i$

$$f(a_i b_i) = (e_1, e_2, \ldots, e_{i-1}, a_i b_i, e_{i+1}, \ldots, e_n)$$

$$= (e_1, e_2, \ldots, e_{i-1}, a_i, e_{i+1}, \ldots, e_n)(e_1, e_2, \ldots, e_{i-1}, b_i, e_{i+1}, \ldots, e_n)$$

(by the definition of composition in G)

$$= f(a_i) f(b_i) \qquad \text{(by the definition of } f\text{)}$$

Thus $f$ is a homomorphism.

Let $f(a_i) = f(b_i)$ for $a_i, b_i \in G_i$

$\Rightarrow$ $(e_1, e_2, \ldots, e_{i-1}, a_i, e_{i+1}, \ldots, e_n) = (e_1, e_2, \ldots, e_{i-1}, b_i, e_{i+1}, \ldots, e_n)$

$\Rightarrow$ $a_i = b_i$

Thus $f$ is one-one.

For each $(e_1, e_2, \ldots, e_{i-1}, a_i, e_{i+1}, \ldots, e_n) \in H_i$, $\exists\, a_i \in G_i$ such that

$$f(a_i) = (e_1, e_2, \ldots, e_{i-1}, a_i, e_{i+1}, \ldots, e_n) \in H_i$$

Thus $f$ is onto.

Hence $f$ is an isomorphism and $H_i \cong G_i$, for each $i$.

***(iii)*** Let $g = (g_1, g_2, \ldots, g_n) \in G$

then $g = (g_1, e_2, e_3, \ldots, e_n)(e_1, g_2, e_3, \ldots, e_n) \ldots (e_1, e_2, \ldots, e_{n-1}, g_n) \in H_1 H_2 \ldots H_n$

thus g can be expressed as product of elements of $H_1, H_2, \ldots H_n$.

For uniqueness, let, if possible,

$$g = (g_1, e_2, e_3, \ldots, e_n)(e_1, g_2, e_3, \ldots, e_n) \ldots (e_1, e_2, \ldots, e_{n-1}, g_n)$$

$$= (g_1', e_2, e_3, \ldots e_n)(e_1, g_2', e_3, \ldots e_n)\ldots(e_1, e_2, \ldots e_{n-1}, g_n') \text{ (where } g_i, g_i' G_i)$$

$\Rightarrow \quad (g_1, g_2, \ldots, g_n) - (g_1', g_2', \ldots, g_n')$

$\Rightarrow \quad g_1 = g_1'$, $g_2 = g_2'$, ..., $g_n = g_n'$

Thus, $g \in G$ can be written uniquely as product of elements of $H_1, H_2, \ldots H_n$.

**Theorem 2.** *Let $G_i$ and $G_2$ be groups, then*

*(i)* $G_1 \times G_2 \cong G_2 \times G_1$

*(ii)* *If* $H_1 = \{(a, e_2) | \ a \in G_1\}$ and $H_2 = \{(e_1, b) | \ b \in G_2\}$, *where $e_1$ and $e_2$ are identity elements of $G_1$ and $G_2$ respectively, then $H_1$ and $H_2$ are normal subgroups of $G_1 \times G_2$.*

*(iii)* $H_1 \cong G_1$ *and* $H_2 \cong G_2$

*(iv)* *The factor (quotient) group $(G_1 \times G_2) | \ H_1$ is isomorphic to $G_2$, and $(G_1 \times G_2) | \ H_2$ is isomorphic to $G_1$.*

**Proof :** ***(i)*** Let us define a mapping

$$f : G_1 \times G_2 \to G_2 \times G_1$$

such that $f(x_1, x_2) = (x_2, x_1) \ \forall (x_1, x_2) \in G_1 \times G_2$.

Obviously $\phi$ is well-defined.

Now, for $(x_1, x_2), (y_1, y_2) \in G_1 \times G_2$,

$$f\left[(x_1, x_2)(y_1, y_2)\right] = f(x_1 y_1, x_2 y_2)$$

$$= (x_2 y_2, x_1 y_1) \qquad \text{(by the definition of } f \text{)}$$

$$= (x_2, x_1)(y_2, y_1)$$

$$= f(x_1, x_2) f(y_1, y_2)$$

thus $f$ is a homomorphism.

Now, let

$$f(x_1, x_2) = f(y_1, y_2)$$

for $(x_1, x_2), (y_1, y_2) \in G_1 \times G_2$

$\Rightarrow \quad (x_2, x_1) = (y_2, y_1)$

$\Rightarrow \quad x_2 = y_2, \ x_1 = y_1$

$\Rightarrow \quad (x_1, x_2) = (y_1, y_2)$

thus $f$ is one-one.

For each $(x_2, x_1) \in G_2 \times G_1$, $x_2 \in G_2$, $x_1 \in G_1$, $\exists\ (x_1, x_2) \in G_1 \times G_2$ such that $f(x_1, x_2) = (x_2, x_1) \in G_2 \times G_1$

Thus $f$ is onto.

Hence f is an isomorphism and $G_1 \times G_2 \cong G_2 \times G_1$.

***(ii)*** and *(iii)* are particular cases ($n = 2$) of theorem 1.

***(iv)*** Let us define

$$\phi : G_1 \times G_2 \to G_2$$

such that

$$\phi(x_1, x_2) = x_2, \forall (x_1, x_2) \in G_1 \times G_2 .$$

Obviously $\phi$ is well-defined.

Now, for $(x_1, x_2), (y_1, y_2) \in G_1 \times G_2$

$$\begin{aligned} \phi(x_1, x_2)(y_1, y_2) &= \phi(x_1 y_1,\ x_2 y_2) \\ &= x_2\, y_2 \\ &= \phi(x_1,\ x_2)\phi(y_1,\ y_2) \end{aligned}$$

thus $\phi$ is a homomorphism.

For each $x_2 \in G_2$, $(x_1, x_2) \in G_1 \times G_2$, such that $\phi(x_1, x_2) = x_2$, thus $\phi$ is onto and hence $\phi$ is an epimorphism.

Now,

$$\begin{aligned} \ker\phi &= \{(g_1, g_2) \in G_1 \times G_2 \mid \phi(g_1, g_2) = e_2 \in G_2\} \\ &= \{(g_1, g_2) \in G_1 \times G_2 \mid g_2 = e_2 \in G_2\} \\ &= H_1 \end{aligned}$$

Hence by the fundamental theorem on homomorphism, we have

$$\frac{G_1 \times G_2}{\ker\phi} \cong f(G_1 \times G_2)$$

*i.e.*

$$\frac{G_1 \times G_2}{H_1} \cong G_2$$

Similarly, the other result can be proved.

**Theorem 3.** *Let G be the external direct product of groups $G_1, G_2, \ldots G_n$. Let*

$$H_i = \{e_1, e_2, \ldots, e_{i-1}, x_i, e_{i+1}, \ldots, e_n \mid x_i \in G_i\}$$

*then*

*(i)* $$\frac{G}{G_i} \cong G_i \times G_2 \times \ldots \times G_{i-1} \times G_{i+1} \times \ldots \times G_n$$

*(ii) If $x \in H_i$, $y \in H_j$ for same $i \neq j$,* then $xy = yx$.

**Proof :** ***(i)*** We have already shown that $G_i \simeq H_i$ and $H_i \lhd G$ (theorem 1), so factor group $\frac{G}{G_i}$ exists, as $H_i$ can be identified by $G_i$. Let us define a mapping

$$f : G \to G_1 \times G_2 \times \ldots \times G_{i-1} \times G_{i+1} \times \ldots \times G_n$$

Such that $f(g_1, g_2, \ldots, g_n) = (g_1, g_2, \ldots, g_{i-1}, g_{i+1}, \ldots, g_n)$

Obviously $f$ is well-defined. Now, for $(a_1, a_2, \ldots, a_n), (b_1, b_2, \ldots, b_n) \in G$, we have

$$\begin{aligned} f\left[(a_1, a_2, \ldots, a_n)(b_1, b_2, \ldots, b_n)\right] &= f\left(a_1 b_1, a_2 b_2, \ldots, a_n b_n\right) \\ &= \left(a_1 b_1, a_2 b_2, \ldots, a_{i-1} b_{i-1}, a_{i+1} b_{i+1}, \ldots, a_n b_n\right) \\ &= \left(a_1, a_2, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n\right)\left(b_1, b_2, \ldots, b_{i-1}, b_{i+1}, \ldots, b_n\right) \\ &= f\left(a_1, a_2, \ldots, a_n\right) f\left(b_1, b_2, \ldots, b_n\right) \end{aligned}$$

thus $f$ is homomorphism. Also, $f$ is onto.

Hence $f$ is an epimorphism.

Now,

$$\begin{aligned} \ker f &= \left\{(a_1, a_2, \ldots, a_n) \middle| f(a_1, a_2, \ldots, a_n) = (e_1, e_2, \ldots, e_{i-1}, e_{i+1}, \ldots, e_n)\right\} \\ &= \left\{(a_1, a_2, \ldots, a_n) \in G \middle| a_j = e_j, \forall j \neq i\right\} \\ &= H_i \cong G_i \end{aligned}$$

Thus, by the fundamental theorem on homomorphism, we have

$$\frac{G}{\ker f} \cong f(G),$$

*i.e.* $$\frac{G}{G_i} \cong G_1 \times G_2 \times \ldots \times G_{i-1} \times G_{i+1} \times \ldots \times G_n$$

***(ii)*** Let $x \in H_i, y \in H_j$, then $\exists\ x_i \in G_i, y_j \in G_j$ such that

$$\begin{aligned} x &= (e_1, e_2, \ldots, e_{j-1}, x_i, e_{i+1}, \ldots, e_n) \text{ and} \\ y &= (e_1, e_2, \ldots, e_{j-1}, y_j, e_{j+1}, \ldots, e_n), \text{ then for } i < j \text{ and } i \neq j, \text{ we have} \\ xy &= (e_1, e_2, \ldots, e_{i-1}, x_i, e_{i+1}, \ldots, e_{j-1}, y_j, e_{j+1}, \ldots, e_n) \\ &= (e_1, e_2, \ldots, e_{j-1}, y_j, e_{j+1}, \ldots, e_n)\ (e_1, e_2, \ldots, e_{i-1}, x_i, e_{i+1}, \ldots, e_n) \\ &= yx \end{aligned}$$

**Theorem 4.** *Let $G_1$ and $G_2$ be two groups. Let $H_1$ and $H_2$ be normal subgroups of $G_1$ and $G_2$ respectively then*

*(i)* $H_1 \times H_2 \lhd G_1 \times G_2$

*(ii)* $\dfrac{G_1 \times G_2}{H_1 \times H_2} \cong \dfrac{G_1}{H_1} \times \dfrac{G_2}{H_2}$

**Proof :** ***(i)*** Since $H_1$ is a subgroup of $G_1$ and $H_2$ is a subgroup of $G_2$, therefore $H_1 \times H_2$ is a subgroup of $G_1 \times G_2$. Now, let $(h_1, h_2) \in H_1 \times H_2$ and $(g_1, g_2) \in G_1 \times G_2$, then

$$\begin{aligned}(g_1, g_2)(h_1, h_2)(g_1, g_2)^{-1} &= (g_1 h_1, g_2 h_2)\left(g_1^{-1}, g_2^{-1}\right) \\ &= \left(g_1 h_1 g_1^{-1}, g_2 h_2 g_2^{-1}\right) \in H_1 \times H_2\end{aligned}$$

as $g_1 h_1 g_1^{-1} \in H_1, \quad g_2 h_2 g_2^{-1} \in H_2 \qquad \left(\because H_1 \lhd G_1, H_2 \lhd G_2\right)$

Thus $H_1 \times H_2 \lhd G_1 \times G_2$.

***(ii)*** Let us define a mapping

$$f : \frac{G_1 \times G_2}{H_1 \times H_2} \to \frac{G_1}{H_1} \times \frac{G_2}{H_2}$$

Such that

$$f\left[(H_1 \times H_2)(g_1, g_2)\right] = (H_1 g_1, H_2 g_2), \text{ for } g_1 \in G_1, g_2 \in G_2$$

Now, for $(H_1 \times H_2)(x_1, x_2), (H_1 \times H_2)(y_1, y_2) \in \dfrac{G_1 \times G_2}{H_1 \times H_2}$

where $x_1, y_1 \in G_1$ and $x_2, y_2 \in G_2$, we have

$$\begin{aligned}f[(H_1 \times H_2)(x_1, x_2)(H_1 \times H_2)(y_1, y_2)] &= f[(H_1 \times H_2)(x_1, x_2)(y_1, y_2)] \\ &= f[(H_1 \times H_2)(x_1 x_2, y_1 y_2)] \\ &= [(H_1 x_1 y_1, H_2 x_2 y_2)] \qquad \text{(by definition of } f\text{)} \\ &= [(H_1 x_1 H_1 y_1, H_2 x_2 H_2 y_2)] \\ &= [(H_1 x_1, H_2 x_2)(H_1 y_1, H_2 y_2)] \\ &= f[(H_1 \times H_2)(x_1, x_2) f (H_1 \times H_2)(y_1, y_2)],\end{aligned}$$

thus $f$ is a homomorphism.

Now, let $f[(H_1 \times H_2)(x_1, x_2)] = f[(H_1 \times H_2)(y_1, y_2)]$

$\Rightarrow \quad (H_1 x_1, H_2 x_2) = (H_1 y_1, H_2 y_2)$

$\Rightarrow \quad \left(x_1 y_1^{-1}, x_2 y_2^{-1}\right) \in H_1 \times H_2$

$\Rightarrow \quad (x_1, x_2)(y_1, y_2)^{-1} \in H_1 \times H_2$

$\Rightarrow \quad (H_1 \times H_2)(x_1, x_2) = (H_1 \times H_2)(y_1, y_2)$

So $f$ is one-one.

Now, for each

$$(H_1 x_1, H_2 x_2) \in \frac{G_1}{H_1} \times \frac{G_2}{H_2},$$

$\exists x_1 \in G_1, x_2 \in G_2$, so that

$(x_1, x_2) \in G_1 \times G_2$ and $(H_1 \times H_2)(x_1, x_2) \in \dfrac{G_1 \times G_2}{H_1 \times H_2}$

such that $f[(H_1 \times H_2)(x_1, x_2)] = (H_1 x_1)(H_2 x_2)$

Thus $f$ is onto.

Hence $f$ is an isomorphism and

$$\frac{G_1 \times G_2}{H_1 \times H_2} \cong \frac{G_1}{H_1} \times \frac{G_2}{H_2}.$$

## 1.4 Internal direct product

**Definition :** A group $G$ is said to be an internal direct product of its subgroups $H$ and $K$ if

*(i)* For $h \in H, k \in K \Rightarrow hk = kh$.

*(ii)* $G = HK$ and every element of $G$ can be uniquely expressed as product of an element of $H$ by an element of $K$.

Generalization of the above definition is as follows :

A group $G$ is said to be an internal direct product of its subgroups $H_1, H_2, ...., H_n$ if,

*(i)* $a_i a_j = a_j a_i$, for $a_i \in H_i$, $a_j \in H_j$ and $i \neq j$.

*(ii)* Each $g \in G$ can be uniquely expressed as $g = h_1 h_2 .... h_n$, where $h_i \in H_i$ $(1 \leq i \leq n)$.

## 1.5 Theorems on internal direct product

**Theorem 1.** *Let $G$ be a group and let $H_1, H_2, ...., H_n$ be the subgroups of $G$. Then $G$ is an internal direct product of $H_1, H_2, ...., H_n$ if and only if the following conditions are satisfied :*

*(i)* $H_i \triangleleft G$ *for* $i = 1, 2, ...., n$

*(ii)* $H_i \cap \left( \prod_{j \neq i} H_j \right) = \{e\}$

*(iii)* $G = H_1 H_2 ... H_n$

**Proof :** ***(i)*** Let $H_1, H_2, ...., H_n$ be subgroups of $G$ satisfying given three conditions *(i)*, *(ii)* and *(iii)*. Let $l$ and $m$ be any two integers such that $1 \leq l < m \leq n$. Obviously $l \neq m$. Now,

$$H_m \cap \left( \prod_{j \neq m} H_j \right) = \{e\} \qquad \text{[by } (ii)\text{]} \qquad .....(1)$$

Now, since $l \neq m$, $H_l \subseteq \left( \prod_{j \neq m} H_j \right)$, therefore

$$H_m \cap H_l = \{e\} \qquad \text{[by } (i)\text{]} \qquad .....(2)$$

Let $a_m \in H_m$ and $a_l \in H_l$ , then

consider the element

$$a_m^{-1}a_l^{-1}a_m a_l = \left(a_m^{-1}a_l^{-1}a_m\right)a_l \in H_l$$

$$\left(\because H_l \lhd G, a_m^{-1}a_m^{-1}a_m \in H_l \text{ and } a_l \in H_l\right)$$

Similarly,

$$a_m^{-1}a_l^{-1}a_m a_l = a_m^{-1}\left(a_l^{-1}a_m a_l\right) \in H_m$$

Thus, $a_m^{-1}a_l^{-1}a_m a_l \in H_m \cap H_l$

so $a_m^{-1}a_l^{-1}a_m a_l = e$ [by (2)]

$\Rightarrow$ $(a_l a_m)^{-1} a_m a_l = e$

$\Rightarrow$ $a_m a_l = a_l a_m$ .....(3)

Now, given that $G = H_1 H_2 \ldots H_n$ [by *(iii)*], so for each $g \in G, \exists\, h_1 \in H_1, h_2 \in H_2, \ldots h_n \in H_n$, such that $x = h_1 h_2 \ldots h_n$. Now, we shall show the uniqueness of this product. Let, if possible $x = h_1 h_2 \ldots h_n = h_1' h_2' \ldots h_n'$, where $h_i' \in H_i\,(1 \le i \le n)$.

$$h_R\left(h_k'\right)^{-1} = \left(h_1' h_1^{-1}\right)\left(h_2' h_2^{-1}\right)\ldots\left(h_{k-1}' h_{k-1}^{-1}\right)\left(h_{k+1}' h_{k+1}^{-1}\right)\ldots\left(h_n' h_n^{-1}\right) \in H_k \cap \left(\prod_{j\neq k} H_j\right) \quad \text{[by (3)]}$$

By *(ii)*, $h_k\left(h_k'\right)^{-1} = e \Rightarrow h_k = h_k'$, $k = 1, 2, \ldots n$

Thus, $G$ is the internal direct product of $H_1, H_2, \ldots H_n$.

**Converse :** Let $G$ is the internal direct product of its subgroups $H_1, H_2, \ldots H_n$.

*(i)* Let $a_k \in H_k$ and let $g = h_1 h_2 \ldots h_n$, $h_i \in H_i\,(1 \le i \le n)$ and $g$ be any element of $G$. Since $G$ is an internal direct product, therefore $a_k h_i = h_i a_k \;\forall i \neq k$, so

$$g^{-1}a_k g = \left(h_1 h_2 \ldots h_n\right)^{-1} a_k \left(h_1 h_2 \ldots h_n\right) = h_k^{-1}a_k h_k \in H_k$$

thus, $H_k \lhd G, k = 1, 2, \ldots n.$

*(ii)* Let $a \in H_i \cap \left(\prod_{j\neq i} H_j\right)$

$\Rightarrow$ $a \in H_i$ and $a \in \prod_{j\neq i} H_j$

$\Rightarrow$ $a = a_i$ and $a = a_1 a_2 \ldots a_{i-1}\, a_{i+1} \ldots a_n$ for same $a_k \in H_k$, $k = 1, 2, \ldots n$.

$\Rightarrow$ $a_i = a_1 a_2 \ldots a_{i-1}\, a_{i+1} \ldots a_n$

$\Rightarrow$ $e\, e \ldots e\, a_i\, e \ldots e = a_1 a_2 \ldots a_{i-1}\, e a_{i+1} \ldots a_n$

$\Rightarrow$ $a_i = e$ (by the uniqueness of expression as product of members of $H_1, H_2 \ldots, H_n$)

$\Rightarrow$ $a = e$,

so $H_i \cap \left(\prod_{j\neq i} H_j\right) = \{e\}$.

*(iii)* Since $G$ is an internal direct product of $H_1, H_2, \ldots H_n$, therefore $G = H_1 H_2 \ldots H_n$ (by the definition)

**Theorem 2.** *Let $G_i$ and $G_2$ be two groups. Let $G = G_1 \times G_2$*

$$H_1 = \{(a, e_2) \mid a \in G_1\} = G_1 \times \{e_2\}$$

and $$H_2 = \{(e_1, b) \mid b \in G_2\} = \{e_1\} \times G_2$$

*then $G$ is an internal direct product of $H_1$ and $H_2$.*

**Proof :** By the definition of $H_1$ and $H_2$, it is obvious that $H_1$ and $H_2$ are subgroups of $G$.

Let $x = (a, e_2) \in H_1$ and $y = (e_1, b) \in H_2$, then

$$\begin{aligned} xy &= (a, e_2)(e_1, b) \\ &= (ae_1, e_2 b) \\ &= (a, b) \end{aligned}$$

and

$$\begin{aligned} yx &= (e_1, b)(a, e_2) \\ &= (e_1 a, be_2) \\ &= (a, b) \end{aligned}$$

Hence $xy = yx$ i.e. every element of $H_1$ commutes with every element of $H_2$. Now we shall show that each element of $G$ can be expressed uniquely as a product of an element of $H_1$ by an element of $H_2$. Let $g = (g_1, g_2)$ by any element of $G$ then

$$g = (g_1, g_2) = (g_1 e_1, e_2 g_2) = (g_1, e_2)(e_1, g_2) \qquad \ldots\ldots(1)$$

So $g \in G$ can be expressed as a product of an element of $H_1$ by an element of $H_2$.
Now, let if possible

$$\begin{aligned} (g_1, g_2) &= (g_1', e_2)(e_1, g_2') = (g_1' e_1, e_2 g_2') \\ &= (g_1', g_2') \end{aligned}$$

$\Rightarrow$ $$g_1 = g_1' \text{ and } g_2 = g_2'$$

which shows uniqueness of the expression.

Hence $G$ is an internal direct product of $H_1$ and $H_2$.

**Theorem 3.** *Let $G$ be a group. $H$ and $K$ are two subgroups of $G$ such that $H$ and $K$ are normal in $G$ and $H \cap K = \{e\}$, then*

*(i) $HK$ is the internal direct product of $H$ and $K$*

*(ii) $HK \cong H \times K$.*

**Prof :** Since $H \lhd G$ and $K \lhd G$, therefore $HK$ is a subgroup of $G$. Now let $h \in H$, $k \in K$, then

$$k^{-1}h^{-1}kh = k^{-1}(h^{-1}kh) \in K \qquad (k^{-1} \in K, h^{-1}kh \in K \text{ as } K \lhd G)$$

and $$k^{-1}h^{-1}kh = (k^{-1}h^{-1}k)h \in H \qquad (h \in H, k^{-1}h^{-1}k \in H \text{ as } H \lhd G)$$

thus $$k^{-1}h^{-1}kh \in H \cap K$$

so $k^{-1}\, h^{-1}\, kh = e$ $\quad [\because\ H \cap K = \{e\}]$

$\Rightarrow$ $(hk)^{-1}\,(kh) = e$

$\Rightarrow$ $kh = hk$

*i.e.* every element of $H$ commutes with every element of $K$. Now, let $x \in HK$, then $x = hk$ for some $h \in H,\ k \in K$.

Let, if possible $x = hk = h_1 k_1$

for $h_1 \in H,\ k_1 \in K$

$\Rightarrow$ $h_1^{-1} h = k_1 k^{-1}$ .....(1)

Since $h_1^{-1} \in H \Rightarrow k_1 k_1^{-1} \in H$ [by 1]

and since $k_1 k^{-1} \in K \Rightarrow h_1^{-1} h \in K$ [ by 1],

thus $h_1^{-1} h \in H \cap K$ and $k_1 k_1^{-1} \in H \cap K$

so $h_1^{-1} h = e$ and $k_1 k_1^{-1} = e$ which gives $h_1 = h$ and $k_1 = k$

Hence each element of $HK$ is unique product of an element of $H$ by an element of $K$. Thus $HK$ is the internal direct product of $H$ and $K$.

***(ii)*** Let us define a mapping

$f : HK \to H \times K$

such that $f(hk) = (h, k),\ h \in H,\ k \in K.$

Obviously $f$ is well defined, one-one and on to as representation $hk \in HK$ is unique.

Now, for $h_1 k_1,\ h_2 k_2 \in HK$, where $h_1, h_2 \in H$ and $k_1, k_2 \in K$, we have

$$\begin{aligned} f(h_1 k_1 h_2 k_2) &= f\left[h_1 (k_1 h_2) k_2\right] \\ &= f\left[h_1 (h_2 k_1) k_2\right] \\ &= f(h_1 h_2 k_1 k_2) \\ &= (h_1 h_2, k_1 k_2) \qquad \text{(by the definition of } f) \\ &= (h_1, k_1)(h_2, k_2) \\ &= f(h_1 k_1)\ f(h_2 k_2) \end{aligned}$$

so $f$ is homomorphism.

Hence f is an isomorphism and $HK \cong H \times K$

**Theorem 4.** *If HK is the internal direct product of H and K, then*

$$\frac{HK}{K} \cong H \quad \textit{and} \quad \frac{HK}{H} \cong K$$

**Proof :** Let us define a mapping

$f : HK \to H$

such that $f(hk) = h,\quad h \in H, k \in K.$

Obviously $\phi$ is well-defined.

For $h_1 k_1 \in HK, h_2 k_2 \in HK$, where $h_1, h_2 \in H, k_1, k_2 \in K$, we have

$$\begin{aligned} f(h_1 k_1\ h_2 k_2) &= f\left[h_1 (k_1 h_2) k_2\right] \\ &= f\left[h_1 (h_2 k_1) k_2\right] \\ &= f(h_1 h_2\ \ k_1 k_2) \\ &= h_1 h_2 \\ &= f(h_1 k_1) f(h_2 k_2) \end{aligned}$$

Thus $f$ is a homomorphism.

Now, for $h \in H, hk \in HK$, where $k \in K$, such that $f(hk) = h$, so $f$ is onto.

Thus $f$ is an epimorphism.

Now,
$$\begin{aligned} \ker f &= \{hk \in HK \mid f(hk) = e \in H\} \\ &= \{h \in H, k \in K \mid h = e\} \\ &= \{k \in K\} \\ &= K \end{aligned}$$

Thus by the fundamental theorem on homomorphism, we have

$$\frac{HK}{\ker f} \cong f(HK)$$

*i.e.*
$$\frac{HK}{K} \cong H.$$

Similarly, we can show that

$$\frac{HK}{H} \cong K.$$

***Ex.*1.** ***Let G be a cyclic subgroup of order 6 generated by a ∈ G i.e. G = [a].***

Let $\{e, a^2\ a^4\} = H$ (say) and $\{e, a^3\} = K$ (say). Obviously $H$ and $K$ are susgroups of $G$ of order 3 and 2 respectively. We observe that

*(i)* $HK = \{ee, ea^3, a^2 e, a^2 a^3, a^4 e, a^4 a^3\}$

$= \{e, a, a^2, a^3, a^4, a^5\}$ (since $a^6 = 1$)

$= G$

*(ii)* Since $G$ is cyclic therefore $G$ is abelian, hence $H$ and $K$ are normal subgroups of $G$.

*(iii)* $H \cap K = \{e\}$

Thus $G$ is the internal direct product of $H$ and $K$.

### Self-learning exercise-1

*(i)* Define external direct product.

*(ii)* Define internal direct product.

*(iii)* Which of the following statement is false.

*(a)* External direct product and internal direct product of same factors are isomorphic.

*(b)* If $G$ is an internal direct product of $H$ and $K$ then it is not necessary that $H$ and $K$ are normal subgroups of $G$.

(c) If $H$ and $K$ are two sub-groups of $G$ such that $G$ is an internal direct product of $H$ and $K$ then $H \cap K = \{e\}$.

*(iv)* If $o(H) = 2$ and $o(K) = 3$, then find $o(H \times K)$.

## 1.6 Summary

In this unit we studied about external direct product of groups and internal direct product and some theorems to understand their properties.

## 1.7 Answers to self learning exercises

### Self-learning exercise-1

**1.** See text **2.** See text

**3.** *(a)* True *(b)* False *(c)* True **4.** 6

## 1.8 Exercises

**1.** Let $G_1, G_2, G_3$ be groups, then show that

*(i)* $(G_1 \times G_2) \times G_3 \cong G_1 \times G_2 \times G_3 \cong G_1 \times (G_2 \times G_3)$

*(ii)* $o(G_1 \times G_2) = o(G_1) \times o(G_2)$

**2.** If $G$ is the internal direct product of its subgroup $H_1, H_2, \ldots H_n$, $(n > 1)$ and $H = H_2 H_3 \ldots H_n$, then show that $G$ is internal direct product of $H_1$ and $H$ also.

**3.** Let $G$ be a group and let $G$ be internal direct product of $H_1, H_2, \ldots H_n$. Let $M$ be external direct product of $H_1, H_2, \ldots H_n$. Then show that $G \cong M$.

**4.** If $H, K$ are subgroups of a group $G$ such that $G = H \times K$, show that $H \cong \frac{G}{K}$ and $K \cong \frac{G}{H}$

**5.** Show that $S_3$ can not be written as internal direct of two non-trivial subgroups.

**6.** If $H, K$ are normal subgroups of $G$, show that $\frac{G}{H \cap K}$ is isomorphic to a subgroup of $\frac{G}{H} \times \frac{G}{K}$.

□□□

# UNIT 2 : Isomorphism Theorems, Conjugacy and the Class equation of a Group

**Structure of the Unit**



## 2.0 Objectives

In this unit we shall study about

1. Group homomorphism, isomorphism and related important theorems.
2. Conjugate elements, conjugate class and class-equation of a finite group.

## 2.1 Introduction

We have already studied about the concept of homomorphism of a group at graduation level. Homomorphism is a special mapping which preseres the group operation and some group properties. Isomorphism is a special kind of homomorphism, which is one-one and onto also. There is structural similarity between two isomorphic groups.

A special relation called conjugacy is an equivalence relation on a group. Equivalence classes related with this relation are called conjugate classes. Class-equation of a finite group $G$ is a relation between order of a group $G$, conjugate classes and normalizer of elements of the group $G$.

## 2.2 Group homomorphism

**Definition :** Let $(G, *)$ and $(G', *')$ be two groups. A mapping $f$ from $G$ to $G'$ is said to be a group homomorphism if

$$f(a * b) = f(a) *' f(b), \quad \forall\, a, b \in G.$$

This mapping preserves the operations of $G$ and $G'$ although these operations are different. When there is no ambiguity, we shall write all abstract groups multiplicatively. Thus, the above condition can be rewritten as

$$f(ab) = f(a)\, f(b), \quad \forall\, a, b \in G.$$

**Kernel of homomorphism :** Let $f$ be a group homorphism of a group $G$ to group $G'$. The Kernel of homomorphism $f$ is the set defined as follows :

$$\text{Ker}\,(f) = \{x \in G \mid f(x) = e'\}$$

where $e'$ is the identity element of $G'$.

Thus Kernel of $f$ is the set of elements of $G$ which are mapped to identity element of $G'$.

**Image of homomorphism :** Let $f$ be a group homomorphism of a group $G$ to a group $G'$. Then image of $f$ is the set defined as follows :

$$Im\,(f) = f(G) = \{f(x) \in G' \mid x \in G\}$$

If $f$ is onto then $f(G) = G'$ and $G'$ is called homomorphic image of $G$. Also $f$ is **epimorphism.**

If homomorphism $f$ is one-one then it is called **monomorphism.**

If domain and codomain of $f$ are same, that is, if $G = G'$, then homomorphism is called **endomorphism.**

## 2.3 Isomorphism

**Definition :** A homorphism of a group $G$ to a group $G'$ is called isomorphism if it is bijective, that is, one-one and onto.

An isomorphism from $G$ to $G$ is called an **automorphism.**

The set of all homomorphism from a group $G$ to group $G'$ is represented by Hom $(G, G')$ and the set of all automorphism is denoted by Aut$(G)$.

If there exists an isomorphism between two groups $G$ and $G'$, then we say that $G$ and $G'$ are isomorphic groups and symbolically it is written as

$$G \cong G'.$$

***Ex.*1.** *Let $(Z, +)$ be the additive group of integers. Let a mapping*

$$f : Z \to Z$$

*defined as* $f(z) = mz, \quad \forall\, z \in Z,$

*where $m$ is any fined integer, such that $m \neq 0$. Show that $f$ is a monomorphism.*

**Sol.** For any $z_1, z_2 \in Z$, we have

$$f(z_1 + z_2) = m(z_1 + z_2) \qquad \text{(by the definition of } f\text{)}$$
$$= mz_1 + mz_2 \qquad \text{(distributive-law)}$$
$$= f(z_1) + f(z_2)$$

Thus $f$ is homomorphism.

Now, let $f(z_1) = f(z_2)$

$\Rightarrow$ $mz_1 = mz_2$

$\Rightarrow$ $z_1 = z_2 \qquad (\because \ m \neq 0)$

thus $f$ is one-one.

Hence $f$ is monomorphism.

***Ex.2.*** *Let $(R, +)$ be the additive group of real number and $(R^+, \cdot)$ be a multiplicative group of positive real numbers. Show that a mapping $\phi : R \to R^+$, defined by $\phi(x) = e^x$ is an isomorphism.*

**Sol.** Obviously $\phi$ is well-defined. For $x_1, x_2 \in R$, we have

$$\phi(x_1 + x_2) = e^{x_1 + x_2}$$
$$= e^{x_1} \, e^{x_2}$$
$$= \phi(x_1) \, \phi(x_2)$$

Thus $\phi$ is a homomorphism.

Now, let $\phi(x_1) = \phi(x_2) \Rightarrow e^{x_1} = e^{x_2}$

$$\Rightarrow x_1 = x_2,$$

thus $\phi$ is one-one.

Now, for every $x \in R^+$, $\exists \log_e x \in R$

such that $\phi(\log_e x) = e^{\log_e x} = x,$

thus $\phi$ is onto.

Consequently $\phi$ is an isomorphism.

***Ex.3.*** *Show that $\phi : R_0 \to R_0$, given by $\phi(x) = x^2$, $\forall \, x \in R_0$ is an homomorphism, where $R_0$ is a multiplicative group of non-zero real numbers. Also find Kernel of $\phi$.*

**Sol.** For $x_1, x_2 \in R_0$, we have

$$\phi(x_1 \, x_2) = (x_1 \, x_2)^2 = x_1^2 \, x_2^2$$
$$= \phi(x_1) \, \phi(x_2)$$

then $\phi$ is a homomorphism.

Kernel of $\phi$ = Ker($\phi$) = $\{x \in R_0 \mid \phi(x) = 1\}$ $\qquad$ {1 is identity in $R_0$}

$$= \{x \in R_0 \mid x^2 = 1\}$$
$$= \{1, -1\}$$

**Some important results of Homomorphism and Isomorphism :**

Here, we are writing some important results without proof because these have been studied in under graduate classes.

**Theorem 1.** *If $f: G \to G'$ is a homomorphism then*

*(i)* $f(e) = e'$, *where e and e′ are identities of G and G′ respectively.*

*(ii)* $f(x^{-1}) = [f(x)]^{-1}, \ \forall x \in G.$

*(iii)* $f(x^n) = [f(x)]^n, \ \forall x \in G$ *and* $\forall n \in Z.$

*(iv) Ker f is a normal subgroup of G.*

*(v) f is a monomorphism iff Ker* $f = \{e\}$

**Theorem 2.** *(Natural homomorphism) A group is homomorphic to its quotient group.*

**Theorem 3.** *(Fundamental theorem of homomorphism) Every homomorphic image of a group G is isomorphic to some quotient group of G.*

**Theorem 4.** *(Double quotient theorem) Let H and N be two normal subgroups of G such that H C N, then*

$$N/H \lhd G/H \quad \text{and} \quad G/H/N/H \cong G/N$$

**Theorem 5.** *(The diamond isomorphism theorem) Let H and N be two subgroups of G such that N is normal in G. Then* $H \cap N$ *is a normal subgroup of H and*

$$\frac{H}{H \cap N} \cong \frac{HN}{N}$$

**Proof :** Since $N \lhd G$, therefore

$$Nx = xN, \ \forall x \in G$$

$$\Rightarrow \quad Nx = xN, \ \forall x \in H \text{ as }, H \subset G$$

$$\Rightarrow \quad NH = HN$$

thus $HN$ is a subgroup of $G$.

Now, we shall show that $N \lhd HN$.

Let $n \in N$, then $\quad en \in HN, \quad (\because \ e \in H)$

$$\Rightarrow \quad n \in HN$$

thus $N \subset HN$ and since $N \lhd G$, therefore $N \lhd HN$ also. Hence the quotient group $\frac{HN}{N}$ exists.

Let as consider a mapping

$$\phi : H \to \frac{HN}{N}$$

defined by $\quad \phi(h) : Nh, \quad \forall x \in H$

$$\because \quad h \in H \Rightarrow he = h \in HN \Rightarrow Nh \in \frac{HN}{N}$$

Let $h_1, h_2 \in H$, then

$$\phi(h_1h_2) = Nh_1h_2 = Nh_1Nh_2 = \phi(h_1)\,\phi(h_2)$$

thus $\phi$ is a homomorphism.

For each $Nx \in \frac{HN}{N}$, $x \in HN$, then $\exists\ h \in H$ and $n \in N$ such that $x = hn$. Thus for each $Nx$,

$$\begin{aligned} Nx &= N(hn) = (Nh)\,n = (hN)\,n \qquad (\because\ N \lhd G,\ Na = aN\ \forall\ a \in G) \\ &= h(Nn) = hN \\ &= Nh = \phi(h) \end{aligned}$$

Thus $\phi$ is onto, so

$$\phi(H) = \frac{HN}{N}$$

Hence $\phi$ is an epimorphism.

Now, $\quad \text{Kernel of } \phi = \text{Ker}(\phi) = \{x \in H \mid \phi(x) = n\} \qquad (N \text{ is identity in } \frac{HN}{N})$

$$\begin{aligned} &= \{x \in H \mid Nx = N\} \\ &= \{x \in H \mid x \in N\} \\ &= H \cap N \end{aligned}$$

then by the fundamental theorem, we have

$$\frac{H}{Ker(\phi)} \cong \phi(H)$$

or

$$\frac{H}{H \cap N} \cong \frac{HN}{N}$$

this proves result *(ii)*.

Since $H \cap N$ is Kernel of $\phi$, therefore $H \cap N$ is a normal subgroup of $H$. This proves result *(i)*

**Theorem 6.** (Zassenhaus Lemma or Butterfly theorem) Let $H$ and $N$ be two subgroups of $G$ and let $H'$ and $N'$ be two normal subgroups of $H$ and $N$, respectively. Then

*(i)* $(H \cap N')H'$ is normal subgroup of $(H \cap N)H'$,

*(ii)* $(H' \cap N)N'$ is normal subgroup of $(H \cap N)N'$,

*(iii)* $\dfrac{(H \cap N)H'}{(H \cap N')H'} \cong \dfrac{(H \cap N)N'}{(H' \cap N)N'}$

**Proof :** Since intersection of any two subgroups of $G$ is again a subgroup of $G$, therefore $(H \cap N)$ is a subgroup of $G$. Also, since $(H \cap N) \subset N$, therefore $H \cap N$ is a subgroup of $N$. Moreover,

$$(H \cap N) \lhd (H \cap N) \quad \text{and} \quad N' \lhd N$$

$$\Rightarrow \qquad (H \cap N) \cap N' \lhd (H \cap N) \cap N$$

$\Rightarrow$ $(H \cap N') \lhd (H \cap N)$

Similarly, $(H \cap N') \lhd (H \cap N)$

Since product of two normal subgroups is again a subgroup, therefore

$$(H \cap N')\,(H' \cap N) = A \qquad \text{(say)},$$

is a normal subgroup of $(H \cap N)$. So we can form quotient group $(H \cap N)/A$.

Also since $H \cap N \subset H$, therefore $H \cap N$ is a subgroup of $H$ and $H'$ is a normal subgroup of $H$, so that $(H \cap N)H'$ is a subgroup of $H$.

Now consider a mapping

$$f : (H \cap N)H' \to (H \cap N)/A$$

defined by $f(xy) = Ax,\ x \in H \cap N,\ y \in H'$

Let $x_1, x_2 \in H \cap N,\ y_1, y_2 \in \cap H'$ such that

$$x_1 y_1 = x_2 y_2$$

$\Rightarrow$ $x_2^{-1} x_1 = y_2 y_1^{-1} \in (H \cap N) \cap H' = H' \cap N \subset A$

($\because\ x_1, x_2 \in H \cap N \Rightarrow x_2^{-1} x_1 \in H \cap N \Rightarrow y_2 y_1^{-1} \in H \cap N$ similarly

$y_2 y_1^{-1} \in H' \Rightarrow x_2^{-1} x_1 \in H'$)

thus $x_2^{-1} x_1 \in A$

$\Rightarrow$ $Ax_1 = Ax_2$

$\Rightarrow$ $f(x_1 y_1) = f(x_2 y_2)$

thus $f$ is well-defined

For $x_1, x_2 \in (H \cap N),\ y_1, y_2 \in H'$, we have

$$f[(x_1 y_1)(x_2 y_2)] = f\left[x_1 x_2 \left(x_2^{-1} y_1 x_2\right) y_2\right]$$

$$= f\left[x_1 x_2 \left(y_1' y_2\right)\right],\ y_1' = x_2^{-1} y_1 x_2 \quad \text{since } H' \lhd H$$

$$= Ax_1 x_2$$

$$= Ax_1 Ax_2$$

$$= f(x_1 y_1) f(x_2 y_2)$$

thus $f$ is homomorphism.

Now, for each $Ax \in (H \cap N)/L,\ x \in H \cap N$ and we can take any element $y \in H'$ such that

$$f(xy) = Ax$$

so that $f$ is onto.

Thus $f$ is an epimorphism.

$$\text{Ker}\,(f) = \{xy \in (H \cap N)H' \mid f(xy) = L\}$$

$$= \{xy \in (H \cap N)H' \mid Lx = L\}$$

$$= \{xy \in (H \cap N)H' \mid x \in L\}$$

$$= \{xy \in (H \cap N)H' \mid x = ab,\ a \in H \cap N',\ b \in H' \cap N\}$$
$$= \{(ab)y \in (H \cap N)H' \mid a \in (H \cap N'),\ b \in H',\ b \in N\}$$
$$= \{a(by) \in (H \cap N)H' \mid a \in H \cap N',\ b \in H'\}$$
$$= \{ay_1 \mid a \in H \cap N',\ by = y_1 \in H'\}$$
$$(\because\ H \cap N' \subset H \cap N,\ y \in H',\ b \in H \ \Rightarrow\ by \in H')$$
$$= (H \cap N')H'$$

Thus $\text{Ker}(f) = (H \cap N')H'$ and hence it is normal subgroup of $(H \cap N)H'$. Also, by the fundamental homomorphism theorem, we have

$$\frac{(H \cap N)H'}{Ker(f)} \cong \text{Homorphic Image of } f$$

that is
$$\frac{(H \cap N)H'}{(H \cap N')H'} \cong \frac{(H \cap N)}{L} \qquad \text{.....(1)}$$

Similarly we can show that

$$\frac{(H \cap N)N'}{(H' \cap N)N'} \cong \frac{(H \cap N)}{L} \qquad \text{.....(2)}$$

By (1) and (2), we get

$$\frac{(H \cap N)H'}{(H \cap N')H'} \cong \frac{(H \cap N)N'}{(H' \cap N)N'}$$

This completes the proof of the theorem.

### Self-learning exercise-1

**1.** Write whether the following statements are true or false :

***(i)*** Homomorphism preserves the group operation.

***(ii)*** A homomorphism is said to be an isomorphism if it is onto only.

***(iii)*** In an epimorphism, function is onto.

***(iv)*** Isomorphism is an equivalence relation.

**2.** Let $R_0$ be the multiplicative group of non-zero real numbers and let $f : R_0 \to R_0$ defined by $f(x) = x^4,\ \forall\ x \in R_0$, be a homomorphism of $R_0$, then find Kernel of $f$.

## 2.4 Conjugate element and conjugate set

Let $G$ be a group and $a \in G$. An element $b \in G$ is said to be conjugate to $a$ if there exists an element $g$ in $G$ such that $b = g\,ag^{-1}$ and it is denoted by $b \sim a$.

From the definition it is clear that if $b \sim a$ then $a \sim b$ as $b = gag^{-1} \Rightarrow a = g^{-1}bg$. Thus $a$ and $b$ are said to be conjugate elements and this relation is called conjugacy relation.

In a similar manner we can define conjugate sets. Two subsets $A$ and $B$ of a group $G$ are said to be conjugate to each other if there exists an element $g$ in $G$ such that

$$B = gAg^{-1} \quad \text{or} \quad A = g^{-1}Bg$$

**Theorem 7.** *Conjugacy on a group G is an equivalence relation.*

**Proof :** We shall show that the conjugacy relation is reflexive, symmetric and transitive. Let $a, b, c \in G$.

***(i)* Reflexive :** Let $a \in G$. Since identity element $e \in G$,

therefore $a = eae^{-1}$, so $a \sim a$, $\forall\, a \in G$,

thus, this relation is reflexive.

***(ii)* Symmetric :** Let $a, b \in G$ such that $a \sim b$, then $\exists\ x \in G$ such that

$$a = xbx^{-1}$$

$$\Rightarrow \quad x^{-1}ax = x^{-1}(xbx^{-1})\ x$$

$$\Rightarrow \quad x^{-1}ax = (x^{-1}x)\ b(x^{-1}x)$$

$$\Rightarrow \quad x^{-1}ax = ebe, \quad e \text{ is identity in } G$$

$$\Rightarrow \quad b = x^{-1}\ ax = x^{-1}a(x^{-1})^{-1}, \text{ for } x^{-1} \in G$$

which shows that $b \sim a$. Hence, this relation is symmetric

***(iii)* Transitive :** Let $a \sim b$ and $b \sim c$, then there exist $x$ and $y$ in $G$ such that

$$a = xbx^{-1} \text{ and } b = ycy^{-1}$$

$$\Rightarrow \quad a = x(ycy^{-1})\ x^{-1}$$

$$\Rightarrow \quad a = (xy)\ c(y^{-1}x^{-1})$$

$$\Rightarrow \quad a = (xy)\ c(xy)^{-1}$$

$$\Rightarrow \quad a \sim c \qquad (\because\ xy \in G)$$

Thus, this relation is transitive. Consequently it is an equivalence relation.

## 2.5 Conjugate class

**Definition :** Let $G$ be a group and $a \in G$. Then set of all elements of $G$ that are conjugate to a is called the conjugate class of a and denoted by $c[a]$. Thus

$$\begin{aligned} c[a] &= \{x \in G \mid x \sim a\} \\ &= \{x \in G \mid x = yay^{-1}, y \in G\} \\ &= \{yay^{-1} \mid y \in G\} \end{aligned}$$

**Note :**

1. Conjugate class of any element a in $G$ is always non-empty. Since $e \in G$ and $a = eae^{-1}$, thus $a \sim a$ and hence $a \in G\ [a]$.
2. If $G$ is an abelian group, then

$$gag^{-1} = gg^{-1}a = ea = a, \quad \forall\, g \in G$$

thus a has only one conjugate element that is a itself. Thus

$$C[a] = \{a\}$$

3. Since an equivalence relation defined on a set partitions the set, therefore, conjugacy relation defined on $G$ decomposes $G$ into mutually disjoint equivalence classes. Thus

$$G = \bigcup_{a \in G} C[a]$$

**Theorem 8.** *Any two conjugate classes of a group are either disjoint or identical.*

**Proof :** Let $G$ be a group and $a, b \in G$.

Let $C[a]$ and $C[b]$ be conjugate classes of $a$ and $b$ respectively. Now, two cases arise,

**Case *(i)* :** $C[a] \cap C[b] = \phi$

**Case *(ii)* :** $C[a] \cap C[b] \neq \phi$

If case *(i)* is true, then both the classes are disjoint and there is nothing to prove.

Let us suppose that case *(ii)* is true, that is, at least one element is common in $C[a]$ and $C[b]$. Let it be $x$. We shall show that both classes are identical, that is, $C[a] = C[b]$, therefore, $x \sim a$ and $x \sim b$, so there exist $g, h \in G$ such that

$$x = gag^{-1} \quad \text{and} \quad x = hbh^{-1}$$

So, $gag^{-1} = hbh^{-1}$

$\Rightarrow$ $g^{-1}(gag^{-1})g = g^{-1}(hbh^{-1})\, g$

$\Rightarrow$ $(g^{-1}g)\, a\, (g^{-1}g) = (g^{-1}h)\, b(h^{-1}g)$

$\Rightarrow$ $eae = (g^{-1}h)\, b(g^{-1}h)^{-1}$

$\Rightarrow$ $a = yby^{-1}$, where $y = g^{-1}h \in G$

$\Rightarrow$ $a \sim b$ .....(3)

Now, let $r \in C[a]$, then $r \sim a$ and by (3) $a \sim b$, gives $r \sim b$.

Thus $r \in C[b]$ and hence $C[a] \subset C[b]$.

Similarly we can obtain $C[a] \subset C[b]$.

Consequently, $C[a] = C[b]$.

## 2.6 Normalizer of an element in a group

**Definition :** Let $G$ be a group and $a \in G$. Then the normalizer of $a$ is a set consists of those elements of $G$ which commute with $a$. It is denoted by $N(a)$. Thus

$$N(a) = \{x \in G \mid ax = xa\}.$$

It is easy to verify that $N(a)$ is a subgroup of $G$.

## 2.7 Centre of a group

**Definition :** Let $G$ be a group. The centre of $G$, denoted by $Z(G)$ is a set consists of those elements of $G$, which commute with every element of $G$, that is,

$$Z(G) = \{x \in G \mid xg = gx, \quad \forall g \in G\}.$$

**Note : 1.** $Z\,G)$ is a normal subgroup of $G$.

**2.** $G$ is abelian iff $G = Z(G)$.

**3.** By the definitions of $N(a)$ and $Z(G)$, it is obvious that $Z(G) \subset N(a)$

**4.** $a \in Z(G) \Leftrightarrow \quad ax = xa, \quad \forall x \in G$

$\Leftrightarrow \quad a = xax^{-1}, \quad \forall x \in G$

$\Leftrightarrow \quad C[a] = \{a\}$

## 2.8 Class-equation for the finite group

We shall obtain the class-equation for the finite group $G$ with the help of following theorems.

**Theorem 9.** *Let $G$ be a finite group and $b \in G$.*

*Then* $$o\left(C[a]\right) = \frac{o(G)}{o\left[N(a)\right]} = \left[G : N(a)\right],$$

*that is, the number of elements conjugate to 'a' in G is equal to the index of the normalizer of a in G.*

**Proof :** Let $G$ be a group and $N(a)$ be the normalizer of $a \in G$, then

$$N(a) = \{x \in G \mid ax = ax\}$$

Also, let $C(a)$ be the conjugate class of a, then

$$C(a) = \{yay^{-1} \mid y \in G\}$$

We know that the index of $N(a)$ in $G$ is the number of distinct cosets of $N(a)$ in $G$, that is, if $A = \{xN(a) \mid x \in G\}$, the set of all distinct cosets of $N(a)$ in $G$, then

$$o(A) = \left[G : N(a)\right] = \frac{o(G)}{o\left[N(a)\right]} \qquad \text{.....(1)}$$

Now, we shall show that

$$o(C[a]) = o(A)$$

Let us consider a mapping

$$\phi : A \to C[a],$$

defined by $\quad \phi[xN(a)] = xax^{-1}, \qquad \forall x \in G$

Let $x, y \in G$, $xN(a), yN(a) \in A$, such that

$$xN(a) = yN(a)$$

$\Rightarrow \quad y^{-1}x \in N(a)$

$\Rightarrow \quad (y^{-1}x)a = a(y^{-1}x)$ [by the definition of $N(a)$]

$\Rightarrow \quad (yy^{-1})(xax^{-1}) = (yay^{-1})(xx^{-1})$

$\Rightarrow \quad e(xax^{-1}) = (yay^{-1})e$ ($e$ is identity in $G$)

$$\Rightarrow \quad xax^{-1} = yay^{-1}$$

$$\Rightarrow \quad \phi\,[xN(a)] = \phi\,[yN(a)],$$

thus $\phi$ is well defined.

Again, let

$$\phi[xN(a)] = \phi[yN(a)],$$

$$\Rightarrow \quad xax^{-1} = yay^{-1}$$

$$\Rightarrow \quad y^{-1}\,(xax^{-1})x = y^{-1}(yay^{-1})\,x$$

$$\Rightarrow \quad (y^{-1}x)\;a\;(x^{-1}x) = (y^{-1}y)\;a(y^{-1}x)$$

$$\Rightarrow \quad (y^{-1}x)ae = ea\;(y^{-1}x)$$

$$\Rightarrow \quad (y^{-1}x)a = a\;(y^{-1}x)$$

$$\Rightarrow \quad y^{-1}x \in N\,(a)$$

$$\Rightarrow \quad xN\,(a) = yN\,(a),$$

thus $\phi$ is one-one.

Now, for each $z \in C\,[a]$, $z = xax^{-1}$, for same $x \in G$ and then $xN\,(a) \in A$ such that

$$\phi\,(xN\,(a)) = xax^{-1} = z,$$

thus $\phi$ is onto.

Consequently $\phi$ is bijection. Hence

$$o\,(A) = o\,(C\,[a]) \qquad \text{.....(2)}$$

From (1) and (2), we have

$$o\,(C\,[a]) = \left[G : N\left(a\right)\right] = \frac{o\left(G\right)}{o\left[N\left(a\right)\right]}$$

$$= \text{index of } N\,(a) \text{ in } G.$$

**The class-equation :**

**Theorem 10.** *Let G be a finite group, then*

$$o\left(G\right) = \sum_{a \in D} \frac{o\left(G\right)}{o\left[N\left(a\right)\right]} = \sum_{i=1}^{n} \frac{o\left(G\right)}{o\left[N\left(a_i\right)\right]}$$

*where D be a set of distinct elements* $a_1, a_2, \ldots, a_n$ *taken one from each of the conjugate classes of G.*

**Proof :** Since $G$ is finite, therefore the number of distinct conjugate classes of $G$ will be finite, say $n$, that is, $o(D) = n$, where $D = \{a_1, a_2, \ldots, a_n\}$. Thus $C[a_1], C[a_2], \ldots, C[a_n]$ are distinct conjugate classes, which partition the group $G$.

So that $$G = \bigcup_{a \in D} C[a] = \bigcup_{i=1}^{n} C[a_i]$$

$$\Rightarrow \quad o(G) = \sum_{i=1}^{n} \left(C[a_i]\right)$$

$$= \sum_{i=1}^{n} \frac{o(G)}{o\left[N(a_i)\right]} \qquad \text{(by the theorem 15)}$$

$$= \sum_{a \in D} \frac{o(G)}{o\left(N(a)\right)} \qquad \text{.....(1)}$$

We know that, $x \in Z(G) \Leftrightarrow C[x] = \{x\}$. Thus number of singleton conjugate classes in $G$ is same as number of elements in $Z(G)$. Therefore equation (1) can be written as

$$o(G) = o\left[Z(G)\right] + \sum_{\substack{a \in D \\ a \notin Z(G)}} \frac{o(G)}{o\left[N(a)\right]} \qquad \text{.....(2)}$$

Here $Z(G)$ is the center of $G$. Equation (2) is known as class-equation for finite group $G$.

### Self-learning exercise-2

**1.** State whether the following statements are true or false :

***(i)*** If $a$ is conjugate to $b$, then $b$ is conjugate to $a$.

***(ii)*** If $a \sim b$ and $b \sim c$, then it is not necessary that $a \sim c$.

***(iii)*** The centre of a group is abelian.

***(iv)*** Normalizer of an element of $a$ group $G$ is always $a$ normal subgroup of $G$.

***(v)*** If $x \in Z(G)$, then $C[x] = \{x\}$

**2.** Find the partition of $G$ obtained from the relation conjugacy on $G$.

**3.** Write class-equation for the finite group $G$.

## 2.9 Summary

In this unit we studied homomorphism and isomorphism of the groups, and their important results. We discussed conjugacy relation, conjugate elements and conjugate class of an element of a group. We observed that conjugacy is an equivalence relation, which partitions the group. We also obtained the class equation of a finite group and discussed.

## 2.10 Answers to self-learning exercises

### Self-learning exercise-1

**1.** ***(i)*** true ***(ii)*** false ***(iii)*** true ***(iv)*** true

**2.** Ker $(f) = \{1, -1\}$

## Self-learning exercise-2

**1.** ***(i)*** true ***(ii)*** false ***(iii)*** true ***(iv)*** false ***(v)*** true

**2.** Set of all distinct conjugate classes is partition of $G$.

**3.** $o(G) = o[Z(G)] + \sum_{\substack{a \in D \\ a \notin Z(G)}} \frac{o(G)}{o[N(a)]}$

where symbols have their usual meaning.

## 2.11 Exercises

**1.** Show that $f : Z \to G$, defined by

$$f(x) = \begin{cases} 1, \text{ if } x \text{ is even} \\ -1, \text{ if } x \text{ is odd} \end{cases}$$

is an epimorphism, where $Z$ is additive group of integers and $G = \{1, -1\}$ is a multiplicative group.

**2.** Let $S_n$ be the symmetric group of order $n$ and $G = \{1, -1\}$ be a multiplicative group. Then show that the mapping $S_n \to G$, defined by

$$f(x) = \begin{cases} 1, \text{ if } x \text{ is even permutation} \\ -1, \text{ if } x \text{ is odd permutation} \end{cases}$$

is an epimorphism and find its Kernel also. **[Ans.** Ker($f$) = $A_n$, alternating group]

**3.** If $f$ is a homomorphism of $a$ group $G$ onto a group $G'$ and $g$ is a homomorphism of $G'$ onto $G''$, then show that $g \, o \, f$ is a homomorphism of $G$ onto $G''$.

**4.** Let $f$ be a homomorphism of $a$ group $G$ onto $G'$. Let $x \in G$ such that $f(x) = x' \in G'$. Then show that $Kx = f^{-1}(x')$, where $K$ = Kernel of $f$.

**5.** Show that subgroup $N$ of $a$ group $G$ is normal if and only if it is the Kernel of same homomorphism.

**6.** List all the conjugate classes in the symmetric group $S_3$ and verify the class equation

**[Ans.** $C[e] = \{e\}$, $C[(123)] = \{(123), (132)\}$, $C[(23)] = \{(23), (12), (13)\}$]

**7.** If in $a$ finite group $G$ an element $a$ has exactly two conjugates, prove that $G$ has $a$ normal subgroup $N$, other than $G$ and $\{e\}$.

**8.** If $x \in G$, where $G$ is a finite group, then show that the number of elements in $C[a]$ is a divisor of the order of $G$.

□ □ □

# UNIT 3 : Commutators, Derived subgroups, Solvable Groups and Composition Series

**Structure of the Unit**



## 3.0 Objectives

In this unit we shall study about Commutators, derived subgroup and a special class of groups called solvable groups. This unit also introduces subnormal series, composition series and a very important theorem called Jordan Holder Theorem.

## 3.1 Introduction

Commutator of any two elements $x$ and $y$ of a group $G$ is an element of the group which can be expressed in terms of $x$, $y$ and their inverses. A group generated by commutators is called derived subgroup of $G$. Study of solvable group is needed in the theory of polynomial equations. Series of subgroups of a group $G$, following some conditions, called subnormal series and composition series are closely related to solvable groups.

## 3.2 Commutator

**Definition :** Let $x$ and $y$ be any two elements of a group $G$. The element $x^{-1} y^{-1} xy$, denoted by $[x, y]$, is called commutator of $x$ and $y$, *i.e.*

$$[x, y] = x^{-1} y^{-1} xy$$

**Note :**

1. Commutator of $x$ and $x$ is $e$, the identity of $G$, that is,

$$[x, x] = x^{-1} x^{-1} x x = e$$

2. Commutator of $y$ and $x$ is inverse of commutator of $x$ and $y$, that is,

$$[x, y]^{-1} = (x^{-1}y^{-1}xy)^{-1} = y^{-1}x^{-1} yx = [y, x]$$

3. The product of two commutators need not be a commutator.

4. The set $C = \{[x, y] \mid x, y \in G\}$ of all commutators in $G$ may or may not be a subgroup of $G$.

5. Commutator of $x$ and $y$ can also be defined as $xy\, x^{-1} y^{-1}$

## 3.3 Derived subgroup

**Definition :** Let $G$ be a group. The subgroup generated by the set $C$ of all commutators of elements of $G$ is said to be **derived subgroup** of $G$. It is denoted by $G'$ or $G^{(1)}$ or $[G, G]$. The subgroup $G^{(1)}$ is the first derived subgroup of $G$. We can define higher derived subgroups of $G$. A subgroup generated by the set of all commutators of elements of $G^{(1)}$ is called first derived subgroup of $G^{(1)}$ and second derived subgroup of $G$ and denoted by $(G^{(1)})'$ or $G^{(2)}$. Similarly, the $n^{\text{th}}$ derived subgroup of $G$ will be $G^{(n)} = (G^{(n-1)})'$

The derived subgroups are also known as commutator subgroups. The series $G = G^{(0)} \supset G^{(1)} \supset G^{(2)} \supset \ldots\ldots \supset G^{(n)} \ldots\ldots$ is called **derived or commutator series** of group $G$.

**Theorem 1.** *Let $G^{(1)}$ be the first derived subgroup of the group $G$, then $G^{(1)} \triangleleft G$ and quotient group $G/G^{(1)}$ is abelian.*

**Proof :** Let $a \in G^{(1)}$ and $g \in G$, then

$$g^{-1}ag = (aa^{-1})\,(g^{-1}ag)$$
$$= a\,(a^{-1}\, g^{-1}ag) \in G^{(1)}$$

$(\because\ a \in G^{(1)}, a^{-1}\, g^{-1}\, ag \in G^{(1)}$, being the commutator of $a$ and $g)$

Thu $\quad a \in G^{(1)}, g \in G \Rightarrow\ g^{-1}\, ag \in G^{(1)}$,

so $G^{(1)}$ is a normal subgroup of $G$. Hence quotient group $G/G^{(1)}$ exists.

Let $x, y \in G$, then $G^{(1)}\, x, G^{(1)}\, y \in G/G^{(1)}$ and $x^{-1}\, y^{-1}\, xy \in G^{(1)}$

$\Rightarrow \quad (yx)^{-1}\, xy \in G^{(1)}$

$\Rightarrow \quad xyG^{(1)} = (yx)\, G^{(1)}$

$\Rightarrow \quad G^{(1)}\, xy = G^{(1)} yx \qquad [\because\ G^{(1)} \triangleleft G$, left cost = right cost$]$

$\Rightarrow$ $G^{(1)}x\, G^{(1)}y = G^{(1)}y\, G^{(1)}x$

Thus $G/G^{(1)}$ is abelian.

**Note :** In general, $G^{(n)} \lhd G^{(n-1)}, G^{(n)} \lhd G$ and $G^{(n-1)}/G^{(n)}$ is abelian.

**Theorem 2.** *Let $G$ be a group. Then it is abelian iff $G^{(1)} = \{e\}$, $e$ being the identity elment in $G$.*

**Proof :** First suppose that $G$ is abelian. Let $x, y \in G$, then the commutator of $x$ and $y$ is

$$x^{-1}y^{-1}\, xy = x^{-1}\, y^{-1}\, yx$$
$$= x^{-1}ex = e$$

Thus, the set of commutators consists of single element $e$, so $G^{(1)} = \{e\}$.

Conversely, let $G^{(1)} = \{e\}$. Let $x,y \in G$, then being the commutator of $x$ and $y$

$x^{-1}y^{-1}xy \in G^{(1)}$

$\Rightarrow$ $x^{-1}y^{-1}xy = e$

$\Rightarrow$ $(yx)^{-1}\, xy = e$

$\Rightarrow$ $xy = yx,$

Hence $G$ is abelian.

**Theorem 3.** *Let $G$ be a group and $H$ be a subgroup of $G$. Then $H \lhd G$ and $G/H$ is abelian iff $[G, G] \subset H$.*

**Proof :** Let $[G,G] = G'$ be the derived subgroup of $G$ and $H$ be a subgroup of $G$.

First suppose that $H$ is normal in $G$ and quotient group $G/H$ is abelian.

Let $x,y \in G$, then $xH, yH \in G/H$,

then $(xH)\,(yH) = (yH)\,(xH)$

$\Rightarrow$ $xyH = yxH$

$\Rightarrow$ $(yx)^{-1}\,(xy) \in \mathrm{H}$

$\Rightarrow$ $x^{-1}y^{-1}\, xy \in H.$

So that $C \subset H$

Where $C$ is the set of all commutators. Since $G'$ is generated by $C$, therefore $G'$ is the smallest subgroup containing $C$.

Hence $G' \subset H$

**Conversely,** Let $G' \subset H$, then $C \subset H$. Let

$g \in G, h \in H$, then

$g^{-1}hg = (hh^{-1})(g^{-1}hg)$

$= h\,(h^{-1}g^{-1}hg) \in H$

$(\because \quad h \in H,\ C \subset H)$

thus, $g \in G, h \in H \Rightarrow g^{-1}hg \in H$, Hence $H \lhd G$.

Again let $x,y \in G$, then $xH, yH \in G/H$. and

$x^{-1} y^{-1} xy \in H$ $(\because\ C \subset H)$

$\Rightarrow$ $(yx)^{-1} xy \in H$

$\Rightarrow$ $xyH = yxH$

$\Rightarrow$ $(xH)\ (yH) = (yH)\ (xH)$

Thus $G/H$ is abelian.

## 3.4 Subnormal series

**Definition :** Let $G$ be a group. A finite series of subgroups $H_i \left(l \le i \le n\right)$ of the group $G$, written as $G = H_0 \supset H_1 \supset H_2 \supset ..... \supset H_n = \{e\}$ is said to be a subnormal series if $H_{i+1} \lhd H_i$, for i = 0, 1, 2, .... ($n$–1).

Since $H_{i+1} \lhd H_i$, therefore we can form $n$ quotient groups $H_i/H_{i+1}$, i = 0, 1, 2,...,($n$–1).

These quotient groups are called factors of the subnormal series. The number $n$ is called length of this series.

If each subgroup $H_i$ is also normal in $G$, then the above series is called normal series.

A group $G$ always has a subnormal series

$$G \supset \{e\}$$

which is a normal series also. If $G$ is simple then it is the only subnormal (normal) series of $G$.

## 3.5 Solvable group

**Definition :** A group $G$ is said to be solvable if it has a subnormal series as

$$G = H_0 \supset H_1 \supset H_2 \supset ..... \supset H_n = \{e\}$$

such that each of its factors $H_i/H_{i+1}$ is an abelian group. This series, then is referred to as a solvable series for group $G$.

***Ex.1. Show that every finite abelian group is solvable :***

**Sol.** Let $G$ be a finite abelian group.

Let $G = H_0$ and $H_1 = \{e\}$, then

$H_1 \lhd H_0$ and since $G$ is abelian therefore its quotient group is abelian.

Then $G$ has a subnormal series

$$G = H_0 \supset H_1 = \{e\}$$

such that its factor $G/\{e\}$ or $H_0/H_1$ is abelian. Hence $G$ is solvable.

***Ex.2. Show that symmetric group $s_4$ is solvable.***

**Sol.** We know that alternating group $A_4$ is a normal subgroup of $S_4$.

Let $V_4 = \{(1), (12)(34), (13)(24), (14)(23)\}$, then it is easy to verify that $V_4 \lhd A_4$.

Thus, we have

$$S_4 \supset A_4 \supset V_4 \supset \{(1)\}$$

obviously it is a subnormal series for $S_4$.

Factors of this series $S_4/A_4$, $A_4/V_4$, $V_4/\{1\}$, are abelian because

$$o\left(\frac{S_4}{A_4}\right) = 2,\ o\left(\frac{A_4}{V_4}\right) = 3 \text{ and } o\left(\frac{V_4}{\{1\}}\right) = 4$$

**Note :** (1). If $o(G)$ = prime or (prime)$^2$, then $G$ is abelian

(2). We can easily show that symmetric groups $S_2$ and $S_3$ are also solvable.

***Ex.3. Show that $S_n$ is non solvable for $n \geq 5$.***

**Sol.** Alternating group $A_n$ is simple for $n \geq 5$, that is, it has no proper normal subgroup. This $A_n$ has only $\{(1)\}$ as a normal subgroup. So $S_n$ has only one subnormal series for $n \geq 5$,

$$S_n \supset A_n \supset \{e\}$$

where $e = (1)$, identity permutation in $S_n$. Factor $S_n/A_n$ is abelian as it has order 2, but the factor $A_n/\{e\}$ is not abelian.

So $S_n$ has no solvable series and hence $S_n$ is not solvable for $n \geq 5$.

**Theorem 4.** *A group $G$ is solvable iff $G^{(n)} = \{e\}$, for some $n \in N$*

**Proof :** First suppose that $G$ is solvable and hence it has a solvable series

$$G = H_0 \supset H_1 \supset H_2 \supset \ldots \supset H_n = \{e\} \qquad \ldots..(1)$$

Here each $H_i$ is a subgroup of $G$, $H_{i+1} \lhd H_i$ and $H_i/H_{i+1}$ is abelian for $i = 0, 1, 2, \ldots. n-1$.

We shall now show that $G^{(i)} \subset H_i$ .....(2)

where $G^{(i)}$ is $i^{\text{th}}$ derived subgroup of $G$.

Since $G/H_1$ is abelian therefore $G^{(1)} \subset H_1$. [by theorem (3)]

Thus statement (2) is true for $i = 1$.

Let (2) is true for $i = r$, that is,

$$G^{(r)} \subset H_r$$

$$\Rightarrow \quad \left(G^{(r)}\right)' \subset \left(H_r\right)'$$

$$\Rightarrow \quad G^{(r+1)} \subset \left(H_r\right)' \qquad \ldots..(3)$$

Since $H_r/H_{r+1}$ is abelian , therefore $(H_r)' \subset H_{r+1}$ [by theorem (3)] ....(4)

by (3) and (4), we have

$$G^{(r+1)} \subset H_{r+1}.$$

Thus, the statement (2) is true for every value of $i$. So, we have

$$G^{(n)} \subset H_n = \{e\} \qquad \text{[by (1)]}$$

but $\{e\} \subset G^{(n)}$ always, so $G^{(n)} = \{e\}$.

**Conversely,** suppose that there exists some $n \in N$ such that $G^{(n)} = \{e\}$.

We now that $G^{(i+1)} \lhd G^{(i)}$ and $G^{(i)}/G^{(i+1)}$

is abelian, therefore $G \supset G^{(1)} \supset G^{(2)} \supset .... \supset G^{(n)} = \{e\}$

is a solvable series for $G$, Hence $G$ is solvable.

**Note :** The above result is known as characteristic property of a solvable group and it is considered as alternate definition of a solvable group.

**Theorem 5.** *Every subgroup of a solvable group is solvable.*

**Proof :** Let $H$ be a subgroup of a group $G$,

then $H \subset G \Rightarrow H' \subset G'$,

then it is easy to show that $H^{(i)} \subset G^{(i)}, \forall i \in N$ .....(1)

Since $G$ is solvable, therefore there exists some $n \in N$ such that

$G^{(n)} = \{e\}$ [by theorem (4)]

Since $H^{(n)} \subset G^{(n)}$ [by (1)]

so that $H^{(n)} \subset \{e\}$, *but* $\{e\} \subset G^{(n)}$ always,

thus $H^{(n)} = \{e\}$.

Hence $H$ is solvable.

**Theorem 6.** *Every homomorphic image of a solvable group is solvable.*

**Proof :** Let $G$ be a solvable group and let $f$ is an epimorphism of $G$ on to $G_1$, then $f(G) = G_1$.

Since $G$ is solvable, then $\exists$ some $n \in N$ such that $G^{(n)} = \{e\}$,

$\Rightarrow$ $f[G^{(n)}] = f\{e\}$

$\Rightarrow$ $[f(G)]^{(n)} = \{e_1\}$ $\{f(e) = e_1$, identity of $G_1\}$

(by induction we can verify that $f[G^{(i)}] = [f(G)]^{(i)}$)

Hence $f(G)$ is solvable.

**Corollary :** ***Every quotient group of a solvable group is solvable.***

**Proof :** There exists a natural homomorphism $p$ of $G$ onto $G/H$, then $G/H$ is a homomorphic image of a solvable group $G$ under $p$. Hence, by above theorem, $G/H$ is solvable.

**Theorem 7.** *Let $G$ be a group and $N \lhd G$. If $N$ and $G/N$ are solvable then $G$ is solvable.*

**Proof.** Let us suppose that $N$ and $G/N$ are solvable. Then there exists a slovable series for $N$ as

$$N = N_0 \supset N_1 \supset N_2 \supset ..... N_k = \{e\} \quad .....(1)$$

Also a solvable series for $G/N$ as

$$\frac{G}{N} = \frac{G_0}{N} \supset \frac{G_1}{N} \supset \frac{G_2}{N} \supset ..... \supset \frac{G_m}{N} = \{N\} \quad .....(2)$$

such that $N_{i+1} \lhd N_i, N_i/N_{i+1}$ is abelian for $i = 0, 1, 2.....k-1$ and

$$\frac{G_{j+1}}{N} \lhd \frac{G_j}{N}, \quad \frac{(G_j/N)}{(G_{j+1}/N)}$$

is abelian for $j = 0, 1, 2,..., m-1$.

From series (2), we observe that each $G_i$ is a subgroup of $G$ and $N$ is a normal subgroup of each $G_i$.

Now, since $\frac{G_{j+1}}{N} \lhd \frac{G_j}{N}$, therefore $G_{j+1} \lhd G_j$, for all $j = 0, 1, 2, ..., m-1$.

Also, $$\frac{G_m}{N} = \{N\} \Rightarrow G_m = N$$

Then $$G = G_0 \supset G_1 \supset G_2 \supset ... G_m = N = N_0 \supset N_1 \supset N_2 \supset ... \supset N_k = \{e\}$$

is a solvable series for $G$.

Hence $G$ is solvable.

## 3.6 Refinement of a subnormal series

**Definition :** Let $\{H_i\}$ and $\{K_i\}$ be two subnormal series of the group $G$ such that

$$G = H_0 \supset H_1 \supset H_2 \supset ... H_k = \{e\} \quad .....(1)$$

and $$G = K_0 \supset K_1 \supset K_2 \supset ... K_m = \{e\} \quad .....(2)$$

The series $\{K_j\}$ is called the refinement of the series $\{H_i\}$ if

$$\{H_0, H_1, H_2, ..., H_k\} \subset \{K_0, K_1, K_2, ..., K_m\}.$$

***Ex.*4.** *The subnormal series $Z \supset 9Z \supset 18Z \supset 72Z \supset \{0\}$ is a refinement of the subnormal series $Z \supset 9Z \supset \{0\}$ of group $(Z,+)$.*

***Ex.*5.** *The subnoraml series $Z \supset 4Z \supset 8Z \supset 24Z \supset 72Z \supset \{0\}$ is refinement of the subnormal series $Z \supset 4Z \supset 8Z \supset \{0\}$ of the group $(Z, +)$.*

## 3.7 Equivalent subnormal series

**Definition :** Two subnoraml series of a group $G$ are said to be equivalent or isomorphic if they are of same length and have isomorphic factors, that is, if there exists a one to one correspondence between set, of factors $\left\{\frac{H_0}{H_1}, \frac{H_1}{H_2}, \frac{H_2}{H_3}, ..., \frac{H_{n-1}}{H_n}\right\}$ and $\left\{\frac{K_0}{K_1}, \frac{K_1}{K_2}, \frac{K_2}{K_3}, ..., \frac{K_{n-1}}{K_n}\right\}$.

**Theorem 8.** *(Schreir) Any two subnormal series for the group G have equivalent refinements.*

**Proof :** Let $$G = H_0 \supset H_1 \supset ... \supset H_i \supset H_{i+1} \supset ... \supset H_m = \{e\} \quad .....(1)$$

and $$G = K_0 \supset K_1 \supset ... \supset K_j \supset K_{j+1} \supset ... \supset K_n = \{e\} \quad .....(2)$$

be two subnormal series for $G$. For $H_i$, $H_{i+1}$, $K_j$, $K_{j+1}$, by Zassenhaus lemma, we have

$$(H_i \cap K_{j+1})H_{i+1} \lhd (H_i \cap K_j)H_{i+1} \quad .....(3)$$

and $$(H_{i+1} \cap K_j)K_{i+1} \lhd (H_i \cap K_j)K_{i+1} \quad .....(4)$$

Consider the chain of subgroups, for each $i$,

$$H_i = (H_i \cap K_0) H_{i+1} \supset (H_i \cap K_1) K_{i+1} \supset \ldots$$

$$\ldots \supset (H_i \cap K_j) H_{i+1} \supset (H_i \cap K_{j+1}) H_{i+1} \supset \ldots$$

$$\ldots \supset (H_i \cap K_{n-1}) H_{i+1} \supset (H_i \cap K_n) H_{i+1} = H_{i+1},$$

since $K_n = \{e\}, (H_i \cap K_n) = \{e\}$

Here each subgroup is normal subgroup to the preceding one [by (3)]

Let $H(i, j) = (H_i \cap K_j) H_{i+1}$,

so $H(i, j+1) = (H_i \cap K_{j+1}) H_{i+1}$,

and $H(i, j+1) \lhd H(i, j)$

Also $H(i, 0)\ H_i$ and $H(i,n) = H_{i+1}$.

Thus $H = H_0 = H(0,0) \supset H(0,1) \supset \ldots \supset H(0, n-1) \supset \ldots$

$$\ldots \supset H(0,n) = H_1 = H(1,0) \supset H(1,1) \supset \ldots$$

$$\ldots \supset H(1, n-1) \supset H(2,0) \supset \ldots \supset H(m-1, n-1) \supset (H_{m-1}, n) = H_m = \{e\} \quad \ldots..(5)$$

is a subnormal series such that it is a refinement of series (1).

Similarly, the subnormal series,

$$G = K_0 = K(0, 0) \supset K(1, 0) \supset \ldots, K(m-1, 0) \supset \ldots$$

$$\ldots \supset K(m, 0) = K_1 = K(0, 1) \supset K(1, 1) \supset \ldots$$

$$\ldots \supset K(m-1, n-1) \supset K(m, n-1) = K_n = \{e\} \quad \ldots..(6)$$

is a refinement of series (2), Here

$$K(i,j) = (H_i \cap k_j)\ K_{j+1},\ K(0, j) = K_j$$

and $K(m, j) = K_{j+1}$.

Series (5) and (6) are of same length and by Zassenhaus lemma

$$\frac{H(i,j)}{H(i,j+1)} = \frac{(H_i \cap k_j) H_{i+1}}{(H_i \cap k_{j+1}) H_{i+1}} \cong \frac{(H_i \cap k_j) K_{j+1}}{(H_{i+1} \cap k_j) K_{j+1}}$$

$$= \frac{K(i,j)}{K(i+1, j)}$$

Thus (5) and (6) have isomorphic factors.

Hence (5) and (6) are equivalent.

## 3.8 Maximal normal subgroup

**Definition :** A normal subgroup $M$ of a group $G$ is said to be a maximal normal subgroup if $M \neq G$ and $N$ is normal in $G$ such that $M \subset N \subset G$, then either $M = N$ or $N = G$. That is, no proper normal subgroup $N$ of $G$ contains $M$.

**Note :**

**1.** If $M$ is a normal subgroup of $G$, then every normal subgroup of $G/M$ is of the form $K/M$ where $K \lhd G$, such that $M \lhd K$. Thus, $M$ is maximal if $G/M$ is simple.

**2.** A group is called simple if it does not possess a proper normal subgroup.

***Ex.*6.** *$A_n$is a maximal normal subgroup of $S_n$*

***Ex.*7.** *Each subgroup generated by prime integer is maximal normal subgroup of the group $(Z, +)$.*

## 3.9 Composition series

**Definition :** A subnormal series $G = H_0 \supset H_1 \supset H_2 \supset ... \supset H_n = \{e\}$ of a group $G$ is called a composition series if all factors $\frac{G_i}{G_{i+1}}$, $i = 0, 1, 2, ... n - 1$, are simple, that is, each $G_{i+1}$ is maximal normal subgroup of $G_i$. If $\{H_i\}$ is a normal series then it is called principal or chief series.

**Note :**

**1.** A composition series can not have any further refinement.

**2.** Every cyclic group of prime order has only one composition series.

**3.** Composition series is not necessarily unique.

***Ex.*8.** *$S_n \supset A_n \supset \{(1)\}$ is composition series of $S_n$ for $n \geq 5$.*

**Theorem 9.** *Every finite group G has a composition series.*

**Proof :** If order of the group $G$ is 2 or 3, then $G \supset \{e\}$ is the composition series of $G$. We shall prove this theorem by mathematical induction on order of $G$. Let us assume that the theorem is true for all those groups whose order is less than the order of $G$. If $G$ is simple, then the theorem is true as $G \supset \{e\}$ is the only composition series for $G$. Let $G$ is not simple. Then there exists a proper normal subgroup $H$ of $G$. Since $o(H) < o(G)$, then by our assumption $H$ has a composition series

$$H \supset H_1 \supset H_2 \supset ... \supset H_n = \{e\}.$$

Now, if $H$ is maximal in $G$ then

$$G \supset H \supset H_1 \supset H_2 \supset ... \supset H_n = \{e\}$$

is a composition series for $G$.

If $H$ is not maximal then $o(G/H)$ is less than $o(G)$, so $G/H$ has a composition series

$$\frac{G}{H} \supset \frac{G_1}{H} \supset \frac{G_2}{H} \supset ... \supset \frac{G_m}{H} = \{H\}$$

Also, $$\frac{G_i/H}{G_{i+1}/H} \cong \frac{G_i}{G_{i+1}}$$

Since, $\frac{G_i/H}{G_{i+1}/H}$ is simple

$\Rightarrow$ $\dfrac{G_i}{G_{i+1}}$ is simple

$\Rightarrow$ $G_{i+1}$ is maximal in $G_i$

Thus $G \supset G_1 \supset G_2 \supset \ldots \supset G_m = H \supset H_1 \supset H_2 \supset \ldots \supset H_n = \{e\}$

is a composition series for $G$.

**Theorem 10.** *An infinite abelian group does not have a composition series.*

**Proof :** Let $G$ be an infinite abelian group. Let, if possible, theorem is not true, that is, $G$ has a composition series. Let

$$G = G_0 \supset G_1 \supset G_2 \supset \ldots \supset G_k = \{e\} \qquad \ldots..(1)$$

Since $G$ is abelian and $\dfrac{G_n}{G_{n+1}}$ is simple for $n = 0, 1, 2, \ldots.. k-1$, therefore $\dfrac{G_n}{G_{i+1}}$ is simple abelian group for all $n$.

So, $o\left(\dfrac{G_n}{G_{n+1}}\right)$ = prime number = $p_n$ (say)

for $n = 0, 1, 2, \ldots., k-1$.

Now, $o(G) = \dfrac{o(G_o)}{o(G_1)}\dfrac{o(G_1)}{o(G_2)}\dfrac{o(G_2)}{o(G_3)}\ldots\dfrac{o(G_{k-1})}{o(G_k)}$ $[\because\ o(G_k) = 1]$

$$= o\left(\frac{G_o}{G_1}\right)o\left(\frac{G_1}{G_2}\right)o\left(\frac{G_2}{G_3}\right)\ldots o\left(\frac{G_{k-1}}{G_k}\right)$$

$$= p_0 \cdot p_1 \cdot p_2 \cdot p_{k-1}$$

= finite number.

It implies that $G$ is finite, which is a contradiction as $G$ is infinite. So our assumption is not true, and $G$ has no composition series.

**Theorem 11.** (*Jordan-Holder Theorem*) *Any two composition series for a group G are equivalent.*

**Proof :** Let $G$ be a group and let

$$G \supset H_1 \supset H_2 \supset \ldots \supset H_l = \{e\} \qquad \ldots..(1)$$

and $$G \supset K_1 \supset K_2 \supset \ldots \supset K_m = \{e\} \qquad \ldots..(2)$$

be two composition series for $G$. Then these series are also subnormal series of $G$. By schreir refinement theorem they have equivalent refinement. But (1) is a composition series so it's refinement is the series itself. Similarly (2) is also a composition series, so its refinement is the series it self. Hence $l = m$ and (1) and (2) are equivalent.

## Self-learning exercise-1

**1. Define following :**

*(i)* Derived subgroup

*(ii)* Subnormal series

*(iii)* Solvable group

*(iv)* Equivalent subnormal series

**2. State whether the following statements are true or false :**

*(i)* Commutator of $a$ and $b$ = commutator of $b$ and a always.

*(ii)* $G$ is abelian $\Leftrightarrow [G, G] = \{e\}$

*(iii)* $G^{(1)} \lhd G$, that is, derived subgroup of $G$ is a normal subgroup of $G$.

*(iv)* In a subnormal series of $G$, all the subgroups of $G$ must be normal subgroup of $G$.

*(v)* Every subgroup of a solvable group is solvable.

*(vi)* $4Z$ is maximal normal subgroup of $(Z, +)$

*(vii)* $S_6$ is a solvable group.

*(viii)* $A_n$ is maximal normal subgroup of $S_n$.

*(ix)* $(Z, +)$ has no composition series.

*(x)* Any two composition series for a finite group $G$ are equivalent.

## 3.10 Summary

In this unit we have studied commutators and the subgroup generated by commutators is called derived subgroup. Properties of the derived subgroups have been discussed with the help of theorems. Subnormal series, solvable series and composition series have also been studied. This unit also introduces solvable group which is useful in theory of equations.

## 3.11 Answers to self learning exercises

### Self learning exercise-1

**1.** See text.

**2.** *(i)* false *(ii)* true *(iii)* true *(iv)* false

*(v)* true *(vi)* false *(vii)* false *(viii)* true

*(ix)* true *(x)* true.

## 3.12 Exercises

1. Show that $S_2$ is solvable.
2. Prove that a solvable group has always a nontrivial abelian normal subgroup.
3. Show that $S_5$ is not solvable.
4. Show that a normal subgroup $M$ of a group $G$ is maximal iff the quotient group $G/M$ is simple.
5. Show that each subgroup generated by prime integer is maximal in $(Z,+)$.
6. A group $G$ is solvable if and only if $G$ has a subnormal series with factor groups of prime order.
7. If $H$ is a simple normal subgroup of a group $G$ and $G/H$ has a composition series, then $G$ has a composition series.
8. Show that a direct product of solvable group is solvable.
9. If $G/H$ is abelian, then show that $H \supset G^{(1)}$.
10. If $H$ is a proper normal subgroup of a group $G$ which has a composition series, then show that there exists a composition series containing $H$.

□ □ □

# UNIT 4 : Euclidean Ring

## Structure of the Unit



## 4.0 Objectives

In this unit we shall study a special kind of a ring called Euclidean ring. We shall also discuss division in commutative ring, divisors, units, associates and prime elements of a ring.

## 4.1 Introduction

We have studied in under graduate classes about ring, subring, commutative ring, ring with unity and integral domains. Euclidean ring is a commutative ring without zero divisors in which a special mapping, called Euclidean valuation is defined. Ring $(Z, +, .)$ of integers and the ring of Gaussian integers are examples of Euclidean ring. Study of Euclidean ring requires the study of divisors, that is, divisibility in a ring. Concept of divisibility is applicable only on commutative ring. This unit also introduces an important result known as unique factorization theorem.

## 4.2 Divisibility in a ring

**Divisor :** Let $R$ be a commutative ring. A non-zero element $a$ of $R$ is called divisor of an element $b \in R$ if there exists $c \in R$ such that $b = ac$. We can also say that a divides $b$ or $a$ is a factor of $b$. Symbolically we write it as $a \mid b$.

If a does not divide $b$, then we write $a \nmid b$.

**Note :** Using the definition of divisor, we can easily obtain the following results :

for $x, y, z \in R$, $x \neq 0$

***(i)*** $x \mid y,\ y \mid z \Rightarrow x \mid z$

***(ii)*** $x \mid y,\ x \mid z \Rightarrow x \mid y \pm z$

***(iii)*** $x \mid y, \Rightarrow x \mid my,\ m \in R$

***(iv)*** $x \mid y, \Rightarrow (-x) \mid y$ and $x \mid (-y)$

***Ex.*1.** *Let* $(Z, +, .)$ *be a ring of integers, then* 2 *is divisor of* 10, *because we can write*

$$10 = 2.5, \qquad 5 \in Z.$$

But $2 \nmid 11$, that is, 2 does not divide 11, because

$$11 = 2.11/2, \quad \text{but} \quad 11/2 \notin Z.$$

***Ex.*2.** *In ring* $(Q, +, .)$ *of rational numbers,* 2 *divides* 11 *because* $11/2 \in Q$ *such that*

$$11 = 2.11/2$$

In $Q$, 11 also divides 2 as

$$2 = 11.\ 2/11, \quad \text{and} \quad 2/11 \in Q.$$

**Greatest common divisor :** Let $R$ be commutative ring. Let $a, b \in R$, then an element $c \in R$ is called greatest common divisor (g.c.d.) of $a$ and $b$ if $c$ is a common divisor of $a$ and $b$ and if any other element $x$ divides both $a$ and $b$ then $x$ must divide $c$. Symbolically, $c$ is a g.c.d. of $a$ and $b$ if $c \mid a$, $c \mid b$ and if $x \mid a$, $x \mid b$ then $x \mid c$.

***Ex.*3.** *In a ring* $(Z, +, .)$, 3 *is a g.c.d. of* 9 *and* 15, *and* 5 *is a g.c.d. of* 10 *and* 15.

**Unit :** Let $R$ be a commutative ring with unity element 1 (identity element for the second operation in $R$). An element $x \in R$ is called unit in $R$ if it has multiplicative inverse in $R$, that is, if there exists $y \in R$ such that $xy = 1$.

Obviously, if $x$ is unit then $y$ is also unit in $R$.

***Ex.*4.** *In a field every non-zero element is a unit, so in* $(R, +, .)$ *and* $(Q, +, .)$ *every non-zero element is a unit. In* $(Z, +, .)$, 1 *and* $-1$ *are the only units.*

**Associates :** Let $R$ be an integral domain. Two elements $x$ and $y$ of $R$ are said to be associates if $x$ divides $y$ and $y$ divides $x$.

***Ex.*5.** *In* $(Z, +, .)$, 2 *is an associate of* 2 *and* $-2$.

**Prime element :** Let $R$ be an integral domain. A non-zero, non-unit element $p$ of $R$ is said to be prime or irreducible if the only divisors of $p$ are either units or its associates, that is, if $p = xy$ where $x, y \in R$, then either $x$ is a unit or $y$ is a unit in $R$. A non-zero element is said to be composite or reducible if it is neither a unit nor a prime.

If any two elements of $R$ have 1 as their g.c.d., then they are called relatively prime.

***Ex.*6.** *In* $(Z, +, .)$, $\pm 2, \pm 3, \pm 5, \ldots$ *are prime elements while* $\pm 4, \pm 6, \ldots$ *are composite elements.* 4 *and* 9 *are relatively prime in Z.*

**Theorem 1.** *Let D be an integral domain. Let x and y be two non-zero elements of D, then x any y are associates if and only if* $x = ay$, *where a is a unit element in D.*

**Proof :** First suppose that $x$ and $y$ are associates in an integral domain $D$. Then by the definition of associates, $x$ divides $y$ and $y$ divides $x$. Now

$$x \mid y \Rightarrow \exists\ b \in D \quad \text{such that} \quad y = bx \qquad \ldots..(1)$$

$$y \mid x \Rightarrow \exists\ c \in D \quad \text{such that} \quad x = cy \qquad \text{.....(2)}$$

We have to show that $c$ is a unit in $D$.

Now, $x = cy$

$\Rightarrow$ $x.1 = c\ (bx)$

$\Rightarrow$ $x.1 = (cb)\ x$

$\Rightarrow$ $x\ (1 - cb) = 0$

$\Rightarrow$ $1 - cb = 0$

$(\because\ x \neq 0$ and $D$ is an integral domain so it is without zero divisors $)$

$\Rightarrow$ $cb = 1$

$\Rightarrow$ $b$ and $c$ both are units in $D$.

Conversely, let $x = ay$, where $a$ is unit in $D$, that is, $a^{-1}$ exists in $D$.

Now, $x = ay$

$\Rightarrow$ $y \mid x$ .....(1)

Again, $x = ay$

$\Rightarrow$ $a^{-1}\ x = a^{-1}(ay)$

$\Rightarrow$ $a^{-1}\ x = (a^{-1}a)y$

$\Rightarrow$ $a^{-1}\ x = y$

$\Rightarrow$ $x \mid y$ .....(2)

By (1) and (2), $x$ and $y$ are associates.

## 4.3 Euclidean ring

**Definition :** A commutative ring $R$ without zero divisors is said to be a Euclidean ring if for every $a\ (\neq 0) \in R$, there is defined a non-negative integer $d(a)$ such that

***(i)*** for all $a, b \in R$, with $b \neq 0$, there exist $q$ and $r$ in $R$ such that

$a = qb + r$, where either $r = 0$ or $d(r) < d(b)$ (Division Algorithm)

***(ii)*** for all $a(\neq 0), b(\neq 0) \in R$,

$$d(a) \leq d(ab)$$

Here, mapping $d : R - \{0\} \rightarrow W$ is called Euclidean valuation of $R$, where $W$ is the set of Whole numbers.

***Ex.*7.** *Ring $(Z, +, .)$ of integers is a Euclidean ring for the Euclidean valuation d, defined by $d(a) = |\ a\ |$, $\forall\ a(\neq 0) \in Z$.*

**Sol.** Let $a, b$ be two non-zero elements of $Z$, then

$$d\ (ab) = |\ ab\ |$$

$$= |\ a\ |\ |\ b\ |$$

$$\geq |\ a\ | = d\ (a) \qquad (\because\ |\ b\ | \geq 1)$$

hence, $d(ab) \geq d(a)$

Let $a \in Z$, $b (\neq 0) \in Z$, then by division algorithm in $Z$, there exist $q$ and $r$ in $Z$ such that

$$a = qb + r,$$

where $0 \leq r < |b|$, that is either $r = 0$ or $d(r) < d(b)$.

Also, $Z$ is a commutative ring without zero-divisors. Consequently $Z$ is a Euclidean ring.

***Ex.*8.** *Every field is a Euclidean ring.*

**Sol.** Fields are trivial example of Euclidean ring. Necessarily field $F$ is commutative ring without zero divisors. Let us define a mapping

$$d : F - \{0\} \to W$$

such that, $d(a) = 0, \quad \forall\, a(\neq 0) \in F.$

Let $a (\neq 0)$, $b (\neq 0) \in F$, then, since $F$ is without zero divisors, $ab \neq 0$, so

$$d(ab) = 0 = d(a)$$

Again for $a \in F$, $b(\neq 0) \in F$, there exists

$ab^{-1} = q$ (say) in $F$ such that

$a = (ab^{-1})b + 0$ or $a = qb + r$, where $r = 0$

So field $F$ is a Euclidean ring.

***Ex.*9.** *The ring of Gaussian integers is a Euclidean ring.*

**Sol.** We know that the set of all Gaussian integers, given by

$$Z(i) = \{a + ib \mid a, b \in Z\}$$

is a commutative ring without zero divisors.

Let us define *a* mapping

$$d : Z(i) - \{0\} \to W$$

by $d(a + ib) = a^2 + b^2$, for $a + ib\ (\neq 0) \in Z(i)$

obviously $d$ is well defined.

For $a_1 + ib_1 (\neq 0)$, $a_2 + ib_2 (\neq 0) \in Z(i)$, we have

$$\begin{aligned} d(a_1 + ib_1)(a_2 + ib_2) &= d[(a_1a_2 - b_1b_2) + i(a_1b_2 + a_2b_1)] \\ &= (a_1a_2 - b_1b_2)^2 + i(a_1b_2 + a_2b_1)^2 \\ &= \left(a_1^2 + b_1^2\right)\left(a_2^2 + b_2^2\right) \\ &= d(a_1 + ib_1)\, d(a_1 + ib_2) \qquad \text{.....(1)} \\ &\geq d(a_1 + ib_1) \qquad \left[\because\ d\left(a_2 + ib_2\right) = a_2^2 + b_2^2 \geq 1\right] \end{aligned}$$

Hence, $d[(a_1 + ib_1)(a_2 + ib_2)] \geq d(a_1 + ib_1)$

Now, let $a + ib \in Z(i)$, $c + id (\neq 0) \in Z(i)$, then

$$(a + ib)(c + id)^{-1} = \frac{a + ib}{c + id} = p + iq \qquad \text{(say)}$$

If $p, q \in Z$, then $p + iq \in Z(i)$ and

$$(a + ib) = (p + iq)(c + id) + 0, \qquad \text{.....(2)}$$

So division algorithm holds true.

If $p, q$ are not integers, then $p + q \notin Z(i)$, then choose integers $m$ and $n$ such that

$$|p - m| \le \frac{1}{2} \text{ and } |q - n| \le \frac{1}{2}. \qquad \text{.....(3)}$$

If we write $(a + ib) = (c + id)(m + in) + r,$

then $r = [(a + ib) - (c + id)(m + in)] \in Z(i)],$

as $a + ib, c + id, m + in \in Z(i)$.

Now,
$$\begin{aligned} d(r) &= d[(a + ib) - (c + id)(m + in)] \\ &= d[(p + iq)(c + id) - (c + id)(m + in)] && \text{[by (2)]} \\ &= d[(c + id)\{(p + iq) - (m + in)] \\ &= d[(c + id)\{(p - m) + i(q - n)] \\ &= d[(c + id)\, d\{(p - m) + i(q - n)] && \text{[by (1)]} \\ &= d[(c + id)[(p - m)^2 + (q - n)^2] \\ &= d(c + id)\left(\frac{1}{4} + \frac{1}{4}\right) && \text{[by (3)]} \\ &= \frac{1}{2} d(c + id) \end{aligned}$$

thus $d(r) < d(c + id)$

So, division algorithm holds true in $Z(i)$. Consequently $Z(i)$ is a Euclidean ring.

***Ex.* 10.** *Every ring of polynomials $F[x]$ over a field $F$ is a Euclidean ring.*

**Sol.** We know that $F(x)$, the set of all polynomials over a field $F$, is an integral domain, therefore it is a commutative ring without zero divisors. Let us define a mapping $d$ on the set of non-zero polynomials of $F[x]$ to set of whole numbers $W$, defined by

$$d(p(x)) = \deg p(x), \qquad \forall\, p(x) (\neq 0) \in F[x]$$

Since $\deg p(x) \in W$, for $p(x) \neq 0$, therefore, $d(p(x) \in W$, so $d$ is well defined.

Let $a(x), b(x)$ be two non-zero polynomials of $F(x)$, then $a(x)\, b(x) \neq 0$ as $F(x)$ is without zero divisors, here 0 is a zero polynomial.

Now,
$$\begin{aligned} d[a(x)\, b(x)] &= \deg [a(x)\, b(x)] \\ &= \deg a(x) + \deg b(x) \\ &\ge \deg [a(x)] = d[a(x)] \end{aligned}$$

so $d[a(x)\, b(x)] \ge d[a(x)]$.

Again by the division algorithm of polynomials, for $a(x) \in F(x)$, $b(x) (\neq 0) \in F(x)$, there exist polynomials $q(x), r(x) \in F(x)$ such that

$$a(x) = q(x)\, b(x) + r(x)$$

where either $r(x) = 0$ or $\deg r(x) < \deg b(x)$,

that is, $d[r(x)] < d[b(x)]$.

Hence $F(x)$ is a Euclidean ring.

**Theorem 2.** *Every ideal $I$ in a Euclidean ring $R$ is of the form $aR$ for some $a \in I$.*

**Proof :** Let $d$ be the Euclidean valuation. If $I$ is the zero ideal then $I = 0R$, $0 \in I$, then the theorem is true. Let us suppose that $I$ is not a zero ideal, that is, there exists at least one non-zero element in it. So, we can choose a non-zero element $a \in I$ such that

$$d(a) \le d(x), \qquad \forall\, x\ (\neq 0) \in I.$$

Now, we shall show that $I = aR = \{ar \mid r \in R\}$.

Let $x \in I$, then by division algorithm, $\exists\ q, r \in R$ such that

$$x = qa + r,$$

where either $r = 0$ or $d(r) < d(a)$ .....(1)

Now, $r = x - qa \in I$ $(\because\ x \in I$ and $q \in R, a \in I \Rightarrow qa \in I)$

Since $d(a)$ is minimum in $I$, so $d(a) < d(r)$ .....(2)

thus, $x = qa$, that is, every element $x$ of $I$ can be written as $qa$ for same $a \in I$, and $q \in R$. So $I$ can be written as $aR$. Hence $I$ is a principal ideal of $R$.

**Theorem 3.** *Every Euclidean ring is a principal ideal domain.*

**Proof :** Let $R$ be a Euclidean ring. Then it is commutative and without zero divisors. In order to show that it is an integral domain, we have to show that it has unity element. We know that $R$ is an ideal of itself. So, by the theorem 2, $\exists\ a \in I$ such that $R = aR$, that is,

$$R = \{ar \mid r \in R\}$$

thus, every element of $R$ can be expressed as some multiple of a. Since $a \in R$ so a also can be written as multiple of itself. So

$$a = ab, \text{ for some } b \in R \qquad .....(1)$$

Now, let $x$ be any element of $R$ then

$$x = ac \quad \text{for some} \quad c \in R \qquad .....(2)$$

$\Rightarrow$ $bx = b(ac)$

$= (ba)c$ (by associativity in $R$)

$= (ab)c$ (by commutativity in $R$)

$= ac$ (by (1))

$= x$ (by (2))

Thus, $bx = x = xb, \qquad \forall\, x \in R$

So, $b$ is the unity element in $R$. Hence $R$ is an integral-domain, that is, every Euclidean ring is an integral domain. Also, by theorem 2 every ideal of $R$ is a principal ideal. Consequently $R$ is a principal ideal domain.

**Theorem 4.** *Let R be a Euclidean ring. a and b be two non-zero elements in R, then a and b have greatest common divisor c which can be written as* $(ma + nb)$, *for some* $m, n \in R$.

**Proof :** We know that the subset $I$, given by

$$I = \{pa + qb \mid p, q \in R\} \qquad \text{.....(1)}$$

of ring $R$ is an ideal of $R$ generated by the set $\{a, b\}$. By the theorem 3, $R$ is a principal ideal domain, so $I$ is also a principal ideal. Let $I$ be generated by a single element $c \in R$.

Since $I$ is generated by $c$, therefore $c \in I$ also, then by (1),

$$c = ma + nb \text{ for some } m, n \in R \qquad \text{.....(2)}$$

Also, $a, b \in I$ so $\exists\ \lambda, \mu \in R$ such that

$$a = \lambda c \quad \text{and} \quad b = \mu c \qquad (\because\ I = [c])$$

$$\Rightarrow \qquad c \mid a \quad \text{and} \quad c \mid b$$

$\therefore$ $c$ is a common divisor of $a$ and $b$, let $d$ be any common divisor of $a$ and $b$, then

$$d \mid a \quad \text{and} \quad d \mid b$$

$$\Rightarrow \qquad d \mid ma \quad \text{and} \quad d \mid nb$$

$$\Rightarrow \qquad d \mid (ma + nb) \Rightarrow d \mid c \qquad \text{(by (2))}$$

Thus, $c$ is a greatest common divisor of $a$ and $b$ which can be expressed as $(ma + nb)$ for some $m, n \in R$.

**Theorem 5.** *Let R be a Euclidean ring. Let* $x, y, z \in R$ *such that x and y are relatively prime and x divides yz, then x divides z.*

**Proof :** Since $x$ and $y$ are relatively prime, therefore 1 is the greatest common divisor of $x$ and $y$. By the previous theorem 4, 1 can be expressed as

$$1 = mx + ny \text{ for some } m, n \in R$$

$$\Rightarrow \qquad 1 \cdot z = (mx + ny) \cdot z$$

$$z = mxz + nyz \qquad \text{.....(1)}$$

Now, given that $x \mid yz \Rightarrow x \mid nyz$ .....(2)

also $x \mid x \Rightarrow x \mid mxz$ .....(3)

By (2) and (3), we have

$$x \mid (mxz + nyz)$$

$$\Rightarrow \qquad x \mid z \qquad \text{(by (1))}$$

**Theorem 6.** *Let R be a Euclidean ring and* $p \in R$ *be a prime element such that* $p \mid ab$, $a, b \in R$, *then either p divides a or p divides b.*

**Proof :** Let us suppose that $p \mid ab$ and $p \nmid a$, then we shall show than $p \mid b$. Since $p$ is prime and $p \nmid a$, then $p$ and $a$ are relatively prime, therefore 1 is the greatest common divisor of $p$ and $a$. Then by the theorem 5, we have

$$p \mid ab \Rightarrow p \mid b.$$

**Corollary :** If $p$ is a prime element of $a$ Euclidean ring $R$ such that $p$ divides the product $a_1$, $a_2$, ..., $a_n$, $a_i \in R$ ($i = 1, 2, ..., n$), then $p$ divides at least one of $a_i$.

**Proof :** This follows immediately from theorem 6.

**Theorem 7.** *Let R be a Euclidean ring. Let a and b be two non-zero elements of R such that b is unit in R, then d (ab) = d (a).*

**Proof :** Let $a$ ($\neq 0$), $b$ ($\neq 0$) $\in R$ and $b$ is unit then $b$ has multiplicative inverse in $R$. By the definition of Euclidean ring, we have

$$d(a) \leq d(ab) \qquad .....(1)$$

Again, for the elements $(ab)$ and $(ab)\, b^{-1}$, we have

$$d(ab) \leq d\{(ab)\, b^{-1}\}$$

($\because$ $a \neq 0, b \neq 0 \Rightarrow ab \neq 0$ and also $\{(ab)\, b^{-1}\} \neq 0$)

$$\Rightarrow \quad d(ab) \leq d\{a\,(bb^{-1})\}$$

$$= d(a) \qquad (\because \; bb\text{–}1 = 1)$$

so $$d(ab) \leq d(a) \qquad .....(2)$$

then by (1) and (2)

$$d(ab) = d(a)$$

**Theorem 8.** *Let a and b be two non-zero elements of a Euclidean ring R such that b is not a unit in R, then d (ab) > d (a).*

**Proof :** Let $b$ is not $a$ unit in a Euclidean ring $R$. Since $a \neq 0$, $b \neq 0$ then $ab \neq 0$ as $R$ is without zero divisors.

By the division algorithm in Euclidean ring $R$, for $a \in R$, $ab$ ($\neq 0$) $\in R$, $\exists$ q, $r \in R$ such that

$$d = q(ab) + r \qquad .....(1)$$

where either $r = 0$ or $d(r) < d(ab)$.

Now, if $r = 0$, then by (1)

$$a = q(ab)$$

$$\Rightarrow \quad a - q(ab) = 0$$

$$\Rightarrow \quad a - q(1 - qb) = 0$$

$$\Rightarrow \quad 1 - qb = 0, \quad \text{since } a \neq 0 \qquad (\because \; R \text{ is without zero divisors})$$

$$\Rightarrow \quad qb = 1$$

$\Rightarrow q$ and $b$ are units in $R$ which is a contradiction as $b$ is not $a$ unit, so $r \neq 0$, therefore

$$d(r) < d(ab)$$

$$\Rightarrow \quad d(a - qab) < d(ab) \qquad \text{(by (1))}$$

$$\Rightarrow \quad d[a(1 - qb)] < d(ab) \qquad .....(2)$$

Also, $$d(a) \leq d[a(1 - qb)] \qquad .....(3)$$

($\because$ $1 - qb \in R$ and $d(x) \leq d(xy)$)

by (2) and (3), we have

$$d(a) < d(ab).$$

**Theorem 9.** *Let R be a Euclidean ring. A non-zero element $a \in R$ is a unit iff $d(a) = d(1)$, where 1 is the unity element of R.*

**Proof :** First suppose that $a(\neq 0) \in R$ is a unit in $R$, then a is invertible, that is $a^{-1}$ exists.

Now, since $aa^{-1} = 1$

$\Rightarrow$ $d(aa^{-1}) = d(1)$ .....(1)

By the definition of Euclidean ring, we have

$$d(a) \leq d(aa^{-1})$$

$$d(a) \leq d(1) \quad \text{(by (1))} \quad .....(2)$$

Also $a \cdot 1 = a$

$\Rightarrow$ $d(a \cdot 1) = d(a)$ .....(3)

Again by the definition of Euclidean ring, we have

$$d(a \cdot 1) \geq d(1)$$

$\Rightarrow$ $d(a) \geq d(1)$ (by (3)) .....(4)

by (2) and (4), we have

$$d(a) = d(1)$$

**Conversely :** Let us suppose that $d(a) = d(1)$, then we have to show that $a$ is $a$ unit in $R$. Let, if possible $a$ is not $a$ unit in $R$, then by the theorem 8, we have

$$d(1) < d(1 \cdot a)$$

$\Rightarrow$ $d(1) < d(a)$

which is $a$ contradiction as $d(a) = d(1)$. So $a$ is $a$ unit in $R$.

**Theorem 10.** *Let R be a Euclidean ring, then every non-zero element in R is either a unit or can be written as the product of a finite number of prime elements of R.*

**Proof :** Let $a$ be a non- zero element of $R$. If a is unit, then the theorem is true. Now, if $a$ is prime, then the theorem is again true. Let us suppose that $a$ is neither $a$ unit nor $a$ prime in $R$. Let $d$ be the corresponding Euclidean valuation of $R$. We shall prove the theorem by mathematical induction on $d(a)$. Let us suppose that the theorem is true for all element $r \in R$ for which $d(r) < d(a)$. Then we shall prove the theorem for $a$, which is a non-unit and reducible element of $R$. Since $a$ is reducible, therefore

$$a = xy \quad .....(1)$$

where neither $x$ nor $y$ is $a$ unit in $R$.

Now $x, y$ are not unit in $R$, then by theorem 8, we have

$$d(x) < d(xy) \quad \text{and} \quad d(y) < d(xy)$$

$\Rightarrow$ $d(x) < d(a)$ and $d(y) < d(a)$

Thus, by induction assumption $x$ and $y$ can be written as a product of finite number of prime elements of $R$. Let us suppose that

$$x = x_1 x_2 \ldots x_n \quad \text{and} \quad y = y_1 y_2 \ldots y_m$$

where $x_i (i = 1, 2, ...., n)$ and $y_i (j = 1, 2, ...., m)$ are prime elements of $R$. Then by (1)

$$a = x_1 x_2 \ldots x_n, y_1 y_2 \ldots y_m,$$

that is, a has been expressed as a product of finite number of prime elements of $R$.

**Theorem 11.** *(Unique Factorization Theorem) Let R be a Euclidean ring. Let a be a non-zero, non-unit element of R such that*

$$a = x_1 x_2 .... x_m = y_1 y_2 ... y_n$$

*where* $x_i (1 \le i \le m)$ *and* $y_j (1 \le j \le n)$ *are prime elements of R. Then* $m = n$ *and each* $x_i$ *is an associate of some* $y_j$ *and each* $y_j$ *is an associate of some* $x_i$.

**Proof :** Given that $a = x_1 x_2 ... x_m = y_1 y_2 ... y_n$ .....(1)

so $x_1 \mid x_1 x_2 ... x_m \Rightarrow x_1 \mid y_1 y_2 ... y_n$

$\Rightarrow$ $x_1$ divides at least one $y_1, y_2, ... y_n$ $(\because$ $x_1$ is prime)

Let $x_1$ divides $y_j$ for same $j (1 \le j \le n)$. But $x_1$ and $y_j$ both are prime elements of $R$. So they must be associates, that is,

$$y_j = u_1 x_1 \quad \text{where } u_1 \text{ is a unit in } R.$$

So, by (1), we have

$$x_1 x_2 ... x_m = y_1 y_2 ... y_{j-1} (u_1 x_1) y_{j+1} ... y_n$$

$$= u_1 x_1 y_1 y_2 ... y_{j-1} y_{j+1} ... y_n$$

$\Rightarrow$ $x_2 x_3 ... x_m = u_1 y_1 y_2 ... y_{j-1} y_{j+1} ... y_n$ (by cancellation law)

proceed similarly for $x_2$.

If $n > m$, then after $m$ steps, the L.H.S. reduces to 1 and R.H.S. becomes a product of some units of $R$ and $(n - m)$ prime elements. The product of some units and some prime numbers can not be equal to 1. Therefore $n > m$ is not possible. So

$$n \le m \qquad .....(2)$$

Interchanging the role of $x_i$ and $y_j$, proceeding as above, we get

$$n \ge m \qquad .....(3)$$

By (2) and (3), we get $m = n$. We have also shown that each $x_i$ is an associate of some $x_j$ and vice-versa.

## 4.4 Unique factorization domain

**Definition :** Let $R$ be an integral domain. Then $R$ is said to be a unique factorization domain (UFD) if any non-zero element of $R$ is either $a$ unit or it can be expressed as the product of a finite number of prime elements and this product is unique up to associates. Thus, if $a \in R$ is a non-zero, non-unit element, then

***(i)*** $a = x_1 x_2 ... x_m$, $x_i (1 \le i \le m)$ are prime in $R$

***(i)*** If $a = y_1 y_2 ... y_n$, also, where $y_j (1 \le i \le n)$ are prime in $R$, then $m = n$ and each $x_i$ is an associate of some $x_i$ and vice-versa.

***Ex.*11.** *Every field F is a unique factorization domain since every non-zero element is invertible with respect to multiplication, that is every non-zero element in F is necessarily a unit. Rings Z and Z (i) are also unique factorization domain.*

**Theorem 12.** *Every Euclidean ring R is a unique factorization domain.*

**Proof :** By theorem 10, every non-zero, non-unit element can be expressed as the product of a finite number of prime elements of $R$ and by theorem 11, this factorization is unique. Thus, every Euclidean ring is a unique factorization domain.

**Note :** Since every Euclidean ring is an integral domain, therefore every Euclidean domain is UFD.

### Self-learning exercise-1

State which of the following statements are true :

**1.** 2 is a divisor of 3 in $(Z, +, .)$.

**2.** 5/2 is a divisor of 11 in $(R, +, .)$.

**3.** 11 is an associate of 11 and –11 in $(Z, +, .)$.

**4.** 5 is a unit element in $(R, +, .)$ but not in $(Z, +, .)$.

**5.** Every ring is a Euclidean ring

**6.** Every Euclidean ring is an integral domain.

**7.** Every ideal of a Euclidean ring is not necessarily a principal ideal.

**8.** A non-zero, non-unit element in a Euclidean ring $R$ can be written as the product of a finite number of prime elements in $R$.

## 4.5 Summary

In this unit we studied about divisors, units, associates and prime elements of a ring which follow certain properties. We also studied Euclidean ring and properties of Euclidean ring. Every Euclidean ring is necessarily an integral domain and unique factorization domain. Every field is always an Euclidean ring so it is UFD also.

## 4.6 Answers to self-learning exercises

### Self-learning exercise-1

| | | | |
|---|---|---|---|
| **1.** false | **2.** true | **3.** true | **4.** true |
| **5.** false | **6.** true | **7.** false | **8.** true |

## 4.7 Exercises

**1.** Show that $1, -1, i, -i$ are units in the ring of Gaussian integers. Also show that $(1 + i)$ is a prime element in it.

**2.** Prove that in a Euclidean ring $R$, ideal generated by $\{a, b\}$, where $a(\neq 0)$, $b(\neq 0) \in R$, is a principal ideal generated by $c \in R$, where $c$ is a greatest common divisor of $a$ and $b$.

3. Show that in a commutative ring $R$ with unity, the product of two units is again a unit in $R$ and the set of all units of $R$ forms an abelian group.
4. Show that in a Euclidean ring $R$, 1 is an associate of any unit.
5. Show that in Euclidean ring $R$, any two greatest common divisors of elements $a$ and $b \in R$, are associates.
6. Prove that the relation of divisibility in an integral domain is reflexive, transitive but not symmetric.
7. In a commutative ring $R$ with unity, the relation 'is an associate of $b$' is an equivalence relation on $R$.
8. Show that the ideal $I = [p]$ is a maximal ideal of a Euclidean ring $R$ if and only if $p$ is prime element of $R$.
9. Prove that every non-zero element in a Euclidean ring $R$ is a unit in $R$ or can be written as a product of prime elements of $R$ and this product is unique up to associates.

□□□

# UNIT 5 : Modules

**Structure of the Unit**



## 5.0 Objectives

In pervious units we have studied algebraic structures, such as groups and rings. These structures involve only binary operations in which the product of two elements in the system is again an element in the system.

In this unit, we introduce a new algebraic structure-module. A module is an additive abelian group; a product (scalar multiplication) is defined to associate elements of a ring to the elements of the module.

## 5.1 Introduction

In this unit, we introduce notion of modules and its properties. We define submodule, direct sum, quotient module and cyclic module. Homomorphism is special kind of mapping from an algebraic structure to similar algebraic structure which preserves the structure. We also study homomorphism of modules in this unit.

## 5.2 Modules

Let $(R, +, .)$ be a ring and let $M$ be a non empty set. Then $M$ is called a **left R-module** or simply a **left module over $R$** if the following are satisfied :

***(i)*** $(M, +)$ is an abelian group

***(ii)*** $\forall\ r \in R$ and $\forall\ m \in M \Rightarrow r\,m \in M$

This law satisfies the following conditions :

**(1)** $r\,(m + n) = r\,m + r\,n$ ;

**(2)** $(r + s)\,m = r\,m + s\,m$ ;

**(3)** $(r\,s)\,m = r\,(s\,m)$

for all $r, s \in R$ and all $m, n \in M$.

If $R$ has unity element 1 such that $1\,m = m\ \forall\ m \in M$,

then $M$ is called a **unital module**.

In the left R-module the ring elements appears on the left. If in the above definition we replace $r\,m$ by $m\,r$, then we have a right **R-module** or simply a **right module over** $R$.

In general a left R-module is not a right R-module if $R$ is not commutative ring. However if the ring $R$ is commutative and $M$ is a left R-module it can be made into a right R-module by defining $m\,r = r\,m$ for all $r \in R$ and all $m \in M$. Since we have

***(i)*** $(m + n)\,r = r\,(m + n)$, as defined

$= rm + rn$, sine $M$ is left R-module

$= mr + nr$

***(ii)*** $m\,(r + s) = (r + s)\,m$

$= r\,m + s\,m$

$= m\,r + m\,s$, and

***(iii)*** $m\,(r\,s) = (r\,s)\,m$

$= (s\,r)\,m$, since $R$ is commutative

$= s\,(r\,m)$

$= s\,(m\,r)$

$= (m\,r)\,s$

for all $r, s \in R$ and all $m, n \in M$

**Remark .** **(1)** If $R$ is a field, then a unital module $M$ is a vector space over $R$.

**(2)** We shall simply say "R-module $M$" in place of "left R-module $M$".

***Ex.*1.** *A ring R is an R-module over its subring :*

**Sol.** Let $S$ be a subring of a ring $R$. Since $R$ is a ring, therefore it is an additive abelian group. Taking the multiplication in $R$ as scalar multiplication we can see that

$\forall\, m \in R$, and $\forall\, r \in S \Rightarrow r\,m \in R$, further

*(i)* $r\,(m + n) = r\,m + r\,n$ for all $m,n \in R$, $r \in S$ follows from the left distributive law in $R$.

*(ii)* $(r + s)\,m = r\,m + s\,m$ for all $m \in R$ and $r, s \in S$ is a consequence of right distributive law in $R$ and

*(iii)* $(r\,s)\,m = r\,(s\,m)$ for all $r, s \in S$ and $m \in R$ follows from associativity in $R$. Hence $R$ is an R-module over $S$.

***Ex.2.*** *Let R be any ring. Then R is a module over itself since the scalar multiplication of a ring element on a module element is just the usual multiplication in the ring R and the three axioms are simply the distributive and associative laws in R .*

***Ex.3.*** *Let n be a positive integer. Let R be any ring. Then the set of n-tuples* $R^n = \{(r_1,r_2,...r_n): r_i \in R,\ i \in \underline{n}\}$ *is an R-module under the termwise operations defined by* $(r_1, r_2,...r_n) + (s_1, s_2, ... s_n) = (r_1 + s_1, ... r_n + s_n)$ *and* $r\,(r_1,...r_n) = (rr_1,...rr_n)$ *for all* $(r_1, r_2, ... r_n), (s_1, s_2, ... s_n) \in R^n$ *and all* $s \in R$.

**Sol.** Since $R^n$ is clearly an abelian group for addition defined above and for the remaining axiom, we have

*(i)* $$\begin{aligned} r[(r_1,r_2,...r_n) + (s_1,s_2,...s_n) &= r\,(r_1 + s_1, r_2 + s_2,...r_n + s_n) \\ &= [r\,(r_1 + s_1), r\,(r_2 + s_2)..., r\,(r_n + s_n)] \\ &= (rr_1 + rs_1, rr_2 + rs_2,..., rr_n + rs_n) \\ &= (rr_1,rr_2, ...rr_n) + (rs_1, rs_2, ...rs_n) \\ &= r\,(r_1,r_2,...r_n) + r\,(s_1,s_2 ,..., s_n) \end{aligned}$$

*(ii)* $$\begin{aligned} (r + s)\,(r_1,r_2,...r_n) &= [(r + s)r_1\ (r + s)r_2, ...\ (r + s)s_n] \\ &= (rr_1 + sr_1, rr_2 + sr_2, ... rr_n + sr_n) \\ &= (rr_1, rr_2, ... rr_n) + (sr_1, sr_2, ... sr_n) \\ &= r\,(r_1,r_2,...r_n) + s\,(s_1,s_2,...s_n) \end{aligned}$$

*(iii)* $$\begin{aligned} rs\,(r_1,r_2,...r_n) &= [(rs)r_1,(rs)r_2, ...\ (rs)\,r_n] \\ &= [r\,(sr_1), r\,(sr_2), ...\ r\,(sr_n)] \\ &= r\,(sr_1, sr_2,... sr_n) \\ &= r\,[\,s\,(r_1, r_2,... r_n)] \end{aligned}$$

for all $n$-tuples and $r, s \in R$.

***Ex.*4.** *Every additive abelian group is a module over the ring Z of integers.*

**Sol.** Let $(M, +)$ be an abelian group. For any integer $a \in Z$ and $m \in M$, we define am as follows :

$$am = \begin{cases} \underbrace{m + m + .... + m}_{(a \text{ times})} & \text{if } a > 0 \\ 0 & \text{if } a = 0 \\ \underbrace{(-m) + (-m) + ... + (-m)}_{(a - \text{times})} & \text{if } a < 0 \end{cases}$$

Hence 0 is the identify of the additive group $M$. Clearly $a\, m \in M$ and for the remaing axioms, we have

(1) If $a > 0$, then

$$\begin{aligned} a(m + n) &= (m + n) + (m + n) + .....+ (m + n)\ (a - \text{times}) \\ &= (\underbrace{m + m + ..... + m}_{a - \text{times}}) = (\underbrace{n + n + ..... + n}_{a - \text{times}}) \\ &= a\, m + a\, n \end{aligned}$$

If $a = 0$, Then

$$a(m + n) = 0 = 0 + 0 = am + an$$

If $a < 0$, then

$$\begin{aligned} a(m + n) &= \{-(m + n)\} + \{-(m + n)\} + ... \{-(m + n)\} \\ &= (-m - n) + (-m - n) + ... + (-m - n) \\ &= \underbrace{\{(-m) + (-m) + ... + (-m)\}}_{a - \text{times}} \\ &\quad + \underbrace{\{(-n) + (-n) + ... + (-n)\}}_{a - \text{times}} \end{aligned}$$

Similarly we can show the remaining axioms.

Hence $M$ is $a$ module over $Z$.

## 5.3 Elementary properties

**Theorem 1.** *Let R be a ring and M be an R-module. Then*

***(i)*** $r\, 0 = 0 \ \forall\, r \in R$

***(ii)*** $0m = 0 \ \forall\, m \in M$

***(iii)*** $(-r)\, m = -(r\, m) = r\, (-m) \ \forall\, r \in R,\ m \in M$

***(iv)*** $r\, (m - n) = r\, m - r\, n \ \forall\, r \in R, \forall\, m, n \in M$

***(iv)*** $(r - s)\, m = r\, m - s\, m \ \forall\, r, s \in R, \forall\ m \in M$

**Proof :** ***(i)*** Since $0 + 0 = 0$, 0 is the identity of $M$.

$\Rightarrow$ $r(0 + 0) = r\,0 \;\forall\, r \in R$

$\Rightarrow$ $r\,0 + r\,0 = r\,0$

$\Rightarrow$ $r\,0 + r\,0 = r\,0 + 0$

$\Rightarrow$ $r\,0 = 0$ [using cancellation law in $M$ ]

***(ii)*** Here $0 + 0 = 0$

$\Rightarrow$ $(0 + 0)\,m = 0\,m \;\forall\, m \in M$

$\Rightarrow$ $0\,m + 0\,m = 0\,m + 0$

$\Rightarrow$ $0\,m = 0$

***(iii)*** For $r \in R$, $r + (-r) = 0$

$\Rightarrow$ $[r + (-r)]\,m = 0\,m$

$\Rightarrow$ $r\,m + (-r)\,m = 0$ [by *(ii)*]

$\Rightarrow$ $(-r)\,m = -(r\,m)$

Similarly

$m + (-m) = 0 \;\Rightarrow\; r[m + (-m)] = r\,0 = 0$

$\Rightarrow$ $[rm + r(-m)] = 0$

$\Rightarrow$ $[r(-m)] = -(r\,m)$

Hence $[r(-m)] = -(r\,m) = (-r)\,m$ for all $r \in R,\ m \in M$.

***(iv)*** For $r \in R,\ m, n \in M$,

$$\begin{aligned} r(m - n) &= r[m + (-n)] \\ &= r\,m + r(-n) \\ &= r\,m - r\,n \end{aligned}$$ [from *(iii)*]

***(v)*** For $r\,s \in R,\ m \in M$,

$$\begin{aligned} (r - s)\,m &= [r + (-s)]\,m \\ &= r\,m + (-s)\,m \\ &= r\,m - s\,n \end{aligned}$$ [from *(iii)*]

**Theorem 2.** *Let $N_1, N_2, \ldots N_k$ be R-modules over a ring R. Then show that*

$$N_1 \times N_2 \times \ldots \times N_k = \{(n_1, n_2, \ldots n_k) : n_i \in N_i\}$$

*with operations defined as*

$$(n_1, n_2, \ldots n_k) + (m_1, m_2, \ldots m_k) = (n_1 + m_1, n_2 + m_2, \ldots n_k + m_k)$$

*and* $r(n_1, n_2, \ldots n_k) = (r\,n_1, r\,n_2, \ldots r\,n_k)$

*for all* $(n_1, n_2, \ldots n_k), (m_1, m_2, \ldots m_k) \in N_1 \times N_2 \times \ldots \times N_k$

*and all* $r \in R$, *is an R-module.*

**Proof :** Since $N_1, N_2, \ldots N_k$ are additive abelian groups therefore their direct product $N_1 \times N_2 \times \ldots \times N_k$ is also an additive abelian group for the addition defined in the theorem. The remaining axioms are as follows :

$$
\begin{aligned}
\textit{(i)}\ r[(n_1, n_2, \ldots n_k) + (m_1, m_2, \ldots m_k)] &= r\,(n_1 + m_1, n_2 + m_2, \ldots n_k + m_k) \\
&= [\,r\,(n_1 + m_1), r\,(n_2 + m_2), \ldots r\,(n_k + m_k)] \\
&= (r\,n_1 + r\,m_1, r\,n_2 + r\,m_2, \ldots r\,n_k + r\,m_k) \\
&= (r\,n_1, r\,n_2, \ldots r\,n_k) + (r\,m_1, r\,m_2, \ldots r\,m_k) \\
&= r\,(n_1, n_2, \ldots n_k) + r\,(m_1, m_2, \ldots m_k) \\
\textit{(ii)}\qquad (r + s)\,(n_1, n_2, \ldots n_k) &= [(r + s)\,n_1, (r + s)\,n_2, \ldots (r + s)\,n_k] \\
&= (r\,n_1 + s\,n_1, r\,n_2 + s\,n_2, \ldots r\,n_k + s\,n_k) \\
&= (r\,n_1, r\,n_2, \ldots r\,n_k) + (s\,n_1, s\,n_2, \ldots s\,n_k) \\
&= r\,(n_1, n_2, \ldots n_k) + s\,(n_1, n_2, \ldots n_k) \\
\textit{(iii)}\qquad (r\,s)\,(n_1, n_2, \ldots n_k) &= [(r\,s)\,n_1, (r\,s)\,n_2, \ldots (r\,s)\,n_k] \\
&= [r\,(s\,n_1), r\,(s\,n_2), \ldots r\,(s\,n_k)] \\
&= r\,(s\,n_1, s\,n_2, \ldots s\,n_k) \\
&= r\,[s\,(n_1, n_2, \ldots n_k)\,]
\end{aligned}
$$

Hence $N_1 \times N_2 \times \ldots \times N_k$ is an R-module.

## 5.4 Sub-modules

Let $M$ be an R-module over a ring $R$, A non void subset $N$ of $M$ is said to be a submodule of $M$ if $N$ itself is an R- module under the operations of addition and scalar multiplication given for $M$ restricted to $N$.

Therefore an R-submodule $N$ of $M$ is a subgroup of $M$ which is closed under the scalar multiplication $r\,m \in N$, for all $r \in R$ and $m \in N$. Remaining axioms for the scalar multiplication then hold in $N$ as they hold in $M$.

Every R-module $M$ has the two submodules $M$ and $\{0\}$. These are called the trivial submodules or improper submodules. Any other submodule of $M$ is known as proper submodule. An R-module $M$ is said to be an **irreducible submodule** if its only submodules are $\{0\}$ and $M$.

We now state and prove a theorem which gives us a criterion for a non-void subset to be a submodule.

**Theorem 3.** *Let R be a ring and let M be an R-module. A non-void subset N of M is a sub-module of M if and only if.*

***(i)*** $x - y \in N$, for all $x, y \in N$

***(ii)*** $r\,x \in N$, for all $x \in N, r \in R$.

**Proof.** If $N$ is a submodule of $M$, then $N$ is an abelian group under addition and is closed under scalar multiplication. Therefore *(i)* and *(ii)* hold.

Conversely, let $N$ be non empty subject of $M$ such that *(i)* and *(ii)* hold.

*(i)* implies that $N$ is additive subgroup of $M$ and therefore $N$ itself is an abelian group under addition.

*(ii)* implies that $N$ is closed under scalar multiplication. The remaining axioms for scalar multiplication hold for all elements in $N$ as they hold in $M$.

Hence $N$ is a submodule of $M$.

**Theorem 4.** *The necessary and sufficient condition for a non-void subset N of an R-module M over a ring R with unity to be a submodule of M is that* $r\,n + s\,y \in N$ *for all* $r, s \in R$ *and all* $x, y \in N$.

**Proof.** Let $N$ be a submodule of $M$, then for all $r, s \in R$ and $x, y \in N$, $r\,x, s\,y \in N$, Hence $r\,x + s\,y \in N$, $(\because\ N$ is additive subgroup of $M)$

Conversely, suppose $N$ is a non-void subset of $M$ such that $r\,x + s\,y \in N$ for all $r, s \in R$ and $x, y \in N$.

Taking $r = 1$(unity of $R$) and $s = -1$ (as $1 \in R \Rightarrow -1 \in R$)

$\Rightarrow \qquad 1.\,x + (-1)\,y \in N$

$\Rightarrow \qquad x - y \in N$ .....(1)

Again, taking $s = 0$, we see that if $r \in R$ and $x, y \in N$, then

$r\,x + 0\,y \in N$

$\Rightarrow \qquad r\,x \in N$ .....(2)

(1) and (2) $\Rightarrow$ $N$ is a submodule of $M$.

**Theorem 5.** *If* $M_1$ *and* $M_2$ *are submodules of an R-module M, then*

*(i)* $M_1 \cap M_2$ is a submodule of $M$, and

*(ii)* $M_1 + M_2 = \{m_1 + m_2 : m_1 \in M_1, m_2 \in M_2\}$ is a submodule of $M$.

**Proof.** ***(i)*** Sine $M_1$ and $M_2$ are submodules of $M$, it follows that $M_1 \cap M_2$ is also a submodule of $M$. Movere since $M$ is additive abelian, $M_1 \cap M_2$ is also additive abelian subgroup of $M$. To complete the proof, it is sufficient to show that $M_1 \cap M_2$ is closed for scalar multiplication.

Let $r \in R$ and $x \in M_1 \cap M_2$.

then $x \in M_1$ and $x \in M_2$.

Now $r \in R$ and $x \in M_1 \Rightarrow r\,x \in M_1$, since $M_1$ is submodule of $M$.

Also $r \in R$ and $x \in M_2 \Rightarrow r\,x \in M_2$, since $M_2$, is submodule of $M$.

Hence $r\,x \in M_1 \cap M_2$ Thus $M_1 \cap M_2$ is a submodule of $M$.

***(ii)*** Let $m = x_1 + x_2$, $n = y_1 + y_2$ be any two elements of $M_1 + M_2$ then $x_1, y_1 \in M_1$ and $x_2, y_2 \in M_2$.

Now, $m - n = (x_1 + x_2) - (y_1 + y_2)$

$= (x_1 - y_1) + (x_2 - y_2)$

Since $M_1$ and $M_2$ are submodules of $M$. Therefore $x_1 - y_1 \in M_1$ and $x_2 - y_2 \in M_2$.

Thus $m - n = (x_1 - y_1) + (x_2 - y_2) \in M_1 + M_2$

Also, let $r \in R,\ m \in M_1 + M_2$

We have $r\,m = r\,(x_1 + x_2) = rx_1 + rx_2$

Again since $M_1$ and $M_2$ are submodules of $M$, therefore $r\,x_1 \in M_1$ and $rx_2 \in M_2$. Thus $rm \in M_1 + M_2$

Hence $M_1 + M_2$ is submodule of $M$.

**Note :** The above results can be generalised to an arbitrary intersection or sum of submodules. Also it is easy to see that the union of two submodules is a submodule if and only if one is contained in other.

## 5.5 Direct sum

Let $M$ be an R-module and let $M_1, M_2,... M_n$ be submodules of $M$. Then $M$ is called the **direct sum** of $M_1, M_2,...M_n$ if every element $m \in M$ is uniquely expressible as

$$m = m_1 + m_2 +.....+ m_n \text{ where}$$

$$m_1 \in M_1,\ m_2 \in M_2\ ,...m_n \in M_n.$$

Symbolically it is denoted as $M = M_1 \oplus M_2 \oplus ..., \oplus M_n$.

**Theorem 6.** *Let $M$ be an R-module and $N_1, N_2,... N_k$ be submodules of $M$. Then the following statements are equivalent :*

*(i)* $M = N_1 \oplus N_2 \oplus ..... \oplus N_k$

*(ii)* If $n_1 + n_2 +.....+ n_k = 0$, then $n_1 = n_2 = ..... = n_k = 0$ for $n_i \in N_i$

*(iii)* $N_i \cap (N_1 + ..... + N_{i-1} + N_{i+1} +..... + N_k) = \{0\}$

**Proof.** *(i)* $\Rightarrow$ *(ii)* Let $M$ is the direct sum of $N_1, N_2,.....N_k$ and $n_1 + n_2 + ..... + n_k = 0$ for $n_i \in N_i$, $i = 1,2..., k$. Since every element of $M$ has unique expression, $0 \in M$, which is written as $0 = 0 + 0 + ..... + 0$ implies that $n_1 = n_2 = .....= n_k = 0$.

*(ii)* $\Rightarrow$ *(iii)* Let $x \in N_i \cap (N_1 +..... + N_{i-1} + N_{i+1} +..... + N_k)$ then $x \in N_i$ and

$x \in (N_1 +..... + N_{i-1} + N_{i+1} + ..... N_k)$

so these exist $n_1 \in N_1, n_2 \in N_2, ....., n_{i-1} \in N_{i-1}, n_{i+1} \in N_{i+1}, .... n_k \in N_k$

such that $x = n_1 + n_2 + ..... + n_{i-1} + n_{i+1}, + ...+ n_k$

$\Rightarrow$ $n_1 + n_2 +.....+ n_{i-1} + (-x) + n_{i+1} + ...+ n_k = 0$

$\Rightarrow$ $n_1 = n_2 = ..... = -x = ..... = n_k = 0$

Since $x \in N_i \Rightarrow -x \in N_i$

*i.e.* $-x$ is the $i$th element in the sum. Hence $x = 0$ and

$$N_i \cap (N_i + ..... + N_{i-1} + N_{i+1} + .... + N_k = \{0\}$$

So *(ii)* $\Rightarrow$ *(iii)*. Finally, to see that *(iii)* $\Rightarrow$ *(i)*, let us assume that for $m \in M$, we have two different representations.

$$m = m_1 + m_2 + ..... + m_k = n_1 + n_2 + ..... + n_k$$

Then $$0 = (m_1 - n_1) + (m_2 - n_2) + ..... + (m_k - n_k)$$

$\Rightarrow$ $$(m_i - n_i) = (n_1 - m_1) + ..... + (n_{i-1} + m_{i-1}) + (n_{i+1} + m_{i+1}) + ... + (n_k - m_k)$$

Now $(m_i - n_i) \in N_i$ and $[(n_1 - m_1) + ..... + (n_{i-1} - m_{i-1}) + (n_{i+1} - m_{i+1}) + ..... + (n_k + m_k)]$

$$\in (N_1 + N_{i-1} + N_{i+1} + .... + N_k)$$

So, $$(m_i - n_i) \in N_i \cap (N_1 + ..... + N_{i-1} + N_{i+1} + .... + N_k = \{0\}$$

$\Rightarrow m_i = n_i \ \forall \ i$. Thus $m \in M$ has unite representations and hence $M$ is the direct sum of $N_1$, $N_2$, ..... $N_k$.

## 5.6 Quotient module

Let $M$ be an R-module and let $N$ be submodule of $M$. Then the set

$$M/N = \{N + x : x \in M\}$$

is called **quotient module**.

We define operations of addition and scalar multiplication as

$$(N + x) + (N + y) = N + (x + y)$$

and $$r\,(N + x) = N + r\,x$$

for all $x, y \in M$ and $r \in R$.

**Theorem 7.** *Let $M$ be an R-module and let $N$ be a submodule of $M$. Then the set $M/N = \{N + x : x \in M\}$ is an R-module for addition and scalar multiplication defined as follows* :

*(i)* $(N + x) + (N + y) = N + (x + y)$

*(ii)* $r\,(N + x) = N + r\,x$

for all $N + x, N + y \in M/N$ and $r \in R$.

**Proof.** Since $M$ is an abelian group under addition +, the quotient group $M/N = \{N + x : x \in M\}$ is defined and is an abelian group with binary operation

$$(N + x) + (N + y) = N + (x + y),$$

for all $N + x, N + y \in M/N$.

To see that the scalar multiplication of the ring element $r$ on the coset $N + x$ is well defined, suppose

$$N + x = N + y \ \Rightarrow \ x - y \in N$$

Since $N$ is an R-submodule, $r\,(x - y) \in N$

$\Rightarrow$ $$r\,x - r\,y \in N$$

$\Rightarrow$ $$N + r\,x = N + r\,y$$

and thus scalar multiplication is well defined. It remains to prove the following axioms for an R-module.

$$
\begin{aligned}
\textit{(i)}\quad r\,[(N+x)+(N+y)] &= r\,[N+(x+y)] \\
&= N + r\,(x+y) \\
&= N + r\,x + r\,y \\
&= (N + r\,x) + (N + r\,y) \\
&= r\,(N+x) + r\,(N+y) \\
\textit{(ii)}\qquad (r+s)\,(N+x) &= N + (r+s)\,x \\
&= N + r\,x + sx \\
&= (N + r\,x) + (N + s\,x) \\
&= r\,(N+x) + r\,(N+x) \\
\textit{(iii)}\qquad (r\,s)\,(N+x) &= N + (r\,s)\,x \\
&= N + r\,(sx) \\
&= r\,[N + (s\,x)] \\
&= r\,[s\,(N+x)]
\end{aligned}
$$

for all $r, s \in R$ and for all $N+x,\ N+y, \in M/N$.

## 5.7 Module homomorphisms

Homomorphism is special kind of mapping from an algebraic structure to similar algebraic structure which preserves the binary operation. Here we consider homomorphism of modules.

**Definition** (module homomorphism ) :– Let $M$ and $M'$ *be* R-modules. A mapping $f : M \to M'$ is called an R-module homomorphism if

*(i)* $f(x+y) = f(x) + f(y)$, for all $x, y \in M$ and

*(ii)* $f(r\,x) = r\,f(x)$ for all $r \in R,\ x \in M$ and

if $R$ is a ring with unity then we can combine *(i)* and *(ii)* as $f(r\,x + s\,y) = r\,f(x) + s\,f(y)$.

Homomorphism is called **module isomorphism** if $f$ is one and onto. Modules $M$ and $M'$ are said to be **isomorphic** if there is some R-module isomorphism of $M$ onto $M'$ and we denote $M \cong M'$.

An R-module homomorphism is a **monomorphism** if it is injective, an **epimorphism** if it is surjective.

An R-module homomorphism $f : M \to M$ from an R-module $M$ into itself is known as an **endomorphism.** Further an isomorphism $f$ from $M$ onto itself is called an **automorphism**.

**Theorem 8.** *If $M$ and $M'$ are two R-modules and if $f : M \to M'$ is a homomorphism, then*

*(i)* $f(0) = 0 \in M'$

*(ii)* $f(-x) = -f(x)$

*(iii)* $f(x-y) = f(x) - f(y)$ for all $x, y \in M$.

**Proof.** *(i)* We have

$$0 + 0 = 0$$

$$\Rightarrow \quad f(0 + 0) = f(0)$$

$$\Rightarrow \quad f(0) + f(0) = f(0) + 0$$

$$\Rightarrow \quad f(0) = 0 \in M'.$$ using cancellation law in $M'$.

*(ii)* for all $x \in M$, we have

$$x + (-x) = 0$$

$$\Rightarrow \quad f[x + (-x)] = f(0) = 0$$

$$\Rightarrow \quad f(x) + f(-x)] = 0 \in M'$$

$$\Rightarrow \quad f(-x) = -f(x)$$

*(iii)* For all $x, y \in M$, we have

$$f(x - y) = f[x + (-y)]$$
$$= f(x) + f(-y)$$
$$= f(x) - f(y)$$ using (*ii*) above.

**Kernel of Homomorphism :** Let $M$ and $M'$ be two R-modules and $f : M \to M'$ be a module homomorphism. The set of all elements of $M$ which are mapped to zero of $M'$ is known as **Kernel of homomorphism** $f$ and is denoted by Ker $(f)$. Symbolically Ker $(f) = \{x \in M \mid f(x) = 0 \in M'\}$.

Since $f(0) = 0 \in M'$, therefore $0 \in$ Ker $(f)$.

**Theorem 9.** *Let $f : M \to M'$ be an R-module homomorphism, then*

*(i)* Ker $(f) = \{x \in M \mid f(x) = 0 \in M'\}$ is a submodule of $M$.

*(ii)* Im $(f) = \{f(x) \mid x \in M'\}$ is a submodule of $M'$.

**Proof.** *(i)* Since $f(0) = 0 \in M'$. It follows that Ker $(f)$ is non-empty subset of $M$.

Let $x, y \in$ Ker $(f)$ then $f(x) = 0, f(y) = 0$.

Since $f(x - y) = f[x + (-y)]$
$= f(x) + f(-y)$
$= f(x) - f(y)$
$= 0$

therefore $x - y \in$ Ker $(f)$.

Thus $x, y \in$ Ker $(f) \Rightarrow (x - y) \in$ Ker $(f)$

Again let $r \in R$ and $x \in$ Ker $(f)$, then

$f(rx) = rf(x) = r0 = 0 \in M'$ and therefore $rx \in$ Ker $(f)$.

Hence Ker $(f)$ is a submodule of $M$.

*(ii)* Since $0 \in M$ and $f(0) = 0 \in M'$, therefore $0 \in$ Im $(f)$.

Thus Im $(f)$ is a non-empty subject of $M'$.

Let $x', y' \in \text{Im}(f)$ so that there exist $x, y \in M$,

such that $f(x) = x'$ and $f(y) = y'$.

Now $\quad x' - y' = f(x) - f(y) = f(x - y) \in \text{Im}(f)$

so $\quad x' - y' \in \text{Im}(f)$

Thus $\quad x', y' \in \text{Im}(f) \Rightarrow x' - y' \in \text{Im}(f)$

Again let $r \in R$ and $x' \in \text{Im}(f)$, then

$$r x' = r f(x) = f(r x) \in \text{Im}(f)$$

$\Rightarrow \quad r x' \in \text{Im}(f)$

It follows that $\text{Im}(f)$ is a submodule of $M'$.

**Theorem 10.** *If $f : M \to M'$ is an R-module homomorphism, then $f$ is a monomorphism if and only if Ker $(f) = \{0\}$.*

**Proof.** First let $f$ be a monomorphism from $M$ to $M'$.

Then $f$ is one-one.

Let $\quad x \in \text{Ker}(f)$,

then $\quad f(x) = 0 \in M'$

$\Rightarrow \quad f(x) = f(0)$

$\Rightarrow \quad x = 0 \qquad$ ($f$ being one-one)

Hence Ker $(f)$ contains only $0 \in M$ i.e. Ker $(f) = \{0\}$.

Conversely, let Ker $(f) = \{0\}$, then let $x, y \in M$ be such that $f(x) = f(y)$,

then $\quad f(x) = f(y)$

$\Rightarrow \quad f(x) - f(y) = 0 \in M'$

$\Rightarrow \quad f(x - y) = 0 \in M'$

$\Rightarrow \quad x - y \in \text{Ker}(f) = \{0\}$

$\Rightarrow \quad x - y = 0$

$\Rightarrow \quad x = y$

Thus $f$ is one-one. Since $f$ is a homomorphism. So it is a monomorphism.

**$\text{Hom}_R(M, M')$ :** We shall now consider the set of all homomorphism from an R-module to another R-module. Let $M$ and $M'$ be two R-modules. We denote by $\text{Hom}_R(M, M')$, the set of all homomorphism from $M$ into $M'$. In this set we shall introduce two operations *(i)* internal composition, *(ii)* scalar multiplication. Here, we assume that R is a ring with unity.

Let $f, g \in \text{Hom}_R(M, M')$, we define the sum $f + g$ by the rule $(f + g)(x) = f(x) + g(x)$ for all $x \in M$.

Now, $x, y \in M$ and $r, s \in R$, we have

$$(f + g)(r x + s y) = f(r x + s y) + g(r x + s y)$$
$$= r f(x) + s f(y) + r g(x) + s g(y)$$

$$= r[f(x) + g(x)] + s[f(y) + [g(y)]$$
$$= r(f + g)(x) + s(f + g)(y)$$

Thus $f + g$ is a module homomorphism from $M$ into $M'$ and therefore

$$f, g \in \text{Hom}_R(M, M') \Rightarrow f + g \in \text{Hom}_R(M, M')$$

The above binary operation in $\text{Hom}_R(M, M')$ is known as point wise addition of morphism.

**Theorem 11.** *Let $M$ and $M'$ be two R-modules. Then the set $\text{Hom}_R(M, M')$ is an abelian group under pointwise addition of morphism.*

**Proof.** We have already seen that point wise addition is a binary composition in $\text{Hom}_R(M, M')$ *i.e.* if $f.g \in \text{Hom}_R(M, M')$.

Then $f + g \in \text{Hom}_R(M, M')$ defined by $(f + g)(x) = f(x) + g(x)$ for all $x \in M$.

Now let $f, g, h \in \text{Hom}_R(M, M')$, then for any $x \in M$,

$$[(f + g) + h](x) = (f + g)(x) + h(x)$$
$$= [f(x) + g(x)] + h(x)$$
$$= f(x) + [g(x) + h(x)]$$
$$= f(x) + (g + h)(x)$$
$$= [f + (g + h)](x)$$

Hence, $(f + g) + h = f + (g + h)$

So, point-wise addition in $\text{Hom}_R(M, M')$ is associative.

Now, to prove commutativity, we see that for any $x \in M$

$$(f + g)(x) = f(x) + g(x)$$
$$= g(x) + f(x)$$
$$= (g + f)(x)$$

so, $f + g = g + f$

The zero map $\hat{0}: M \to M'$ such that $\hat{0}(x) = 0, \forall x \in M$ is an element of $\text{Hom}_R(M, M')$.

Also for any $x \in \text{Hom}_R(M, M')$

$$(\hat{0} + f)(x) = \hat{0}(x) + f(x) = 0 + f(x) = f(x) \Rightarrow \hat{0} + f = f$$

Thus the zero map is the identity element in $\text{Hom}_R(M, M')$.

Let for any $f \in \text{Hom}_R(M, M')$, define the map

$$-f: M \to M' \text{ by } (-f)(x) = -f(x) \ \forall x \in M.$$

Then it is obvious that $-f \in \text{Hom}_R(M, M')$.

Also for any $x \in M$,

$$[f + (-f)](x) = f(x) + (-f)(x)$$
$$= f(x) - f(x) = 0 = \hat{0}(x)$$

Therefore $f + (-f) = \hat{0}$

Thus each element in Hom $_R$ $(M, M')$ has its additive inverse in Hom $_R$ $(M, M')$.

Hence Hom $_R$ (M, $M'$) is an abelian group.

**Theorem 12.** *Let R* is a commutative *ring and* $r \in R, f \in Hom_R (M, M')$, *then* $rf \in Hom_R (M, M')$ *defined by* $(rf)(x) = rf(x)$ *for all* $x \in M$ *and with this scalar multiplication the abelian group* $Hom_R (M, M')$ *is an R-module.*

**Proof.** Let $x, y \in M$ and $s_1, s_2, \in R$, then

$$\begin{aligned}(rf)(s_1 x + s_2 y) &= rf(s_1 x + s_2 y)\\ &= r[s_1 f(x) + s_2 f(y)]\\ &= r s_1 f(x) + r s_2 f(y)\\ &= s_1 f(rx) + s_2 f(ry)\\ &= s_1 rf(x) + s_2 rf(y)\\ &= s_1 (rf)(x) + s_2 (rf)(y)\end{aligned}$$

Hence $rf: M \to M'$ is an R-module homomorphism. So $rf \in Hom_R (M, M')$.

We need to verify now the remaining R-module axioms

*(i)* $$\begin{aligned}[r(f+g)](x) &= r(f+g)(x)\\ &= r[f(x) + g(x)]\\ &= rf(x) + rg(x)\\ &= (rf)(x) + (rg)(x)\\ &= (rf + rg)(x)\end{aligned}$$

so, $r(f+g) = rf + rg)$

*(ii)* $$\begin{aligned}[(r+s)f](x) &= (r+s)f(x)\\ &= rf(x) + sf(x)\\ &= (rf)(x) + (sf)(x)\\ &= (rf + sf)(x)\end{aligned}$$

so $$\begin{aligned}(rs)f](x) &= (rs)f(x)\\ &= r[sf(x)]\\ &= r(sf)(x)\end{aligned}$$

so $(rs)f = r(sf)$

Hence $\text{Hom}_R (M, M')$ is an R-module.

**Theorem 13.** *Let M be an R-module and let N be a submodule of M. Then the natural projection map* $p : M \to M / N$ *defined by* $p(x) = N + x$ *for all* $x \in M$ *is an R-module with Kernel N.*

**Proof.** Let $x, y \in M$ and $r, s \in R$, then

$$p(rx+sy) = N + (rx+sy)$$
$$= (N+rx) + (N+sy)$$
$$= r(N+x) + s(N+y)$$
$$= r\,p(x) + s\,p(y)$$

Thus $\phi$ is an R-module homomorphism from $M$ to $M/N$.

Let $K$ be the Kernel of $p$. Then $K = \{x \in M \mid \phi(x) = N\}$

Now we shall prove that $K = N$

$$x \in K \Leftrightarrow p(x) = N$$
$$\Leftrightarrow N + x = N$$
$$\Leftrightarrow x \in N$$

$$\therefore \quad K = N$$

## 5.8 Isomorphism theorems

All the isomorphism theorems stated for groups also hold for R-modules. The proofs are similar to the corresponding theorems for groups.

**Theorem 14.** (*Fundamental theorem an module homomorphism*) *Let $M$, $M'$ be R-modules and $f : M \to M'$ be an R-module homomorphism. Then Ker $(f)$ is a submodule of $M$ and $M/Ker\,(f) \cong Im\,(f)$. Equivalently, every homomorphic image of an R-module is isomorphic to some quotient module.*

**Proof.** Let $K$ be the Kernel of homomorphism $f$.

By def. $K = \{x \in M : f(x) = 0\}$

Since $f(0) = 0 \in M'$

$\therefore \quad 0 \in K$ and so $K$ is non-empty set.

Let $x, y \in K$ and $r \in R$ be arbitrary, then $f(x) = 0 = f(y)$ and $x,y \in M$.

$\Rightarrow x - y \in M$ as $(M, +)$ is an abelian group.

$$f(x-y) = f(x) - f(y) = 0 - 0 = 0$$

$$\therefore \quad x, y \in K \ \forall\ x,y \in K$$

$\Rightarrow K$ is additive subgroup of $M$.

$$r \in R, x \in K \quad \Rightarrow \quad r \in R, x \in M, f(x) = 0$$

$\Rightarrow rx \in M, f(x) = 0$ by definition of R-modules.

$\Rightarrow \quad f(rx) = r\,f(x) = r.\,0 = 0$

$\Rightarrow \quad f(rx) = 0, rx \in M \Rightarrow rx \in K.$

Hence $K$ is submodule of $M$.

Therefore $M/K$ is well defined. We define a mapping $\phi = M/K \to Im\,(f)$ *s.t.* $\phi(K+x) = f(x)$ for all $x \in M$.

Now this mapping is well defined, since for any $x, y \in M$

$$K + x = K + y \Rightarrow x - y \in K$$

$$\Rightarrow \quad f(x - y) = 0 \in M'$$

$$\Rightarrow \quad f(x) - f(y) = 0$$

$$\Rightarrow \quad f(x) = f(y)$$

$$\Rightarrow \quad \phi(K + x) = \phi(K + y)$$

We shall now show that $\phi$ is an isomorphism.

For any $K + x, K + y \in M/K$, we have

$$\phi(K + x) = (K + y) \Rightarrow f(x) = f(y)$$

$$\Rightarrow \quad f(x - y) = 0 \in M'$$

$$\Rightarrow \quad x - y \in K$$

$$\Rightarrow \quad K + x = K + y$$

$\therefore$ is one-one.

$\phi$ is onto, because for any $x' \in \text{Im}(f)$, there exists $x \in M$ such that $f(x) = x'$, which implies that for each $x' \in \text{Im}(f)$, these exists $K + x \in M/K$ such that

$$\phi(K + x) = f(x)$$

Finally for any $K + x, K + y \in M/K$ and $r, s \in R$, we have

$$\begin{aligned}\phi[r(K + x) + r(K + y)] &= \phi(K + rx)(K + ry) \\ &= \phi(K + rx + sy) \\ &= f(rx + sy) \\ &= rf(x) + sf(y) \\ &= r\phi(K + x) + s\phi(K + y)\end{aligned}$$

$\therefore$ $\phi$ is an $R$-module homomorphism.

Hence $f$ is an isomorphism from $\dfrac{M}{\text{Ker}(f)}$ onto $\text{Im}(f)$ and thus $\dfrac{M}{\text{Ker}(f)} \cong \text{Im}(f)$.

**Theorem 15.** *Let $M_1$ and $M_2$ are sub-modules of an R-module M, then*

$$\frac{M_1 + M_2}{M_2} \cong \frac{M_1}{M_1 \cap M_2}$$

**Proof.** Since $M_1$ and $M_2$ are submodules of an R-module $M$, therefore by theorem 5, $M_1 + M_2$ and $M_1 \cap M_2$ are sub-modules of $M$. Since $M_2 \subset M_1 + M_2$ and $M_1 \cap M_2 \subset M$, therefore $\dfrac{M_1 + M_2}{M_2}$ and $\dfrac{M_1}{M_1 \cap M_2}$ are defined.

We define a mapping $\phi : M_1 \rightarrow \frac{M_1 + M_2}{M_2}$

*s.t.* $\phi(x) = M_2 + x \neq$ for all $x \in M_1$.

Since $x \in M_1 \Rightarrow x \in M_1 + M_2$

$\Rightarrow \quad \phi(x) - M_2 + x \in \frac{M_1 + M_2}{M_2}$

Therefore the mapping $\phi$ is well defined.

We shall now show that $\phi$ is a module homomorphism from $M_1$ onto $\frac{M_1 + M_2}{M_2}$ with Kernel $M_1 \cap M_2$.

For any $x,y \in M_1$ and $r, s \in R$, we have

$$\phi(r x + s y) = M_2 (r x + s y) = (M_2 + r x) + (M_2 + s y)$$
$$= r (M_2 + x) + r (M_2 + y) = r \phi(x) + s \phi(y)$$

$\therefore$ $\phi$ is a module homomorphism from $M_1$ to $\frac{M_1 + M_2}{M_2}$.

To prove $\phi$ is onto, let $M_2 + x \in \frac{M_1 + M_2}{M_2}$, then $x \in M_1 + M_2$ and there exists a unique representation for $x$ as $x_1 + x_2 = x, x_1 \in M_1, x_2 \in M_2$.

Thus for each $M_2 + x \in \frac{M_1 + M_2}{M_2}$ these exists $x_1 \in M_1$ such that

$$\phi(x_1) = M_2 + x_1 = (M_2 + x_2) + x_1 = M_2 + (x_2 + x_1) = M_2 + x$$

Here $\phi$ is onto.

We shall now show that Ker $(\phi) = M_1 \cap M_2$.

Let $x_1 \in M_1$ be such that $x_1 \in$ Ker $(\phi)$, then

$x_1 \in$ Ker $(\phi) \Rightarrow \phi(x_1) = M_2$

$\Rightarrow \quad M_2 + x_1 = M_2$

$\Rightarrow \quad x_1 \in M_2$

$\Rightarrow \quad x_1 \in M_1 \cap M_2 \qquad (\therefore x_1 \in M_1)$

$\therefore \quad$ Ker $(\phi) \subset M_1 \cap M_2$

Again Let $x \in M_1 \cap M_2$. Then $x_1 \in M_1$ and $x_1 \in M_2$

$x \in M_2 \Rightarrow M_2 + x = M_2$

$\Rightarrow \quad \phi(x) + M_2$

$\Rightarrow \quad x \in$ Ker $(\phi)$

$\therefore \quad M_1 \cap M_2 \subset$ Ker $(\phi)$.

Thus Ker $(\phi) = M_1 \cap M_2$.

Hence $\phi : M_1 \rightarrow \frac{M_1 + M_2}{M_2}$ is a homomorphism with Ker $(\phi) = M_1 \cap M_2$ and $Im\ (\phi) = \frac{M_1 + M_2}{M_2}$. Therefore by fundamental theorem on Module homomorphism, we have

$$\frac{M_1}{M_1 \cap M_2} \cong \frac{M_1 + M_2}{M_2}.$$

## 5.9 Generation of modules

### 5.9.1 Submodule generated by a subset

Let $M$ be an R-module. For any subset A of $M$, $N = < A >$ is called a submodule of $M$ **generated** by the subset $A$, if $N$ contains $A$ and any submodule $K$ of $M$ which contains $A$ also contains $N$. Thus the submodule $N$ generated by $A$ is the smallest submodule of $M$ which contains $A$.

If $N$ and $K$ are submodules of $M$, then the smallest submodule of $M$ containing $N \cup K$ is called the submodule generated by $N$ and $K$.

**Theorem 16.** *The submodule S generated by N and K is the submodule*

$$N + K = \{a + b \mid a \in N, b \in K\}$$

**Proof.** Clearly $N + K$ is a submodule of $M$. Also $N \subseteq N + K$ and $K \subseteq N + K$, so that

$$S \subseteq N + K.$$

Conversely, for any $a \in N$, $b \in K$, we have $a, b \in S \Rightarrow a + b \in S$

Thus $N + K \subseteq S$

consequently S = $N + K$.

### 5.9.2 Finitely generated module.

An R-module $M$ is said to be finitely generated if it is generated by some finite subset. Thus $M$ is finitely generated if there exist finite elements $a_1, a_2, ..., a_n \in M$ such that every element $m \in N$ can be expressed as

$$m = r_1 a_1 + r_2 a_2 + ..... + r_n a_n$$

Where $r_1, r_2, ..... s_n \in R$

In this case we write

$$M = < a_1, a_2, ... a_n >.$$

## 5.10 Cyclic module

An R-module $M$ is called **cyclic** if there exists an element $m \in M$ such that

$$M = < m > = \{r m : r \in R\}$$

Thus a cyclic R-module is generated by a single element. Here $m$ is called generator of $M$.

A submodule $N$ of $M$ is cyclic if it is generated by one element $a \in M$,

*i.e.* $$N = < a > = Ra = \{r\, a \mid a \in R\}.$$

Note that if $R$ is a ring with unity and $M$ is an R-module, then for any finite subset $A = \{a_1\, a_2 \ldots\ldots\, a_n\}$ of $M$, the set

$$R\, A = \{r_1\, a_1 + r_2\, a_2 + \ldots + r_n\, a_n \mid r_i \in R\}.$$

**Theorem 17.** *Let R be a ring with unity and M be an R-module. Let N be finitely generated submodule of M generated by a subset* $A = \{a_1\, a_2 \ldots\ldots\, a_n\}$ *of M. Then*

$$N = RA = Ra_1 + Ra_2 + \ldots\ldots + Ra_n.$$

**Proof.** $N = RA$ is the smallest submodule of $M$ which contains $A$.

Now since $1 \in R$, so

$$1\, a_i = a_i \in R\, a_i,\ i = 1, 2, \ldots n.$$

and each $R\, a_i = < a_i >$ is a cyclic submodule of $M$.

Now, let $$N_i = R\, a_i,\ i = 1, 2, \ldots\, n.$$

Then $N_1 + N_2 + \ldots + N_n$ is just the submodule generated by the set $N \cup N_2 \cup \ldots\ldots \cup N_n$, and is the smallest submodule of $M$ containing $N_i$, $i = 1, 2, \ldots\, n$. But $N_1, N_2, \ldots N_n$ are generated by sets

$A_1 = \{a_1\}, A_2 = \{a_2\} \ldots, A_n = \{a_n\}$ respectively. So $N_1 + N_2 + \ldots\ldots + N_n$ is generated by $A_1 \cup A_2 \cup \ldots \cup A_n$ *i.e.* by the set $A = \{a_1, a_2, \ldots a_n\}$.

Hence $N = R\, A = N_1 + N_2 + \ldots + N_n$

$= Ra_1 + R\, a_2 + \ldots + R\, a_n.$

**Minimal generating set :**

A submodule $N$ of an R-module $M$ may have many different generating sets. If $N$ is finitely generating submodule then a generating set containing minimum number of elements is called a **minimal generating set** for $N$. The number of elements in a minimal generating set is called the rank of the submodule $N$.

**Theorem 18.** *Let R be an Euclidean ring. Then any finitely generated R-module N is the direct sum of a finite number of cyclic submodules.*

**Proof.** To prove the theorem we use mathematical induction on the rank of $N$. If the rank of $N$ is 1, then $N$ is generated by a single element, hence it is cyclic and theorem is proved. Let us assume that rank of $N$ is $n$. Suppose that the theorem is true for all R-modules of rank $(n-1)$ then each module of rank $(n-1)$ is direct sum of finite number of its cyclic submodules.

Now we shall prove that the theorem is also true for the module $N$ whose rank is $n$. Assume $\{a_1, a_2, \ldots a_n\}$ is a minimal generating set of module $N$ such that, if $t_1\, a_1 + t_2\, a_2 + \ldots + t_n\, a_n = 0$ implies that $t_1\, a_1 = t_2\, a_2 = \ldots = t_n\, a_n = 0$, for $t_i \in R$ then we see that $N$ is the direct sum of $N_1, N_2, \ldots N_n$ and each $N_i$ is a cyclic submodule generated by $a_i$. So in this case the theorem is true.

Now, let $\{a_1, a_2,...a_n\}$ be a minimal generating set for $N$ such that $t_1 a_1 + t_2 a_2 +... + t_n a_n = 0$ in which not all of $t_1 a_1, t_2 a_2,..., t_n a_n$ are zero.

Let in all such relations for minimal generating sets these exists $s_1 \in R$ occurring as a coefficient such that the Euclidean valuation $d(s_1)$ is the smallest positive integer.

Let it be $\{b_1, b_2,...b_n\}$.

So, that $s_1 b_1 + s_2 b_2 +... + s_n b_n = 0$ .....(1)

Since $R$ is a Euclidean ring and $s_1, r_1 \in R$, there exist $m_1, t \in R$

such that $r_1 = m_1 s_1 + t,$

where either $t = 0$ or $d(t) < d(s_1)$.

Multiplying equation (1) by $m_1$ and subtracting it from $r_1 b_1 + r_2 b_2 +... + r_n b_n = 0$. We obtain $t_1 b_1 + (r_2 - m_1 s_2) b_2 +... + (r_n - m_1 s_n) b_n = 0$. If $t \neq 0$, then $d(t) < d(s_1)$, which contradicts the fact that $d(s_1)$ is the smallest, so $t$ must be zero and $r_1 = m_1 s_1$, hence $s_1 \mid r_1$.

We also prove that $s_1 \mid s_i$ for $i = 2, 3, .....,n$. Since $s_1, s_2, \in R$ and $R$ is a Euclidean ring, so these exist $m_2, t \in R$ such that $s_2 = m_2 s_1 + t$, where either $t = 0$ or $d(t) < d(s_1)$.

Now $b' = b_1 + m_2 b_2, b_2, b_3,... b_n$ also generate $N$.

And

$$\begin{aligned} & s_1 b'_1 + t b_2 + s_3 b_3 + ... + s_n b_n \\ &= s_1(b_1 + m_2 b_2) + t b_2 + s_3 b_3 + ... + s_n b_n \\ &= s_1 b_1 + s_2 b_2 + s_3 b_3 + ... + s_n b_n \\ &= 0 \end{aligned}$$

So that, $t$ occurs as a coefficient in some relation for a minimal generating set. Hence if $t \neq 0$, then $d(t) < d(s_1)$ which contradicts the choice of $s_1$.

Therefore $t = 0$, and hence $s_1 \mid s_2$. Similarly, it can be shown for other $s_i$, *i.e.* $s_1 \mid s_i$, $i = 2, 3,...n$. We write $s_i = m_i s_1$ .

The following set $\{b_1^* = b_1 + m_2 b_2 + m_3 b_3 + ... + m_n b_n, b_2, b_3,... b_n\}$ generates the module $N$. If $N_1$ is the cyclic submodule generated by $b_1^*$ and if $N_2$ is the submodule generated by $b_2, b_3, ... b_n$ then $N = N_1 + N_2$ since $b_1^*, b_2, ... b_n$ generate $N$.

Let $x \in N_1 \cap N_2$ so $x \in N_1$ and $x = r_1 b_1^*$ for some $r_1 \in R$.

Also $x \in N_2$, hence $x = r_2 b_2 + ... + r_n b_n$, for $r_2, r_3, ... r_n \in R$

Thus $s_1 b_1^* + (-r_2) b_2 + ... (-s_n) b_n = 0,$

that is, $r_1 b_1 + (r_1 m_2 - r_2) b_2 + ... + (r_1 m_n - r_n) = 0$

Thus $s_1 \mid r_1$, *i.e.* $r_1 = t s_1$, for some $t \in R$.

Thus, $x = s_1 b_1^* = t s_1 b_1^*$

$= t [s_1 (b_1 + m_2 b_2 + ... + m_n b_n)]$

$= t (s_1 b_1 + s_2 b_2 + ... + s_n b_n)$

$= 0$

hence, $N_1 \cap N_2 = \{0\}$ and so, $N = N_1 \oplus N_2$

Again $N_2$ is generated by $b_2, b_3, ... b_n$, so its rank is $(n-1)$ and by the induction assumption $N_2$ is the direct sum of cyclic submodules.

Hence $N$ is the direct sum of a finite number of cyclic submodules.

### Self-learning exercise-1

State whether the following statements are true or false :

***(i)*** *(a)* There are two binary operations defined in an R-module.

*(b)* There is one internal and one external operations defined in an R-module.

*(c)* Sum of two submodules is a submodule of an R-module.

*(d)* Union and intersection of two submodules are always submodules of an R-module.

*(e)* A submodule generated by a subset A of an R-module $M$ is the smallest submodule of $M$ containing A.

## 5.11 Summary

In this unit we have studied a new algebraic structure called 'module'. One binary operation 'addition' is defined in it and an external mapping called scalar multiplication is defined to associate elements of a ring to the elements of the module. We also studied sum modules, homomorphism of modules, generation of submodules and cyclic submodules.

## 5.12 Answers to self-learning exercises

### Self-learning exercise-1

**1.** *(a)* False *(b)* True *(c)* True *(d)* False *(e)* True

## 5.13 Exercises

**1.** Define module. Prove that every abelian group $G$ is module over the ring of integers.

**2.** Show that a left ideal $M$ in a ring $R$ is an R-module.

**3.** Show that the range of homomorphism of an R-module $M$ is a submodule of $M$.

**4.** If $T : M \to N$ is homomorphism, then $T$ is isomorphism if and only if $K(T) = \{0\}$, where $M$ and $N$ are R-module.

**5.** Prove that any unital irreducible R-module is cyclic.

**6.** Prove that every ring $R$ is an R-module over itself.

**7.** Let $M$ be an R-module. If $x \in M$. Let $\lambda(x) = \{r \in R : rx = 0\}$. Show that $\lambda(x)$ is a left ideal of $R$.

**8.** Suppose that $R$ is a ring with unity and that $M$ is a module over $R$ but is not unital. Prove that these exists an $x \neq 0$ in $M$ such that $rx = 0$ for all $r \in R$.

**9.** If $\lambda$ is a left ideal of $R$ and if $M$ is an R-module, show that for $x \in M$, $\lambda(x) = \{rx : r \in \lambda\}$ is a submodule of $M$.

□□□

# UNIT 6 : Linear Transformation of Vector Spaces

**Structure of the Unit**



## 6.0 Objectives

In under graduate classes we have studied that a vector space is an additive abelian group, with a scalar multiplication, defined to associate a scalar from field to vector of the vector space. We also formulated the concept of homomorphism which we call linear transformation for vector spaces. In order to define a linear transformation between any two vector spaces, it is necessary to suppose that both vector spaces are defined over the same field. We know that a vector space is always defined over a field.

## 6.1 Introduction

In this unit we shall discuss the linear transformation of vector spaces and algebra of linear transformations, dual space, dual basis and their properties. We shall also obtain the relation between basis and its dual basis, linear map and its dual map.

## 6.2 Linear transformation

Let $V$ and $V'$ be vector spaces over the same field $F$. A function $t : V \to V'$ is said to be a linear transformation if it satisfies the following conditions :

***(i)*** $t(u + v) = t(u) + t(v), \quad \forall\, u, v \in V,$

***(ii)*** $t(\alpha u) = \alpha t(u), \quad \forall\, u \in V$ and $\alpha \in F.$

In other words, a function $t : V \to V'$ is said to be a linear transformation if it preserves the following two basic operations on vector spaces :

***(i)*** vector addition,

***(ii)*** scalar multiplication.

However, the conditions *(i)* and *(ii)* are equivalent to the following single condition

$$t(\alpha u + \beta v) = \alpha t(u) + \beta t(v) \quad \forall\, u \in V \text{ and } \alpha, \beta \in F. \qquad .....(1)$$

Generalising the equation (1), we get

$$t(\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n) = \alpha_1(v_1)\, \alpha_2(v_2) + ... + \alpha_n t(v_n)$$

$$= \sum_{i=1}^{n} \alpha_i t(v_i)$$

$\forall\, v \in V$ and $\alpha_i, \in F,\ i = 1, 2, .... n.$

The following have the same meaning : linear transformation, linear mapping, vector space homomorphism, linear function.

**Note :**

1. It $\alpha = 0$, then we have

   $$t(0u) = 0\ t(u) = \mathbf{0}'$$

   Thus $\quad t(\mathbf{0}) = \mathbf{0}'$

   It shows that a linear transformation maps the zero vector into zero vector, $\mathbf{0}'$ being the zero vector of $V'$.

2. If $t : V \to V'$ is onto, then $V'$ is called a homomorphic image of $V$.

3. Since $F$ is a vector space over it self, a linear transformation $t$ from $V$ to $F$, is called a linear function, and when a linear transformation $t$ from $V$ to $V$ itself, it is called a linear operator.

4. A linear transformation $t : V \to V'$ is called an isomorphism of $V$ onto $V'$, if the map $t$ is bijective i.e. one-one and onto.

5. Two vector space $V$ and $V'$ over the same field $F$ are called isomorphic, if there exists an isomorphism of $V$ onto $V'$, we denote it by $V \cong V'$.

6. If $t$ from $V$ to $V'$ is an isomorphism then it is easy to see that these exist an isomorphism $t^{-1}$(inverse of $t$) of $V'$ onto $V$.

The other similar terms, such as monomorphism, epimorphism, endomorphism and automorphism are also analogous to the corresponding concepts in groups and modules and have their usual meaning.

### 6.2.1 Kernel of a linear transformation :

Let $V$ and $V'$ be two vector spaces over the same field $F$ and $t : V \to V'$ be a linear transformation. Then the Kernel of $t$ written as Ker $(t)$ and is defined as the set of all vectors of $V$ which are mapped onto the zero vector $\mathbf{0}'$ of $V'$, that is

$$\text{Ket } (t) = \{u \in V' : t\ (u) = \mathbf{0}' \in V'\}$$

Ket$(t)$ is also known as null space of $t$.

### 6.2.2 Image of a linear transformation :

Let $V$ and $V'$ be any two vector spaces over $a$ field $F$ and $t : V \to V'$ be a linear transformation. The range of $t$, written as $im\ (t)$, is the set of all vectors of $V'$, which are the images of all the vectors of $V$, that is

$$im\ (t) = \{t\ (u) \in V' : u \in V\}$$

$im\ (t)$ is also known as range space of $t$.

***Ex.*1.** *Show that the following mapping is linear* $t : R^3 \to R^2$ *given by*

$$t\ (x, y, z) = (z, x + y), \quad \forall\ (x, y, z) \in R^3$$

**Sol.** Given that $t : R^3 \to R^2$ such that

$$t\ (x, y, z) = (z, x + y), \quad \forall\ (x, y, z) \in R^3$$

Let $u = (x, y, z), v = (x_1, y_1, z_1) \in R^3,$

then $u + v = (x + x_1, y + y_1, z + z_1), ku = (kx, ky, kz), k \in R.$

So that
$$\begin{aligned} t\ (u + v) &= (z + z_1, x + x_1, y + y_1) \\ &= (z + x + y) + (z_1, x_1 + y_1) \\ &= t\ (x, y, z)\ t\ (x_1, y_1, z_1) \\ &= t\ (u) + t\ (v), \end{aligned}$$

and
$$\begin{aligned} t\ (ku) &= (kz, kx + ky) \\ &= k\ (z, x + y) \\ &= kt\ (x, y, z) \\ &= kt\ (u) \end{aligned}$$

Hence $t$ is linear.

***Ex.*2.** *Show that the following are not linear*

***(i)*** $t : R^2 \to R^2$ *defined by* $t(x, y) = (x^3, y^3)$

***(ii)*** $t : R^2 \to R$ *defined by* $t(x, y) = |\ x - y\ |$

**Sol.** *(i)* The map $t : R^2 \to R^2$ is given by

$$t\ (x, y) = (x^3, y^3)$$

Let $u = (x_1, y_1), v = (x_2, y_2) \in R^2,$

then $u + v = (x_1 + x_2, y_1 + y_2)$

so that $t(u+v) = t(x_1 + x_2, y_1 + y_2)$

$t(u+v) = ((x_1 + x_2)^3, (y_1 + y_2)^3)$

thus $t(u+v) \neq t(u) + t(v)$

which shows that $t$ is not linear.

*(ii)* A map $t : R^2 \to R$ is given by

$$t(x, y) = |x - y|$$

Let $u = (x_1, y_1),\ v = (x_2, y_2) \in R^2$

then $t(u+v) = t(x_1 + x_2,\ y_1 + y_2)$

$= |(x_1 + x_2) - (y_1 + y_2)|$

$= |(x_1 - y_1) + (x_2 - y_2)|$

$\neq |x_1 - y_1| + |x_2 - y_2|$

Thus $t(u+v) \neq t(u) + t(v)$,

and so $t$ is not linear.

**Theorem 1.** *Let $t : V \to V'$ be a linear transformation. Then*

*(i) Ker $(t)$ is a vector sub space of $V$, and*

*(ii) im $(t)$ is a vector sub space of $V'$.*

**Proof :** *(i)* Given that $t : V \to V'$ be a linear transformation, then from the definition of Ker $(t)$, we have

$$\text{Ker}(t) = \{u \in V : t(u) = \mathbf{0} \in V'\}$$

where $\mathbf{0}$ is the zero vector of vector space $V'$.

Since $t(\mathbf{0}) = \mathbf{0} \in V'$

$\Rightarrow$ $\mathbf{0} \in \text{Ker}(t)$ [By definition of Ker $(t)$]

$\Rightarrow$ Ker$(t)$ is a non-empty subset of $V$.

Let $u, v \in \text{Ker}(t),\ \alpha, \beta \in F$, then

$t(u) = \mathbf{0},\ \ t(v) = \mathbf{0}$

Now, $t(\alpha u + \beta v) = \alpha t(u) + \beta t(v)$ [$\because$ $t$ is linear]

$= \alpha\mathbf{0} + \beta\mathbf{0}$

$= \mathbf{0} \in V'$

Thus $\alpha u + \beta v \in \text{Ker}(t),\ \forall u, v \in \text{Ker}(t),\ \alpha, \beta \in F$

which shows that Ker $(t)$ is a vector sub space of $V$.

*(ii)* From the definition of image space im $(t)$, we have

$$im(t) = \{t(u) \in V' : u \in V\}.$$

Since $t(\mathbf{0}) = \mathbf{0} \in V'$, when ever $\mathbf{0} \in V$ so that $im(t)$ is a non-empty sub set of $V'$.

Now, let $u', v' \in im(t)$ and $\alpha, \beta \in F$

Then there exist $u$ and $v$ in $V$ such that

$$u' = t(u), \quad v' = t(v)$$

Now, $\alpha u' + \beta v' \in V$, and so

$$\alpha u' + \beta v' = \alpha t(u) + \beta t(v)$$
$$= t(\alpha u + \beta v) \in im(t)$$

So, $\alpha u' + \beta v' \in im(t)$, $\forall\, u', v' \in im(t)$, and $\alpha, \beta \in F$.

Thus $im(t)$ is a vector sub space of $V'$.

**Theorem 2.** *Let $t : V \to V'$ be a linear transformation, then*

*(i) $t$ is monomorphism if and only if $Ker(t) = \{\mathbf{0}\}$,*

*(ii) If the set $\{v_1, v_2,..., v_n\}$ is linearly dependent then the set $\{(v_1), t(v_2),..., t(v_n)\}$ is also linearly dependent.*

*(iii) If the set $\{(v_1), t(v_2),..., t(v_n)\}$ is linearly independent then the set $\{v_1, v_2,..., v_n\}$ is linearly independent.*

*(iv) If the set $\{v_1, v_2,..., v_n\}$ spans $V$ then the set $\{(v_1), t(v_2),..., t(v_n)\}$ spans $V'$.*

**Proof : *(i)*** First suppose that $t : V \to V'$ be a monomorphism, *i.e.* $t$ one-one

To prove that

$$\text{Ker}(t) = \{\mathbf{0}\}.$$

Let $u \in \text{Ker}(t)$ be an arbitrary vector, and so

$$t(u) = \mathbf{0} \in V'$$

$\Rightarrow$ $t(u) = t(\mathbf{0})$ $\quad [\because\ (t) = \{\mathbf{0}\}]$

$\Rightarrow$ $u = \mathbf{0}$ $\quad [\because\ t \text{ is one-one}]$

Thus $\text{Ker}(t) = \{\mathbf{0}\}$.

Conversely suppose that $\text{Ker}(t) = \{\mathbf{0}\}$.

To prove that $t$ is monomorphism.

Let $u, v \in V$ such that $t(u) = t(v)$

$\Rightarrow$ $t(u) - t(v) = \mathbf{0}$

$\Rightarrow$ $t(u - v) = \mathbf{0}$

$\Rightarrow$ $u - v \in \text{Ker}(t)$

$\Rightarrow$ $u - v = \mathbf{0}$ $\quad [\because\ \text{Ker}(t) = \{\mathbf{0}\}]$

$\Rightarrow$ $u = v$

Thus $t$ is monomorphism.

***(ii)*** Given that $\{v_1, v_2,..., v_n\}$ be linearly dependent set, then there exist scalars $\alpha_1, \alpha_2,..., \alpha_n \in F$, not all zero, such that

$$\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n = \mathbf{0}$$

$\Rightarrow$ $t(\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n) = t(\mathbf{0})$

$\Rightarrow \quad \alpha_1 t(v_1) + \alpha_2 t(v_2) + ... + \alpha_n t(v_n) = \mathbf{0}$ [$\because$ $t$ is linear]

which show that $\{t(v_1), t(v_2),..., t(v_n)\}$ is linearly dependent set.

***(iii)*** Given that $\{t(v_1), t(v_2),..., t(v_n)\}$ is linearly independent set. To prove that $\{v_1, v_2,..., v_n\}$ is also linearly independent set. Let there exist scalars $\alpha_1, \alpha_2 ..... \alpha_n \in F$, such that

$$\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n = \mathbf{0}$$

$\Rightarrow \quad t(\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n) = t(\mathbf{0})$

$\Rightarrow \quad \alpha_1 t(v_1) + \alpha_2 t(v_2) + ... + \alpha_n t(v_n) = \mathbf{0}$

$\Rightarrow \quad \alpha_1 = \alpha_2 = ... = \alpha_n = \mathbf{0}$ [$\because$ $\{t(v_1), t(v_2),..., t(v_n)\}$ in linearly independent]

Thus $\{v_1, v_2,..., v_n\}$ is linearly independent set.

***(iv)*** Given that $\{v_1, v_2,.., v_n\}$ spans $V$, so for any $v \in V$, these exist scalars $\alpha_1, \alpha_2,..., \alpha_n \in F$ such that

$$v = \alpha_1 v_1 + \alpha_2 v_2 + .....+ \alpha_n v_n$$

$\Rightarrow \quad t(v) = (\alpha_1 v_1 + \alpha_2 v_2 + .....+ \alpha_n v_n)$

$$= \alpha_1 t(v_1) + \alpha_2 t(v_2) + .....+ \alpha_n t(v_n) \quad [\because \ t \text{ is linear}]$$

which shows that each $t(v) \in im(t)$ is in the linear combination of vectors $t(v_1), t(v_2),..., t(v_n)$.

Hence $\{t(v_1), t(v_2),..., t(v_n)\}$ spans $im(t)$.

**Theorem 3.** *Let $V$* and *$V'$ be vector spaces over a field $F$ and $B = \{b_1, b_2,..., b_n\}$ be a basis for $V$. Then there exists a unique linear transformation, $t : V \to V'$ for any list $b_1', b_2', ..., b_n'$ of vectors in $V'$, such that*

$$t(b_i) = b_i', \qquad i = 1, 2,..., n.$$

**Proof :** Given that $B = \{b_1, b_2,..., b_n\}$ be basis for $V$, so each vector $v \in V$, can be uniquely expressed as

$$v = \alpha_1 b_1 + \alpha_2 b_2 + ... + \alpha_n b_n, \quad \text{for} \quad \alpha_i \in F, i = 1, 2,..., n$$

Now, we define a map $t : V \to V'$ by

$$t(v) = t(\alpha_1 b_1 + \alpha_2 b_2 + ... + \alpha_n b_n), \qquad \forall\, v \in V$$

$$= \alpha_1 b_1' + \alpha_2 b_2' + ... + \alpha_n b_n'$$

$$= \sum_{i=1}^{n} \alpha_i b_i'$$

Let $$u = \sum_{i=1}^{n} \alpha_i b_i, \; w = \sum_{i=1}^{n} \beta_i b_i, \text{ for some } \alpha_i, \beta_i \in F, i = 1, 2,..., n$$

and $\lambda, \mu \in F$

Now $$t(\lambda u + \mu w) = t\left(\lambda \sum_{i=1}^{n} \alpha_i b_i + \mu \sum_{i=1}^{n} \beta_i b_i, \right)$$

$$= t\left(\sum_{i=1}^{n} (\lambda\alpha_i + \mu\beta_i) b_i\right)$$

$$= \sum_{i=1}^{n} (\lambda\alpha_i + \mu\beta_i) b_i'$$

$$= \lambda \sum_{i=1}^{n} \alpha_i\, b_i' + \mu \sum_{i=1}^{n} \beta_i b_i'$$

$$= \lambda t(u) + \mu t(w)$$

Thus $t$ is a linear transformation.

Next, for each $b_i \in B$, we have

$$t(b_i) = t(0 \cdot b_1 + 0 \cdot b_2 + \ldots + 1 \cdot b_i + \ldots + 0 \cdot b_n)$$

$$= 0 \cdot b_1' + 0 \cdot b_2' + \ldots + 1 \cdot b_i' + \ldots + 0 \cdot b_n'$$

$$t(b_i) = b_i', \qquad i = 1, 2, \ldots, n$$

Thus $t(b_i) = b_1', t(b_2) = b_2', \ldots, t(b_n) = b_n'$.

**Uniqueness of linear transformation**

If possible, suppose $s : V \to V'$ be a another linear transformation such that

$$s(b_i) = b_i', \qquad i = 1, 2, \ldots\ldots, n$$

Now, $\forall\, v \in V$, we have

$$s(v) = (\alpha_1 b_1 + \alpha_2 b_2 + \ldots + \alpha_n b_n)$$

$$= \alpha_1 s(b_1) + \alpha_2 s(b_2) + \ldots + \alpha_n s(b_n)$$

$$= \alpha_1 b_1' + \alpha_2 b_2' + \ldots + \alpha_n b_n'$$

$$= t(v), \qquad \forall\, v \in V$$

Thus $s = t$

Hence $t$ is a unique linear transformation such that

$$t(b_i) = b_i', \qquad i = 1, 2, \ldots, n.$$

**Theorem 4.** *Let $V$ and $V'$ be any two vector spaces over the same field $F$ and $B = \{b_1, b_2, \ldots, b_n\}$ be a basis for $V$ and $B' \{b_1', b_2', \ldots, b_n'\}$ be a set of vectors in $V'$. If $t : V \to V'$ be a linear transformation such that*

$$t(b_i) = b_i', \qquad i = 1, 2, \ldots\ldots, n.$$

*Then $t$ is an isomorphism if and only if the set $B'$ is a basis for $V'$.*

**Proof :** We define a map $t : V \to V'$ by

$$t(v) = t\left(\sum_{i=1}^{n} \alpha_i\, b_i\right)$$

$$= \sum_{i=1}^{n} \alpha_t \, b_t', \qquad \forall\, v = \alpha_1 b_1 + .....+ \alpha_n b_n \in V,$$

for some $\alpha_1, \alpha_2, ..., \alpha_n \in F$.

Then it is clear that $t$ is a linear transformation such that

$$t(b_i) = b_i', \qquad i = 1, 2, ....., n.$$

Now first suppose that $t$ is an isomorphism. To prove that $B'$ is a basis for $V'$, we show

***(i) $B'$ is linearly independent :***

Let there exist scalars $\alpha_1, \alpha_2, ..., \alpha_n \in F$

such that $\alpha_1 b_1' + \alpha_2 b_2' + ... + \alpha_n b_n' = \mathbf{0}$

$$\alpha_1 t(b_1) + \alpha_2 t(b_2) + ... + \alpha_n t(b_n) = t(\mathbf{0})$$

$\Rightarrow$ $t(\alpha_1 b_1 + \alpha_2 b_2 + .....+ \alpha_n b_n) = t(\mathbf{0})$

$\Rightarrow$ $\alpha_1 b_1 + \alpha_2 b_2 + .....+ \alpha_n b_n = \mathbf{0}$ [$\because$ $t$ is an isomorphism]

$\Rightarrow$ $\alpha_1 = \alpha_2 = .....= \alpha_n = 0$ [$\because$ $B = \{b_1, b_2, ..., b_n\}$ is a basis for $V$]

Thus $B' = \{b_1', b_2', ..., b_n'\}$ linearly independent.

***(ii) $B'$ spans $V'$:***

Since $t$ is onto, so that for each $v' \in V'$, there exists $v \in V$ such that

$$v' = t(v)$$

$$= t(\alpha_1 b_1 + \alpha_2 b_2 + ... + \alpha_n b_n)$$

where $v = \alpha_1 b_1 + \alpha_2 b_2 + ... + \alpha_n b_n \in V$,

for some $\alpha_1, \alpha_2, ..., \alpha_n \in F$,

$\therefore$ $v' = \alpha_1 t(b_1) + \alpha_2 t(b_2) + ... + \alpha_n t(b_n)$

$$= \alpha_1 b_1' + \alpha_2 b_2' + ... + \alpha_n b_n'$$

Thus each vector $v' \in V'$ is a linear combination of vectors of $B'$. Which shows that $B'$ spans $V'$.

Hence $B' = \{b_1', b_2', ..., b_n'\}$ is a basis for $V'$.

Conversely assume that $B'$ is a basis for $V'$. To prove that $t$ is an isomorphism, we show

***(i) $t$ is one-one :***

Let $v = \sum_{i=1}^{n} \alpha_i \, b_i \in V$, such that

$v \in \text{Ker}\,(t)$

$\Leftrightarrow$ $t(v) = \mathbf{0}$

$\Leftrightarrow$ $t\left(\sum_{i=1}^{n} \alpha_i \, b_i\right) = \mathbf{0}$

$\Leftrightarrow \quad \sum_{i=1}^{n} \alpha_i \, t(b_i) = \mathbf{0}$

$\Leftrightarrow \quad \sum_{i=1}^{n} \alpha_i \, b_i' = \mathbf{0}$

$\Leftrightarrow \quad a_i = 0, \quad i = 1, 2, ..., n \qquad [\because \ B' \text{ is a basis for } V']$

$\Leftrightarrow \quad v = \mathbf{0}$

So Ker ($t$) = {**0**} and thus $t$ is one-one

***(ii) t* is onto :**

Let $v' = \beta_1 b_1' + \beta_2 b_2' + ... \beta_n b_n' \in V'$

$\Rightarrow \quad v' = t(\beta_1 b_1 + \beta_2 b_2 + ... + \beta_n b_n) = t(v)$,

when $v = \beta_1 b_1 + ..... + \beta_n b_n \in V$ for some $\beta_1, ..., \beta_n \in F$.

Thus $t$ is onto.

Hence $t$ is an isomorphism.

**Theorem 5.** *Let $V$ be an $n$-dimensional vector space over a field $F$. Prove that $V$ is isomorphic to the vector space $F^n$, hence also show that any two finite dimensional vector spaces of the same dimension are isomorphic.*

**Proof :** Let $B = \{b_1, b_2, ..., b_n\}$ be a basis for $V$, so that each vector $v \in V$ can be expressed uniquely as

$$v = \alpha_1 b_1 + ... + \alpha_n b_n, \quad \text{for} \quad \alpha_i \in F, \ i = 1, 2, ..., n.$$

Now we define a map $t : V \to F^n$ as

$$t(v) = t(\alpha_1 b_1 + \alpha_2 b_2 + ... + \alpha_n b_n)$$
$$= (\alpha_1, \alpha_2, ..., \alpha_n) \in F^n, \quad \forall \ v \in V,$$

To prove that $t$ is an isomorphism.

***(i) t* is linearly transformation :**

Let $u, v \in V$ and $\lambda, \mu \in F$,

then $u = \alpha_1 b_1 + ..... + \alpha_n b_n$,

and $v = \beta_1 b_1 + ..... + \beta_n b_n, \quad \text{for} \quad \alpha_i, \beta_i \in F, \ i = 1, 2, ....., n$

So, $t(\lambda u + \mu v) = [\lambda(\alpha_1 b_1 + ..... + \alpha_n b_n) + \mu (\beta_1 b_1 + ..... + \beta_n b_n)]$

$= t[(\lambda\alpha_1 + \mu\beta_1) b_1 + ..... + (\lambda\alpha_n + \mu\beta_n) b_n]$

$= (\lambda\alpha_1 + \mu\beta_1, ....., \lambda\alpha_n + \mu\beta_n)$

$= \lambda (\alpha_1, ....., \alpha_n) + \mu (\beta_1, ....., \beta_n)$

$= \lambda t (\alpha_1 b_1 + ..... + \alpha_n b_n) + \mu (\beta_1 b_1 + ..... + \beta_n b_n)$

$= \lambda t(u) + \mu t(v)$

Thus $t$ is a linear transformation.

Now first suppose that $t$ is an isomorphism. To prove that $B'$ is a basis for $V'$, we show

***(ii) $t$ is one-one :***

Let $u = \alpha_1 b_1 + ... + \alpha_n b_n$,

and $v = \beta_1 b_1 + ... + \alpha_n b_n \in V$,

$\alpha_i, \beta_i \in F,\ i = 1, 2, ..., n$

be such that $t(u) = t(v)$

$\Rightarrow \quad t(\alpha_1 b_1 + ... + \alpha_n b_n) = t(\beta_1 b_1 + ... + \beta_n b_n)$

$\Rightarrow \quad (\alpha_1, \alpha_2, ..., \alpha_n) = (\beta_1, \beta_2, ..., \beta_n)$

$\Rightarrow \quad \alpha_1 = \beta_1, \alpha_2 = \beta_2, ..., \alpha_n = \beta_n$

$\Rightarrow \quad u = v$

thus $t$ is one-one.

***(iii) $t$ is onto :***

For each $(\alpha_1, \alpha_2, ..., \alpha_n) \in F^n$, then there exist $v = \alpha_1 b_1 + ... + \alpha_n b_n \in V$ such that

$$(\alpha_1, \alpha_2, ..., \alpha_n) = t(v)$$

which shows that $t$ is onto.

Hence $t : V \to F^n$ is an isomorphism and thus

$$V \cong F^n.$$

Now let $V$ and $V'$ be any two vector spaces of the same dimension $n$, then

$$V \cong F^n \quad \text{leads to} \quad V' \cong F^n$$

Since the relation of an isomorphism in an equivalence relation in vector spaces,

so $V \cong V'$

Hence any two finite dimensional vector spaces of the same dimension are isomorphic.

**Theorem 6.** *Let $V$ be a finite dimensional vector space over a field $F$ and $B = \{v_1, v_2, ..., v_n\}$ be a set of vectors is $V$. Then a map*

$$t : F^n \to V \text{ such that } t(\alpha_1, \alpha_2, ..., \alpha_n) = \alpha_1 v_1 + ... + \alpha_n v_n \ \forall\ (\alpha_1, \alpha_2, ..., \alpha_n) \in F^n,$$

*is a linear transformation, and*

*(i) $t$ is monomorphism iff $B$ is linearly independent,*

*(ii) $t$ is an epimorphism iff $B$ spans $V$, and*

*(iii) $t$ is an isomorphism iff $B$ is a basis for $V$.*

**Proof :** Let $\{\alpha_1, \alpha_2, ..., \alpha_n\}, \{\beta_1, \beta_2, ..., \beta_n\} \in F^n$, and $\lambda, \mu \in F$, then we have

$$\begin{aligned} & t[\lambda(\alpha_1, \alpha_2, ..., \alpha_n) + \mu(\beta_1, \beta_2, ..., \beta_n)] \\ &= t(\lambda\alpha_1 + \mu\beta_1, \lambda\alpha_2 + \mu\beta_2, ..., \lambda\alpha_n + \mu\beta_n) \\ &= (\lambda\alpha_1 + \mu\beta_1) v_1 + ... + (\lambda\alpha_n + \mu\beta_n) v_n \\ &= \lambda(\alpha_1 v_1 + ... + \alpha_n b_n) + \mu(\beta_1 v_1 + ... + \beta_n v_n) \\ &= \lambda t(\alpha_1, \alpha_2, ..., \alpha_n) + \mu t(\beta_1, \beta_2, ..., \beta_n) \end{aligned}$$

Thus $t$ is a linear transformation

***(i)*** First suppose that $t$ is a monomorphism. To prove that the set $B$ is linearly independent, let there exist scalars $\alpha_1, \alpha_2,..., \alpha_n \in F$ such that

$$\alpha_1 v_1 + ... + \alpha_n v_n = \mathbf{0}$$

$\Rightarrow$ $t(\alpha_1, \alpha_2,..., \alpha_n) = t(0, 0,..., 0)$

$\Rightarrow$ $(\alpha_1, \alpha_2,..., \alpha_n) = (0, 0,..., 0)$ $[\because$ $t$ is monomorphism$]$

$\Rightarrow$ $\alpha_1 = 0, \alpha_2 = 0,..., \alpha_n = 0$

Hence $B = \{v_1, v_2,..., v_n\}$ is linearly independent.

Conversely suppose that $B$ is linearly independent. To prove that $t$ is monomorphism, let $(\alpha_1, \alpha_2,..., \alpha_n)$ and $(\beta_1, \beta_2,..., \beta_n) \in F^n$ such that

$$t(\alpha_1, \alpha_2,..., \alpha_n) = t(\beta_1, \beta_2,..., \beta_n)$$

$\Rightarrow$ $\alpha_1 v_1 + ... + \alpha_n v_n = \beta_1 b_1 + ... + \beta_n b_n$

$\Rightarrow$ $(\alpha_1 - \beta_1) v_1 + ... + (\alpha_n - \beta_n) v_n = \mathbf{0}$

$\Rightarrow$ $\alpha_1 - \beta_1 = \alpha_2 - \beta_2 = ... = \alpha_n - \beta_n = 0$

$[\because$ $B = \{v_1, v_2,..., v_n\}$ is linearly independent$]$

So, $\alpha_1 = \beta_1, \alpha_2 = \beta_2,..., \alpha_n = \beta_n$

Thus $(\alpha_1, \alpha_2,..., \alpha_n) = (\beta_1, \beta_2,..., \beta_n)$

Hence $t$ is monomorphism.

***(ii)*** First suppose that $t$ is an epimorphism. To prove that the set $B$ spans $V$.

Since $t$ is an epimorphism (*i.e.* onto), so that for each vector $v \in V$ there exists some $(\alpha_1, \alpha_2,..., \alpha_n) \in F^n$ such that

$$v = t(\alpha_1, \alpha_2,..., \alpha_n)$$
$$= \alpha_1 v_1 + ... + \alpha_n v_n$$

which shows that each $v \in V$ is a linear combination of vectors of $B$.

Thus the set $B$ spans $V$.

Conversely suppose that $B$ spans $V$, *i.e.* each vector $v \in V$ is a linear combination of vectors of $B$, so that there exist scalars $\alpha_1, \alpha_2,..., \alpha_n \in F$ such that

$$v = \alpha_1 v_1 + ... + \alpha_n v_n$$

$\Rightarrow$ $v = t(\alpha_1, \alpha_2,..., \alpha_n)$, for $(\alpha_1, \alpha_2,..., \alpha_n) \in F^n$

Thus $t$ is an epimorphism.

***(iii)*** From *(i)* and *(ii)*, it is clear that $t$ is an isomorphism iff the set $B$ is a basis for $V$.

## Self-learning exercise-1

**1.** Take the correct one :

***(i)*** If $t : V \to V'$ is the zero homomorphism (linear transformation), then $im(t)$ is equal to :

*(a)* $\{\mathbf{0}\}$ *(b)* $V$

*(c)* $V'$ *(d)* none of these

***(ii)*** If $t : R^3 (R) \to R^2 (R)$ is defined by $t(x_1, x_2, ....., x_3) = (x_1, x_2)$, $\forall (x_1, x_2, ....., x_3) \in R^3$, then which of the following alternative is not true for $t$ ?

($a$) $t$ is linear transformation ($b$) $t$ is an isomorphism

($c$) Range $(t) = R^2$ ($d$) Ker$(t) = \{(0, 0, x_3) : x_3 \in R\}$.

**2.** Fill in the blanks

*(i)* A linear transformation is also known as ..... .

*(ii)* If $t : V \to V'$ is a linear transformation, then $t(u + v)$ = ..... .

*(iii)* If $t : V \to V'$ is a linear transformation, then Ker$(t)$ is a sub space of ..... , and $im(t)$ is a sub space of ..... .

*(iv)* If $t : V \to V'$ is monomorphism iff Ker$(t)$ = ..... .

*(v)* If $V$ and $V'$ are finite dimensional vector spaces over the same field $F$, then $V \cong V'$ iff ..... .

**3.** State whether the following statements are true or false :

*(i)* The identity map on a vector space is a linear transformation.

*(ii)* A linear transformation maps zero to zero.

*(iii)* Sum of two linear transformation is a linear transformation.

## 6.3 Dual spaces

Let $V$ be a vector space over a field $F$. Since $F$ it self is a vector space over $F$, then a linear transformation on $V$ to $F$ is called a linear functional on $V$.

If $V$ is a vector space over a field $F$, then its dual space is the vector space $Hom(V, F)$ which is consisting of all linear functionals on $V$. The dual space of $V$ is denoted by $V^*$.

**Theorem 7.** *Let $V$ be n-dimensional vector space over a field $F$ and $B = \{b_1, b_2, ....., b_n\}$ be a basis for $V$, then for any $n$ scalars $\lambda_1, \lambda_2, ..., \lambda_n \in F$ there exists a unique linear functional $f \in V^*$ such that*

$$f(b_i) = \lambda_i, \qquad i = 1, 2, ..., n.$$

**Proof :** We define a function $f : V \to F$ by

$$f(v) = f\left(\sum_{i=1}^{n} \alpha_i b_i\right)$$

$$= \sum_{i=1}^{n} \alpha_i \lambda_i, \qquad \text{for } \lambda_i \in F,$$

and $v = \alpha_1 b_1 + .....+ \alpha_n b_n \in V$, for $\alpha_i \in F$, $i = 1, 2, ..., n$.

Now, suppose, $u = \sum_{i=1}^{n} \alpha_i b_i$, $v = \sum_{i=1}^{n} \beta_i b_i \in V$, $\alpha_i, \beta_i \in F$, $i = 1, 2, ..., n$

and $\lambda, \mu \in F$, we have

$$t(\lambda u + \mu v) = f\left[\lambda \sum_{i=1}^{n} \alpha_i\, b_i + \mu \sum_{i=1}^{n} \beta_i\, b_i\right]$$

$$= f\left[\sum_{i=1}^{n} (\lambda\, \alpha_i + \mu\, \beta_i) b_i\right]$$

$$= \sum_{i=1}^{n} (\lambda\, \alpha_i + \mu\, \beta_i)\lambda_i$$

$$= \lambda \sum_{i=1}^{n} \alpha_i\, \lambda_i + \mu \sum_{i=1}^{n} \beta_i\, \lambda_i$$

$$= \lambda f(u) + \mu f(v)$$

Thus $f$ is a linear functional and $f \in V^*$.

Now, $b_i = 0 \cdot b_1 + \ .....+ 1 \cdot b_i + ..... + 0 \cdot b_n$

$f(b_i) = f(0 \cdot b_1 + \ .....+ 1 \cdot b_i + ..... + 0 \cdot b_n)$

$f(b_i) = \lambda_i, \qquad i = 1, 2, ....., n.$

**Uniqueness of linear functional :**

If possible suppose that $s : V \to F$ be a linear functional such that

$$s(b_i) = \lambda_i, \qquad i = 1, 2,..., n.$$

Now for all $v \in V$, we have

$$s(v) = s\left(\sum_{i=1}^{n} \alpha_i\, b_i\right)$$

$$= \sum_{i=1}^{n} \alpha_i\, s(b_i) \qquad [\because \ s \text{ is linear functional}]$$

$$= \sum_{i=1}^{n} \alpha_i\, \lambda_i$$

$$= f(v)$$

Thus $s = f$

Hence $f$ is a unique linear functional such that

$$f(b_i) = \lambda_i, \qquad i = 1, 2,..., n.$$

**Theorem 8.** *Let $V$ be a vector space over a field $F$ and $B = \{b_1, b_2,..., b_n\}$ be a basis for $V$. Then the dual space $V^*$ has a basis $B^* = \{f_1, f_2,..., f_n\}$ such that*

$$f_i(b_j) = \delta_{ij}; \qquad i, j = 1, 2,..., n$$

*where $\delta_{ij} \in F$ is a Kronecker delta.*

**Proof :** Let $v \in V$, then

$$v = \lambda_1\, b_1 + ... + \lambda_n\, b_n, \quad \text{where } \lambda_1, \lambda_2,..., \lambda_n \in F$$

Now we define a function $f_i : V \to F$ by

$$f_i(v) = f_i(\lambda_1 b_1 + ... + \lambda_n b_n)$$

and $$= \lambda_i, \qquad i = 1, 2, ..., n.$$

To show that $f_i$ is a linear functional :

Let $$u = \sum_{i=1}^{n} \alpha_i \, b_i$$

and $$v = \sum_{i=1}^{n} \beta_i \, b_i \in V, \quad \text{for} \quad \alpha_i, \beta_i \in F, \, i = 1, 2, ..., n$$

Also let $\lambda, \mu \in F$, then we have

$$f_i(\lambda u + \mu v) = f_i\left[\lambda \sum_{i=1}^{n} \alpha_i \, b_i + \mu \sum_{i=1}^{n} \beta_i \, b_i\right]$$

$$= f_i\left[\sum_{i=1}^{n} (\lambda\alpha_i + \mu\beta_i) \, b_i\right]$$

$$= \lambda\alpha_i + \mu\beta_i$$

$$= \lambda f_i \sum_{i=1}^{n} (\alpha_i \, b_i) + \mu f_i \sum_{i=1}^{n} (\beta_i \, b_i)$$

$$= \lambda f_i(u) + \mu f_i(v)$$

Thus $f_i$ is a linear functional and so $f_1, f_2, ....., f_n \in V^*$.

And since $$b_j = 0 \cdot b_1 + .....+ 1 \cdot b_j + ..... + 0 \cdot b_n,$$

we have $$f_i(b_j) = \delta_{ij}, \qquad i, j = 1, 2, ....., n.$$

Now we shall show that $B^* = \{f_1, f_2, ....., f_n\}$ is a basis for dual space $V^*$.

Since dim $V$ = dim $V^*$ = $n$, so in order to prove that $B^*$ is a basis for $V^*$, it is sufficient to show that $B^*$ is linearly independent set.

Let there exist scalars $\alpha_1, \alpha_2, ....., \alpha_n \in F$ such that $\alpha_1 f_1 + \alpha_2 f_2 + ..... + \alpha_n f_n = \hat{0}$, where $\hat{0}$ is a zero map

$$\Rightarrow \qquad \sum_{i=1}^{n} \alpha_i \, f_i = \hat{0}$$

$$\Rightarrow \qquad \left(\sum_{i=1}^{n} \alpha_i \, f_i\right)(b_j) = \hat{0}(b_j), \qquad j = 1, 2, ..., n$$

$$\Rightarrow \qquad \sum_{i=1}^{n} \alpha_i \, f_i(b_j) = 0$$

$$\Rightarrow \qquad \sum_{i=1}^{n} \alpha_i \, \delta_{ij} = 0$$

$\Rightarrow \qquad \alpha_j = 0; \qquad j = 1, 2,..., n.$

Hence $B = \{f_1, f_2,..., f_n\}$ is linearly independent and thus $B^*$ is a basis for $V^*$.

### 6.3.1 Dual basis

Let $B = \{b_1, b_2,..., b_n\}$ be a basis of a vector space $V(F)$ , then the basis $B^* = \{f_1, f_2,..., f_n\}$ of dual space $V^*$ defined by

$$f_i(b_i) = \delta_{ij}, \qquad i, j = 1, 2,..., n$$

is called the dual basis of $B$.

**Theorem 9.** *Let $V$ be a finite dimensional vector space over a field $F$, then for each non-zero vector $v \in V$, there exists a linear functional $f \in V^*$ such that*

$$f(v) \neq 0.$$

**Proof :** Let $B = \{b_1, b_2,..., b_n\}$ be a basis for $V$ and $B^* = \{f_1, f_2,..., f_n\}$ be the dual basis of $B$ in $V^*$, then by definition of dual basis we have

$$f_i(b_i) = \delta_{ij}, \qquad i, j = 1, 2,..., n$$

there exist unique scalars $\alpha_1, \alpha_2,..., \alpha_n \in F$ such that

$$v = \alpha_1 b_1 + ... + \alpha_n b_n$$

$$\Rightarrow \quad f_i(v) = f_i(\alpha_1 b_1 + ... + \alpha_n b_n), \qquad i = 1, 2,..., n$$

$$= f_i\left(\sum_{j=1}^{n} \alpha_j b_j\right)$$

$$= \sum_{j=1}^{n} \alpha_j f_i\left(b_j\right)$$

$$= \sum_{j=1}^{n} \alpha_j \delta_{ij}$$

$$f_i(v) = \alpha_1$$

Now, if $\quad f(v) = 0, \qquad \forall\ f \in V^*$

$\Rightarrow \quad f_i(v) = 0, \qquad i = 1, 2, ....., n$

$\Rightarrow \quad \alpha_i = 0$

$\Rightarrow \quad v = \mathbf{0}$

which is a contradiction, because $v \neq \mathbf{0}$. Hence there exists at least one $f \in V^*$ such that

$$f(v) \neq 0.$$

## 6.4 Second dual of a vector space

Let $V$ be a finite dimensional vector space over a field $F$ and $V^*$ be its dual space, then the dual space of $V^*$ is the vector space Hom($V^*$, $F$). The dual space of $V^*$ is denoted by $V^{**}$ and it is called the second dual of $V$.

**Note :** For a finite dimensional vector space $V$, dim $V$ = dim $V^*$ and so dim $V^*$ = dim $V^{**}$

$\Rightarrow \qquad$ dim $V$ = dim $V^{**}$

Thus $\qquad V \cong V^{**}$

**Theorem 10.** *Let V be a finite dimensional vector space over the field F, then there is a natural isomorphism of V onto V**.*

**Proof :** We define a function, for fixed vector $v \in V$,

$$\phi_v : V^* \to F,$$

by $$\phi_v(f) = f(v), \qquad \forall\ f \in V^*$$

To prove that $\phi_v$ is a linear functional from $V^*$ to $F$,

let $f, g \in V^*$ and $\lambda, \mu \in F$, then we have

$$\begin{aligned}\phi_v(\lambda f + \mu f) &= (\lambda f + \mu f)(v) \\ &= \lambda f(v) + \mu g(v) \\ &= \lambda\, \phi_v(f) + \mu\, \phi_v(g)\end{aligned}$$

Thus $\phi_v$ is a linear functional from $V^*$ of $F$, and so $\phi_v \in V^{**}$, which defines a natural map $t : V \to V^{**}$ by

$$t(v) = \phi_v, \qquad \forall\ v \in V.$$

Now to prove that $t$ is an isomorphism, suppose that $u, v \in V$ and $\lambda, \mu \in F$, we have

$$\begin{aligned}\phi_{\lambda u + \mu v}(f) &= f(\lambda u + \mu v) \\ &= \lambda f(u) + \mu f(v) \\ &= \lambda\, \phi_u(f) + \mu\, \phi_v(f), \qquad \forall\ f \in V^*.\end{aligned}$$

$$\therefore \qquad \phi_{\lambda u + \mu v} = \lambda \phi_u + \mu \phi_v$$

Now, $$\begin{aligned}t(\lambda u + \mu v) &= \phi_{\lambda u + \mu v} \\ &= \lambda \phi_u + \mu \phi_v \\ &= \lambda\, t(u) + \mu\, t(v)\end{aligned}$$

Thus $t$ is a linear map.

***t* is one-one :**

Let there exist $u, v \in V$ such that

$$t(u) = t(v)$$

$$\Rightarrow \qquad t(u - v) = \mathbf{0} \quad \Rightarrow \quad u - v \in \text{Ker}\,(t)$$

Now, for $$\mathbf{0} \neq u,$$

$$\phi_u(f) = f(u) \neq 0 \quad \text{for all} \quad f \in V^*.$$

Thus $$t(u) = \phi_u$$

is not a zero map.

$$\therefore \qquad \text{Ker}\,(t) = \{\mathbf{0}\}$$

$$\Rightarrow \qquad u = v$$

Thus $t$ is one-one.

***t* is onto :**

$\because$ $\dim V = \dim V^*$

$= \dim V^{**}$

*i.e.* $\dim V = \dim V^{**}$

Also $t$ is a monomorphism from $V$ to $V^{**}$ of the same dimension and so $t$ is an epimorphism (*i.e.* onto).

Hence $t$ is an isomorphism. Thus there is a natural isomorphism $t$ from $V$ onto $V^{**}$.

## 6.5 Dual map

Let $V$ and $V'$ be finite dimensional vector spaces over the same field $F$ and $B$, $B'$ be the bases of $V$ and $V'$ respectively and also let $B^*$, $B'^*$ be the dual bases. For each $f' \in V^*$ and for fixed $t \in \text{Hom}(V, V')$, the map $f'ot$ is a linear transformation form $V$ to $F$ *i.e.* $f'ot \in V^*$. Thus the map $f' \to f'ot$ defines a function from $V'^*$ to $V^*$. This is denoted by $t^*$ and is called the dual map of $t$.

Hence if $t : V \to V'$ is a linear transformation, then the map $t^* : V'^* \to V^*$ defined as

$$t^*(f') = f'ot \in V^*, \qquad \forall\ f' \in V'^*,$$

is a linear transformation and it is called dual map of $t$.

**Theorem 11.** *Let $V$ and $V'$ be finite dimensional vector spaces over a field $F$ and $t : V \to V'$ be a linear transformation and $t^*$ be the dual map of $t$, then $t$ and $t^*$ have the same rank.*

**Proof :** Let $\dim V = n$ and $\dim V' = m$.

Let $\text{rank}(t) = r$, then there exist bases $B = \{b_1, b_2, ....., b_n\}$ and $B' = \{b_1', b_2', ...., b_m'\}$ of $V$ and $V'$ respectively, such that

$$t(b_i) = \begin{cases} b_i', & i = 1, 2, ..., r \\ \mathbf{0}, & i = r+1, r+2, ..., n. \end{cases}$$

Let $B^* = \{f_1, f_2, ....., f_n\}$ and $B'^* = \{f_1', f_2', ...., f_m'\}$ be the bases dual to $B$ and $B'$ respectively, them

$$f_i(b_i) = \delta_{ij}; \qquad i, j = 1, 2, ..., n$$

and $$f_p'(b_q') = \delta_{pq}; \qquad p, q = 1, 2, ..., m.$$

Since $t^* : V'^* \to V^*$ be the dual map of $t$ and so,

$$t^*(f') = f'ot \in V^*, \qquad \forall\ f' \in V^*.$$

Now, $$\left[t^*(f_i')\right](b_j) = (f_i'ot)(b_j)$$

$$= f_i'(t(b_j))$$

$$= \begin{cases} f_i'(b_j), & \text{for} \quad j \le r \\ f_i'(\mathbf{0}), & \text{for} \quad j = r+1, r+2, \ldots, n \end{cases}$$

$$= \begin{cases} \delta_{ij} & \text{for} \quad i, j = 1, 2, \ldots, r \\ 0 & \text{otherwise.} \end{cases}$$

$$= \begin{cases} f_i(b_j) & \text{for} \quad i, j = 1, 2, \ldots, r \\ \hat{0}(b_j) & \text{otherwise.} \end{cases}$$

Thus $$t^*(f_i') = \begin{cases} f_i & \text{for} \quad i = 1, 2, \ldots, r \\ \hat{0} & \text{for} \quad i = r+1, r+2, \ldots, n \end{cases}$$

Hence rank $t^* = r$.

## Illustrative examples

***Ex.*1.** *If* $B = \{e_1 = (1, 0), e_2 = (0, 1)\}$ *is the usual basis* $R^2$. *Determine its dual basis.*

**Sol.** Let $B^* = \{f_1, f_2\}$ be the basis dual to $B$.

To determine the dual basis, we suppose

$$f_1(x, y) = ax + by, f_2(x, y) = cx + dy,$$

such that $$f_1(e_1) = 1, f_1(e_2) = 0, f_2(e_1) = 0, f_2(e_2) = 1.$$

Now,
$$\begin{aligned} 1 &= f_1(e_1) \\ &= f_1(1, 0) \\ &= a \cdot 1 + b \cdot 0 \quad \Rightarrow \quad a = 1 \\ 0 &= f_1(e_1) = f_1(0, 1) \\ 0 &= a \cdot 0 + b \cdot 1 \quad \Rightarrow \quad b = 0 \\ 0 &= f_2(e_1) = f_2(1, 0) \\ &= c \cdot 1 + d \cdot 0 \quad \Rightarrow \quad c = 0 \end{aligned}$$

and,
$$\begin{aligned} 1 &= f_2(e_2) \\ &= f_2(0, 1) \\ &= c \cdot 0 + d \cdot 1 \quad \Rightarrow \quad d = 1 \end{aligned}$$

Thus $$f_1(x, y) = x, \qquad f_2(x, y) = y$$

Hence $B^*$ $\{f_1(x, y) = x, f_2(x, y) = y\}$ be the dual basis.

***Ex.*2.** *If* $B = \{b_1 = (-1, 1, 1), b_2 = (1, -1, 1), b_3 = (1, 1, -1)\}$ *is a basis of* $V_3(R)$, *then find the basis dual to* $B$.

**Sol.** Let $B^* = \{f_1, f_2, f_3\}$ be the basis dual to $B$ so that

$$f_i(b_j) = \delta_{ij}; \qquad i, j = 1, 2, 3.$$

Suppose that $f_1(x, y, z) = a_1x + a_2y + a_3z$

$f_2(x, y, z) = b_1x + b_2y + b_3z$

$f_3(x, y, z) = c_1x + c_2y + c_3z.$

**To find $f_1$ :** $f_1(b_1) = 1$

$\Rightarrow$ $f_1(-1, 1, 1) = 1$

$\Rightarrow$ $-a_1 + a_2 + a_3 = 1$ .....(1)

$f_1(b_2) = 0$

$\Rightarrow$ $f_1(1, -1, 1) = 0$

$\Rightarrow$ $a_1 - a_2 + a_3 = 0$ .....(2)

and $f_1(b_3) = 0$

$\Rightarrow$ $f_1(1, 1, -1) = 0$

$\Rightarrow$ $a_1 + a_2 - a_3 = 0$ .....(3)

Solving (2) and (3), we get

$$\frac{a_1}{1-1} = \frac{-a_2}{-1-1} = \frac{a_3}{1+1}$$

$\Rightarrow$ $$\frac{a_1}{0} = \frac{a_2}{2} = \frac{a_3}{2}$$

$\Rightarrow$ $$\frac{a_1}{0} = \frac{a_2}{1} = \frac{a_3}{1} = K(\text{say})$$

$\Rightarrow$ $a_1 = 0, a_2 = K, a_3 = K$

Putting these in (1), we get

$$2K = 1 \quad \Rightarrow \quad K = \frac{1}{2}$$

Therefore, $a_1 = 0,\ a_2 = \frac{1}{2}, a_3 = \frac{1}{2}$

Thus, $f_1(x, y, z) = 0 \cdot x + \frac{1}{2} \cdot y + \frac{1}{2} \cdot z$

$$= \frac{1}{2}(y + z).$$

Similarly, $f_2(x, y, z) = \frac{1}{2}(x + z)$, and $f_3(x, y, z) = \frac{1}{2}(x + y)$

Thus the dual basis,

$$B^* = \left\{\frac{1}{2}(y + z), \frac{1}{2}(x + z), \frac{1}{2}(x + y)\right\}.$$

***Ex.3.*** *Let the linear functional f on R is given by f′(x, y) = 2x – 5y. For each linear transformation t : $R^3 \to R^2$, find its dual map, where*

*(i)* $t(x, y, z) = (x - y, y + z)$,

*(ii)* $t(x, y, z) = (x + y + 2z, 2x + y)$,

*(iii)* $t(x, y, z) = (x + y, 0)$,

**Sol.** By definition of dual map, we have

$$t^*(f') = f' o\ t$$

$$[t^*(f')](x, y, z) = f'[t(x, y, z)]$$

*(i)* $$[t^*(f')](x, y, z) = f'(t(x, y, z))$$
$$= f'(x - y, y + z)$$
$$= 2(x - y) - 5(y + z)$$
$$= 2x - 7y - 5z$$

*(ii)* $$[t^*(f')](x, y, z) = f'(t(x, y, z))$$
$$= f'(x + y + 2z, 2x + y)$$
$$= 2(x + y + 2z) - 5(2x + y)$$
$$= -8x - 3y + 4z$$

*(iii)* $$[t^*(f')](x, y, z) = f'(t(x, y, z))$$
$$= f'(x + y, 0)$$
$$= 2(x + y) - 5.0$$
$$= 2(x + y)$$

### Self-learning exercise-2

**1.** Which is the correct one ?

If $V$ is a finite dimensional vector space and $V^*$ is its dual space, then

(*a*) $\dim V^* = \dim V$ (*b*) $\dim V^* > \dim V$

(*c*) $\dim V^* < \dim V$ (*d*) none of these

**2.** Fill in the blanks :

***(i)*** If $V$ is a finite dimensional vector space, then $V^*$ is ..... to dim $V^{**}$.

***(ii)*** If $V$ is a vector space, then $V^{**}$ is called ..... .

**3.** State whether the followings are true or false :

***(i)*** Each linear transformation is a linear functional,

***(ii)*** Each linear functional is a linear transformation.

## 6.6 Algebra of linear transformations

The set of all linear transformations of $V$ to $V'$ (vector spaces over the same field $F$), is denoted by Hom $(V, V')$.

Now we define addition and scalar multiplication operation in Hom $(V, V')$ as follows :

Let $t_1, t_2 \in Hom(V, V')$, we define a map $t_1 + t_2 : V \to V'$ by

$$(t_1 + t_2)(v) = t_1(v) + t_2(v), \qquad \forall\ v \in V,$$

and a map $(\alpha t_1) : V \to V'$ by

$$(\alpha t_1)(v) = \alpha t_1(v) \qquad \forall\ v \in V, \alpha \in F,$$

**Closurness for addition and scalar multiplication in *Hom* $(V, V')$ :**

Let $u, v \in V$ and $\lambda, \mu \in F$, we have

$$\begin{aligned}(t_1 + t_2)(\lambda u + \mu v) &= t_1(\lambda u + \mu v) + t_2(\lambda u + \mu v)\\ &= \lambda t_1(u) + \mu t_1(v) + \lambda t_2(u) + \mu t_2(v)\\ &= \lambda(t_1 + t_2)(u) + \mu(t_1 + t_2)(v).\end{aligned}$$

Thus $t_1 + t_2$ is a linear transformation of $V$ to $V'$, and so that $t_1, t_2 \in Hom(V, V')$.

Thus addition of linear transformation is closed in $Hom(V, V')$.

Similarly,
$$\begin{aligned}(\alpha t_1)(\lambda u + \mu v) &= \alpha t_1(\lambda u + \mu v)\\ &= \alpha\{\lambda t_1(u) + \mu t_1(v)\}\\ &= \lambda(\alpha t_1)(u) + \mu(\alpha t_1)(v).\end{aligned}$$

Thus $\alpha t_1$ is a linear transformation of $V$ to $V'$ and so that $\alpha t_1 \in Hom(V, V')$.

Thus scalar multiplication is closed in $Hom(V, V')$.

**Theorem 12.** *The set Hom $(V, V')$ of all linear transformations of V to V', forms a vector space over the same field F.*

**Proof :** Let $V$ and $V'$ be the vector spaces over the same field $F$.

To prove that the set $Hom(V, V')$ is a vector space over $F$ under the operations defined as follows :

$$(f + g)(v) = f(v) + g(v),$$
$$(\alpha f)(v) = \alpha(v), \qquad \forall\ f, g \in Hom(V, V'), \text{ and } \alpha \in F.$$

Now left as an exercise for the reader.

**Theorem 13.** *Let V and V' be any two finite dimensional vector spaces over the same field F. Then the vector space Hom $(V, V')$ of all linear transformations of V to V', is also finite dimensional, and*

$$Hom(V, V') = dim\ V \times dim\ V'.$$

**Proof :** Suppose that $V$ be $m$-dimensional and $V'$ be n-dimensional vector spaces over the field $F$ and $B = \{b_1, b_2, ....., b_m\}$, $B' = \{b_1', b_2', ...., b_n'\}$ be the bases of $V$ and $V'$ respectively.

Let $u \in V$, then we can write uniquely, $v = \lambda_1 b_1 + ... + \lambda_m b_m$, for $\lambda_1, \lambda_2, ..., \lambda_m \in F$.

Now we define a map $t_{ij} : V \to V'$ by

$$t_{ij}(v) = t_{ij}\left(\sum_{i=1}^{m} \lambda_i\, b_i\right), \qquad \forall v \in V$$

$$= \lambda_i\, b_i', \qquad \begin{matrix} i = 1,2,...,m \\ j = 1,2,...,n \end{matrix}$$

It can be easily shown that $t_{ij}$ is a linear transformation of $V$ to $V'$ and so that $t_{ij} \in Hom\,(V, V')$. And it is such that

$$t_{ij}(b_r) = \begin{cases} \mathbf{0} & \text{if} \quad i \neq r \\ b_j' & \text{if} \quad i = r \end{cases}$$

that is, $$t_{ij}(b_r) = \delta_{ir}\, b_j' \qquad r = 1, 2,..., m$$

where $\delta_{ir} \in F$ is Kronecker delta.

Now we shall show that the set of these $m\,n$ linear transformations,

$$C = \left\{ t_{ij} \in Hom(V, V'), \quad \begin{matrix} i = 1,2,...,m \\ j = 1,2,...,n \end{matrix} \right\},$$

forms a basis for $Hom\,(V, V')$.

***(i) C* is linearly independent :**

Let there exist scalars $\alpha_{ij} \in F$, $\begin{matrix} i = 1,2,.....,m \\ j = 1,2,.....,n, \end{matrix}$

such that $\sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij}\, t_{ij} = \hat{0}$, where $\hat{0}$ is zero map

$$\Rightarrow \qquad \left(\sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij}\, t_{ij}\right)(b_r) = \hat{0}(b_r), \qquad \forall\ b_r \in B, \qquad r = 1,2,...m$$

$$\Rightarrow \qquad \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij}\, t_{ij}(b_r) = \mathbf{0}$$

$$\Rightarrow \qquad \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij}\, b_{ir}\, b_j' = \mathbf{0}$$

$$\Rightarrow \qquad \sum_{j=1}^{n} \alpha_{rj}\, b_j' = \mathbf{0}$$

$$\Rightarrow \qquad \alpha_{rj} = 0 \qquad [\because\ B' \text{ is a basis for } V']$$

for $r = 1, 2, ..., m,$ and $j = 1, 2, ..., n$

Hence $C$ is linearly independent.

***(ii) C* spans *Hom* (*V*, *V′*) :**

Let $t \in Hom\,(V, V')$ be any arbitrary vector, *i.e.* $t : V \to V'$, so that

$$t(b_r) \in V', \quad \text{for} \quad r = 1, 2, ..., m$$

$$\Rightarrow \quad t(b_r) = \alpha_{r1}\, b_1' + \alpha_{r2}\, b_2' + ... + \alpha_{rn}\, b_n'$$

$$= \sum_{j=1}^{n} \alpha_{rj}\, b_j', \qquad \text{for some } \alpha_{rj} \in F.$$

Now, we take,

$$\left( \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij}\, t_{ij} \right) (b_r)$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij}\, t_{ij} (b_r)$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij}\, \delta_{ij}\, b_j'$$

$$= \sum_{j=1}^{n} \alpha_{rj}\, b_j'$$

$$= t\,(b_r)$$

Thus

$$t = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij}\, t_{ij}$$

which shows that each $t \in Hom\,(V, V')$ is a linear combination of vectors of $C$, and thus $C$ spans $Hom\,(V, V')$.

Hence $C$ is a basis for $Hom\,(V, V')$, so that

$$dim\ Hom\,(V, V') = m\,n$$

$$= dim\ V \times dim\ V'.$$

## 6.7 Rank and nullity of a linear transformation

Let $V$ and $V'$ be any two vector spaces over the field $F$ and $t : V \to V'$ be a linear transformation. Then image of $t$ *i.e.* $im(t)$ is a vector subspace of $V'$ and if it finite dimensional, then the dimension of this subspace is called the rank of $t$.

Similarly, Ker($t$) is a vector subspace of $V$ and if it is finite dimensional, then its dimension is called nullity of $t$.

**Theorem 14.** (*Sylvester's law of nullity*) *Let $V$ and $V'$ be vector spaces over the same field $F$ and $t : V \to V'$ be a linear transformation. If $V$ is finite dimensional, then*

$$dim\ V = rank\ (t) + nullity\ (t).$$

**Proof :** Given that $V$ and $V'$ be vector spaces over the field $F$ and $V$ is finite dimensional.

Let $dim\ V = m$ and $dim$ Ker($t$) = $r$, then

$$r \leq m \qquad [\because \text{ Ker } (t) \text{ is a vector subspace of } V]$$

Let $B_1 = \{b_1, b_2,..., b_r\}$ be the basic for Ker $(t)$.

Since $B_1$ is linearly independent set of vectors in $V$, so that it can be extended to a basis $B = \{b_1, b_2,..., b_r, b_{r+1},..., b_m\}$ of $V$.

We claim that the set

$$B_2 = \{t(b_{r+1}), t(b_{r+2}),..., t(b_m)\} \quad \text{is a basis of } im(t).$$

***(i) $B_2$ is linearly independent :***

Let there exist scalars $\alpha_{r+1}, \alpha_{r+2},..., \alpha_m \in F$, such that

$$\alpha_{r+1}\, t(b_{r+1}) + ... + \alpha_m\, t(b_m) = \mathbf{0}$$

$$\Rightarrow \quad t(\alpha_{r+1}\, b_{r+1} + ... + \alpha_m b_m) = \mathbf{0} \qquad [\because t \text{ is linear}]$$

$$\Rightarrow \quad \alpha_{r+1}\, b_{r+1} + ... + \alpha_m b_m \in \text{Ker } (t).$$

Then, there exist scalars $\alpha_1, \alpha_2,..., \alpha_r \in F$, such that

$$\alpha_{r+1}\, b_{r+1} + ... + \alpha_m b_m = \alpha_1 b_1 + ... + \alpha_r b_r$$

$$\Rightarrow \quad (-\alpha_1)\, b_1 + ... + (-\alpha_r)\, b_r + \alpha_{r+1}\, b_{r+1} + ... + \alpha_m b_m = \mathbf{0}$$

$$\Rightarrow \quad \alpha_1 = \alpha_2 = ... = \alpha_m = 0. \qquad [\because B = \{b_1, b_2,..., b_m\} \text{ is basis of } V]$$

Hence $B_2$ is linearly independent.

***(ii) $B_2$ spans im $(t)$ :***

Let $v' \in im(t)$, then there exists $v \in V$ such that

$$v' = t(v).$$

Since $v \in V$, so that exists scalars $\alpha_1, \alpha_2,..., \alpha_m \in F$ such that

$$v = \alpha_1 b_1 + ... + \alpha_m b_m$$

$$\Rightarrow \quad t(v) = t(\alpha_1 b_1 + ... + \alpha_m b_m)$$

$$\Rightarrow \quad v' = \alpha_1 t(b_1) + ... + \alpha_r t(b_r) + \alpha_{r+1} t(b_{r+1}) + ... + \alpha_m t(b_m)$$

$$= \alpha_{r+1} t(b_{r+1}) + ... + \alpha_m t(b_m)$$

$$[\because b_1, b_2,..., b_r \in \text{Ker } (t), \text{ and so } t(b_1) = t(b_2) = ... = t(b_r) = \mathbf{0}]$$

which shows that $B_2$ spans $im$ $(t)$.

Hence $B_2 = \{t(b_{r+1}) + ... + t(b_m)\}$ is a basis for $im(t)$, and so that

$$\dim im\,(t) = m - r$$

$$= \dim V - \dim \text{Ker } (t)$$

$$\Rightarrow \quad \text{rank } (t) = \dim V - \text{nullity } (t)$$

Thus $\quad \dim V = \text{rank } (t) + \text{nullity } (t).$

**Theorem 15.** *Let $t : V \to V'$ be a linear transformation of rank $r$ and dim $V = m$, dim $V' = n$. Then $r \leq min\ (m, n)$, and there exists a basis $\{b_1, b_2,..., b_m\}$ of $V$ and a basis $\{b_1', b_2',..., b_n'\}$ of $V'$ such that*

$$t(b_1) = b_1',\ t(b_2) = b_2', ..., t(b_r) = b_r',$$
$$t(b_{r+1}) = \mathbf{0},\ t(b_{r+2}) = \mathbf{0}, ..., t(b_m) = \mathbf{0},$$

**Proof :** By Sylvester's law of nullity, we have

$$\dim V = \text{rank}\,(t) + \text{nullity}\,(t)$$
$$m = r + \text{nullity}\,(t) \qquad [\because \dim V = m, \text{rank}\ t = r] \qquad .....(1)$$

Obviously $r \le m$

Thus $r \le min\,(m, n)$. $[\because\ im\,(t)$ is a subspace of $V' \Rightarrow \dim im\,(t) \le \dim V']$

From equation (1), we have

$$\text{nullity}\,(t) = \dim \text{Ker}\,(t) = m - r,$$

So assume that $\{b_{r+1}, b_{r+2}, ..., b_m\}$ be a basis of Ker $(t)$, which is a linearly independent set in $V$, so it can be extended to a basis $\{b_1, b_2, ..., b_r, b_{r+1}, .... b_m\}$ of $V$.

Now $\{b_1, b_2, ..., b_r\}$ is a linearly independent set of vectors, because it is a subset of basis.

Let there exist scalars $\alpha_1, \alpha_2, ..., \alpha_r, \in F$ such that

$$\alpha_1 t(b_1) + \alpha_2 t(b_2) + ... + \alpha_r t(b_r) = \mathbf{0}$$
$$\Rightarrow \quad t(\alpha_1 b_1 + \alpha_2 b_2 + ... + \alpha_r b_r) = \mathbf{0}$$
$$\Rightarrow \quad \alpha_1 b_1 + ... + \alpha_r b_r \in \text{Ker}(t)$$
$$\Rightarrow \quad \alpha_1 b_1 + ... + \alpha_r b_r = \alpha_{r+1} b_{r+1} + ... + \alpha_m b_m \quad \text{for} \quad \alpha_{r+1}, ..., \alpha_m \in F.$$
$$\Rightarrow \quad \alpha_1 b_1 + ... + \alpha_r b_r + (-\alpha_{r+1}) b_{r+1} + ... + (-\alpha_m) b_m = \mathbf{0}$$
$$\Rightarrow \quad \alpha_1 = ... = \alpha_r = (-\alpha_{r+1}) = ... = (-\alpha_m) = 0$$

Hence $\{t(b_1), t(b_2), ..., t(b_r)\}$ is a linearly independent set, then it can be a part of a basis of $V'$, *i.e.* there exist vectors $b_1', b_2', ..., b_r'$ in a basis of $V'$ such that

$$t(b_1) = b_1',\ t(b_2) = b_2', ..., t(b_1) = b_r'$$

and clearly, $t(b_{r+1}) = \mathbf{0}, t(b_{r+2}) = \mathbf{0}, ..., t(b_m) = \mathbf{0}$

$[\because\ \{b_{r+1}, b_{r+2}, ..., b_m\}$ is a basis of Ker$(t)]$

***Ex.*1.** *Show that the map* $t : V_2(R) \to V_3(R)$ *defined by*

$$t(a, b) = (a + b, a - b, b)$$

*is a linear transformation. Find range, rank, null space and nullity of t.*

**Sol.** Given that $t : V_2(R) \to V_3(R)$ defined by

$$t\,(a, b) = (a + b, a - b, b), \qquad \forall\ a, b \in R.$$

Let $x = (a, b)\ y = (c, d) \in V_2\,(R)$,

and let $\lambda, \mu \in R$, then $\lambda x + \mu y \in V_2$

*(i)* $t\,(\lambda x + \mu y) = t\,[\lambda\,(a, b) + \mu\,(c, d)]$

$= t\,(\lambda a + \mu c, \lambda b + \mu d)$

$= (\lambda a + \mu c + \lambda b + \mu d, \lambda a + \mu c - \lambda b - \mu d, \lambda b + \mu d)$

$$= (\lambda (a + b), \lambda (a - b), \lambda b) + (\mu (c + d), \mu (c - d), \mu d)$$
$$= \lambda (a + b, a - b, b) + \mu (c + d, c - d, d)$$
$$= \lambda t (a, b) + \mu t (c, d)$$
$$= \lambda t (x) + \mu t (y).$$

Which shows that $t$ is a linear transformation

*(ii)* To find range space *i.e.* $im(t)$ of $t$ :

$\because$ $\{(1, 0), (0, 1)\}$ is the usual basis of $V_2$.

By definition,

$$t (1, 0) = (1 + 0, 1 - 0, 0) = (1, 1, 0)$$
$$t (0, 1) = (0 + 1, 0 - 1, 1) = (1, -1, 1)$$

Now, $\{(1, 0), (0, 1)\}$ is a basis if $V_2$

$\Rightarrow$ vectors $(1, 0), (0, 1)$ spans $V_2$

$\Rightarrow$ $t (1, 0), t (0, 1)$ spans $im(t)$.

$\Rightarrow$ $(1, 1, 0), (1, -1, 1)$ spans $im(t)$.

Also let $\alpha, \beta \in R$ such that

$$\alpha (1, 1, 0) + \beta (1, -1, 1) = 0 = (0, 0, 0)$$

$\Rightarrow$ $(\alpha + \beta, \alpha - \beta, \beta) = (0, 0, 0)$

$\Rightarrow$ $\alpha + \beta = 0, \alpha - \beta = 0, \beta = 0$

$\Rightarrow$ $\alpha = 0, \beta = 0$

Thus $(1, 1, 0), (1, -1, 1)$ are linearly independent. Hence $\{(1, 1, 0), (1, -1, 1)\}$ is a basis of range space *i.e.* $im(t)$.

$\Rightarrow$ $\dim im(t) = 2$

$\Rightarrow$ $\text{rank } (t) = 2.$

*(iii)* To find null space of $t$ *i.e.* $\text{Ker}(t)$ :

Since $t : V_2 \to V_3.$

by Sylvester's law of nullity, we have

$$\dim V_2 = \text{rank}(t) + \text{nullity}(t)$$

$\Rightarrow$ $2 = 2 + \dim \text{Ker}(t).$

$\Rightarrow$ $\dim \text{Ker } (t) = 0$

$\Rightarrow \text{Ker}(t)$ is a zero space

$\Rightarrow$ $\text{Ker}(t) = \{(0, 0)\}.$

Also, $\dim im(t) = \text{rank}(t) = 2$

$\Rightarrow$ $\dim \text{Ker}(t) = \text{nullity}(t) = 0.$

Further $im(t)$ is a subspace of $V_3(R)$ generated by the vectors $(1, 1, 0), (1, -1, 1)$.

***Ex.2.*** *Let $t$ be a linear operator on a vector space $V(F)$. If $t^2 = 0$, then find the relation between range space of $t$ and null space of $t$. Also given an example of a linear operator on $V_2(R)$ such that $t^2 = 0$ but $t \neq 0$.*

**Sol.** *(i)* Given that $t^2 = 0 \Rightarrow t^2(v) = 0, \quad \forall\, v \in V$

$$t(t(v)) = 0$$

Let $t(v) \in \text{Ker}(t) \quad \forall\, v \in V.$

But $t(v) \in im(t)$, for $v \in V$

Thus $im(t) \subset \text{Ker}(t)$.

$\Rightarrow$ rang space of $(t) \subset$ null space of $t$.

*(ii)* We consider a linear map $t : V_2(R) \to V_2(R)$ defined by

$$t(a, b) = (0, a) \quad \forall\, (a, b) \in V_2$$

clearly $t \neq 0$

Now, $t^2(a, b) = t(t(a, b))$

$$= t(0, a)$$

$$= (0, 0) = \hat{0}(a, b)$$

$\Rightarrow$ $t^2(a, b) = 0\,(a, b)$, for $(a, b) \in V_2$

$\Rightarrow$ $t^2 = 0.$

## 6.8 Summary

In this unit we have studied linear transformation of vector spaces, which plays a very important role in the study of vector spaces. We have seen that the set Hom$(V, V')$ of all linear transformations of $V$ to $V'$ forms a vector space itself over field $F$. We have also studied the concept of dual space, second dual, and dual map.

## 6.9 Answers to self-learning exercises

### Self-learning exercise-1

**1.** ***(i)*** *(a)* ***(ii)*** *(d)*

**2.** ***(i)*** vector space homomorphism ***(ii)*** $t(u) + t(v)$ ***(iii)*** $V, V'$

***(iv)*** $\{0\}$ ***(v)*** $\dim V = \dim V'$

**3.** ***(i)*** T ***(ii)*** T ***(iii)*** T

### Self-learning exercise-2

**1.** $a$ **2.** ***(i)*** equal ***(ii)*** second dual

**3.** ***(i)*** false ***(ii)*** true

## 6.10 Exercises

1. Which of the following functions $t : R^2 \to R^2$ are linear transformations ?

   ***(i)*** $t(a, b) = (1+ a, b)$,

   ***(ii)*** $t(a, b) = (b, a)$

   ***(iii)*** $t(a, b) = (a + b, a)$

2. Find the dual basis of the basis

   $$B = \{(1, -1, 3), (0, 1, -1), (0, 3, -2)\} \text{ for } V_3(R)$$

3. Let $t : V \to V'$ be a linear transformation of vector spaces of the same dimension. Then following statement are equivalent :

   ***(i)*** $t$ is an isomorphism

   ***(ii)*** $t$ is injective, *i.e.* $\text{Ket}(t) = \{\mathbf{0}\}$

   ***(iii)*** $t$ is surjection *i.e.* $t(V) = V'$

   ***(iv)*** $t$ sends a basis of $V$ to a basis of $V'$.

4. Let $V$ and $V'$ be vector spaces over the field $F$, then prove that the set $\text{Hom}(V, V')$ of all linear transformations of $V$ to $V'$ is a vector space over the field $F$.

5. If $V$ is a finite dimensional vector space over the field $F$ and $v_1 \neq v_2$ are in $V$, then prove that there is an $f \in V^*$ such that $f(v_1) \neq f(v_2)$.

□□□

# UNIT 7 : Basic Theory of Field Extensions, Simple Extension, Algebraic and Transcendental Extensions

**Structure of the Unit**



## 7.0 Objectives

After reading this unit you will be able to understand about field extension, simple field extension, algebraic and transcendental elements, algebraic and transcendental field extensions.

## 7.1 Introduction

In this unit, we shall take up the study of theory of finite field extensions. The concept of field has a central place in Algebra. It has wide applications in Linear Algebra and in the Theory of Equations which deals with the study of roots of polynomials.

## 7.2 Basic theory of field extensions

We know that a field $(F, +, .)$ is a commutative ring with unity in which every nonzero element has a multiplicative inverse. We also Know that a field has no proper ideal, that is, if $I$ is an ideal of a field $F$, then either $I = \{0\}$ or $I = F$.

If $1 \in F$ is the unity element in $F$, then the smallest positive integer $n$ such that $n.1 = 0$ is called the characteristic of $F$. If no such positive integer exist, then characteristic of $F$ is said to be zero. The characteristic of a field is either zero or a prime number.

A field $F$ is said to be embedded in a field $K$ if $F$ is isomorphic onto a subset of the field $K$. By virtue of isomorphism, it readily follows that isomorphic image of $F$ is a subfield of $K$. For example, the field $Q$ of rational numbers is embedded in the field $R$ of real numbers which is embedded in the field $C$ of complex numbers.

If $R$ is a ring, we define a polynomial $p(x)$ with coefficients in $R$, as

$$p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n,$$

where $a_i \in R$. if $a_i = 0$ for all $i$, then it is known as a **zero polynomial.** The set of all such polynomials is devoted by $R[x]$ and it is a ring with respect to addition and multiplication of polynomials. $R[x]$ is known as **polynomial ring** over $R$. Further, if $R$ is an integral domain, then so is $R[x]$.

If $p(x)$ is not the zero polynomial, its **Leading coefficient** is $a_n$, where $a_n \neq 0$ and $a_i = 0$ for all $i > n$, then $n$ is known as **degree** of $p(x)$ and we write $deg\ p(x) = n$. $a_n x^n$ is called **leading term.** A polynomial is constant if and only if its degree is zero. If the leading coefficient, that is, $a_n = 1$, then polynomial is known as **monic polynomial.** The notion of prime number can be generalized to polynomials. A nonzero polynomial $f(x) \in F[x]$ is irreducible over field $F$ if $deg\ f(x) \geq 1$ and there is no factorization of the form $f(x) = g(x)\ h(x)$ in $F[x]$ such that $deg\ g(x) < deg\ f(x)$ and $deg\ h(x) < deg\ f(x)$. In other words we can say that whenever $f(x) = g(x)\ h(x)$, where $g(x), h(x) \in F[x]$, then either $g(x) \in F$ or $h(x) \in F$. If a polynomial is not irreducible, then it is called reducible. Note that irreducibility of a polynomial depends on the nature of the field. For example, $x^2 + 2$ is irreducible over $R$ but reducible over $C$.

**Field extensions**

Let $F$ be a field. A field $K$ is called an **extension field** of $F$ if $K$ contains $F$ as a subfield. We known that every field is a victor space over its subfield, so if $K$ is a field extension of a field $F$, then $K$ is a vector space over $F$. As a vector space over $F$, the dimension of $K$ may be finite or infinite. If a vector space $K$ over a field $F$ is finite. Dimensional, then we say that $K$ is a finite field extension of $F$ and dimension of $K$ over $F$ is known as degree of $K$ over $F$ and is denoted by the symbol $[K : F]$. If dimension of $K$ over $F$ is not finite, then we say that $K$ is an infinite extension of $F$.

**Theorem 1.** *If $K$ is a finite field extension of a field $F$ and $L$ is a finite field extension of $K$, then $L$ is a finite field extension of $F$ and* $[L : F] = [L : K]\ [K : F]$.

**Proof :** Let $B = \{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ be a basic for $L$ as a vactor space over $K$, and let $C = \{\beta_1, \beta_2, \ldots, \beta_n\}$ be a basis of $K$ over $F$.

Then $[L : K] = m$ and $[K : F] = n$. We will show that the set of $mn$ products $P = \{\alpha_i \beta_i \mid 1 \leq i \leq m\ ;\ 1 \leq j \leq n\}$ is a basis of vactor space $L$ over $F$, and this will prove the theorem.

Let $\alpha$ be any element of $L$. Since set $B$ is a basis for $L$ over $K$, so there exist $a_1, a_2, \ldots, a_m \in K$ such that

$$\alpha = a_1 \alpha_1 + a_2 \alpha_2 + ... + a_m \alpha_m \text{ or } \alpha = \sum_{i=1}^{m} a_i \alpha_i \quad .....(1)$$

Again, since set $C$ is a basis for $K$ over $F$, so for each $i = 1, 2, ..., m$ there are elements $b_{i1}$, $b_{i2}, ..., b_{in}, \in F$ such that $a_i = b_{i1} \beta_1 + b_{i2} \beta_2 + ... + b_{in} \beta_n$

or $$a_i = \sum_{j=1}^{n} b_{ij} \beta_j \quad .....(2)$$

using (2) in (1) we get $$\alpha = \sum_{i=1}^{m} \left( \sum_{j=1}^{n} b_{ij} \beta_j \right) \alpha_i$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} b_{ij} \left( \alpha_i \beta_j \right)$$

which shows that $P$ spans $L$ as a vector space over $F$. To see that set $P$ is linearly independent, let us assume that

$$\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \left( \alpha_i \beta_j \right) = 0, \ a_{ij} \in F.$$

$$\Rightarrow \quad \sum_{i=1}^{m} \left( \sum_{j=1}^{n} a_{ij} \beta_j \right) \alpha_i = 0 \quad .....(3)$$

Since $B$ is a basis of $L$ over $K$ and $\sum_{j=1}^{n} a_{ij} \beta_j \in K$, so from (3), we have

$$\sum_{j=1}^{n} a_{ij} \beta_j = 0 \quad .....(4)$$

Again, since $C$ is a basis of $K$ over $F$ and $a_{ij} \in F$, so from (4) we have $a_{ij} = 0$ for $1 \le i \le m$, $1 \le j \le n$.

This shows that the set $P$ is linearly independent, and hence $P$ is a basis of $L$ over $F$. Consequently, $[L : F] = m\, n = [L : K]\, [K : F]$

or $[L : F] = [L : K]\, [K : F]$

**Corollary.** If $L$ is a finite extension of a field $F$ and $K$ is a subfield of $L$ containing $F$, then $[K : F]$ divides $[L : F]$

**Proof :** Since $L$ is a finite field extension of $F$, so $[L : F]$ is finite. Again, since any set of elements in $L$ which is linearly independent over $K$ is also linearly independent over $F$ and so $[L : K]$ is finite, since $[K : F]$ is finite. Since $K$ is a subfield of $L$ containing $F$, so $F$ is a subfield of $K$ and hence $[K : F]$ is finite, since every subspace of a finite dimensional vector space is finite. Now by above theorem we have

$$[L : F] = [L : K]\,[K : F]$$

$\Rightarrow$ $[K : F]$ divides $[L : F]$.

Note that if $[L : F]$ is a prime member, then there exist no field $K$ properly contained in $L$ and properly containing $F$. In others wards we can say that, if $[L : F]$ is prime and $K$ is any subfield of $L$ containing $F$, then either $K = L$ or $L = F$.

***Ex.*1.** *Let F be an arbitrary field. Since every field is a vactor space over itself, so F is an extension over itself of degree one as dim F (F) =1.*

***Ex.*2.** *The field* $Q\left(\sqrt{3}\right) = \left\{a + b\sqrt{3} \,\middle|\, a, b \in Q\right\}$ *is a finite extension of Q and* $\left[Q\sqrt{3} : Q\right] = 2$ *as the set* $\left\{1, \sqrt{3}\right\}$ *is a basis of* $Q\left(\sqrt{3}\right)$ *over the field Q of rational numbers.*

***Ex.*3.** *The field* $Q\left(\sqrt{2}, \sqrt{3}\right) = \left\{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{2}\sqrt{3} \,\middle|\, a, b, c, d \in Q\right\}$ *is a finite extension of Q and* $\left[Q\left(\sqrt{2}, \sqrt{3}\right) : Q\right] = 4$, *since the set* $\left\{1, \sqrt{2}, \sqrt{3}, \sqrt{6}\right\}$ *is a basis of* $Q\left(\sqrt{2}, \sqrt{3}\right)$ *over the field Q of rational numbers.*

*Also* $Q\left(\sqrt{2}, \sqrt{3}\right)$ *is a finite extension of* $Q\left(\sqrt{3}\right)$ *and* $\left[Q\left(\sqrt{2}, \sqrt{3}\right) : Q\left(\sqrt{3}\right)\right] = 2$, *since the set* $\left\{1, \sqrt{2}\right\}$ *is a basis of* $Q\left(\sqrt{2}, \sqrt{3}\right)$ *over* $Q\left(\sqrt{3}\right)$.

***Ex.*4.** *The field C of complex numbers is a finite extension of the field R of real numbers. Here* $[C : R] = 2$, *since the set* $\{1, i\}$ *is a basis of vector space C* $(R)$.

## 7.3 Simple field extension

Let $K$ be an extension field of a field $F$ and let a be an element of $K$. We define a family $M$ of subfields of $K$ as follows :

$$M = \left\{L \,\middle|\, L \text{ is a subfield of } K \text{ containing both } F \text{ and a}\right\}.$$

$M \neq \phi$, since $K \in M$. Let $E = \cap_{L \in M} L$, *i.e.,* $E$ is the intersection of all member of the family $M$.

Since, intersection of an arbitrary collection of subfields of a given field is again a subfield of that field and each number of $M$ contain both $F$ and $a$, so $E$ is a subfield of $K$ containing both $F$ and $a$. Also, if $H$ is any subfield of $K$ containing both $F$ and $a$, then $H \in M$ and consequently $E \subset H$. Thus $E$ is the smallest subfield of $K$ containing both $F$ and $a$. The existence of such a subfield leads us to the following definition.

Let $K$ be an extension field of a field $F$ and $a \in K$, then $F(a)$ is the smallest subfield of $K$ containing both $F$ and $a$. We call $F(a)$ the subfield obtained by adjoining $a$ to $F$. This description of $F(a)$ is purely external one. We now give an alternative and more constructive description of $F(a)$. Let

$$B = \{ b_0 + b_1 a + b_1 a^2 + ... + b_n a^n | b_i \in F, n \geq 0\},$$

that is, $B$ consists of polynomials in a with coefficients from $F$.

Also, let $$W = \left\{ \frac{p(a)}{q(a)} \middle| p(a), q(a) \in B \text{ and } q(a) \neq 0 \right\}.$$

Clearly $W$ is a subfield of $K$ containing both $F$ and a hence $F(a) \subseteq W$ .....(1)

Again, since $F(a)$ is a subfield of $K$ containing both $F$ and a, therefore by virtue of closure under addition and multiplication on, $F(a)$ must contain all elements of the set $B$ and hence it must also contain all quotients of such elements.

So, $W \subseteq F(a)$ .....(2)

from (1) and (2) we get

$$F(a) = W = \left\{ \frac{p(a)}{q(a)} \middle| p(a), q(a) \in F[x] \text{ and } q(a) \neq 0 \right\}.$$

An extension $K$ of a field $F$ is called a **simple extension** if $K = F(a)$ for some $a \in K$.

Similarly, if $a_1, a_2, ..... , a_n \in K$, then $F(a_1, a_2, ..... , a_n)$ is the smallest subfield of $K$ containing $a_1, a_2, ..... , a_n$ and $F$. Clearly $F \subset F(a_1) \subset F(a_1, a_2) \subset ..... \subset F(a_1, a_2, ..... , a_n) \subset K$.

***Ex.*1.** *The field C of complex numbers is a simple extension of the field R of real numbers, since* $R(i) = C$.

***Ex.*2.** *Let* $F = Q$ *and* $K = \left\{ \frac{a+b\sqrt{2}}{c+d\sqrt{2}} \middle| a, b, c, d \in Q, c \text{ or } d \neq 0 \right\}$

*Then K is a field and* $K = Q(\sqrt{2})$ *is a simple extension of Q.*

## 7.4 Algebraic and transcendental extension

Let $K$ be a field extension of a field $F$, and let $\alpha$ be any element of $K$. Then $\alpha$ is said to be **algebric** over $F$ if it is the root of some nonzero polynomial $p(x) \in F[x]$. Thus, if $\alpha$ is algebric over $F$, then there exist scalars $a_0, a_1, a_2, ..... , a_n \in F$ not all zero such that

$a_0 + a_1 \alpha + a_2 \alpha^2 + .... + a_n \alpha^n = 0$,

that is, $1, \alpha, \alpha^2, ..... , \alpha^n$ are lineally dependent. If $\alpha$ is not algebraic over $F$, that is, if $\alpha$ is not a root of any such polynomial, then $\alpha$ is called **transcendental** over $F$.

The element $\alpha \in K$ is said to be **algebric of degree *n*** over $F$, if it satisfies a nonzero polynomial over $F$ of degree $n$ but no nonzero polynomial of lower degree.

If every element of $K$ is algebric over $F$, then an extension $K$ over $F$ is called **algebric extension** of $F$. If $K$ is not an algebric extension of $F$, if is called a **transcendental extension** of $F$.

Note that the two properties, algebric and transcendental, depend on the given field $F$. For example, the complex number $2\pi i$ is algebric over the field of real numbers $R$ but transcendental over the field of rational numbers $Q$. Also, every element $\alpha$ of a field $F$ is algebric over $F$, because it is the root of the polynomial $x - \alpha$, which has coefficients in $F$. Hence every field $F$ is an algebric extension of itself.

The two properties for $\alpha$ can be described in terms of the basic homomorphism of $F[x]$ into $K$, *i.e.*,

$$\phi_\alpha : F[x] \to K \text{ defined by}$$
$$\phi_\alpha [f(x)] = f(\alpha) \text{ for } f(x) \in F[x].$$

The element $\alpha$ is transcendental over $F$ if $\phi_\alpha$ is injective, *i.e.,* $\ker \phi_\alpha$ is $\{0\}$ and algebric over $F$ otherwise, that is, if Kernel of $\phi_\alpha$ is not zero.

A polynomial $f(x) \in f[x]$ of least degree with $f(\alpha) = 0$ is called the minimal polynomial of $\alpha$. To ensure uniqueness of minimal polynomial for $\alpha$ over $F$, we must impose further restriction that it should be monic, that is, the coefficient of highest power of $x$ should be one. If the degree of minimal polynomial of $\alpha$ is $n$, then $\alpha$ is said to be an algebric element of degree $n$ over $F$.

***Ex.1.*** *The field R of real numbers is not an algebric extension of the field Q of rational numbers because $\pi$, e are elements of R which are not algebric over Q.*

***Ex.2.*** *The field C of complex numbers is algebric extension over R.*

***Ex.3.*** *$\alpha = 2i$ is algebric over R with minimal polynomial $x^2 + 4$.*

***Ex.4.*** *$\alpha = \sqrt[3]{3}$ is algebric over Q with minimal polynomial $x^3 - 3 \in Q[x]$.*

**Theorem 2.** *Every finite extension of a field is an algebric extension. But the converse is not necessarily tree.*

**Proof :** Let $K$ be a finite extension of a field $F$ and let degree of $K$ over $F$ be $n$, that is $[K : F] = n$. Now we have to show that $K$ is an algebric extension over $F$. For this it is sufficient to show that every element of $K$ is an algebraic element over $F$.

Let $\alpha$ be an arbitrary element of $K$. Since $K$ is a field, so $\alpha, \alpha^2, \ldots\ldots \alpha^n$ are all belong to $K$ but $1 \in K$, so the set $\{1, \alpha, \alpha^2, \ldots\ldots, \alpha^n\}$ of $n + 1$ elements of $K$ is linearly dependent because dimension of $K$ over $F$ is $n$. Since the set $\{1, \alpha, \alpha^2, \ldots\ldots, \alpha^n\}$ is linearly dependent, so there exist elements $a_0, a_1, a_2, \ldots\ldots, a_n$ in $F$ not all zero, such that $a_0 + a_1 \alpha + a_2 \alpha^2 + \ldots\ldots + a_n \alpha^n = 0$.

This shows that $f(x) = a_0 + a_1 x + a_2 x^2 + \ldots\ldots + a_n x^n$ is a nonzero polynomial in $F[x]$ having $\alpha$ as a root, and so $\alpha$ is an algebric element over $F$. Since $\alpha$ is any element of $K$, so every element of $K$ is algebric over $F$ and hence $K$ is an algebric extension of $F$.

However, the converse of above is not necessarily true, because if we consider $K$ be the field of all those complex numbers that are algebraic over $Q$, then $K$ is an algebraic extension of $Q$ that is not finite.

**Theorem 3.** *Let K be a field extension of a field F and let* $\alpha \in K$ *be algebric over F. Then any two minimal monic polynomial for* $\alpha$ *over F are equal.*

**Proof :** Let $p(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots + a_n$,

and $q(x) = x^n + b_1 x^{n-1} + b_2 x^{n-2} + \ldots + b_n$

be two minimal monic polynomials of degree $n$ for $\alpha$ over $F$. Then

$$\alpha^n + a_1 \alpha^{n-1} + a_2 \alpha^{n-2} + \ldots + a_n = 0,$$

and $$\alpha^n + b_1 \alpha^{n-1} + b_2 \alpha^{n-2} + \ldots + b_n = 0.$$

Hence $$\alpha^n + a_1 \alpha^{n-1} + a_2 \alpha^{n-2} + \ldots + a_n = \alpha^n + b_1 \alpha^{n-1} + b_2 \alpha^{n-2} + \ldots + b_n$$

or $$(a_1 - b_1) \alpha^{n-1} + (a_2 - b_2) \alpha^{n-2} + \ldots + (a_n - b_n) = 0,$$

which shows that a is $\alpha$ root of the polynomial

$$h(x) = (a_1 - b_1) x^{n-1} + (a_2 - b_2) x^{n-2} + \ldots + (a_n - b_n)$$

of degree $n-1$ in $F[x]$, which contradicts the fact that the minimal polynomial for $\alpha$ over $F$ is of degree $n$ over $F$. Hence $h(x)$ must be a zero polynomial. This shows that $a_i - b_i = 0$, that is $a_i = b_i$ for $i = 1, 2, \ldots, n$ and hence $p(x) = q(x)$.

**Theorem 4.** *Let K be a field extension of a field F and let* $\alpha \in K$, *where* $a \neq 0$ *and* $\alpha$ *is algebric over F. Then there is an irreducible polynomial* $p(x) \in F[x]$ *such that* $p(\alpha) = 0$. *This irreducible polynomial is uniquely determined upto a constant factor in F and is a polynomial of minimal degree* $\geq 1$ *in* $F[x]$ *having* $\alpha$ *as a zero. If* $f(\alpha) = 0$ *for* $f(x) \neq 0$, *then* $p(x)$ *divides* $f(x)$.

**Proof :** Let $\phi_\alpha$ be the basic homomorphism of $F[x]$ into $K$, that is,

$\phi_\alpha : F[x] \to K$ defined by

$\phi_\alpha [f(x)] = f(\alpha)$ for $f(x) \in F[x]$.

Clearly $\phi_\alpha$ is a ring homomorphism. Let $I = \ker \phi_\alpha = \{f(x) \in F[x] \mid f(\alpha) = 0\}$. Since $\alpha$ is algebric over $F$, so $I$ is a nonzero ideal in $F[x]$. Again, since $F[x]$ is a principal ideal domain, so $I$ must be a principal ideal. Since $I$ is a principal ideal, so there exists $p(x) \in F[x]$ such that $I = \langle p(x) \rangle$.

Now if $f(\alpha) = 0$ for $f(x) \in F[x]$, which $f(x) \neq 0$, then $f(x) \in I = \langle P(x) \rangle$, so there exists $g(x) \in F[x]$ such that $f(x) = p(x)\, g(x)$. This shows that $p(x)$ divides $f(x)$. Clearly, $p(x)$ is a polynomial of minimal degree $\geq 1$ having $\alpha$ as a root, and any other such polynomial of the same degree as $p(x)$ must be of the form $\lambda\, p(x)$ for some $\lambda \in F$.

Now it remains to show that $p(x)$ is irreducible. Let, if possible $p(x)$ be reducible over $F$. Then there exist $q(x)$, $h(x)$ in $F[x]$, such that

$$p(x) = q(x)\, h(x),$$

where deg $q(x)$ and deg $h(x) <$ deg $p(x)$.

Now $p(\alpha) = 0$ implies that $q(\alpha)\, h(\alpha) = 0$,

So either $q(\alpha) = 0$ or $h(\alpha) = 0$,

since $K$ is a field and so it is without zero devises. This contradicts the fact that $p(x)$ is a polynomial of minimal degree $\geq 1$ such that $p(\alpha) = 0$. Hence $p(x)$ is irreducible polynomial.

**Theorem 5.** *Let K be a field extension of a field F. Then an element $a \in K$ is algebric over F if and only if F(a) is finite extension of F, that is, $[F(a) : F]$ is finite.*

**Proof :** Let $F(a)$ be a finite extension of $F$ and let $[F(a) : F] = n$, a positive integer.

Since $F(a)$ is a field and $a \in F(a)$, so $a, a^2,...,a^n$ are all belong to $F(a)$, but $1 \in F(a)$, so the set $\{1, a, a^2,...,a^n\}$ of $n + 1$ elements of $n$-dimensional vector space $F(a)$ over $F$ must be linearly dependent.

Hence, there exist $\alpha_0, \alpha_1, \alpha_2,...,\alpha_n \in F$, not all zero such that

$$\alpha_0 + \alpha_1 a + \alpha_2 a^2 +...+\alpha_n a^n = 0.$$

This show that a satisfied a nonzero polynomial

$$f(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 +...+\alpha_n x^n \in F[x]$$

of degree $n$ and hence $a$ is algebric over $F$.

Conversely, suppose that $a$ is algebric over $F$ and $p(x)$ is the minimal polynomial of a over $F$ such that deg $p(x) = n$. Consider any nonzero element

$$q(a) = \alpha_0 + \alpha_1 a + \alpha_2 a^2 +...+\alpha_m a^m \text{ in } F[a].$$

Then $q(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 +...+\alpha_m x^m \in F[x]$ and $p(x)$ does not devise $q(x)$ and consequently greatest common divisor of $p(x)$ and $q(x)$ is 1. So, there exist $g(x)$ and $h(x) \in F[x]$ such that

$$p(x)\, g(x) + q(x)\, h(x) = 1$$

$$\Rightarrow \quad p(a)\, g(a) + q(a)\, h(a) = 1$$

$$\Rightarrow \quad q(a)\, h(a) = 1, \quad \text{since} \quad p(a) = 0$$

$$\Rightarrow \quad [q(a)]^{-1} = h(a) \in F[a].$$

Hence $F[a]$ is a field. Since $F(a)$ is the field of quotient of $F[a]$, so $F(a) = F[a]$. Again, since $a \in F(a)$ and $F(a)$ is a field, so $1, a, a^2,..., a^{n-1}$ are all in $F(a)$. If for some elements $\alpha_0, \alpha_1, \alpha_2,...,\alpha_{n-1} \in F$ not all zero such that $\alpha_0 + \alpha_1 a + \alpha_2 a^2 +...+ \alpha_{n-1} a^{n-1} = 0$, then a satisfies a nonzero polynomial of degree $< n = \deg p(x)$ of minimal polynomial. This is a contradiction and hence $1, a, a^2,..., a^{n-1}$ are linearly independent over $F$. Let $f(a) = b_0 + b_1 a + b_2 a^2 +...+ b_m a^m$ be any element of $F[a] = F(a)$, then $f(x) = b_0 + b_1 x + b_2 x^2 +...+ b_m x^m \in F[x]$. By division algorithm property in $F[x]$ there exist $q(x)$ and $r(x) \in F[x]$ such that

$$f(x) = q(x)\, p(x) + r(x),$$

where $r(x) = 0$ or $deg\ r(x) < deg\ p(x)$.

$$\Rightarrow \quad f(a) = q(a)\, p(a) + r(a)$$

$$\Rightarrow \quad f(a) = r(a), \text{ since } p(a) = 0.$$

Let $r(x) = c_0 + c_1 n +...+ c_{n-1} a^{n-1}$

since $r(x) = 0$ or $deg\ r(x) < deg\ p(x)$

Thus $f(a) = r(a) = c_0 + c_1 a +...+ c_{n-1} a^{n-1}$,

which shows that $f(a)$ is a linear combination of $1, a, a^2,..., a^{n-1}$. Hence set $\{1, a, a^2,..., a^{n-1}\}$ is a basis of $F(a)$ over $F$ and consequently, $[F(a) : F] = n$, which is finite.

**Theorem 6.** *Let K be a field extension of a field F and* let $\alpha_1, \alpha_2,..., \alpha_n$ be elements in $K$ which are algebric over $F$. *Then* $F(\alpha_1, \alpha_2,..., \alpha_n)$ is a finite extension of $F$ and hence an algebric extension of $F$.

**Proof :** Since $\alpha_1, \alpha_2,..., \alpha_n$ are elements of an extension field $K$ of $F$ which are algebric over $F$, so $F(\alpha_1, \alpha_2,..., \alpha_n)$ is the smallest subfield of $K$ containing $\alpha_1, \alpha_2,..., \alpha_n$ and $F$. By definition of $F(\alpha_1, \alpha_2,..., \alpha_n)$ we have

$$F \subset F(\alpha_1) \subset F(\alpha_1, \alpha_2) \subset ... \subset F(\alpha_1, \alpha_2,...\alpha_{n-1}) \subset F(\alpha_1, \alpha_2,...\alpha_{n-1}, \alpha_n) \subset K.$$

Since every nonzero polynomial over $F$ is a nonzero polynomial over $F(\alpha_1, \alpha_2,...\alpha_{i-1})$ as it is a superfield of $F$, so $\alpha_i$ is algebric over $F$ implies that $\alpha_i$ is algebric over $F(\alpha_1, \alpha_2,...\alpha_{i-1})$ and hence by theorem 5, we get

$[F(\alpha_1, \alpha_2,...\alpha_{i-1})](\alpha_i) = F(\alpha_1, \alpha_2,...\alpha_i)$ is a finite extension of $F(\alpha_1, \alpha_2,...\alpha_{i-1})$.

Let $[F(\alpha_1, \alpha_2,...\alpha_i) : F(\alpha_1, \alpha_2,...\alpha_{i-1})] = m_i$, where $m_i$ is finite, then

$[F(\alpha_1, \alpha_2,...\alpha_n) : F] = [F(\alpha_1, \alpha_2,...\alpha_n) : F(\alpha_1, \alpha_2,...\alpha_{n-1})] \times [F(\alpha_1, \alpha_2,...\alpha_{n-1}) : F(\alpha_1, \alpha_2,...\alpha_{n-2})] \times ... \times [F(\alpha_1) : F] = m_n m_{n-1}...m_1$ = finite number, since each $m_i$ is finite.

Thus $F(\alpha_1, \alpha_2,...\alpha_n)$ is a finite extension of $F$.

Since every finite extension of a field is an algebric extension, so $F(\alpha_1, \alpha_2,...\alpha_n)$ is an algebric extension of $F$.

**Theorem 7.** *Let $F \subset K \subset L$ be three fields. If $L$ is an algebric extension of $K$ and if $K$ is an algebric extension of $F$, then $L$ is an algebric extension of $F$.*

**Proof :** Here we have to show that $L$ is an algebric extension of $F$. For this we shall show that every element of $L$ is algebric over $F$. Let $\alpha$ be any arbitrary element of $L$. Since $L$ is an algebric extension of $K$, so there exists a nonzero monic polynomial

$$x^n + a_1 x^{n-1} + a_2 x^{n-2} + ... + a_n \in K[x], a_i \in K$$

such that $$\alpha^n + a_1 \alpha^{n-1} + a_2 \alpha^{n-2} + ... + a_n = 0.$$

Again, since $K$ is an algebric extension of $F$ and $a_1, a_2,...a_n$ are elements of $K$, so each of $a_1, a_2,...a_n$ is algebric over $F$ and hence by theorem 6, $E = F(a_1, a_2,...a_n)$ is a finite extension of $F$. Since a satisfies the polynomial $x^n + a_1 x^{n-1} + a_2 x^{n-2} + ... + a_n$ whose coefficients $a_i \in E$ for $i = 1, 2,...,n$, so a is algebric over $E$ and hence by theorem 6 $[E(\alpha) : E]$ is finite.

Since $E$ is a finite extension of $F$ and $E(\alpha)$ is a finite extension of $E$, so by transitive property of finite field extension $E(\alpha)$ is a finite extension of $F$. Since every finite extension is an algebric extension, so $E(\alpha)$ is an algebric extension of $F$ and hence $\alpha$ is algebric element over $F$. But $\alpha \in L$, so $L$ is an algebric extension of $F$.

**Theorem 8.** *If F is a field and p (x) be an irreducible polynomial of positive degree over a field F, then there is an extension K = F [x] / p <x > of F such that [K : F] = deg p (x) and p (x) has a root in K.*

**Proof :** Let $p(x) = \alpha_0 + \alpha_1 x + ... + \alpha_n x^n$, where $\alpha_i \in F$ be an irreducible polynomial in $F[x]$ of degree $n$. Since $p(x)$ is irreducible, the principal ideal $I = < p(x) >$ is a nonzero prime ideal. Again, since $F[x]$ is a principal ideal domain, $I$ is a maximal ideal and hence $K = F[x] / I$ is a field.

Let $L = \{I + \alpha \mid \alpha \in F\}$,

then $L \subset K$. Now we shall show that $L$ is a subfield of $K$ isomorphic to $F$. If is easy to verify that $L$ is a subfield of $K$. Consider a mapping

$\phi : F \to L$ defined by $\phi(\alpha) = I + \alpha$ for all $\alpha \in F$. Then for all $\alpha, \beta \in F$, we have

$$\phi(\alpha + \beta) = I + (\alpha + \beta)$$
$$= (I + \alpha) + (I + \beta)$$
$$= \phi(\alpha) + \phi(\beta), \text{ and}$$
$$\phi(\alpha\beta) = I + \alpha\beta$$
$$= (I + \alpha)(I + \beta)$$
$$= \phi(\alpha)\,\phi(\beta)$$

This shows that $\phi$ is a homomorphism. Let

$\phi(\alpha) = \phi(\beta)$, then

$I + \alpha = I + \beta$, this implies $\alpha - \beta \in I$.

Since $I = < p(x) >$ is a principal ideal and $\alpha - \beta \in I$, so there exist some $g(x) \in F[x]$ such that $\alpha - \beta = p(x)\, g(x)$.

Since $\alpha - \beta$ is a constant polynomial,

so $g(x) = 0$ because $p(x)$ is a polynomial of positive degree, its product with any other nonzero polynomial cannot be equal to a constant polynomial.

Now $g(x) = 0$ implies $\alpha - \beta = 0$, that is, $\alpha = \beta$.

Thus $\phi(\alpha) = \phi(\beta)$ implies $\alpha = \beta$ and hence $\phi$ is one-one.

$\phi$ is clearly onto, since for any $I + \alpha \in L$, $\alpha \in F$

such that $\phi(\alpha) = I + \alpha$.

Thus $\phi$ is an isomorphism and hence $L \cong F$.

Since $L$ is a subfield of $K$ and $L \cong F$, so $F$ may be regarded as a subfield of $K$, that is, $K$ is an extension of $F$.

Let $m = I + x \in K$, we claim that $m$ is a root of

$p(x) = \alpha_0 + \alpha_1 x + ... + \alpha_n x^n$ in $K$.

Now in $K$, we have

$$\begin{aligned} p(m) &= (I+\alpha_0) + (I+\alpha_1)\, m + \ldots + (I+\alpha_n)\, m^n \\ &= (I+\alpha_0) + (I+\alpha_1)(I+x) + \ldots + (I+\alpha_n)(I+x)^n \\ &= (I+\alpha_0) + (I+\alpha_1)(I+x) + \ldots + (I+\alpha_n)(I+x^n) \\ &= (I+\alpha_0) + (I+\alpha_1 x) + \ldots + (I+\alpha_n x^n) \\ &= I + (\alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n) \\ &= I + p(x) \\ &= I, \text{ since } I = < p(x) >. \end{aligned}$$

But $I + 0 = I$ is the zero element in $K$, and hence $m$ is a root of $p(x)$ in $K$.

Now it remains to show that $[K : F] = deg\ p(x)$.

We claim that the set $S = \{I + 1, I + x, \ldots, I + x^{n-1}\}$ is a basis of $K$. Let

$$a_0(I+1) + a_1(I+x) + \ldots + a_{n-1}(I+x^{n-1}) = I$$

$$\Rightarrow \quad (I+a_0) + (I+a_1 x) + \ldots + (I+a_{n-1} x^{n-1}) = I$$

$$\Rightarrow \quad I + (a_0 + a_1 x + \ldots + a_{n-1} x^{n-1}) = I$$

$$\Rightarrow \quad a_0 + a_1 x + \ldots + a_{n-1} x^{n-1} \in I = < p(x) >,$$

so there exists $g(x) \in F[x]$ such that

$$a_0 + a_1 x + \ldots + a_{n-1} x^{n-1} = p(x)\, g(x).$$

Since left hand side is a polynomial of degree $n - 1$, so

$$g(x) = 0 \quad \text{because if} \quad g(x) \neq 0,$$

then $deg\ p(x)\, g(x) \geq deg\ p(x) = n$.

Now $g(x) = 0$ implies

$$a_0 + a_1 x + \ldots + a_{n-1} x^{n-1} = 0,$$

which is possible only when $a_0 = a_1 = \ldots a_{n-1}$ and hence the set $S$ is linearly independent.

Let $I + g(x)$ be any dement of $K$, then $g(x) \in F[x]$. By division algorithm property in $F[x]$ there exist $q(x)$ and $r(x)$ in $F[x]$ such that

$$g(x) = p(x)\, q(x) + r(x),$$

where either $r(x) = 0$ or $deg\ r(n) < deg\ p(x)$.

Now

$$\begin{aligned} I + g(x) &= I + [p(x)\, q(x) + r(x)] \\ &= [I + p(x)\, q(x)] + [I + r(n)] \\ &= I + r(x), \quad \text{since} \quad p(x)\, q(x) \in I \\ &= I + (a_0 + a_1 x + \ldots + a_{n-1} x^{n-1}), \quad \text{since} \quad deg\ r(x) < deg\ p(x) = n. \\ &= (I + a_0) + (I + a_1 x) + \ldots + (I + a_{n-1} x^{n-1}) \\ &= a_0(I+1) + a_1(I+x) + \ldots + a_{n-1}(I+x^{n-1}) \end{aligned}$$

$\Rightarrow$ $I + g(x)$ is expressible as a linear combination of elements of $S$. But $I + g(x)$ be any element of $K$, so every element of $K$ is expressible as a linear combination of element of $S$ and hence $S$ spans $K$. This shows that $S$ is a basis of $K$ over $F$ and consequently $dim\ K(F) = n$, *i.e.,* $[K : F] = deg\ p(x)$.

**Corollary :** Let $p(x)$ be a polynomial of positive degree over a field $F$. Then there is an extension field $K$ of $F$ such that $p(x)$ has a root in $K$ and $[K : F] \leq deg\, p(x)$.

**Proof :** Let $deg\, p(x)$ be $n$. If $p(x)$ is irreducible over $F$, then by above theorem there exists an extension $K$ of $F$ such that $p(x)$ has a root in $K$ and hence $[K : F] = n = deg\, p(x)$. Again, if $p(x)$ is not irreducible over $F$, then there exists irreducible factor $g(x)$ of $p(x)$ in $F[x]$ such that $p(x) = g(x)\, q(x)$, for some $q(x) \in F[x]$. Clearly $deg\, g(x) < deg\, p(x)$ and again by above theorem there exists an extension $K$ of $F$ such that $p(x)$ has a root in $K$, that is, $p(\alpha) = 0$ for some $\alpha \in K$. Therefore $p(a) = g(a)\, q(a)\, 0$, which shows that a is algebric over $F$ of degree $n$ and hence $[K : F] \leq n$.

**Theorem 9.** *Let $f(x)$ be a polynomial of degree $n \geq 1$ over a field $F$, then there exists a finite extension $K$ of $F$ in which $f(x)$ has $n$ roots such that $[K : F] \leq \lfloor n$.*

**Proof :** We shall prove the theorem by using induction on the $deg\, f(x) = n$. If $deg\, f(x) = 1$, then $f(x)$ is of the form $\alpha x + \beta$, $\alpha, \beta \in F$ and $\alpha \neq 0$. In this case $F$ can be considered as extension of itself with $[K : F] = 1$ and $x = -\frac{\beta}{\alpha} \in F$ is a root of $f(x)$. Thus the theorem is true when $deg\, f(x) = 1$.

Let us assume that the theorem holds, for all polynomials of degree less then $n$ over the field $F$. Since $f(x)$ is a polynomial of degree $n$ over $F$, so by above corollary there exists an extension $L$ of $F$ such that $f(x)$ has a root $\alpha$ in $L$ such that $[L : F] \leq deg\, f(x) = n$. Again, since $\alpha \in L$ is a root of $f(x)$ in $L$, so over $L[x]$ we have

$$f(x) = (x - \alpha)\, g(x), \text{ where } deg\, g(x) = n - 1.$$

Since $g(x)$ is a polynomial of degree $n - 1 < n$ over the field $L$, so by our induction hypothesis there exists an extension $K$ of $L$ of degree at most $\lfloor n-1$ in which $g(x)$ has $n - 1$ roots in $K$.

Since $K$ is a finite extension of $L$ and $L$ is a finite extension of $F$, so by transitive property of finite field extension $K$ is a finite extension of $F$ and

$$[K : F] = [K : L][L : K] \leq \lfloor n-1 \,.\, n = \lfloor n \quad \text{or} \quad [K : F] \leq \lfloor n \,.$$

Thus we have shows that $K$ is a finite extension of $F$ of degree at most $\lfloor n$ in which $f(x)$ has $n$ roots.

**Theorem 10.** *Let $K$ be an extension of a field $F$. Then the elements in $K$ which are algebric over $F$ from a subfield of $K$.*

**Proof :** Let $S$ be a collection of all those elements of $K$ which are algebric over $F$. Clearly $S$ is a subset of $K$. Here we have to show that $S$ is a subfield of $K$. For this it is sufficient to show that for all $a$, $b, \in S$ implies $a \pm b$, $ab$ and $\frac{a}{b}(b \neq 0)$ are all in $S$. Let $a$ and $b$ be any two elements of $S$. Then $a$ and $b$ are algebric over $F$ implies that $b$ is also algebraic over $F(a)$, since $F(a)$ is a super field of $F$. Since $b$ is algebric over $F(a)$, so by Theorem 5 $[F(a)](b) = F(a, b)$ is a finite extension of $F(a)$ and hence $[F(a, b) : F(a)]$ is finite. Again, since a is algebric over $F$, so $[F(a) : F]$ is finite. Since $F \subset F(a) \subset F(a, b)$, therefore by transitive property of finite field extension, $F(a, b)$ is a finite extension of $F$. Since every finite field extension is an algebric field extension, so $F(a, b)$ is an algebric field extension of $F$.

Again, since $F(a, b)$ is a field and $a, b$ are elements in $F(a, b)$, therefore $a \pm b$, $ab$ and $\frac{a}{b}(b \neq 0)$ are all in $F(a, b)$. Consequently $a \pm b$, $ab$ and $\frac{a}{b}(b \neq 0)$ are all algebric over $F$ and hence $a \pm b$, $ab$ and $\frac{a}{b}(b \neq 0)$ are all elements of $S$ as required.

**Corollary :** Let $K$ be an extension of a field $F$. If $a$ and $b$ in $K$ are algebric over $F$ of degree $m$ and $n$ respectively, then $a \pm b$, $ab$ and $\frac{a}{b}(b \neq 0)$ are algebric over $F$ of degree atmost $mn$.

**Proof :** Since $a$ is algebric of degree $m$ and $b$ is algebric of degree $n$ over $F$, therefore by theorem 8, we get

$$[F(a) : F] = m \quad \text{and } [F(b) : F] = n.$$

Now $b$ is algebric of degree $n$ over $F$ implies $b$ is algebric of degree atmost $n$ over $F(a)$ as $F(a)$ is a super field of $F$. Thus

$$[(F(a))(b) : F(a)] = [F(a, b) : F] \leq n.$$

By transitive property of field extension we get

$$[F(a, b) : F] = [F(a, b) : F(a)]\,[F(a) : F] \leq m\,n = \text{finite, since } m, n \text{ are finite.}$$

This shows that $F(a, b)$ is a finite extension of $F$ and hence it is also an algebric extension of $F$. Since $F(a, b)$ is a field containing $a$ and $b$, it follows that $a \pm b$, $ab$, $\frac{a}{b}(b \neq 0)$ are all in $F(a, b)$. If follows that $a \pm b$, $ab$ and $\frac{a}{b}(b \neq 0)$ are algebric of degree atmost $mn$ over $F$.

## Self-learning exercise-1

**1.** Consider the set $Q(\sqrt[3]{5}) = \{a + b\sqrt[3]{5} + c\sqrt[3]{25} : a, b, c \in Q\}$, then :

(A) $\left[Q\sqrt[3]{5} : Q\right] = 2$ (B) $\left[Q\sqrt[3]{5} : Q\right] = 3$

(C) $\left[Q\sqrt[3]{5} : Q\right] = 1$ (D) None of these

**2.** If $L$ is a field extension of a field $F$ such that $[L : F] = p$, $a$ prime number and if $K$ is any subfield of $L$ containing $F$, then :

(A) $K = L$ (B) $K = F$

(C) $K \subset F \subset L$ (D) $K = L$ or $K = F$

**3.** If $w$ is one of the imaginary cube roots of unity, then $Q(w)$ is a finite field extension of $Q$.

(True / False)

**4.** The field $C$ of complex numbers is not an algebric extension of $R$. (True / False)

**5.** The field $R(i = \sqrt{-1})$ is a simple field extension of the field $R$ of real numbers and $C = R(i)$.

(True / False)

## 7.5 Some important examples

**Ex.1.** *If $K$ is a finite extension of degree $n = [K : F]$ over $F$, then show that every element $u$ of $K$ has over $F$ a degree which is a divisor of $n$.*

**Sol.** It is given that $K$ is a finite extension of $F$ such that $[K : F] = n$. Since $u \in K$, so $F(u)$ is a simple field extension containing both $u$ and $F$ such that $F \subset F(u) \subset K$. Again, since $[K : F]$ is finite, so $[F(u) : F]$ is also finite as every subspace of a finite dimensional vector space is finite. Since every finite field extension is an algebraic extension so $F(u)$ is an algebric extension of $F$ and hence $u$ is an algebric element of $K$ over $F$ and consequently $[F(u) : F]$ = degree of $u$.

By theorem1 of this unit we have $[K : F] = [K : F(u)]\ [F(u) : F]$

$\Rightarrow$ $n = [K : F(u)] \times$ degree of $u$ over $F$

$\Rightarrow$ degree of $u$ over $F$ is a divisor of $n$.

**Ex.2.** *Let $Q$ be the field of rational numbers, then show that* $Q\left(\sqrt{2}, \sqrt{3}\right) = Q\left(\sqrt{2} + \sqrt{3}\right)$.

**Sol.** By definition of simple field extension, we have

$$Q\left(\sqrt{2}, \sqrt{3}\right) = \left\{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{2}\sqrt{3} \,\middle|\, a, b, c, d \in Q\right\}$$

and

$$Q\left(\sqrt{2}, \sqrt{3}\right) = \left\{a + b\left(\sqrt{2} + \sqrt{3}\right) \,\middle|\, a, b \in Q\right\}.$$

Since $Q\left(\sqrt{2}, \sqrt{3}\right)$ is a field containing both $\sqrt{2}$ and $\sqrt{3}$,

so $\sqrt{2} + \sqrt{3} \in Q\left(\sqrt{2}, \sqrt{3}\right)$

$\Rightarrow$ $Q\left(\sqrt{2} + \sqrt{3}\right) \subseteq Q\left(\sqrt{2}, \sqrt{3}\right)$ .....(1)

Since $Q\left(\sqrt{2} + \sqrt{3}\right)$ is a field containing $\sqrt{2} + \sqrt{3}$, therefore

$$\left(\sqrt{2} + \sqrt{3}\right)^{-1} \in Q\left(\sqrt{2} + \sqrt{3}\right)$$

which implies $\sqrt{3} - \sqrt{2} \in Q\left(\sqrt{2} + \sqrt{3}\right)$.

Now $$\sqrt{3} = \frac{1}{2}2\sqrt{3} = \frac{1}{2}\left(2 + \sqrt{3} + \sqrt{3} - \sqrt{2}\right) \in Q\left(\sqrt{2} + \sqrt{3}\right)$$

and $$\sqrt{2} = \frac{1}{2}2\sqrt{2} = \frac{1}{2}\left[\left(2 + \sqrt{3}\right) - \left(\sqrt{3} - 2\right)\right] \in Q\left(\sqrt{2} + \sqrt{3}\right)$$

Therefore $\sqrt{2}, \sqrt{3} \in Q\left(\sqrt{2} + \sqrt{3}\right)$

$\Rightarrow$ $Q\left(\sqrt{2}, \sqrt{3}\right) \subseteq Q\left(\sqrt{2} + \sqrt{3}\right)$ .....(2)

From (1) and (2), we get

$$Q\left(\sqrt{2}, \sqrt{3}\right) = Q\left(\sqrt{2} + \sqrt{3}\right).$$

**Ex.3.** *Let K be an extension of a field F. Let α be an algebric element of odd degree over F. Show that $\alpha^2$ is algebric over F and that F (α) = F ($\alpha^2$).*

**Sol.** Since α is an algebric element of odd degree over $F$, so $F(\alpha)$ is the extension of $F$ containing both α and $F$ such that $[F(\alpha) : F]$ = odd number.

Since $F(\alpha)$ is finite extension of $F$, so $F(\alpha)$ is an algebric extension of $F$. Again, since $F(\alpha)$ is a field, so $\alpha \in F(\alpha)$, thus $\alpha^2$ is algebric over $F$ and $F(\alpha^2) \subseteq F(\alpha)$.

Since a is α root of the polynomial $x^2 - \alpha^2$ with coefficient in $F$ and $F(\alpha^2)$ is a super field of $F$, so α is a root of the polynomial $x^2 - \alpha^2$ with coefficients in $F(\alpha^2)$, so α is algebric of degree at most 2 over $F(\alpha^2)$ that is, $[F(\alpha) : F(\alpha^2)] \leq 2$.

By theorem 1 of this unit we have

$$[F(\alpha) : F] = [F(\alpha) : F(\alpha^2)]\,[F(\alpha^2) : F]$$

$\Rightarrow$ $[F(\alpha) : F(\alpha^2)]$ divides $[F(\alpha) : F]$

But $[F(\alpha) : F(\alpha^2)] \leq 2$ and $[F(\alpha) : F]$ is odd, then we must have $[F(\alpha) : F(\alpha^2)] = 1$ and hence $F(\alpha) = F(\alpha^2)$.

## 7.6 Summary

In this unit we have discussed about simple field extension, algebraic and transcendental field extension and some important results on these topics.

## 7.7 Answers to self-learning exercises

**Self-learning exercise-1**

**1.** (B) **2.** (D) **3.** True **4.** False **5.** True.

## 7.8 Exercises

1. Show that a finite field extension of prime degree is a simple extension.
2. Let $K$ be a field extension of a field $F$. Let α, β, $\in K$ be algebric over $F$ of degree $m$ and $n$ respectively, and let $m$, $n$ be relatively prime. Then prove that $[F(\alpha, \beta) : F] = m\,n$.
3. Let $K$ be a finite field extension of a field $F$. If for any two subfields $L_1$ and $L_2$ of $K$ containing $F$, either $L_1 \subset L_2$ or $L_2 \subset L_1$, show that $K$ is a simple extension of $F$.
4. Let $R$ be the field of real numbers and $Q$ the field of rational numbers. In $R$, $\sqrt{2}$ and $\sqrt{3}$ are both algebric over $Q$. Exhibit a polynomial of degree 4 over $Q$ satisfied by $\sqrt{2} + \sqrt{3}$.

   [**Ans.** $f(x) = x^4 - 10x^2 + 1$]
5. If $K$ is an extension of a field $F$ of prime degree, then prove that any element in $K$ but not in $F$ generates all of $K$ over $F$.

□ □ □

# UNIT 8 : Splitting Fields, Normal Extension, Separable and Inseparable Extensions and Automorphism of Extensions

**Structure of the Unit**



## 8.0 Objectives

After reading this unit you will be able to under stand about splitting fields of a polynomial, normal extension, separable and inseparable extensions and automorphism of extensions. You will be also to know that if $F$ is a field, then every polynomial $f(x) \in F[x]$ has a splitting field, and irreducible polynomial $f(x) \in F[x]$ is separable if and only if $f'(x) \neq 0$ and every non-constant polynomial over a field of characteristic zero is separable.

## 8.1 Introduction

In this unit we shall take up the study of splitting fields of polynomial $f(x) \in F[x]$, where $F$ is a field, normal extension, separable and inseparable extensions and automorphism of extensions. We will prove some important results related to these topics. In the end of the unit we will discuss some important problems related to these topics.

## 8.2 Splitting fields

Let $K$ be a field extension of a field $F$ and Let $p(x)$ be a polynomial in $F[x]$ of degree $n \geq 1$. Then $K$ is said to be a **splitting field** of $p(x)$ if $p(x)$ can be factored into $n$ linear factors over $K$, that is,

$$p(x) = a(x-\alpha_1)(x-\alpha_2)...(x-\alpha_n)$$

where a is a nonzero element of $F$ and $\alpha_i \in K$ ; $i = 1, 2,..., n$, and there does not exist any proper subfield $L$ of $K$ containing $F$ such that $p(x)$ can be factored into $n$ linear factors over $L$.

In other words we can say that $K$ is a splitting field of $p(x)$, if $K$ contains all $n$ roots $\alpha_1, \alpha_2,..., \alpha_n$ of $p(x)$ and $K = F(\alpha_1, \alpha_2,..., \alpha_n)$.

***Ex.1.*** *If w is imaginary cube root of unity, then splitting field of polynomial*

$$x^3 - 2 \in Q[x] \text{ is } Q\left(w, \sqrt[3]{2}\right).$$

***Ex.2.*** *If w is imaginary cube root of unity, then* $x^3 - 1 \in Q[x]$ *splits over the field C of complex numbers but its splitting field is* $Q(w)$.

**Theorem 1.** *Let F be a field, then every polynomial of positive degree in* $F[x]$ *has a splitting field.*

**Proof.** Let $p(x) \in F[x]$ be a polynomial of degree $n \geq 1$. Now we have to show that $p(x)$ has a splitting field. We shall prove the theorem by induction on $n$. If $n = 1$, then $p(x)$ is a linear, that is, $p(x) = \alpha x + \beta$ with $\alpha, \beta \in F$ and $\alpha \neq 0$. Thus we have

$$p(x) = \alpha\left(x - \frac{\beta}{\alpha}\right).$$

Hence $K = F$ is a splitting field of $p(x)$ in this case.

Now assume that the theorem is true for all polynomials of degree less-then $n$.

If $\quad deg.\ p(x) = n > 1, \quad$ then let

$$p(x) = p_1(x)\, p_2(x)...p_l(x) \qquad .....(1)$$

where each of the polynomial on right hand side of (1) is irreducible over $F[x]$. If for each $l$, $p_l(x)$ is a polynomial of degree one, then $F$ itself is a splitting field of $p(x)$. Now suppose that at least one of them say $f_i(x)$ is of degree $\geq 2$, $1 \leq i \leq l$. These there exists an extension $F(\alpha_1)$ containing a root $\alpha_1$ of $f_i(x)$. Thus in the field $F(\alpha_1)$ we have

$$p(x) = (x - \alpha_1) f(x),$$

where $deg\, f(x) = n-1$ and $f(x) \in F(\alpha_1)[x]$.

Now, since degree of $f(x)$ is less-then $n$, so by over assumption there exists splitting field

$$[F(\alpha_1)](\alpha_2, \alpha_3,...,\alpha_n) = F(\alpha_1, \alpha_2,...,\alpha_n) \text{ of } f(x).$$

Clearly, then $K = F(\alpha_1, \alpha_2,...,\alpha_n)$ is a splitting field of $p(x) \in F[x]$.

**Theorem 2.** *Let F be a field, and let f (x) be a nonzero polynomial in F [x]. Then splitting field of f (x) is an algebraic extension of F.*

**Proof.** Let degree of $f(x)$ be $n \geq 1$ and let $E$ be the splitting field of $f(x)$. Also, let $\alpha_1, \alpha_2,...,\alpha_n$ be the roots of $f(x)$. Then $E = F(\alpha_1, \alpha_2,...,\alpha_n)$. Now we have to show that $E$ is an algebraic extension of $F$. Let

$$E_1 = F(\alpha_1)$$
$$E_2 = E_1(\alpha_2) = F(\alpha_1)(\alpha_2) = F(\alpha_1, \alpha_2)$$
$$E_3 = E_2(\alpha_2) = F(\alpha_1, \alpha_2)(\alpha_3) = F(\alpha_1, \alpha_2, \alpha_3)$$
..... .....
..... .....
..... .....
$$E = E_n = E_{n-1}(\alpha_n) = F(\alpha_1, \alpha_2...\alpha_{n-1})(\alpha_n) = F(\alpha_1, \alpha_2,...,\alpha_n).$$

Since each $\alpha_i \in E$ ; $i = 1, 2, ..., n$ is a root of a nonzero polynomial $f(x)$ over F, therefore $\alpha_i$, $i = 1, 2,...,n$ are all algebraic over $F$. Since each of the fields $F, E_1 E_2,..., E_n = E$ is obtained on adjoining an algebraic element to its predecessor, so that each of the degree $[E_1 : F]$, $[E_2 : E_1]$,... $[E_n : E_{n-1}]$ is finite. We have

$$[E : F] = [E : E_{n-1}][E_{n-1} : E_{n-2}],...,[E_1 : F]$$

is finite. Thus $E$ is a finite extension of $F$ and consequently $E$ is an algebraic extension of $F$, since every finite extension is an algebraic extension.

## 8.3 Normal extension

Let $K$ be an algebraic extension of a field $F$. Then $K$ is said to be **normal extension** of $F$, if the splitting field of the minimal polynomial $f(x) \in F[x]$ for each element of $K$ is contained in $K$. In other words we can say that if $K$ is a normal extension of $F$ and $\alpha \in K$ and if $f(x)$ is the minimal polynomial of $\alpha$ over $F[x]$, then $f(x)$ can be expressed as a product of linear factors in $K[x]$.

***Ex.1.*** *The field of real numbers R is not a normal extension of the field of rational numbers Q, because* $2^{1/3} \in R$ *and* $f(x) = x^3 - 2 \in Q[x]$ is the minimal polynomial for $2^{1/3}$, but it does not split into linear factors in $R$ as other two roots of $f(x)$ are imaginary.

***Ex.2.*** *If* $\alpha = \cos\frac{\pi}{4} + i\sin\frac{\pi}{4} = e^{i\pi/4}$, *then* $Q(\alpha)$ *is a normal extension of Q, because the minimal polynomial for* $\alpha$ *is*

$$f(x) = x^4 + 1 \in Q[x] \text{ and}$$
$$f(x) = x^4 + 1 = (x^2 - i)(x^2 + i)$$

$$= \left(x-\sqrt{i}\right)\left(x+\sqrt{i}\right)\left(x-i^{3/2}\right)\left(x+i^{3/2}\right)$$

where $\alpha^2 = i, \sqrt{i} = \alpha, i^{3/2} = \alpha^{3/2}$

and thus $x^4 + 1 = (x-\alpha)(x+\alpha)\left(x-\alpha^3\right)\left(x+\alpha^3\right).$

This shows that $Q(\alpha)$ is the splitting field of $f(x)$

**Theorem 3.** *A finite algebraic extension K of a field F is a normal extension of F if K is splitting field of some polynomial over F.*

**Proof.** It is given that $K$ is a finite algebraic extension of $F$, so $K = F(\alpha_1, \alpha_2 ... \alpha_n)$, where each $\alpha_i \in K$ ; $i = 1, 2, ..., n$ is algebraic over $F$. Let $f_1(x), f_2(x), ..., f_n(x)$ be the minimal polynomials of $\alpha_1$, $\alpha_2 ... \alpha_n$ over $F$ respectively. Since $K$ is a normal extension of $F$, so the splitting field of each

$$f_1(x), f_2(x), ..., f_n(x)$$

is contained in $K$. Hence $K$ is the splitting field of the polynomial $f(x) = f_1(x)\ f_2(x) \ldots f_n(x)$.

**Theorem 4.** *Let K be a normal extension of a field F and L is an intermediate field, so that $F \subset L \subset K$, then K is also a normal extension of L.*

**Proof.** It is given that $K$ is a normal extension of a field $F$ and $L$ is a field such that $F \subset L \subset K$. Now we have to show that $K$ is a normal extension of $L$. For this it is sufficient to show that the splitting field of the minimal polynomial for each element of $K$ over $L$ is contained in $K$.

Let $\alpha$ be any element of $K$ and let $p(x)$ and $q(x)$ be the minimal polynomials for $\alpha$ over $F$ and $L$ respectively. Since $F \subset L$, Therefore $f(x) \in F[x]$ implies $f(x) \in L[x]$. Also $q(x)$ is the minimal polynomial of $\alpha$ over $L$, then $q(x)$ is a divisor of $p(x)$. This shows that each root of $q(x)$ in $K$ is also a root of $p(x)$ in $K$. But each root of $p(x)$ is in $K$, so it follows that each root of $q(x)$ is in $K$. Hence $K$ is a normal extension of $L$.

## 8.4 Separable and inseparable extensions

At first we define roots of a polynomial in an extension field. Let $K$ be an extension field of a field $F$. Then, an element $\alpha \in K$ is said to be a **root** of a polynomial $f(x) \in F[x]$, if $f(x) = 0$. If $\alpha$ is a root of $f(x)$, then $(x - \alpha)$ divides $f(x)$ over $K$. $\alpha$ is said to be a root of **multiplicity m** if $(x - \alpha)^m$ is a divisor of $f(x)$ but $(x - \alpha)^{m+1}$ is not a divisor of $f(x)$.

If $m = 1$, then $\alpha$ is called a **simple root**, otherwise it is called a **multiple root.**

We now define derivative of a polynomial over a field. Let

$$f(x) = a_0 + a_1 x + a_2 x^2 + ... + a_n x^n$$

be a polynomial of degree $n$ over a field $F$. Then **derivative** of $f(x)$ is denoted by $f'(x)$ and is defined by

$$f'(x) = a_1 + 2a_2 x + ... + na_n x^{n-1}.$$

Note that the properties of derivative which are true in calculus are not necessarily true here.

Since $F$ is an arbitrary field, so it may be a field of finite characteristic or it may be a field of characteristic zero or infinite. If $F$ is a field of characteristic $n \neq 0$, then the derivative of a polynomials $f(x) = x^n$ over $F$ is $nx^{n-1} = (n \cdot 1)\, x^{n-1} = 0$, since the characteristic of the field is $n$, so $n \cdot 1 = 0$. Thus the derivative of a non-constant polynomial may be zero. Again, if $F$ is a field of characteristic zero, that is, $F$ is an infinite field and $f(x)$ is a polynomial of degree $n \geq 1$, then $f'(x)$ is a polynomial of degree $n - 1$.

**Theorem 5.** *Let $f(x)$ and $g(x)$ be any two polynomials over a field $F$ and $\alpha \in F$, then*

*(i)* $(f(x) + g(x))' = f'(x) + g'(x)$,

*(ii)* $(\alpha f(x))' = \alpha f'(x)$,

*(iii)* $(f(x)\, g(x))' = f'(x)\, g(x) + f(x)\, g'(x)$.

**Theorem 6.** *Let $K$ be an extension of a field $F$ and let $f(x) \in F[x]$ be a polynomial of degree $n \geq 2$. Then $\alpha \in K$ is a multiple root of $f(x)$ if and only if $\alpha$ is a common root of $f(x)$ and $f'(x)$.*

**Proof.** Let $\alpha$ be a multiple root of $f(x)$ of multiplicity $m \geq 2$, then

$$f(x) = (x - a)^m \; g(x),\ g(x) \in K[x] \text{ and } g(\alpha) \neq 0.$$

Differentiating above with respect to $x$, we get

$$f'(x) = m(x - \alpha)^{m-1} g(x) + (x - \alpha)^m g'(x)$$

$$\Rightarrow \qquad f'(\alpha) = 0,$$

which shows that $\alpha$ is a root of $f'(x)$. Thus $\alpha$ is a common root of $f(x)$ and $f'(x)$.

Conversely, suppose that $\alpha$ is a common root of $f(x)$ and $f'(x)$. Now we have to show that $\alpha$ is a multiple root of $f(x)$. Let, if possible $\alpha$ is the simple root of $f(x)$, then

$$f(x) = (x - \alpha)\, g(x),$$

where $g(x) \in K[x]$ and $g(\alpha) \neq 0$.

Taking the derivative of above we get

$$f'(x) = g(x) + (x - \alpha)\, g'(x)$$

$$\Rightarrow \qquad f'(\alpha) = g(\alpha) \neq 0,$$

which shows that $\alpha$ is not a root of $f'(x)$. Hence $\alpha$ is a multiple root of $f(x)$.

**Theorem 7.** *Let $F$ be a field and let $f(x)$ be an irreducible polynomial in $F[x]$. Then $f(x)$ has a multiple root in some field extension if and only if $f'(x) = 0$.*

**Proof.** Let $\alpha$ be a multiple root of $f(x)$ in some field extension $K$ of $F$. Then by theorem 6, $\alpha$ is a common root of $f(x)$ and $f'(x)$. Since $f(x)$ is an irreducible polynomial over $F$ such that $f(\alpha) = 0$ and $f'(x)$ is another polynomial over $F$ such that $f'(\alpha) = 0$, so $f(x)$ divides $f'(x)$.

If $f'(x) \neq 0$, then $deg\, f'(x) < deg\, f(x)$ and hence $f(x)$ cannot divide $f'(x)$, which contradicts the fact that $f(x)$ divides $f'(x)$ and hence $f'(x) = 0$.

Conversely suppose that $f'(x) = 0$ and $\alpha \in K$ be a root of $f(x)$. Then $f(\alpha) = 0$ and $f'(\alpha) = 0$. This shows that $\alpha$ is a common root of $f(x)$ and $f'(x)$ in $K$ and hence by theorem 6, $\alpha$ is a multiple root of $f(x)$ in $K$.

**Theorem 8.** *Let $F$ be a field and let $f(x)$ be an irreducible polynomial in $F[x]$, Then if the characteristic of $F$ is zero, then $f(x)$ has no multiple root in any field extension of $F$.*

**Proof.** Let degree of $f(x)$ be $n \geq 1$. Since characteristic of $F$ is zero and $f(x)$ is an irreducible polynomial of degree $n \geq 1$, so $deg\, f'(x) = n - 1 \geq 0$ and hence $f'(x) \neq 0$. Hence by above theorem $f(x)$ can not have a multiple root in any field extension of $F$.

**Theorem 9.** *If $F$ is a field of characteristic $p \neq 0$, then the polynomial*

$$f(x) = x^{p^n} - x \in F[x], \text{ for } n \geq 1,$$

*has distinct roots.*

**Proof.** Since $f(x) = x^{p^n} - x$.

$\therefore$ $f'(x) = p^n x^{p^n - 1} - 1$

or $f'(x) = p^n \cdot 1 x^{p^n - 1} - 1$ .....(1)

Since $F$ is a field of characteristic $p$, so

$p \cdot 1 = (1 + 1 + 1 + ...+ \text{upto } p \text{ terms}) = 0$ and consequently $p^n \cdot 1 = 0$.

Therefore form (1) we have

$$f'(x) = -1. \qquad .....(2)$$

From (1) and (2) we see that $f(x)$ and $f'(x)$ have no nontrivial common factor. Hence by theorem 6, $f(x)$ has no multiple root.

Let $f(x)$ be an irreducible polynomial over a field $F$. Then $f$ is said to be **separable** over $F$ if all the roots of $f(x)$ in its extension field $K$ are simple, that is, $f(x)$ has no multiple roots in its extension field $K$ over $F$.

An arbitrary polynomial over $F$ is said to be **separable** over $F$ if each of its irreducible factor is separable. A polynomial which is not separable is known as inseparable.

By using theorem 7 and 8 on multiple roots of a polynomial and the definition of separable polynomial, we observe that :

***(i)*** an irreducible polynomial $f(x) \in f[x]$ is separable if and only if $f'(x) \neq 0$,

***(ii)*** every non-constant polynomial over a field of characteristic zero is separable.

Let $K$ be a field extension of a field $F$. Then an algebraic element $a \in K$ over a field $F$ is said to **separable** over $F$, if the minimal polynomial of a over $F$ is separable, that is, if it satisfies a polynomial over $F$ having no multiple roots.

An algebraic extension $K$ of a field $F$ is said to be a **separable extension,** if every element of $K$ is separable over $F$, otherwise $K$ is said to be an inseparable extension.

As observed above every polynomial over a field of characteristic zero is separable, we see that every algebraic extension of a field of characteristic zero is a separable extension.

**Theorem 10.** *An irreducible polynomial $f(x)$ over a field of characteristic $p > 0$, is inseparable if and only if $f(x) \in F[x^p]$, that is, $f(x)$ is a polynomial in $x^p$.*

**Proof.** Let $f(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + ... + \alpha_n x^n$ with $\alpha_n \neq 0$ be an irreducible polynomial in $F[x]$ of degree $n \geq 1$. Then

$$f'(x) = \alpha_1 + 2\alpha_2 x + ... + n\alpha_n x^{n-1}.$$

We know that $f(x)$ is inseparable if and only if $f'(x) = 0$. But $f'(x) = 0$ if and only if

$$\alpha_1 = 0, 2\alpha_2 = 0, ..., n\alpha_n = 0.$$

Since the characteristic of $F$ is $p > 0$, it follows that each nonzero element is of order $p$. So, for any $k$, $1 \leq k \leq n$, $k\alpha_k = 0$ and $\alpha_k \neq 0$ implies that $p$ is a divisor of $k$, *i.e.,* $\alpha_k = 0$ or $k = mp$ for some positive integer $m$. This means, in $f(x)$ if any term is $\alpha_k x^k$ with $\alpha_k \neq 0$, then it is of type

$$\alpha_{mp} x^{mp} = \alpha_{mp} \left(x^p\right)^m.$$

So, that

$$f(x) = a_0 + x_1 x^p + a_2 x^{2p} + ... + a_m x^{mp}$$

for some positive integer $m$. Thus $f(x) \in f[x^p]$.

**Theorem 11.** *A polynomial $f(x)$ is separable if and only it is relative prime to its derivative, i.e., $(f(x), f'(x)) = 1$.*

**Proof.** By theorem 6, we know that a polynomial $f(x)$ has a multiple root a if and only if a is also a root of $f'(x)$, that is, $f(x)$ and $f'(x)$ are both divisible by the minimal polynomial for a. Thus by definition of separable polynomial $f(x)$ is separable if and only if it is relative prime to its derivative.

**Theorem 12.** *If $F$ is a finite field of characteristic $p$, then $a \to a^p$ is an automorphism of $F$.*

**Proof.** Let $a$ and $b$ be any two elements of $F$. Then by binomial theorem we have

$$(a + b)^p = a^p + {}^pc_1 a^{p-1} b + {}^pc_2 a^{p-2} b^2 + ... + {}^pc_p b^p.$$

Since $p$ is characteristic of field $F$, so for each $m$ with $1 \leq m \leq p-1$, $p$ divides ${}^pc_m$. So that

$${}^pc_m a^{p-m} b^m = 0 \text{ for } m = 1, 2, ..., p-1.$$

Hence $(a + b)^p = a^p + b^p$. Now

$$\sigma(a + b) = (a + b)^p = a^p + b^p = \sigma(a) + \sigma(b)$$

and $$\sigma(a\,b) = (a\,b)^p = a^p\,b^p = \sigma(a) \cdot \sigma(b).$$

Thus $\sigma$ is an endomorphism. Farther

$$\text{Ker } \sigma = \{a \in F \mid \sigma(a) = 0\}$$

$$= \{a \in F \mid a^p = 0\} = \{0\}$$

and hence σ is a monomorphism. Since $F$ is a finite field and $\sigma : F \to F$ is one-one, so σ is also onto. Hence σ is an automorphism.

**Theorem 13.** *Any algebraic extension of a finite field F is a separable extension.*

**Proof.** Let $K$ be any algebraic extension of $F$. Also let $f(x)$ be any irreducible polynomial over $F$. Now we have to show that $f(x)$ is separable. Let, if possible $f(x)$ be inseparable over $F$, so by theorem 10, $f(x) \in f[x^p]$. Let

$$f(x) = b_0 + b_1 x^p + b_2 x^{2p} + \ldots + b_m x^{mp}$$

for some $b_i \in F$, $0 \le i \le m$. From above theorem we see that $a \to a^p \ \forall \ a \in F$ is an automorphism of $F$, so we can find $a_i \in F$ such that $b_i = a_i^p \ \forall \ i$ and consequently

$$f(x) = a_0^p + a_1^p x^p + a_2^p x^{2p} + \ldots + a_m^p x^{mp}$$

$$= \left(a_0 + a_1 x + a_2 x^2 + \ldots + a_m x^m\right)^p,$$ which shows that $f(x)$ is not irreducible. This is a contradiction and hence $f(x)$ is separable. Thus by definition $K$ is a separable extension of $F$.

## 8.5 Perfect field

A field $F$ is called perfect field if all finite extensions of F are separable.

**Theorem 14.** *Every field of characteristic zero is perfect.*

**Proof.** Let $F$ be a field of characteristic zero. We have to show that $F$ is perfect. For this we shall show that every finite extension of $F$ is separable. Let $K$ be any finite extension of $F$. In order to show that $K$ is separable over $F$, it is sufficient to show that minimal polynomial for each element of $K$ over $F$ is separable.

Let α be an arbitrary element of $K$ and let

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n, \text{ with } a_n \neq 0$$

be a minimal polynomial for α over $F$. Then

$$f'(x) = a_1 + 2a_2 x + \ldots + n a_n x^{n-1}.$$

Let, if possible $f(x)$ be inseparable, then

$$f'(x) = 0,$$

which shows that $ma_m = 0$ for all $m \ge 1$. But $F$ is a field of characteristic zero, so $ma_m = 0$ for all $m \ge 1$ is possible only if $a_m = 0$ for $m \ge 1$. This gives

$$f(x) = a_0,$$

which is a constant polynomial having a multiple root. But this gives a contradiction because no constant polynomial has a root. Hence $f(x)$ has no multiple root, so is separable. Therefore a is separable over $F$ and hence 3 $K$ is a separable extension of $F$. Consequently $F$ is perfect.

**Theorem 15.** *A field of characteristic $p \neq 0$ such that each element of the field is the $p^{th}$ power of some member of the same is perfect.*

**Proof.** Let $F$ be a field of characteristic $p \neq 0$. We have to show that $F$ is perfect. For this we shall show that every finite extension of $F$ is separable. In order to show that $K$ is separable over $F$, it is sufficient to show that minimal polynomial for each element of $K$ over $F$ is separable. Since the minimal polynomial for each element of $K$ over $F$ is irreducible so we need to show that each irreducible polynomial over $F$ is separable.

Let $f(x)$ be any irreducible polynomial over $F$. Suppose $f(x)$ is inseparable over $F$, so by theorem 10 $f(x) \in F[x^p]$, that is,

$$f(x) = a_0 + a_1x^p + a_2x^{2p} + \ldots + a_mx^{mp}$$

for some $a_i \in F$, $0 \leq i \leq m$. Then by given condition, corresponding to the elements $a_0, a_1, \ldots, a_n$ in $F$ there exist $b_0, b_1, \ldots, b_n$ in $F$ such that

$$a_0 = (b_0)^p,\ a_1 = (b_1)^p, \ldots,\ a_m = (b_m)^p. \text{ Therefore}$$

$$f(x) = b_0^p + b_1^p x^p + \ldots + b_m^p x^{mp}$$

$$= \left(b_0 + b_1x + \ldots + b_mx^m\right)^p.$$

This shows that $f(x)$ is not irreducible, which is a contradiction. Thus, every irreducible polynomial over $F$ is separable and hence, $F$ is a perfect field.

## 8.6 Automorphism of extensions

Let $K$ be a field. Then a mapping $f: K \to K$ is called an automorphism of $K$ if

*(i)* $f$ is one-one,

*(ii)* $f$ is onto,

*(iii)* $f(a+b) = f(a) + f(b)$ and $f(a\,b) = f(a)f(b)$ for all $a, b, \in K$.

It can be easily seen that

$$f(0) = 0, f(1) = f(1), f(-a) = -f(a) \text{ and}$$

$$f(a^{-1}) = [f(a)]^{-1} \text{ for every } 0 \neq a \in K.$$

If $K$ is a field extension of a field $F$, then an automorphism $f$ of $K$ **fixes** $F$ point-wise if $f(a) = a$ for every $a \in F$. In this case $f$ is known as ***F*-automorphism.**

Let *Aut* $(K)$ be the set of all automorphisms of $K$, then it can be easily seen that Aut $(K)$ is a group with respect to operation composition of functions. The identity map $I_K$ on $K$ is the identity element of *Aut* $(K)$ because

$f\ 0\ I_K = f = I_K\ 0\ f$ for every $f$ Aut $(K)$. We know that if $\sigma \in$ Aut $(K)$,

then $\sigma^{-1} \in$ Aut $(K)$

such that $\sigma\ 0\ \sigma^{-1} = I_k = \sigma^{-1}\ 0\ \sigma$.

Hence every element of Aut $(K)$ is invertible.

**Theorem 16.** *Let $K$ be a field extension of a field $F$ and $f(x) \in F[x]$. If $\phi : K \to K$ is an automorphism such that $\phi(a) = a\ \forall\ a \in F$, and if $\alpha \in K$ is a root of $f(x)$, then $\phi(\alpha)$ is also a root of $f(x)$.*

**Proof.** Let $f(x) = b_0 + b_1x + b_2x^2 + ... + b_nx^n$, $b_i \in F$ for $i = 0, 1, 2, ..., n$. Since $\alpha$ is a root of $f(x)$, so

$$b_0 + b_1\alpha + b_2\alpha^2 + ... + b_n\alpha^n = 0$$

$$\Rightarrow \quad \phi(b_0 + b_1\alpha + b_2\alpha^2 + ... + b_n\alpha^n) = \phi(0)$$

or $$\phi(b_0) + \phi(b_1\alpha) + \phi(b_2\alpha^2) + ... + \phi(b_n\alpha^n) = 0$$

or $$\phi(b_0) + \phi(b_1)\,\phi(\alpha) + \phi(b_2)\,\phi(\alpha^2) + ... + \phi(b_n)\,\phi(\alpha^n) = 0$$

or $$b_0 + b_1\,\phi(\alpha) + b_2\,\phi(\alpha^2) + ... + b_n\,\phi(\alpha^n) = 0, \quad \text{since } \phi \text{ fixes } F.$$

or $$b_0 + b_1\,\phi(\alpha) + b_2\,[\phi(\alpha)]^2 + ... + b_n\,[\phi(\alpha)]^n = 0.$$

This shows that $\phi(\alpha)$ is a root of $f(x)$.

**Theorem 17.** *Let $K$ be a field extension of a field $F$. Then the set $G(K \mid F)$ of all automorphism of $K$ which heave every element of $F$ fixed is a subgroup of the group Aut $(K)$.*

**Proof.** Since $G(K \mid F)$ is a set of all automorphism of $K$ which heave every element of $F$ fixed, so

$$G(K \mid F) = \{\phi \in Aut(K) \mid \phi(a) = a\ \forall\ a \in F\}.$$

Since the identity automorphism $I_K$ of $K$ is such that $I_K(x) = x\ \forall\ x \in K$, so it leaves every element of $F$ fixed and hence $I_K \in G(K \mid F)$ and consequently $G(K \mid F) \neq \phi$. Let $\phi, \psi$ be any two elements of $G(K \mid F)$. Then $\phi(a) = a$ and $\psi(a) = a$ for all $a \in F$. Now for all $a \in F$, we have

$$\begin{aligned}(\phi\ 0\ \psi^{-1})(a) &= \phi[\psi^{-1}(a)] \\ &= \phi[\psi(a)], \quad \text{since } \psi(a) = a, \text{ so } \psi^{-1}(a) = a\ \forall\ a \in F \\ &= \phi(a) \\ &= a\end{aligned}$$

This implies $\phi\ o\ \psi^{-1} \in G(K/F)$ and $G(K/F)$ is a subgroup of Aut $(K)$.

## 8.7 Fixed field of a group of automorphisms

Let $H$ be a subgroup of the group of all automorphisms of a field $K$. Then the set

$L = \{a \in K \mid \phi(a) = a \text{ for all } \phi \in H\}$ is called the fixed field of $H$. We shall show that it is a subfield of $K$.

**Theorem 18.** *Let H be a subgroup of all automorphisms of a field K. Then the fixed field of H is a subfield of K.*

**Proof :** Let $L$ be a fixed field of $H$. Then by definition of fixed field we have

$$L = \{a \in K \mid \phi(a) = a \text{ for all } \phi \in H\}.$$

Since $\phi(1) = 1$ and $(0) = 0$, so 1 and 0 are in $L$ and hence $L \neq \phi$. Let $x$ and $y$ be any two elements of $L$, and $\phi$ is any element of $H$. Then

$$\begin{aligned}\phi(x - y) &= \phi[x + (-y)] \\ &= \phi(x) + \phi(-y) \\ &= \phi(x) - \phi(y) \\ &= x - y.\end{aligned}$$

So $x - y \in L$. Again, let $x \in L$, $0 \neq y \in L$, then

$$\begin{aligned}\phi(xy^{-1}) &= \phi(x)\,\phi(y^{-1}) \\ &= \phi(x)[\phi(y) - 1] \\ &= ab^{-1}.\end{aligned}$$

So $ab^{-1} \in L$. Hence $L$ is a subfield of $K$.

**Theorem 19.** *If K is a field and if* $\phi_1, \phi_2, \ldots, \phi_n$ *are distinct automorphisms of K, then it is impossible to find elements* $b_1, b_2, \ldots, b_n$ *not all zero, in K such that* $b_1\phi_1(a) + b_2\phi_2(a) + \ldots + b_n\phi_n(a) = 0$, *for all* $a \in K$.

**Proof :** Let, if possible, we can find elements $b_1, b_2, \ldots, b_n$, not all zero, in $K$ such that

$$b_1\phi_1(a) + b_2\phi_2(a) + \ldots + b_n\phi_n(a) = 0, \ \forall\ a \in K \qquad \ldots\ldots(1)$$

Among all relations of type (1) we can field a relation which has as few nonzero terms as possible. After counting again let this minimal relation be

$$b_1\phi_1(a) + b_2\phi_2(a) + \ldots + b_m\phi_m(a) = 0 \qquad \ldots\ldots(2)$$

where $b_1, b_2, \ldots b_n$ are all different form zero. If $m = 1$, then from (2), we get

$$b_1\phi_1(a) = 0, \ \forall\ a \in K$$

$$\Rightarrow \quad b_1\phi_1(1) = 0, \text{ since } 1 \in K$$

$$\Rightarrow \quad b_1 \cdot 1 = 0, \text{ since } \phi_1(1) = 1$$

$$\Rightarrow \quad b_1 = 0,$$

which contradicts over assumption that $b_1 \neq 0$. So, $m > 1$. Since $\phi_1 \neq \phi_m$, so there exists an element $c \in K$ such that $\phi_1(c) \neq \phi_m(c)$. Since $ca \in K$ for all $a \in K$, so relation (2) will hold good for $ca$, that is,

$$b_1 \phi_1(ca) + b_2\phi_2(ca) + ... + b_m\phi_m (ca) = 0 \text{ for all } a \in K$$

or $$b_1\phi_1(c)\, \phi_1(a) + b_2\phi_2(c)\, \phi_2(a) + ... + b_m\phi_m (c)\, \phi_m (a) = 0 \text{ for all } a \in K \quad .....(3)$$

Multiplying relation (2) by $\phi_1 (c)$ and subtracting the result from (3) we get

$$b_2 [\phi_2(c) - \phi_1(c)]\, \phi_2(a) + ... + b_m [\phi_m (c) - \phi_1 (c)]\, \phi_m (a) = 0 \quad .....(4)$$

If we put $a_i = b_i [\phi_i(c) - \phi_1(c)]$, for $i = 2,...,m$, then $a_i$ are in $K$, $a_m = b_m [\phi_m(c) - \phi_1(c)] \neq 0$, since $b_m \neq 0$, and $\phi_m(c) - \phi_1(c) \neq 0$. Thus relation (4) may be written as

$$a_2 \phi_2(a) + ... + a_m \phi_m (a) = 0, \text{ for all } a \in K.$$

This relation has $(m - 1)$ terms, which contradicts over assumption that (2) is a minimal relation. This proves the theorem.

**Theorem 20.** *If $K$ is a finite extension of a field $F$, then the group $G(K \mid F)$ of $F$ automorphisms of $K$ is finite and $o[G(K \mid F)] \leq [K : F]$.*

**Proof :** Let $[K : F] = n$ and let $b_1, b_2,...,b_n$ be a basis of $K$ over $F$. Suppose we can find $(n + 1)$ distinct automorphisms $\phi_1, \phi_2,...,\phi_{n+1}$ in $G(K \mid F)$. Consider the following system of homogeneous linear equations in $n + 1$ unknowns $x_1, x_2,...,x_{n+1}$, with coefficients in $K$.

$$\phi_1 (b_1)\, x_1 + \phi_2 (b_1)\, x_2 + ... + \phi_{n+1} (b_1)\, x_{n+1} = 0$$

$$\phi_1 (b_2)\, x_1 + \phi_2 (b_2)\, x_2 + ... + \phi_{n+1} (b_2)\, x_{n+1} = 0$$

.....

.....

.....

$$\phi_1 (b_i)\, x_1 + \phi_2 (b_i)\, x_2 + ... + \phi_{n+1} (b_i)\, x_{n+1} = 0$$

.....

.....

.....

$$\phi_1 (b_n)\, x_1 + \phi_2 (b_n)\, x_2 + ... + \phi_{n+1} (b_n)\, x_{n+1} = 0$$

In this system of linear homogenous equations, the number of equations is less then the number of unknowns.

Hence this system must have a nontrivial solutions (not all zero) $x_1 = a_1, x_2 = a_2,..., x_{n+1} = a_{n+1}$ in $K$. Thus, we have

$$a_1 \phi_1 (b_i) + a_2 \phi_2 (b_i) + ... + a_{n+1} \phi_{n+1} (b_i) = 0 \text{ for } i = 1, 2,...,n. \quad .....(1)$$

Let $u$ be any element of $K$. Since the set $\{b_1, b_2,...,b_n\}$ is a basis of $K$ over $F$, therefore there exist $c_i \in F$, $i = 1, 2,...n$ such that

$$u = c_1 b_1 + c_2 b_2 + ... + c_n b_n$$

Now, $\phi_1 (u) = \phi_1 (c_1 b_1 + c_2 b_2 + ... + c_n b_n)$

or $\phi_1 (u) = \phi_1 (c_1 b_1) + \phi_1(c_2 b_2) + ... + \phi_1 (c_n b_n)$

or $\phi_1 (u) = \phi_1 (c_1)\, \phi_1(b_1) + \phi_1(c_2)\, \phi_1(b_2) + ... + \phi_1 (c_n)\, \phi_1 (b_n)$

or $\phi_1 (u) = c_1\, \phi_1(b_1) + c_2\, \phi_1(b_2) + ... + c_n\, \phi_1 (b_n)$,

since $\phi_1$ leaves every element of $F$ fixed.

Similarly

$$\phi_2 (u) = c_1\, \phi_2(b_1) + c_2\, \phi_1(b_2) + ... + c_n\, \phi_2 (b_n)$$

Multiplying above equations by $a_1, a_2, ... a_{n+1}$ respectively and adding we get

$$a_1 \phi_1 (u) + a_2 \phi_2(u) + ... + a_{n+1}\, \phi_{n+1} (u) = c_1 [a_1 \phi_1 (b_1) + a_2 \phi_2(b_1) + ...$$

$$+ a_{n+1}\, \phi_{n+1} (b_1)] + ... + c_n [a_1 \phi_1 (b_n) + a_2 \phi_2(b_n) + ... + a_{n+1}\, \phi_{n+1} (b_n)]$$

$$= c_1 \times 0 + ... + c_n \times 0, \qquad \text{[using (1)]}$$

$$= 0.$$

Thus we conclude that, if $\phi_1, \phi_2, ..., \phi_{n+1}$ are distinct automorphisms of $K$, we can find $a_1, a_2, ... a_{n+1}$ in $K$, not all zero, such that $a_1 \phi_1 (u) + a_2\, \phi_2 (u) + ... + a_{n+1}\, \phi_{n+1} (u) = 0$ for all $u \in K$. By theorem 19 it is not possible. So, there can not be $n + 1$ distinct automorphisms in $G (K \mid F)$ and consequently

$$o [G (K \mid F)] \leq n$$

Hence, $o [G (K \mid F)] \leq [K : F]$.

## 8.8 Some important examples

***Ex.1.*** *Prove that if the complex number z is a root of the polynomial f (x) with real coefficients, then $\bar{z}$, the complex conjugate of z is also a root of f (x).*

**Sol .** We know that the field $C$ of complex numbers is an extension of the field $R$ of real numbers. Let $f (x) \in R [x]$ and $z = x + iy$ be a root of $f (x)$ in $C$. We have to prove that $\bar{z} = x - iy$ is also a root of $f (x)$ in $C$.

Consider a mapping $\phi : C \to C$ defined by $\phi (x + iy) = \overline{x + iy} = x - iy \ \forall\ x + iy \in C$. It is easy to see that $\phi$ is an automorphism of $C$. Let a be any element of $R$, then as an element of $C$, it can be written as $a + io$. Then

$$\phi (a) = \phi (a + io) = a - io = a.$$

Thus $\phi$ leaves every element of $R$ fixed. Therefore by theorem 16, if $z$ is a root of $f (x)$ in $C$, then $\phi (z) = \bar{z}$ is also a root of $f (x)$ in $C$.

***Ex.2.*** *Let K be the field of complex numbers and let F be the field of real numbers. Find G (K | F) and fixed field of G (K | F).*

**Sol .** Let $\phi$ be an automorphism of $K$. Then

$$[\phi(i)]^2 = \phi(i)\,\phi(i)$$
$$= \phi(i^2)$$
$$= \phi(-1)$$
$$= -1$$

Therefore $\phi(i) = \pm\sqrt{-1} = \pm i$. If in addition $\phi$ leaves every real number fixed, then for any $x + iy$ $K$, we have

$$\phi(x + iy) = \phi(x) + \phi(iy)$$
$$= \phi(x) + \phi(i) \cdot \phi(y)$$
$$= x \pm iy.$$

Now, if we take $\phi_1(x + iy) = x + iy$ and $\phi_2(x + iy) = x - iy$, then it is easy to see that $\phi_1$ and $\phi_2$ are automorphisms of $K$ such that $\phi_1(x) = x$ and $\phi_2(x) = x$ for all $x \in F$. Hence $G(K \mid F) = \{\phi_1, \phi_2\}$. Now, let $L$ be the fixed field of $G(K \mid F)$, Then

$$L = \left\{x + iy \in K \mid \phi(x + iy) = x + iy \;\forall\; \phi \in G(K \mid F)\right\}.$$

$$\phi_2(x + iy) = x + iy$$
$$\Rightarrow \quad x - iy = x + iy$$
$$\Rightarrow \quad 2iy = 0$$
$$\Rightarrow \quad y = 0.$$

Thus fixed field of $G(K \mid F)$ is $F$.

***Ex.3.*** *Let K be extension of the field of rational numbers Q. Show that any automorphism of K must leave every element of Q fixed.*

**Sol .** Let $\phi$ be any automorphism of $K$ and let a be any element of $Q$. Now we have to show that $\phi(a) = a$. Following cases arise :

**Case 1.** When $a = 0$. In this cas $\phi(a) = 0$.

**Case 2.** When a is a positive integer.

Then
$$\phi(a) = \phi(1 + 1 + 1 + \ldots \text{ to } a \text{ terms})$$
$$= \phi(1) + \phi(1) + \ldots \text{ to } a \text{ terms}$$
$$= 1 + 1 + \ldots \text{ to } a \text{ terms}$$
$$= a$$
$$\Rightarrow \quad \phi(a) = a$$

**Case 3.** When a is a negative integer.

Then $a = -m$, where m is a positive integer. In this case

$$\phi(a) = \phi(-m) = \phi[(-1) + (-1) + \ldots \text{ to } m \text{ terms}]$$

$$= \phi(-1) + \phi(-1) + \ldots \text{ to } m \text{ terms}$$

$$= (-1) + (-1) + \ldots \text{ to } m \text{ terms}$$

$$= -m = a$$

$$\Rightarrow \qquad \phi(a) = a$$

**Case 4.** When a is a rational number of the form $\frac{m}{n}, m, n$ are integers and $n \neq 0$.

In this case

$$\phi(a) = \phi\left(\frac{m}{n}\right)$$

$$= \phi(mn^{-1})$$

$$= \phi(m)\,\phi(n^{-1})$$

$$= \phi(m)\,[\phi(n)]^{-1}$$

$$= mn^{-1}, \text{ by case 1, 2 and 3.}$$

$$= \frac{m}{n} = a$$

Hence f (a) = a for all $a \in Q$.

## Self-learning exercise-1

**1.** The splitting field of polynomial $x^2 + 1 \in R[x]$ is the field $R$ of real numbers. [True | False]

**2.** If $F$ is a field, then polynomial $x^2 + x + 1 \in F[x]$ has a splitting field. [True | False]

**3.** The field R of real numbers is a normal extension of the field Q of rational numbers. [True | False]

**4.** Every non-constant polynomial over a field of characteristic zero is separable. [True | False]

**5.** A polynomial $f(x)$ is inseparable if and only if $f(x)$ and $f'(x)$ are relative prime. [True | False]

**6.** Every field of characteristic zero is perfect. [True | False]

## 8.9 Summary

In this unit we have discussed about splitting fields, normal extension, separable and inseparable extensions, automorphism of extension and some important results on these tropics.

## 8.10 Answer to self-learning exercise

**Self-learning exercise –1**

| | | |
|---|---|---|
| **1.** False | **2.** True | **3.** False |
| **4.** True | **5.** False | **6.** True |

## 8.11 Exercises

1. If $F$ is a field, then prove that every polynomial $f(x) \in F[x]$ has a splitting field.
2. Let $F$ be a field such that characteristic of $F$ be $p > 0$. Also let $f(x) = x^p - \alpha$ be a polynomial in $F[x]$ with no root in $F$. Then prove that $f(x)$ is an inseparable polynomial.
3. Show that a polynomial $f(x) \in F[x]$ is separable if and only if it is relative prime to its derivative.
4. Show that the splitting field of a polynomial $x^p - 1 \in Q[x]$, $p$ prime, is of degree $p - 1$ over $Q$.
5. Let $K = Q(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in Q\}$ and $F = Q$. Find $G(K \mid F)$ and fixed field of $G(K \mid F)$.

□□□

# UNIT 9 : Galois Theory

**Structure of the Unit**



## 9.0 Objectives

After reading this unit you will be able to under stand about Galois extension and Galois group. Fundamental theorem of Galois group, Extensions by radicals and solvability of polynomial equations. You will be also able to know that the general polynomial equation of degree $n$ is not solvable by radicals for $n \geq 5$.

## 9.1 Introduction

In this unit we shall take up the study of Galois theory and prove Fundamental theorem of Galois theory. We will also prove some important results related to Galois theory.

## 9.2 Galois extension and Galois group

Let $K$ be a finite field extension of a field $F$. Then $K$ is said to be a **Galois extension** of $F$, if it is both normal and separable, that is, if $K$ is a splitting field of some separable polynomial $p(x) \in F[x]$.

**Theorem 1.** *Let K be a Galois extension of a field F. Then the set of all F-automorphisms of K is a group with respect to operation composition of functions.*

**Proof.** Let $G(K \mid F)$ be a collection of all $F$-automorphisms of $K$. Now we have to show that $G(K \mid F)$ is a group for composition of automorphisms. Let $\sigma$ and $\tau$ be any two elements of $G(K \mid F)$. Then $\sigma\tau$ is also an automorphism of $K$ onto $K$. Now for any $a \in F$, we have

$$(\sigma\ \tau)(a) = \sigma(\tau(a))$$
$$= \sigma(a), \text{ since } \tau(a) = a\ \forall\ a \in F$$
$$= a, \text{ since } \sigma(a) = a\ \forall\ a \in F$$

$\Rightarrow \qquad (\sigma\ \tau)\ (a)\ a\ \forall\ a \in F$

and hence $\sigma\ \tau \in G\ (K \mid F)$, which shows that composition of $F$-automorphisms is a binary composition in $G\ (K \mid F)$. Since the identity automorphism $I_K$ of $K$ leaves every element of $F$ fixed, so $I_K \in G\ (K \mid F)$ is the identity element in $G\ (K \mid F)$ because for every $\sigma\ a \in F$,, we have

$$\sigma\ I_K = \sigma = I_K\ \sigma.$$

Let $\sigma$ be any element of $G\ (K \mid F)$. Then $\sigma^{-1}$ is an automorphism of $K$ onto $K$ such that

$$\sigma\ \sigma^{-1} = I_K = \sigma^{-1}\ \sigma\ .$$

Now for any $a \in F$, we have $\sigma^{-1}\ (a) = a$, since $\sigma\ (a) = a\ \forall\ a \in F$. This shows that $\sigma^{-1} \in G\ (K \mid F)$ and hence every element of $G\ (K \mid F)$ is invertible.

We know that composition of functions is associative, so for any $\sigma, \tau, \phi \in G\ (K \mid F)$ and for any $a \in F$, we have

$$\begin{aligned}\sigma\ (\tau\ \phi)\ (a) &= \sigma\ (\tau\ (\phi\ (a))) \\ &= \sigma\ (\tau\ (a)), \text{ since } \phi\ (a)\ a\ \forall\ a \in F \\ &= \sigma\ (a), \text{ since } \tau\ (a) = a\ \forall\ a \in F \\ &= a, \text{ since } \sigma\ (a) = a\ \forall\ a \in F\end{aligned}$$

and

$$\begin{aligned}\{(\sigma\ \tau)\ \phi\}(a) &= (\sigma\ \tau)\ \phi\ (a) \\ &= (\sigma\ \tau)\ a \\ &= \sigma\ (\tau\ (a)) \\ &= \sigma\ (a) \\ &= a \text{ and hence } \sigma\ (\tau\ \phi) = (\sigma\ \tau)\ \phi,\end{aligned}$$

which shows that composition in $G\ (K \mid F)$ is associative. Hence $G\ (K \mid F)$ is a group

**Note that** if $K$ is a Galois extension of a field $F$, then the group $G\ (K \mid F)$ of all $F$-automorphisms of $K$ is called the **Galois group** of $K$ over $F$.

**Theorem 2.** *Let K be a Galois extension of a field F. Then an element of K which remains invariant for each member of the Galois group G (K | F) is necessarily a member of F.*

**Proof.** Since $K$ is a Galois extension of field $F$, so $K$ is finite extension of $F$ which is both normal and separable. Let $\alpha$ be an arbitrary element of $K$ which remains invariant under every member $G\ (K \mid F)$, *i.e.* $\sigma\ (\alpha) = \alpha\ \forall\ \sigma \in G\ (K \mid F)$. Now we have to show that $\alpha \in F$. Since $K$ is a finite normal extension of $F$, so $K$ is the splitting field of some polynomial $f\ (x) \in F\ [x]$.

Again, since each finite extension is algebraic, therefore $K$ is algebraic over $F$ and $\alpha \in K$ is algebraic over $F$. Let $p\ (x)$ be the minimal polynomial for $\alpha$ over $F$. As $K$ is normal over $F$ and one root $\alpha$ of $p\ (x)$ in $K$, therefore each root of $p\ (x)$ belongs to $K$, *i.e.*, $K$ is the splitting field of $p\ (x)$ also. Then $K$ is the splitting field of $f\ (x)\ p\ (x) \in F\ [x]$.

Let, if possible *deg p* $(x) \geq 2$. Since $K$ is a separable over $F$ and all the roots of $p(x)$ are distinct in $K$, so there exists an element $\beta \in K$ with $\beta \neq \alpha$ such that $\beta$ is a root of $p(x)$ over $F$.

Thus $\alpha$ and $\beta$ are two distinct roots of an irreducible polynomial $p(x) \in F[x]$, so there exists an $F$-isomorphism $\sigma : F(\alpha) \to F(\beta)$ such that $\sigma(\alpha) = (\beta)$. $\sigma$ can be extended to an $F$-automorphism $\bar{\sigma} : K \to K$, *i.e.*, there exists $\bar{\sigma}\, G(K \mid F)$ with $\bar{\sigma}(\alpha) = \beta \neq \alpha$ so that we arive to a contradiction. Hence the $deg\ p(x) \ngeq 2$, accordingly *deg p* $(x) = 1$, *i.e.*, $p(x) = x - \alpha$ where $\alpha \in F$ because $p(x) \in F[x]$ and $p(\alpha) = 0$.

**Theorem 3.** *The order of the Galois group* $G(K \mid F)$ *is equal to the degree of K over F, i.e.,*

$$o\,[G(K \mid F)] = [K : F].$$

**Proof.** Since $K$ is a finite separable extension of $F$, therefore it is a simple extension of $F$. Therefore there exists an element $a \in K$ such that

$$K = F(a).$$

Let $f(x)$ be a minimal polynomial for $a$ over $F$ and let $deg\, f(x) = n$. Then, we have

$$[K : F] = n.$$

As $K$ is reparable over $F$, the roots of $f(x)$ are all simple. Let $a = a_1, a_2, \ldots, a_n$ be $n$ distinct root of $f(x)$ in $K$. Then $K = F(a_i)$ for each $i = 1, 2, \ldots, n$.

Now for each $a_i$, there exists *a* $F$-automorphism $\phi_i$ of $K$ such that $\phi_i(a_i) = a_i$ and since $a_i$ generates $K$ over $F$, so each $\phi_i$ is unique. Form Theorem 16 of Unit-8 we know that if $\phi$ is a $F$-automorphism of $K$ and $a$ is a root of $f(x)$ in $K$, then $\phi(a)$ is also a root of $f(x)$ in $K$, and hence $\phi(a) = a_i$ for some $i$ and consequently $\phi_i = \phi$. Hence the Galois group consists of $\phi_1, \phi_2, \ldots, \phi_n$. Therefore

$$o\,[G(K \mid F)] = [K : F].$$

**Theorem 4. (*Artin*)** *Let G be a finite group of automorphisms of a field K. Let F be the fixed field of G, i.e.*

$$F = \{x \in K \mid \phi(x) = x, \textit{for all } \phi \in G\}.$$

*Then K is a Galois extension of F with* $G(K \mid F) = G$.

**Proof.** Let $o(G) = n$ and let $\phi_1, \phi_2, \ldots, \phi_n$ be the distinct elements of $G$. If $a \in K$, let $\alpha_1, \alpha_2, \ldots, \alpha_m$ be the distinct elements among $\phi_1(a), \phi_2(a), \ldots, \phi_n(a)$ $(m \leq n)$. If $\sigma \in G$, then $\sigma(\alpha_1), \sigma(\alpha_2), \ldots, \sigma(\alpha_m)$, are distinct. Moveover $\sigma\phi_1, \sigma\phi_2, \ldots, \sigma\phi_n$ are all the elements of $G$. Hence $\sigma(\alpha_1), \sigma(\alpha_2), \ldots, \sigma(\alpha_m)$ is a permutation of $\alpha_1, \alpha_2, \ldots, \alpha_m$. Consider $f(x) \in K[x]$ defined by $f(x) = \prod_{i=1}^{m}(x - \alpha_i)$. Then

$$\sigma(f(x)) = \prod_{i=1}^{m}(x - \sigma(\alpha_i))$$

$$= \prod_{i=1}^{m}(x - \alpha_i) = f(x) \text{ for all } \sigma \in G.$$

Hence $f(x) \in F[x]$. Since all the roots of $f(x)$ lie in $K$ are distinct, so $f(x) \in F[x]$ is a separable polynomial. Moreover, $f(x) \in F[x]$ is irreducible for if $g(x)$ is the minimal polynomial of $a$ over $F$, then $g(\phi_i(a)) = \phi_i(g(a)) = 0$, *i.e.* $g(\alpha_i) = 0$ for all $i$. This implies $f(x)$ divides $g(x)$ and since $g(x)$ is irreducible, $f(x) = g(x)$.

Since $f(a) = 0$, $a$ is algebraic and separable over $F$ and $[F(a) : F] \leq n = 0(G)$. Hence $K$ is a separable extension of $F$.

To show that $K$ is a finite extension of $F$, let $N$ be a finite extension of $F$ such that $N \subset K$. Since $N$ is separable extension of $F$, $N = F(\alpha)$ for some $\alpha \in N$. Thus $[N : F] = [F(\alpha) : F] \leq n$. Choose $N$ is a finite extension of $F$ such that $[N : F]$ has maximum value. We claim that $N = K$, showing that $K$ is finite extension of $F$. Write $N = F(\alpha)$. Let $\beta \in K$, and $M$ be the subfield of $K$ generated by $N$ and $\beta$. Then $M$ is a finite extension of $F$ and $\beta$. Then $M$ is a finite extension of $F$ and, by maximality, $[M : F] = [N : F]$. But $[M : F] = [M : N][N : F]$ and hence $[M : N] = 1$, *i.e.* $M = N$. This shows that $\beta \in N$, for any $\beta \in K$, *i.e.* $K \subset N$ and hence $K = N$.

It remains to show that $K$ is normal extension of $F$. Since $K$ is a separable extension of $F$, so $K = f(\alpha)$ for some $\alpha \in K$. Since $\phi_1, \phi_2, ..., \phi_n$ are distinct, $\phi_1(\alpha), \phi_2(a), ..., \phi_n(\alpha)$ are distinct. Hence

$$g(x) = \prod_{i=1}^{n} (x - \phi_i(\alpha)) \in F[x]$$

is of degree $n$ and it is the minimum polynomial of $\sigma_1(\alpha) = \alpha$ ($\sigma_1$ = identity) over $F$. Thus $K$ is the splitting field of $g(x) \in F[x]$ showing that $K$ is a normal extension of $F$ such that $[K : F] = n$. Clearly $G \subset G(K \mid F)$. By theorem 3, $G(K \mid F)$ has order $[K : F] = n = 0(G)$. Hence $G = G(K \mid F)$.

**Theorem 5.** *Let K be a Galois extension of a field F and let characteristic of F be zero. Then the fixed field under the Galois group G (K | F) is F itself.*

**Proof.** Since $K$ is a Galois extension of field $F$, so $K$ is finite, normal and separable extension of field $F$. Again, since every finite separable extension is a simple extension of $F$ and hence $K = F(a)$ for some $a \in K$.

Let $f(x) \in F[x]$ be a minimal polynomial of a and let $E$ be the splitting field of $f(x)$. Since $K$ is a normal extension of $F$, so the splitting field of every polynomial over $F$ is contained in $K$. Thus we have $E \subseteq K$ .....(1)

Also $a \in E$, because $E$ being the splitting field of $f(x)$ for which $f(a) = 0$. Now $E$ is a splitting field containing $F$ and $a$ and $K = F(a)$ is the smallest field containing $F$ and $a$, so

$$K \subseteq E \qquad .....(2)$$

Form (1) and (2) we get $E = K$. This shows that $K$ is the splitting field of the minimal polynomial for $a$ over $F$. Let $deg\, f(x) = m$. Then $[K : F] = m$. Hence by theorem 3, $o[G(K \mid F)] = m$. Now, if $K_{G(K \mid F)}$ denotes the fixed field under $G(K \mid F)$, then by theorem 4, we have

$$o[K : K_{G(K \mid F)}] = o[G(K \mid F)] = m = [K : F].$$

Hence, $K_{G(K \mid F)} = F$.

## 9.3 Fundamental theorem of Galois theory

**Theorem 6.** *Let K be a Galois extension of a field F. Then there exists a one-to-one correspondence between the set of all subfields of K containing F and the set of all subgroups of* $G(K \mid F)$. *Further, if E is any subfield of K which contains F, then*

*(i)* $[K : E] = o\,[G(K \mid E)]$ *and* $[E : F]$ = *index of* $G(K \mid E)$ *in* $G(K \mid F)$,

*(ii) E is normal extension of F if and only if* $G(K \mid E)$ *is a normal subgroup of* $G(K \mid F)$,

*(iii) If E is a normal extension of F, then* $G(E \mid F) \cong G(K \mid F) \mid G(K \mid E)$.

**Proof.** Let $\phi$ be a any element of $G(K \mid E)$, where $E$ is any subfield of $K$ containing $F$. Then the $\phi$ leaves every element of $E$ fixed. Since $F \subset E$, so $\phi$ leaves every element of $F$ fixed and hence $\phi \in G(K \mid F)$. Thus $G(K \mid E) \subseteq G(K \mid F)$.

Since both $G(K \mid F)$ and $G(K \mid E)$ are subgroups of the groups $Aut(K)$ of all automorphisms of $K$ and $G(K \mid E) \subseteq G(K \mid F)$, it follows that $G(K \mid E)$ is a subgroup of $G(K \mid F)$ .

Let $H_1$ be the set of all subfields of $K$ containing $F$ and $H_2$ be the set of all subgroups of $G(K \mid F)$. Consider a mapping

$$\psi : H_1 \to H_2 \text{ defined by}$$

$$\psi(E) = G(K \mid E) \text{ for all } E \in S_1.$$

$\psi$ is onto, for, it $H \in H_2$, then $H$ is a subgroup of $G(K \mid F)$ and let

$$K_H = \{x \in K \mid \phi(x) = x \text{ for all } \phi \in H\}$$

be the fixed field of $H$. Then $K_H$ is a subfield of $K$. Now $\phi \in H$ implies $\phi \in G(K \mid F)$, and hence $\phi(\alpha) = \alpha$ for all $\alpha \in F$, therefore, $F \subseteq K_H$ and thus $K_H$ is a subfield of $K$ containing $F$ and hence

$$\psi(K_H) = G(K \mid K_H)$$

$$= H, \text{ by theorem 4.}$$

For one-one, let $E_1$ and $E_2$ be any two elements of $H_1$ such that

$$\psi(E_1) = \psi(E_2)$$

$$\Rightarrow \quad G(K \mid E_1) = G(K \mid E_2)$$

$$\Rightarrow \quad \text{the fixed field of } G(K \mid E_1) = \text{the fixed field of } G(K \mid E_2)$$

$$\text{or} \quad K_{(G(K \mid E_1))} = K_{(G(K \mid E_2))}$$

which implies $E_1 = E_2$, since $K$ is a Galois extension of $F$ and therefore $K$ is also a Galois extension of $E_1$ as well as $E_2$. So by theorem 5, the fixed fields under the Galois groups $G(K \mid E_1)$ and $G(K \mid E_2)$ are respectively $E_1$ and $E_2$. Thus

$$\psi(E_1) = \psi(E_2) \Rightarrow E_1 = E_2$$

and hence $\psi$ is one-one.

***(i)*** Since $K$ is a normal extension of $F$ and $E$ is a subfield of $K$ such that $F \subset E \subset K$, so $K$ is a normal extension of $E$. Therefore, by Theorem 3, we have

$$[K : F] = o\,(G\,(K \mid F)) \text{ and } [K : E] = o\,(G\,(K \mid E)).$$

Moreover, $[K : F] = [K : E]\,[E : F]$

$$\Rightarrow \quad o\,(G\,(K \mid F)) = o\,(G\,(K \mid E))\,[E : F]$$

$$\Rightarrow \quad [E : F] = \frac{o\left(G\left(K \mid F\right)\right)}{o\left(G\left(K \mid E\right)\right)} = \text{index of } G\,(K \mid E) \text{ in } G\,(K \mid F).$$

***(ii)*** First suppose that $E$ is a normal extension of $F$, then we shall show that $G\,(K \mid E)$ is a normal subgroup of $G\,(K \mid F)$ . For any $\sigma \in G\,(K \mid F)$ and $\psi \in G\,(K \mid E)$ we actually show that $\sigma^{-1}\,\psi\,\sigma \in G\,(K \mid E)$.

Let $\alpha$ be an arbitrary element of $E$. Since $E$ is a normal extension of $F$, so that the splitting field of the minimal polynomial of $\alpha$ over $F$ is contained in $E$ and consequently, every conjugate of $\alpha$ over $F$ is again in $E$. Since $\sigma\,(\alpha)$ is conjugate of $\alpha$ for any $\sigma \in G\,(K \mid F)$, then $\sigma\,(\alpha) \in E$. Thus for any automorphism $\psi \in G\,(K \mid E)$, $\psi\,(\sigma\,(\alpha) = \sigma\,(\alpha)$.

Now $\quad (\sigma^{-1}\,\psi\,\sigma)\,(\alpha) = \sigma^{-1}\,(\psi\,(\sigma\,(\alpha))) = \sigma^{-1}\,(\sigma\,(\alpha)) = \alpha$

$$\Rightarrow \quad \sigma^{-1}\,\psi\,\sigma \in G\,(K \mid E) \;\forall\; \sigma \in G\,(K \mid F) \text{ and } \psi \in G\,(K \mid E).$$

Hence $G\,(K \mid E)$ is a normal subgroup of $G\,(K \mid F)$.

Conversely suppose that $G\,(K \mid E)$ is a normal subgroup of $G\,(K \mid F)$, then we shall show that $E$ is a normal extension of $F$. Let $\alpha$ be an arbitrary element of $E$ and let $p\,(x)$ be the minimal polynomial of $\alpha$ over $F$. Let $L$ be the splitting field of $p\,(x)$. Since $K$ is a normal extension of $F$ so that $L \subseteq K$. If $\beta$ is any root of $p\,(x)$ in $L$, then $\beta$ is conjugate of $\alpha$ over $F$. Therefore there exists an $F$-automorphism $\sigma$ of $K$ such that $\sigma\,(\alpha) = \beta$.

Now $G\,(K \mid E)$ is a normal subgroup of $G\,(K \mid F)$, then for $\sigma \in G\,(K \mid F)$ and $\psi \in G\,(K \mid E)$, we have

$$\sigma^{-1}\,\psi\,\sigma \in G\,(K \mid E) \Rightarrow (\sigma^{-1}\,\psi\,\sigma)\,(\alpha) = \alpha$$

$$\Rightarrow\; (\sigma^{-1}\,\psi)\,(\sigma\,(\alpha)) = \alpha$$

$$\Rightarrow\; (\sigma^{-1}\,\psi)\,(\beta) = \alpha, \quad \text{since} \quad \sigma\,(\alpha) = \beta$$

$$\Rightarrow\; \sigma^{-1}\,(\psi\,(\beta)) = \alpha$$

$$\Rightarrow\; \psi\,(\beta) = \sigma\,(\alpha)$$

$$\Rightarrow\; \psi\,(\beta) = \beta$$

$$\Rightarrow\; \psi \in E.$$

Thus $\beta \in L \Rightarrow \beta \in E$ and hence $L \subseteq E$. This shows that every splitting field of the minimal polynomial for $\alpha \in E$ over $F$ is contained in $E$ and hence $E$ is a normal extension of $F$.

*(iii)* Let $E$ be a normal extension of $F$. There $E = F(\alpha)$ for some $\alpha \in E$.

Let $\sigma$ be any element of $G(K \mid F)$.

Let $\sigma'$ be the restriction of $\sigma$ to $E$. Then

$$\sigma'(a) = \sigma(a) \text{ for all } \alpha \in E.$$

Since $\sigma$ leaves every element of $F$ fixed, therefore $\sigma'$ also leaves every element of $F$ fixed and hence $\sigma' \in G(K \mid F)$. Consider a mapping

$$\phi : G(K \mid F) \to G(E \mid F) \text{ defined by}$$

$$\phi(\sigma) = \sigma' \text{ for all } \sigma \in G(E \mid F).$$

Now, for any two $\sigma_1, \sigma_2 \in G(K \mid F)$, we have

$$\phi(\sigma_1 \sigma_2) = (\sigma_1 \sigma_2)' \qquad \text{.....(1)}$$

Since for all $a \in E$, we have

$$\begin{aligned}(\sigma_1 \sigma_2)'(a) &= (\sigma_1 \sigma_2)(a)\\ &= \sigma_1(\sigma_2(a))\\ &= \sigma_1(\sigma_2'(a))\\ &= \sigma_1'(\sigma_2'(a))\\ &= (\sigma_1' \sigma_2')(a)\end{aligned}$$

$$\Rightarrow \qquad (\sigma_1 \sigma_2)' = \sigma_1' \sigma_2' \qquad \text{.....(2)}$$

From (1) and (2) we get

$$\begin{aligned}\phi(\sigma_1 \sigma_2) &= \sigma_1' \sigma_2'\\ &= \phi(\sigma_1)\phi(\sigma_2)\end{aligned}$$

$\Rightarrow$ $\phi$ is a homomorphism.

Let $\psi$ be any element of $G(E \mid F)$, then $\psi(\alpha)$ is conjugate of $\alpha$ over $F$, so there exists an $F$-automorphism $\sigma$ of $K$ such that $\sigma(\alpha) = \psi(\alpha)$. Now $\sigma$ and $\psi$ both are identity mapping on $F$ and $E = F(\alpha)$, so

$$\sigma(a) = \psi(a) \ \forall \ a \in E = F(\alpha)$$

$\therefore$ $\psi = \sigma' = \phi(\sigma)$. Hence $\phi$ is onto.

Further
$$\begin{aligned}\ker(\sigma) &= \left\{\sigma \in G\left(K \mid F\right) \middle| \phi(\sigma) = I\right\}\\ &= \left\{\sigma \in G\left(K \mid F\right) \middle| \sigma' = I\right\}\\ &= \left\{\sigma \in G\left(K \mid F\right) \middle| \sigma'(\alpha) = I(\alpha) \ \forall \alpha \in E\right\}\\ &= \left\{\sigma \in G\left(K \mid F\right) \middle| \sigma(\alpha) = \alpha \ \forall \alpha \in E\right\}\\ &= G\left(K \mid F\right).\end{aligned}$$

Hence by the Fundamental Theorem of Group Homomorphism, we get

$$G(E \mid F) \cong G(K \mid F) \mid G(K \mid E).$$

**Note that** if $K$ is a finite extension of a field $F$, then two elements $\alpha$ and $\beta$ of $K$ are said to be **conjugate** over $F$ if they have the same minimal polynomial over $F$.

***Ex.*1.** *Show that the Galois group of* $x^4 + 1 \in Q[x]$ *is the Klein four-group.*

**Sol.** Let $E = Q(\alpha)$, where $\alpha = e^{\frac{i\pi}{4}}$. Since the roots of $x^4 + 1$ are $\alpha$, $\alpha^3$, $\alpha^5$ and $\alpha^7$, $E$ is the splitting field of $x^4 + 1$ over $Q$. Again, since $x^4 + 1$ is irreducible over $Q$, $[E : Q] = 4$. Also the characteristic of $Q$ is zero, so $E$ is *a* normal separable extension of $Q$. Thus $o(G(E \mid Q)) = [E : Q] = 4$. If $\sigma \in G(E \mid Q)$ and $\beta \in E$, then $\beta = a_0 + a_1\alpha + a_2\alpha^2 + a_3\alpha^3$ and $\sigma(\beta) = a_0 + a_1\sigma(\alpha) + a_2\sigma(\alpha^2) + a_3\sigma(\alpha^3)$. Hence $\sigma$ is determined by its effect on $\alpha$. Since $\alpha$ is a root of $x^4 + 1$ therefore $\sigma(\alpha)$ is also a root of $x^4 + 1$. Again, since there are four elements in $G(E \mid Q)$, it follows that $G(E \mid Q) = \{\sigma_1, \sigma_3, \sigma_5, \sigma_7\}$, where $\sigma_1(\alpha) = \alpha$, $\sigma_3(\alpha) = \alpha^3$, $\sigma_5(\alpha) = \alpha^5$ and $\sigma_7(\alpha) = \alpha^7$. Note that $\sigma_3(\sigma_3(\alpha)) = \alpha^9 = \alpha$, $\sigma_5(\sigma_5(\alpha)) = \alpha^{25} = \alpha$ and $\sigma_7(\sigma_7(\alpha)) = \alpha^{49} = \alpha$, so $\sigma_3^2 = \sigma_1 =$ identity, $\sigma_5^2 = \sigma_1$ and $\sigma_7^2 = \sigma_1$. Therefore, $G(E \mid Q)$ is the Klein four-group, since every element except the identity has order two.

***Ex.*2.** *Let F be a field of characteristic* $\neq 2$. *Let* $x^2 - a \in F[x]$ *be an irreducible polynomial over F. Then its Galois group is of order* 2.

**Sol.** Clearly, if $\alpha$ is one root of $x^2 - a$, then $-\alpha$ is the other root. So $\alpha \neq -\alpha$ because characteristic of $F \neq 2$. Thus, $x^2 - a$ is separable over $F$. The splitting field $F(\alpha)$ of $x^2 - a$ over $F$ is a finite, separable, and normal extension of degree 2 over $F$. Thus, $o(G(F(\alpha) \mid F)) = 2$.

***Ex.*3.** *The group* $G(Q(\alpha), Q)$, *where* $\alpha^5 = 1$ *and* $\alpha \neq 1$, *is isomorphic to the cyclic group of order* 4.

**Sol.** $\alpha^5 = 1 \Rightarrow \alpha^5 - 1 = 0$

$$\Rightarrow (\alpha - 1)(1 + \alpha + \alpha^2 + \alpha^3 + \alpha^4) = 0$$

$$\Rightarrow 1 + \alpha + \alpha^2 + \alpha^3 + \alpha^4 = 0, \text{ since } \alpha \neq 1.$$

Thus $\alpha$ is a root of a polynomial $f(x) = 1 + x + x^2 + x^3 + x^4 \in Q[x]$. Since $f(x)$ is irreducible over $Q$, so $[Q(\alpha) : Q] = 4$. Again, since all the roots of polynomial $x^5 - 1 \in Q[x]$ are $1, \alpha, \alpha^2, \alpha^3, \alpha^4$, $\alpha^5$, so $Q(\alpha)$ is the splitting field of $x^5 - 1 \in Q[x]$ and hence *a* normal extension of $Q$. Thus,

$$o[G(Q(\alpha) \mid Q)] = [Q(\alpha) : Q] = 4.$$

This shows that there are four $Q$-automorphisms of $Q(\alpha)$. Since $\{1, \alpha, \alpha^2, \alpha^3\}$ is the basis of $Q(\alpha)$ over $Q$, then any element of $Q(\alpha)$ is

$$a_0 + a_1\alpha + a_2\alpha^2 + a_3\alpha^3,\ a_i \in Q.$$

Let $G(Q(\alpha) \mid Q) = \{\sigma_1, \sigma_2, \sigma_3, \sigma_4\}$. The four $Q$-automorphism of $Q(\alpha)$ are as follows :

$$\sigma_1(a_0 + a_1\alpha + a_2\alpha^2 + a_3\alpha^3) = a_0 + a_1\alpha + a_2\alpha^2 + a_3\alpha^3,$$

$$\sigma_2 (a_0 + a_1\alpha + a_2\alpha^2 + a_3\alpha^3) = a_0 + a_1\alpha^2 + a_2\alpha^4 + a_3\alpha^6$$
$$= a_0 + a_1\alpha^2 + a_2\alpha^4 + a_3\alpha,$$
$$\sigma_3 (a_0 + a_1\alpha + a_2\alpha^2 + a_3\alpha^3) = a_0 + a_1\alpha^3 + a_2\alpha^6 + a_3\alpha^9$$
$$= a_0 + a_1\alpha^3 + a_2\alpha + a_3\alpha^4,$$
$$\sigma_4 (a_0 + a_1\alpha + a_2\alpha^2 + a_3\alpha^3) = a_0 + a_1\alpha^4 + a_2\alpha^8 + a_3\alpha^{12}$$
$$= a_0 + a_1\alpha^4 + a_2\alpha^3 + a_3\alpha^2.$$

Clearly $\{\sigma_1, \sigma_2, \sigma_3, \sigma_4\}$ forms a cyclic group of order 4 generated by $\sigma_2$ and $\sigma_3$. Hence the result.

## 9.4 Extensions by radicals and solvability

We know that there are formulas, giving the roots of a quadratic, cubic and quartic polynomials, which can be written in terms of radicals of rational expressions of the coefficients. However, there is no such "quintic formula" *i.e.* there is no general solution of a fifth-degree equation by radicals.

An extension field $E$ of a field $F$ is an extension by radicals or **radical extension** of $F$, if there are elements $\alpha_1, \alpha_2, \ldots, \alpha_r \in E$ and positive integers $n_1, n_2, \ldots, n_r$ such that $E = F(\alpha_1, \alpha_2, \ldots, \alpha_r)$, $\alpha_1^{n_1} \in F$, and $\alpha_i^{n_i} \in F(\alpha_1, \alpha_2, \ldots, \alpha_{i-1})$, for $1 < i \leq r$.

A polynomial $p(x) \in F[x]$ is **solvable by radicals** over $F$ if the splitting field of $p(x)$ over $F$ is a radical extension of $F$. Thus, if $E$ is the splitting field of $p(x) \in F[x]$ over $F$, then $p(x)$ is solvable by radicals over $F$, if there exists a finite chain of extensions

$$F \subset F_1 = F(\alpha_1) \subset F_2 = F_1(\alpha_2) \subset, \ldots, \subset F_r = F_{r-1}(\alpha_r),$$

Where $\alpha_1^{n_1} \in F$, $\alpha_2^{n_2} \in F_1, \ldots, \alpha_r^{n_r} \in F_{r-1}$, such that $F \subset E \subset F_r$, *i.e.* the roots of $p(x)$ all lie in extension field $F_r$.

**Theorem 7.** *Let n be a positive integer, and let F be a field containing all the $n^{th}$ roots of unity. Let K be the splitting field of $x^n - a \in F[x]$ over F. Then $K = F(\alpha)$, where $\alpha$ is any root of $x^n - a$, and Galois group $G(K \mid F)$ is abelian.*

**Proof.** If $w = \cos\left(\frac{2\pi}{n}\right) + \sin\left(\frac{2\pi}{n}\right)$ and $\alpha$ is a root of $x^n - a$, then $\alpha, \alpha w, \ldots, \alpha w^{n-1}$ are all distinct root of $f(x) = x^n - a$. Thus, the splitting field $K$ of $f(x)$ over $F$ is $K = F(\alpha)$, since it is the smallest field containing $F$ and $\alpha$. Let $\sigma_1$ and $\sigma_2$ be any two elements of $G(K \mid F)$, then $\sigma_1, \sigma_2$ are automorphisms of $K = F(\alpha)$ leaving every element of $F$ fixed. Since $\alpha \in K$ is a root of $x^n - a$, so $\sigma_1(\alpha)$ and $\sigma_2(\alpha)$ are also roots of $x^n - a$. Therefore, for some $i$ and $j$, $\sigma_1(\alpha) = \alpha w^i$ and $\sigma_2(\alpha) = \alpha w^j$.

Now
$$(\sigma_1 \sigma_2)(\alpha) = \sigma_1(\sigma_2(\alpha))$$
$$= \sigma_1(\alpha w^j)$$
$$= \sigma_1(\alpha)\, \sigma_1(w^j), \text{ since } \sigma_1 \text{ is an automorphism}$$
$$= \alpha w^i w^j, \text{ since } \sigma_j \in F$$

$$= \alpha\, w^{i+j}$$

Similarly we can show that

$$(\sigma_2\, \sigma_1)\,(\alpha) = \alpha\, w^{i+j}$$

Hence $\sigma_1\, \sigma_2 = \sigma_2\, \sigma_1 \;\forall\; \sigma_1, \sigma_2 \in G\,(K \mid F)$. This shows that $G\,(K \mid F)$ is abelian.

**Theorem 8.** *Let F be a field of characteristic zero containing all $n^{th}$ roots of unity. If $f\,(x) \in F\,[x]$ is solvable by radicals over F, then the Galois group of $f\,(x)$ over F is solvable.*

**Proof.** Let $f\,(x) \in F\,[x]$ be solvable by radicals over $F$, and let $K$ be the splitting field of $f\,(x)$ over $F$. Then there exists a finite chain of extensions

$$F \subset F_1 = F\,(\alpha_1) \subset F_2 = F_1\,(\alpha_2) \subset \ldots \subset F_r = F_{r-1}\,(\alpha_r) \qquad \ldots..(1)$$

where $\alpha_1^{n_1} \in F, \alpha_2^{n_2} \in F_1, \ldots, \alpha_r^{n_r} \in F_{r-1}$ such that $F \subset K \subset F_r$ ; $\alpha_1, \alpha_2, .., \alpha_r \in K$ and $n_1, n_2, .., n_r \in N$. Now we have to show that the Galois group $G\,(K \mid F)$ is solvable.

Since F has all $n^{th}$ roots of unity, so $F_r$ may be assumed normal extension of $F$. Since $F_r$ is a normal extension of $F$, so $F_r$ is also a normal extension of each intermediate field $F_i$ and further, each $F_i$ is a normal extension of $F_{i-1}$. Hence by the Fundamental Theorem of Galois Theory we have

$$G\,(F_r \mid F_i) \lhd G\,(F_r \mid F_{i-1}) \qquad \ldots..(2)$$

and
$$G\,(F_i \mid F_{i-1}) \cong G\left(F_r \mid F_{i-1}\right) \Big| G\left(F_r \mid F_i\right) \qquad \ldots..(3)$$

Now consider the chain of subgroups

$$G\left(F_r \mid F\right) \supset G\left(F_r \mid F_1\right) \supset G\left(F_r \mid F_2\right) \supset \ldots \supset \ldots \supset G\left(F_r \mid F_{r-1}\right) \supset \{e\} \qquad \ldots..(4)$$

where each subgroup in above chain is a normal subgroup of the preceding one by (2).

By the Theorem 7, each $G\,(F_i \mid F_{i-1})$ is an abelian group. So by (3) each quotient group $G\left(F_r \mid F_{i-1}\right) \Big| G\left(F_r \mid F_i\right)$ of chain (4) is abelian, being an isomorphic image of an abelian group. Thus (4) is a solvable series for group $G\,(F_r \mid F)$ and hence $G\,(F_r \mid F)$ is solvable.

Since $F \subset K \subset F_r$, and $K$ is splitting field of $f\,(x)$ over $F$, so $K$ is a normal extension of $F$. By Fundamental Theorem of Galois Theory we have

$$G\,(F_r \mid K) \lhd G\,(F_r \mid F)$$

and
$$G\,(K \mid F) \cong G\left(F_r \mid F\right) \Big| G\left(F_r \mid K\right) \qquad \ldots..(5)$$

Now $G\left(F_r \mid F\right) \Big| G\left(F_r \mid K\right)$ is solvable being quotient group of solvable group $G\,(F_r \mid F)$ and hence by (5) the Galois group $G\,(K \mid F)$ is solvable being isomorphic to a solvable group.

**Note that** the converse of above theorem is also true.

## 9.5 Insolvability of the quintic

At first we introduce some definitions and show that the general polynomial of degree $n$ has $S_n$ as Galois group.

Let $F$ be a field, then $F[x_1, x_2,...,x_n]$ is an integral domain in $n$ indeterminates $x_1, x_2,...,x_n$ over $F$. The associated field of quotients (fractions) is called the **field of rational functions** in $x_1, x_2,...,x_n$ over $F$ and it is denoted by $F(x_1, x_2,...,x_n)$. Let $S_n$ be the symmetric group of degree $n$ acting on the set $\{1, 2,...,n\}$. For each $\sigma \in S_n$ we can define an automorphism $\sigma'$ of $F(x_1, x_2,...,x_n)$ as follows. Consider any rational function $f(x_1, x_2,...,x_n)$. Then

$$\sigma'(f(x_1, x_2,...,x_n)) = f(x_{\sigma(1)}, x_{\sigma(2)},..., x_{\sigma(n)}).$$

Thus $\sigma \to \sigma'$ is a monomorphism of $S_n$ into the group of all $F$-automorphisms of $F(x_1, x_2,...,x_n)$. So we can treat $S_n$ as a group of $F$-automorphisms of $F(x_1, x_2,...,x_n)$. Let $S$ be the fixed field under $S_n$. Each member of $S$ is called a symmetric function.

In other words we can say that a rational function $f(x_1, x_2,...,x_n)$ is called **symmetric function** if it is not changed by any permutation of the variables , *i.e.*

$$f(x_1, x_2,...,x_n) = f(x_{\sigma(1)}, x_{\sigma(2)},..., x_{\sigma(n)}) \text{ for every } \sigma \in S_n.$$

The **elementary symmetric functions** in $x_1, x_2,...,x_n$ are defined by

$$a_1 = x_1 + x_2 +...+ x_n = \sum_{i=1}^{n} x_i$$

$$a_2 = x_1 x_2 + x_1 x_3 +...+ x_2 x_3 + x_2 x_4 +...+ x_{n-1} x_n$$

$$= \sum_{i<j} x_i x_j$$

$$a_3 = \sum_{i<j<k} x_i x_j x_k$$

.....

.....

$$a_n = x_1 x_2 ... x_n.$$

For example when $n = 3$, then $a_1 = x_1 + x_2 + x_3$, $a_2 = x_1 x_2 + x_1 x_3 + x_2 x_3$, $a_3 = x_1 x_2 x_3$ ; when $n = 4$, then $a_1 = x_1 + x_2 + x_3 + x_4$, $a_2 = x_1 x_2 + x_1 x_3 + x_1 x_4 + x_2 x_3 + x_2 x_4 + x_3 x_4$, $a_3 = x_1 x_2 x_3 + x_2 x_3 x_4 + x_1 x_2 x_4 + x_1 x_3 x_4$ and $a_4 = x_1 x_2 x_3 x_4$.

If $x_1, x_2,..., x_n$ be $n$ indeterminates over any field $F$ and $a_1, a_2,..., a_n$ be the elementary symmetric functions in $x_1, x_2,..., x_n$, then the polynomial

$$f(x) = \prod_{i=1}^{n} (x - x_i)$$

$$= x^n - a_1 x^{n-1} + a_2 x^{n-2} - a_3 x^{n-3} +...+ (-1)^n a_n \text{ is called } \boldsymbol{n^{th}} \textbf{ genetic polynomial.}$$

**Note that** if $f(x_1, x_2,..., x_n)$ is a polynomial in $x_1, x_2,..., x_n$ which is symmetric, then it is easy to see that $f$ is actually a polynomial in $a_1, a_2,..., a_n$. For example, the polynomial $x_1^3 + x_2^3 + x_3^3$ is symmetric in $x_1, x_2, x_3$ and it can be written as $a_1^3 - 3a_1a_2 + 3a_3$.

**Theorem 10.** *The general polynomial equation of degree n is not solvable by radicals for* $n \geq 5$.

**Proof.** Let $F$ be a field and $a_1, a_2, \ldots, a_n$ be the elementary symmetric functions in the $n$ indeterminates $x_1, x_2, \ldots, x_n$. Let

$$f(x) = \prod_{i=1}^{n} (x - x_i)$$

$$= (x - x_1)(x - x_2) \ldots (x - x_n)$$

$$= x^n - a_1 x^{n-1} + a_2 x^{n-2} \ldots + (-1)^n a_n$$

be the general polynomial of degree $n$, whose roots are the indeterminate $x_1, x_2, \ldots, x_n$. Note that $f(x)$ has coefficients in $F(a_1, a_2, \ldots, a_n)$, and factors over $F(x_1, x_2, \ldots, x_n)$.

Let $K = F(x_1, x_2, \ldots, x_n)$ and $S = F(a_1, a_2, \ldots, a_n)$. Clearly $K$ is the splitting field of $f(x)$ over $F$. Since the $x_i$ are indeterminate over $F$, each $\sigma \in S_n$ induces an automorphism $\sigma'$ of $K$ defined by $\sigma'(a) = a$ for all $a \in F$ and $\sigma'(x_i) = x_{\sigma(i)}$. Again, since $\prod_{i=1}^{n} (x - x_i) = \prod_{i=1}^{n} \left(x - x_{\sigma(i)}\right)$, we have $\sigma'(a_i) = a_i$, for each $i$. So $\sigma'$ leaves $S$ fixed. Hence $\sigma' \in G(K \mid S)$.

Since $o(S_n) = \lfloor n$, so

$$o\left(G\left(K \mid S\right)\right) \geq \lfloor n \qquad \ldots\ldots(1)$$

Again, since the splitting field of a polynomial of degree $n$ over $S$ has degree atmost $\lfloor n$ over $S$, so we have

$$o\left(G\left(K \mid S\right)\right) \leq \lfloor n \qquad \ldots\ldots(2)$$

From (1) and (2) we get $o\left(G\left(K \mid S\right)\right) = \lfloor n$, and the automorphism $\sigma'$ comparise the full Galois group $G(K \mid S)$. Therefore $G(K \mid S) \cong S_n$.

But $S_n$ is not solvable for $n \geq 5$, so $G(K \mid S)$ is not solvable for $n \geq 5$. Hence by Theorem 8, $f(x)$ is not solvable by radicals over $S$, when $n \geq 5$.

**Note that** if $f(x) \in Q[x]$ be a monic irreducible polynomial over $Q$ of degree $p$, where $p$ is prime and if $f(x)$ has exactly two non-real roots in $C$, then the Galois group of $f(x)$ is isomorphic to $S_p$.

## Self-learning exercise-1

**1.** If $K$ is a finite field extension of a field $F$, then $K$ is a Galois extension of $F$, if :

*(a)* $K$ is normal extension

*(b)* $K$ is separable extension

*(c)* $K$ is both normal and separable extension

*(d)* Note of these

**2.** If $K$ is an extension of a field $F$, then the identity automorphism $I_K$ on a field $K$ leaves every element of $F$ fixed. [T/F]

**3.** The order of the Galois group $G(K \mid F)$ is not equal to the degree of $K$ over $F$. [T/F]

**4.** The field $Q\left(\sqrt[3]{2}\right)$ is not a radical extension of the field $Q$ of rational numbers. [T/F]

**5.** The polynomial $x^5 - 8x + 6$ is not solvable by radical over $Q$. [T/F]

**6.** $Q\left(\sqrt[3]{2}, \sqrt[5]{3}\right)$ is a radical extension of $Q$. [T/F]

## 9.5 Summary

I this unit we have discussed about Galois extension and Galois group, Fundamental theorem of Galois theory, extensions by radicals and solvability of the quintic and important results on these topics.

## 9.6 Answers to self learning exercises

### Self learning exercise-1

**1.** (c) **2.** True **3.** False **4.** False

**5.** True **6.** True

## 9.7 Exercises

**1.** Verify the Fundamental Theorem of Galois Theory for the splitting field $x^4 - 2 \in Q[x]$.

**2.** Show that a finite field is a Galois extension of any of its subfields.

**3.** Let $K$ be the splitting field of $f(x) = x^4 - 10x^2 + 1$ over $Q$. Then find $G(K \mid Q)$.

**4.** Show that the following polynomials :

***(i)*** $3x^5 - 15x + 5 \in Q[x]$

***(ii)*** $x^5 - x - 1 \in Q[x]$ are not solvable by radicals.

□ □ □

# UNIT 10 : Matrices of Linear Maps

**Structure of the Unit**



## 10.0 Objectives

This unit provides a general overview of linear transformation and matrices. The objective of this unit is to establish a relation between linear transformations and matrices. It is possible to represent a linear transformations by a matrix and conversely by taking a particular basis of a vector space $V$ and using the action of a linear transformation from $V$ to another vector space $V'$ on this basis.

## 10.1 Introduction

This unit introduces matrix of a linear map from a vector space $V$ to a vector space $V'$ and linear maps corresponding to matrices defined over a field $F$. This idea is then used to translate properties of linear maps to the corresponding properties of matrices. In the case of finite dimensional vector spaces, calculations with linear transformations are mush easy, because if the images of a linear transformation are known for any basis of vector space $V$ then the images of all vectors in $V$ can be calculated. We use this property in the introduction of matrices to describe linear maps. In this unit we shall study matrices of linear maps, composite maps, and of dual maps.

## 10.2 Matrix of a linear map

Let $V$ and $V'$ be any two finite dimensional vector spaces over a field $F$, and $B = \{b_1, b_2,...,b_n\}$, $B' = \{b_1', b_2', ..., b_m'\}$ be bases (ordered basis) of $V$ and $V'$ respectively. Also let $t : V \to V'$ be a linear transformation (linear map).

Since $b_j \in V \Rightarrow t(b_j) \in V'$, so that there exist scalars $a_{ij} \in F$

such that $t(b_j) = \sum_{i=1}^{m} a_{ij}\, b_i'$, $j = 1, 2, \ldots, n$

Then, we have

$$\begin{aligned}
t(b_1) &= a_{11}\, b_1' + a_{21}\, b_2' + \ldots + a_{m1}\, b_m', \\
t(b_2) &= a_{12}\, b_1' + a_{22}\, b_2' + \ldots + a_{m2}\, b_m', \\
&\vdots \\
&\vdots \\
t(b_j) &= a_{1j}\, b_1' + a_{2j}\, b_2' + \ldots + a_{mj}\, b_m', \\
&\vdots \\
&\vdots \\
t(b_n) &= a_{1n}\, b_1' + a_{2n}\, b_2' + \ldots + a_{mn}\, b_m',
\end{aligned}$$

The matrix of linear map $t$, in denoted by $M_{B'}^{B}(t)$, with respect to the basis $B$ of $V$ and $B'$ of $V'$.

The matrix of the coefficients is

$$M_{B'}^{B}(t) = \begin{bmatrix}
a_{11} & a_{12} & \ldots & a_{1j} & \ldots & a_{1n} \\
a_{21} & a_{22} & \ldots & a_{2j} & \ldots & a_{2n} \\
\vdots & & & & & \\
a_{i1} & a_{i2} & \ldots & a_{ij} & \ldots & a_{in} \\
\vdots & & & & & \\
a_{m1} & a_{m2} & \ldots & a_{mj} & \ldots & a_{mn}
\end{bmatrix}$$

Now, if $B^* = \{f_1, f_2, \ldots, f_n\}$, $B'^* = \{f_1', f_2', \ldots, f_m'\}$ be the basis dual to $B$ and $B'$ respectively, then for all $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$, we have

$$\begin{aligned}
f_i'\left[t(b_j)\right] &= f_i'\left[\sum_{r=1}^{m} a_{rj}\, b_r'\right] \\
&= \sum_{r=1}^{m} a_{rj}\, f_i'(b_r') \\
&= \sum_{r=1}^{m} a_{rj}\, \delta_{ir} \\
&= a_i \\
&= \left[M_{B'}^{B}(t)\right]_{ij}
\end{aligned}$$

Thus $\left[M_{B'}^{B}(t)\right]_{ij} = a_{ij} = f_i'\left[t(b_j)\right]$, $\quad i = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, n,$

and is called matrix of the linear map $t$. This matrix depends on the map $t$ as well as on the particular basis used. Since each vector $v$ of a vector space $V$ over a field $F$ is a linear map $V: F \to V$, the entries of the matrix of a vector $v$ relative to some basis $B = \{b_1, b_2,..., b_n\}$ of $V$, and the multiplicative identity $1 \in F$, as the basis for vector space $F$, in given by

$$\begin{aligned}\left[M(v)\right]_{i1} &= f_i(v(1)), \quad i = 1, 2,...,n \\ &= f_i(v.1) \\ &= f_i(v) \\ &= i^{th} \text{ coordinate of } v \text{ relative to basis } B \text{ of } V,\end{aligned}$$

where, $B^* = \{f_1, f_2,..., f_n\}$ in the basis dual to $B$.

Hence,
$$M(v) = \begin{bmatrix} f_1(v) \\ f_2(v) \\ \vdots \\ f_n(v) \end{bmatrix}$$

This matrix is known as coordinate matrix of vector $v$ relative to basis $B$.

Similarly for each $v \in V$, $t(v) \in V'$ and $M(t(v))$ is coordinate matrix of $t(v)$, relative to basis $B'$ and is given by

$$M(t(v)) = \begin{bmatrix} f_1'(t(v)) \\ f_2'(t(v)) \\ \vdots \\ f_m'(t(v)) \end{bmatrix},$$

where $B'^* = \left\{f_1', f_2', ..., f_m'\right\}$ is basis dual to $B'$.

Given that $A = [a_{ij}]$ over $F$ in which each entry $a_{rs}$ is a transformation from $F$ to $F$, each basis vector $b_r'$ is a linear transformation form $F$ to $V'$. Thus the composite $b_r' a_{rs} f_s$ of the linear transformations

$$f_s: V \to F,$$
$$a_{rs}: F \to F,$$
$$b_r': F \to V',$$

is a transformation, from $V$ to $V'$ for all $r = 1, 2,...m$ and $s = 1, 2,...,n$.

Since the sum of a finite number of linear transformation is a linear transformation and so the sum,

$\sum_{r=1}^{m}\sum_{s=1}^{n} b_r' \; a_{rs} \; f_s$, is a linear transformation from $V$ to $V'$. We denote this sum by $t$, then we have

$$\begin{aligned}
\left[M_{B'}^{B}(t)\right]_{ij} &= f_i'\left(t(b_j)\right) \\
&= f_i'\left[\left(\sum_{r=1}^{m}\sum_{r=1}^{n} b_r' \; a_{rs} \; f_s\right)(b_j)\right] \\
&= f_i'\left[\sum_{r=1}^{m} b_r'\left(\sum_{s=1}^{n} a_{rs} \; f_s \; (b_j)\right)\right] \\
&= f_i'\left[\sum_{r=1}^{m} b_r'\left(\sum_{s=1}^{n} a_{rs} \; \delta_{rj}\right)\right] \\
&= f_i'\left[\sum_{r=1}^{m} b_r' \; a_{rj}\right] \\
&= \sum_{r=1}^{m} f_i'\left(b_r'\right) a_{rj} \\
&= \sum_{r=1}^{m} \delta_{ir} \; a_{rj} \\
&= a_{ij}, \; i = 1, 2, ..., m; \; j = 1, 2, ..., n
\end{aligned}$$

Hence $M_{B'}^{B}(t) = A$.

Thus each $m \times n$ matrix $A = [a_{ij}]$ over field $F$ with basis $B = \{b_1, b_2, ...b_n\}$ and $B' = \{b_1', b_2', ..., b_m'\}$ of $V$ and $V'$ respectively, defines a linear transformation $t : V \to V'$ given by

$$t = \sum_{r=1}^{m}\sum_{s=1}^{n} b_r' \; b_{rs} \; f_s,$$

whose matrix relative to bases $B$ and $B'$ is matrix $A$ itself. This linear transformation is denoted by $M^{-1}(A)$ and is known as map of matrix $A$.

**Theorem 1.** *Let $V$ be finite dimensional vector space over a field $F$ with $B = \{b_1, b_2, ..., b_n\}$ as the basis of $V$, then the matrix of the identity map on $V$ in given by $M_B^B(I_v) = I_n$, identity matrix of order $n$.*

**Proof :** Let $B^* = \{f_1, f_2, ..., f_n\}$ be basis dual to $B$, then

$$\begin{aligned}
\left[M_B^B(I_v)\right]_{ij} &= f_i\left(I_v(b_j)\right) = f_i\,(b_j) \\
&= \delta_{ij}, \; i, j = 1, 2, ..., n
\end{aligned}$$

Thus $M_B^B(I_v) = I_n$.

**Theorem 2.** *Let $V$ and $V'$ be finite dimensional vector spaces over a field $F$ and $B = \{b_1, b_2,..., b_n\}$ and $B' = \{b_1', b_2',..., b_n'\}$* be bases of $V$ and $V'$ respectively, then

***(i)*** The matrix of zero map $\hat{0}$ on $V$ is given by

$$M_B^B\left(\hat{0}\right) = \mathbf{0}_{n\times n} \text{ and}$$

***(ii)*** the matrix of zero map $\hat{0}: V \to V'$ is given by

$$M_{B'}^B\left(\hat{0}\right) = \mathbf{0}_{m\times n}$$

**Proof :** ***(i)*** $$\left[M_B^B\left(\hat{0}\right)\right]_{ij} = f_i\left(\hat{0}\,(b_j)\right) = f_i\left(\mathbf{0}\right) = 0$$

Thus $M_B^B\left(\hat{0}\right) = \mathbf{0}_{n\times n}$, where $B^* = \{f_1, f_2,..., f_n\}$ is the basis dual to $B$.

***(ii)*** $$\left[M_{B'}^B\left(\hat{0}\right)\right]_{ij} = f_i'\left(\hat{0}\,(b_j)\right) = f_i'\left(\mathbf{0}\right) = 0$$

Thus $M_{B'}^B\left(\hat{0}\right) = \mathbf{0}_{m\times n}$, where $B'^* = \left\{f_1', f_2',..., f_m'\right\}$ is the basis dual to $B'$.

***Ex.*1.** *Let $t : R^2 \to R^2$ be a linear transformation defined by*

$$t(a, b) = (2a - 3b, a + b),\ \forall\,(a, b) \in R^2.$$

*Then find the matrix of $t$ relative to the basis $B = \{(1, 0), (0, 1)\}$, $B' = \{(2, 3), (1, 2)\}$.*

**Sol.** Since $t(a, b) = (2a - 3b, a + b),\ \forall\,(a, b) \in R^2$, so that

$$t(1, 0) = (2, 1) = 2\,(1, 0) + 1\,(0, 1)$$

$$t(0, 1) = (-3, 1) = (-3)\,(1, 0) + 1\,(0, 1)$$

Thus $M_B^B(t)$ = Transpose of the coefficient matrix

$$= \begin{bmatrix} 2 & 1 \\ -3 & 1 \end{bmatrix}^T = \begin{bmatrix} 2 & -3 \\ 1 & 1 \end{bmatrix}$$

Similarly $t(2, 3) = (-5, 5)$

$$= 25(1, 2) + (-15)\,(2, 3)$$

$$t(1, 2) = (-4, 3)$$

$$= 18(1, 2) + (-11)\,(2, 3)$$

So that,

$$M_{B'}^{B'}(t) = \begin{bmatrix} 25 & 18 \\ -15 & -11 \end{bmatrix}.$$

***Ex.*2.** *Let $t : R^3 \to R^3$ be a linear transformation such that,*

$$t(a, b, c) = (3a + c, -2a + b, -a + 2b + 4c).$$

*what is the matrix of $t$ in the ordered basis $\{\alpha_1, \alpha_2, \alpha_3\}$, where*

$$\alpha_1 = (1, 0, 1),\ \alpha_2 = (-1, 2, 1),\ \alpha_3 = (2, 1, 1).$$

**Sol.** Given that $t : R^3 \to R^3$ is a linear transformation such that

$$t(a, b, c) = (3a + c, -2a + b, -a + 2b + 4c) \quad \text{.....(1)}$$

and, $B = \{\alpha_1, \alpha_2, \alpha_3\}$ is an ordered basis of $R^3$.

Now, $t(\alpha_1) = t(1, 0, 1) = (4, -2, 3)$ .....(2*a*)

$t(\alpha_2) = t(-1, 2, 1) = (-2, 4, 9)$ .....(2*b*)

and, $t(\alpha_3) = t(2, 1, -1) = (7, -3, 4)$ .....(2*c*)

Let $(x, y, z) = \lambda\,\alpha_1 + \mu\,\alpha_2 + \upsilon\,\alpha_3$ .....(3)

Then $(x, y, z) = \lambda\,(1, 0, 1) + \mu\,(-1, 2, 1) + \upsilon\,(2, 1, 1)$

$= (\lambda - \mu + 2\upsilon, 2\mu + \upsilon, \lambda + \mu + \upsilon)$

so that, $\lambda - \mu + 2\upsilon = x,$

$2\mu + \upsilon = y,$

$\lambda + \mu + \upsilon = z$

Solving for $\lambda$, $\mu$, $\nu$, we get

$$\lambda = \frac{1}{4}(-x - 3y + 5z),$$

$$\mu = \frac{1}{4}(-x + y + z),$$

$$\upsilon = \frac{1}{2}(x + y - z).$$

Putting these in equation (3), we get

$$(x, y, z) = \frac{\alpha_1}{4}(-x - 3y + 5z) + \frac{\alpha_2}{4}(-x + y + z) + \frac{\alpha_3}{4}(2x + 2y - 2z). \quad \text{.....(4)}$$

Putting $x = 4, y = -2, z = 3$ in equation (4), and using 2 (*a*), we get

$$t(\alpha_1) = (4, -2, 3) = \frac{17}{4}\alpha_1 - \frac{3}{4}\alpha_2 - \frac{1}{2}\alpha_3 \quad \text{.....(5)}$$

Putting $x = -2, y = 4, z = 9$ in (4) and using equation 2 (*b*), we get

$$t(\alpha_2) = (-2, 4, 9) = \frac{35}{4}\alpha_1 + \frac{15}{4}\alpha_2 - \frac{7}{2}\alpha_3 \quad \text{.....(6)}$$

Putting $x = 7, y = -3, z = 4$ in equation (4) and using equation 2 (*c*), we get

$$t(\alpha_3) = (7, -3, 4) = \frac{11}{2}\alpha_1 - \frac{3}{2}\alpha_2 + 0\alpha_3 \quad \text{.....(7)}$$

Hence

$$M_B^B(t) = \begin{bmatrix} \frac{17}{4} & \frac{35}{4} & \frac{11}{2} \\ -\frac{3}{4} & \frac{15}{4} & -\frac{3}{2} \\ -\frac{1}{2} & -\frac{7}{2} & 0 \end{bmatrix}$$

## 10.3 Matrices of composite map

**Theorem 3.** *Let V, W, U be vector spaces over the same field F. Let*

$$\{v_j\}_{j=1}^{n}, \{w_i\}_{i=1}^{m}$$

*and* $\{u_r\}_{r=1}^{k}$ *be the bases of V, W and U respectively. If t : V → W, s : W → U are linear transformations, and A and B are the matrix relative to t and s respectively. Then the matrix relative to sot is B A.*

**Proof :** Let $t(v_j) = \sum_{i=1}^{m} a_{ij}\, w_i$

and $s(w_i) = \sum_{r=1}^{k} b_{ri}\, u_r$ , for $a_{ij}, b_{ri} \in F$.

Now,

$$(sot)(v_j) = s\left[t\left(v_j\right)\right]$$

$$= s\left[\sum_{i=1}^{m} a_{ij}\, w_i\right]$$

$$= \sum_{i=1}^{m} a_{ij}\, s\left(w_i\right).$$

$$= \sum_{i=1}^{m} a_{ij}\left(\sum_{r=1}^{k} b_{ri}\, u_r\right)$$

$$= \sum_{r=1}^{k}\left(\sum_{i=1}^{m} b_{ri}\, a_{ij}\right) u_r$$

$$= \sum_{r=1}^{k} c_{rj}\, u_r,$$

where $c_{rj} = \sum_{i=1}^{m} b_{ri}\, a_{ij}$ is the $(r,j)^{th}$ entry of matrix $BA$.

Hence, $M(sot) = BA$

$$= M(s)\, M(t).$$

**Theorem 4.** *A linear transformation t : V → V is invertible iff matrix of t relative to some bases B of V is invertible.*

**Proof :** First let $t : V \to V$ be invertible, that is there exists $t^{-1} : V \to V$, such that

$$t \, o \, t^{-1} = I_v = t^{-1} \, o \, t$$

$$\Rightarrow \quad M_B^B\left(t \, o \, t^{-1}\right) = M_B^B\left(I_v\right) = M_B^B\left(t^{-1} \, o \, t\right)$$

$$\Rightarrow \qquad M_B^B(t)\, M_B^B\left(t^{-1}\right) = I = M_B^B\left(t^{-1}\right) M_B^B(t)$$

Hence the matrix of t relative to basis $B$ *i.e.* $M_B^B(t)$ is invertible.

Conversely suppose that $M_B^B(t)$ is an invertible matrix so that $\left[M_B^B(t)\right]^{-1} \in M_{n\times n}(F)$

Thus there exists a linear transformation $s : V \to V$, such that

$$M_B^B(s) = \left[M_B^B(t)\right]^{-1}$$

$$\Rightarrow \qquad M_B^B(t)\, M_B^B(s) = I = M_B^B(s)\, M_B^B(t)$$

$$\Rightarrow \qquad M_B^B(t \circ s) = M_B^B(I_v) = M_B^B(s \circ t)$$

$$\Rightarrow \qquad t \circ s = I_v = s \circ t$$

Hence $t$ is invertible.

**Theorem 5.** *Let $V$ and $V'$ be $n$ and $m$ dimensional vector spaces over a field $F$. Then for given bases $B$ and $B'$ of $V$ and $V'$ respectively, the function assigning to each linear transformation $t : V \to V'$ its matrix $M_{B'}^B(t)$ relative to bases $B$, $B'$ is an isomorphism between the vector space Hom $(V, V')$ and the space $F^{m\times n}$ of all $m\times n$ matrices over $F$ i.e.*

*Hom* $(V, V') \cong F^{m\times n}$.

**Proof :** Let $B = \{b_1, b_2, ..., b_n\}$ and $B' = \{b_1', b_2', ..., b_m'\}$ be the bases of $V$ and $V'$ respectively and $B^* = \{f_1, f_2, ..., f_n\}$, $B^{*\prime} = \{f_1', f_2', ..., f_m'\}$ be the bases dual to $B$ and $B'$ respectively.

We define a mapping $\phi$ : *Hom* $(V, V') \to F^{m\times n}$, as follows :

$$\phi(t) = M_{B'}^B(t), \ \forall\, t \in Hom\,(V, V').$$

It is clear that $\phi$ is well defined. We shall show that $\phi$ is an isomorphism

***(i)* $\phi$ is a linear transformation :**

Let $t_1, t_2 \in Hom\,(V, V')$ and $\lambda, \mu \in F$, then $\lambda t_1 + \mu t_2 \in Hom\,(V, V')$, and

$$\left[M_{B'}^B(\lambda t_1 + \mu t_2)\right]_{ij} = f_i'\left[(\lambda t_1 + \mu t_2)(b_j)\right], \quad i = 1, 2, ..., m; \ \ j = 1, 2, ..., n$$

$$= f_i'\left[\lambda t_1(b_j) + \mu t_2(b_j)\right]$$

$$= \lambda f_i'\left[t_1(b_j) + \mu f_i'\left(t_2(b_j)\right)\right]$$

$$= \lambda\left[M_{B'}^B(t_1)\right]_{ij} + \mu\left[M_{B'}^B(t_2)\right]_{ij}$$

$$= \left[\lambda M_{B'}^B(t_1)\right]_{ij} + \left[\mu M_{B'}^B(t_2)\right]_{ij}$$

$$= \left[\lambda M_{B'}^B(t_1)\, \mu + M_{B'}^B(t_2)\right]_{ij}$$

Thus

$$M_{B'}^B(\lambda t_1 + \mu t_2) = \lambda M_{B'}^B(t_1) + \mu M_{B'}^B(t_2)$$

$\Rightarrow \qquad \phi(\lambda t_1 + \mu t_2) = \lambda \phi(t_1) + \mu \phi(t_2)$

Thus $\phi$ is a linear transformation.

***(ii)* $\phi$ is one-one :** Let $t_1, t_2 \in Hom\,(V, V')$ be such that

$$M_{B'}^{B}(t_1) = A = [a_{ij}],$$

and $$M_{B'}^{B}(t_2) = B = [b_{ij}],$$

and $$\phi(t_1) = \phi(t_2)$$

$$\Rightarrow \quad M_{B'}^{B}(t_1) = M_{B'}^{B}(t_2)$$

$$\Rightarrow \quad \left[M_{B'}^{B}(t_1)\right]_{ij} = \left[M_{B'}^{B}(t_2)\right]_{ij}, \quad i = 1, 2,..., m; \ \ j = 1, 2,..., n$$

$$\Rightarrow \quad f_i'\left(t_1(b_j)\right) = f_i'\left(t_2(b_j)\right)$$

$$\Rightarrow \quad a_{ij} = b_{ij}$$

$$\Rightarrow \quad b_i' a_{ij} = b_i' b_{ij}$$

$$\Rightarrow \quad \sum_{i=1}^{m} b_i' a_{ij} = \sum_{i=1}^{m} b_i' b_{ij}$$

$$\Rightarrow \quad t_1(b_j) = t_2(b_j)$$

$$\Rightarrow \quad t_1 = t_2 \ \text{ on } B$$

$$\Rightarrow \quad t_1 = t_2 \ \text{ on } V \qquad [\because B \text{ is a basis for } V]$$

Thus $\phi$ is one-one.

***(iii)* $\phi$ is onto :** For each $A = (a_{ij}) \in F^{m \times n}$ there exists a linear transformation $t : V \to V'$ such that

$$\phi(t) = M_{B'}^{B}(t) = A.$$

Thus $\phi$ is onto.

Hence $\phi : Hom\,(V, V') \to F^{m \times n}$ is an isomorphism and so that

$$Hom\,(V, V') \cong F^{m \times n}.$$

## 10.4 Matrices of dual maps

**Theorem 6.** *Let $V$ and $V'$ be finite dimensional vector spaces over a field $F$ with bases $B$ and $B'$ respectively. If $t : V \to V'$ be a linear transformation, then*

$M_{B^*}^{B'^*}(t^*) = \left[M_{B'}^{B}(t)\right]^T$, *where $t^*$ is the dual map of $t$ and $B^*$ and $B'^*$ are the bases dual to $B$ and $B'$ respectively.*

**Proof :** Let $B = \{b_1, b_2,..., b_n\}$ and $B' = \{b_1', b_2',..., b_m'\}$ be the bases of vector space $V$ and $V'$ respectively, and $B^* = \{f_1, f_2,..., f_n\}$, $B'^* = \{f_1', f_2',..., f_m'\}$ be the bases dual to $B$ and $B'$ respectively.

Then, we have

$$t(b_j) = \sum_{i=1}^{m} b_i' \, a_{ij}, \quad a_{ij} \in F, \qquad j = 1, 2, ..., n \qquad .....(1)$$

so that

$$M_{B'}^{B}(t) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

It is given that $t^*$ in the dual map of $t$, so $t^* : V'^* \to V^*$ such that

$$t^*(f') = f'ot, \ \forall \ f' \in V'^*, \qquad .....(2)$$

and

$$f'ot = \sum_{j=1}^{n} f_j \, a_{ij}, \qquad i = 1, 2, ..., m \qquad .....(3)$$

From (2) and (3), we have

$$t^*(f') = \sum_{j=1}^{n} f_j \, a_{ij}$$

$$= \sum_{j=1}^{n} f_j \left(A^T\right)_{ji}, \quad i = 1, 2, ..., m \qquad .....(4)$$

Equation (1) and (4) depict the effect on the basis vector in bases $B$ and $B'^*$ respectively. From (1) and (4), we have

$$\left[M_{B'}^{B}(t)\right]_{ij} = a_{ij}$$

and

$$\left[M_{B^*}^{B'^*}\left(t^*\right)\right]_{ji} = \left(A^T\right)_{ji}, i = 1, 2, ..., m; \ j = 1, 2, ..., n$$

$\Rightarrow$

$$\left[M_{B}^{B'^*}(t)^*\right]_{ij} = \left(A^T\right)_{ij},$$

$\Rightarrow$

$$M_{B^*}^{B'^*}\left(t^*\right) = \left[M_{B'}^{B}(t)\right]^T.$$

**Theorem 7.** *Let $V$, $V'$ and $V''$ be finite dimensional vector spaces over a field $F$ and let $B$, $B'$ and $B''$ be there respectively bases. Then for linear transformations $t : V \to V'$ and $s : V' \to V''$,*

$$M_{B''}^{B}(sot) = M_{B''}^{B'}(s) \, M_{B'}^{B}(t).$$

**Proof :** Let $B = \{b_1, b_2, ..., b_n\}$, $B' = \{b_1', b_2', ..., b_p'\}$ and $B'' = \{b_1'', b_2'', ..., b_m''\}$ be the bases for $V$, $V'$ and $V''$ respectively, and $B^* = \{f_1, f_2, ..., f_n\}$ $B'^* = \{f_1', f_2', ..., f_p'\}$, $B''^* = \{f_1'', f_2'', ..., f_m''\}$ be the bases dual to $B$ $B'$ and $B''$ respectively.

Also let, $M_{B''}^{B'}(s) = p, M_{B'}^{B}(t) = Q$ and $M_{B''}^{B}(sot) = R.$

Since $P$, $Q$ and $R$ be $m \times p$, $p \times n$ and $m \times n$ matrices respectively, so $PQ$ and $R$ of same order $m \times n$ matrices.

Now, we have

$$(PQ)_{ij} = \sum_{r=1}^{p} (P)_{ir} (Q)_{rj}, \; i = 1, 2,..., m; \; j = 1, 2,..., n$$

$$= \sum_{r=1}^{p} \left[ f_i''\left(s(b_r')\right)\right]\left[ f_r'\left(t(b_j)\right)\right]$$

$$= f_i'' \left[ \left( \sum_{r=1}^{p} s\left(b_r'\right) f_r' \right) t(b_j) \right]$$

$$= f_i'' \left[ s\left( \sum_{r=1}^{p} b_r' \, f_r' \right) t(b_j) \right]$$

$$= f_i'' \left[ s(1)\, t\left(b_j\right)\right]$$

$$= f_i'' \left[ s\left[ t\left(b_j\right)\right]\right] \qquad \left[ \because \sum_{r=1}^{p} b_r' f_r' = 1 \right]$$

$$= f_i'' \left[ (sot)\left(b_j\right)\right]$$

$$= \left[ M_{B''}^{B}(sot)\right]_{ij}$$

$$= (R)_{ij}$$

Thus $PQ = R$

Hence

$$M_{B''}^{B'}(s)\, M_{B'}^{B}(t) = M_{B''}^{B}(sot)$$

## Self-learning exercise –1

**1.** True/False statements :

*(i)* If $V$ and $V'$ be finite dimensional vector spaces over the same field $F$, then each linear transformation from $V$ to $V'$ determines a matrix. [T/F]

*(ii)* Each matrix determines a linear transformation. [T/F]

*(iii)* Linear transformation determined by a matrix is independent of the choice of basis. [T/F]

**2.** Fill in the blanks :

*(i)* Matrix of identify map on $V_4(R)$ is the ..... matrix of order ..... .

*(ii)* Linear transformation of identity matrix $I_5(R)$ is the ..... map on ..... .

*(iii)* Matrix of zero map from $V_3(R)$ to $V_2(R)$ is the ..... matrix of order ..... .

## 10.5 Summary

In this unit we have studied matrix of a linear transformation, and some particular matrices, such as matrix of an identity map, matrix of a zero map, matrix of composite map and matrix of a dual map.

## 10.6 Answers to self-learning exercises

**Self-learning exercise-1**

**1.** ***(i)*** T ***(ii)*** T ***(iii)*** F

**2.** ***(i)*** identity, $4 \times 4$ ***(ii)*** identity, $R^5$ $(R)$, ***(iii)*** zero, $2 \times 3$.

## 10.7 Exercises

**1.** Describe the matrix of a linear form $f : V \to F$ relative to a basis $B$ to $V$.

**2.** If the matrix of a linear transformation $t : R^2 \to R^2$, relative to the standard basis

$$B = \{e_1^{(2)}, e_2^{(2)}\} \text{ of } R^2 \text{ is } \begin{bmatrix} 2 & -3 \\ 1 & 1 \end{bmatrix},$$

then find the matrix of $t$ relative to the basis $C = \{(1, 1), (1, -1)\}$.

**3.** If the matrix of a linear map $t$ on $V_3$ $(C)$ with respect to the basis $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)$ is

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

with respect to basis $\{(1, 1, -1), (-1, 0, 1), (1, 2, 1)\}$. ?

**4.** Let $V = R^3$ and $t : V \to V$ be a linear map, defined by $t(x, y, z) = (x + z, -2x + y, -x + 2y + z)$. What is the matrix of $t$ with respect to basis $\{(1, 0, 1), (-1, 1, 1), (0, 1, 1)\}$ ?

□□□

# UNIT 11 : Rank and Nullity of Matrices

**Structure of the Unit**



## 11.0 Objectives

In this unit, we shall study the notion of a rank and nullity of a matrix in general, which plays an important role in the solution of linear equations, and also other concepts such as eigenvalues and eigne vectors of linear maps and their matrices.

## 11.1 Introduction

In unit-6, we have studied vector spaces, dual spaces, linear transformation of vector spaces, their rank, nullity and dual maps. In unit 10, matrices of these linear transformations and their properties have been studied. This unit, further introduces the concept of rank of a matrix through the rank of the corresponding linear transformation and its properties. This unit also provides other related concepts such as invertible matrices, eigenvalues and eigenvectors, change of basis and similar matrices.

## 11.2 Rank and nullity of a matrix

Let $A = [a_{ij}]$ be any $m \times n$ matrix over a field $F$, then we write

$$A = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_i \\ \vdots \\ R_m \end{bmatrix},$$

where $R_i = [a_{i1}, a_{i2}, ..., a_{in}]$ is the $i^{th}$ row of matrix $A$. Similarly, we may write

$$A = [c_1\ c_2, ...\ c_2\ ...\ c_n],$$

where $$c_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix},$$ is the $j^{th}$ column of $A$.

We now consider $m$ rows $R_1, R_2, ..., R_m$ of matrix $A$ as row vectors of the space $F^n$, and $n$ columns $c_1, c_2, ..., c_n$ of matrix $A$ as column vectors in $F^m$.

The vector subspace generated by row vectors is called the **row space** of $A$ and the vector subspace generated by column vectors is called the **column space** of $A$.

**Row rank of a matrix :**

The row rank of a matrix $A$ is the dimension of the vector subspace of $F^n$, generated by $R_1, R_2, ..., R_m$ (row space).

**Column rank of a matrix :**

The column rank of a matrix $A$ is the dimension of the vector subspace $F^m$, generated by $c_1, c_2, ..., c_n$ (column space). The rank the matrix $A$ is also the rank of the linear transformation $t_A : F^n \to F^m$. Similarly, nullity of the matrix $A$ is also the nullity of the linear transformation $t_A : F^n \to F^m$.

**Theorem 1.** *Let $t_A \in \text{Hom}(F^n, F^m)$ be the linear transformation corresponding to an $m \times n$ matrix $A = [a_{ij}]$ over a field F, then the rank of $t_A$, as a linear transformation equals to the column rank of A.*

**Proof :** Since $t_A \in \text{Hom}(F^n, F^m)$ be the linear transformation, then $t_A\ F^n \to F^m$ be such that

$$t_A(X) = AX, \qquad \forall\ X \in F^n,$$

so that $$t_A\left(e_j^{(n)}\right) = Ae_j^{(n)}, \quad \text{for} \quad j = 1, 2, ..., n$$

$$= c_j \in F^m$$

$$= \sum_{i=1}^{m} e_i^{(m)}\, a_{ij} \in F^m,$$

where $c_j$ is the $j^{\text{th}}$ column vector of matrix $A$. Thus $t_A$ is the linear transformation which maps the $j^{\text{th}}$ element of the standard ordered basis of $F^n$ to the $j^{\text{th}}$ column $c_j$ of the matrix $A$. Hence the column space is the image space *i.e.* $im(t_A)$. Therefore the column rank of $A$ is the dimension of $im(t_A)$ *i.e.* the rank of $t_A$.

Thus, rank $(t_A)$ = column rank of $A$.

**Note :** Since, rank $(t_A)$ = rank of $A$.

But rank $(t_A)$ = column rank of $A$,

Hence, rank of $A$ = column rank of $A$.

**Corollary :** For any matrix $A$, the row rank of $A$, equals to the column rank of $A$.

**Proof :** Let $V$ and $V'$ be vector spaces over a field $F$ and $t : V \to V'$ be a linear transformation. Let the matrix of it is $A$. So the matrix of $t^* : V'^* \to V^*$, is $A^T$, where $t^*$ is the dual map of $t$.

Since, rank of $t$ = column rank of $A$ .....(1)

So that rank of $t^*$ = column rank of $A^T$

= Row rank of $A$ .....(2)

But rank of $t$ = rank of $t^*$

From (1) and (2), we obtain

row rank of $A$ = column rank of $A$.

**Theorem 2.** *For any matrix A over a field F, rank* $(A)$ = *rank* $(A^T)$.

**Proof :** Let $V$ and $V'$ be vector space over a field $F$ and $B$, $B'$ be their respective bases. Also let $t : V \to V'$ be a linear transformation, such that

$$M_{B'}^{B}(t) = A$$

or $M(t) = A$,

so rank $(t)$ = rank $(A)$ .....(1)

Let $t^* : V'^* \to V^*$ be the dual map of $t$,

then rank $(t)$ = rank $(t^*)$ .....(2)

From (1) and (2),

rank $(A)$ = rank $(t^*)$ .....(3)

Also, we have $M(t^*) = [M(t)]^T$

$M(t^*) = A^T$,

so that rank $(t^*)$ = rank $(M(t^*))$

$$= \text{rank } (A^T) \qquad \text{.....(4)}$$

From (3) and (4), we get

$$\text{rank } (A) = \text{rank } (A^T).$$

## 11.3 Invertible matrix

Let $A$ be an $n \times n$ square matrix over a field $F$. If there exists an $n \times n$ matrix $B$ such that

$$AB = I_n = BA,$$

then $A$ is called an invertible matrix and $B$ is called inverse of $A$.

**Theorem 3.** *Let $V$ be a vector space over a field $F$ and $B$ be its basis. Then a linear transformation $t : V \to V$ is invertible iff matrix of $t$ relative to basis $B$ is invertible.*

**Proof :** First suppose that $t : V \to V$ is an invertible linear transformation, then there exists a linear transformation $t^{-1} : V \to V$ such that

$$tot^{-1} = I = t^{-1}ot$$

$$\Rightarrow \quad M_B^B\left(tot^{-1}\right) = M_B^B(I) = M_B^B\left(t^{-1}ot\right)$$

$$\Rightarrow \quad M_B^B(t)M_B^B\left(t^{-1}\right) = I = M_B^B\left(t^{-1}\right)M_B^B(t)$$

Thus $M_B^B(t)$ is invertible, and

$$\left[M_B^B(t)\right]^{-1} = M_B^B\left(t^{-1}\right).$$

Conversely suppose that $M_B^B(t)$ be an invertible matrix. Let $\left[M_B^B(t)\right]^{-1}$ be the inverse of $M_B^B(t)$. Let $\left[M_B^B(t)\right]^{-1} \in F^{n \times n}$, so that there exists a linear transformation $s : V \to V$ such that

$$M_B^B(s) = \left[M_B^B(t)\right]^{-1}$$

$$\Rightarrow \quad M_B^B(s)M_B^B(t) = I_n = M_B^B(t)M_B^B(s)$$

$$\Rightarrow \quad M_B^B(sot) = M_B^B(I) = M_B^B(tos)$$

$$\Rightarrow \quad sot = I = tos$$

Thus $t$ is invertible and $t^{-1} = s$.

**Theorem 4.** *An $n \times n$ square matrix $A$ over a field $F$ is invertible iff rank $(A) = n$.*

**Proof :** First let the matrix $A$ be invertible, then there exists an $n \times n$ matrix $B$ over $F$ such that

$$AB = I_n = BA$$

Now, $AB = I_n$

$\Rightarrow$ A has a right inverse

$\Rightarrow$ $t_A : F^n \to F^n$ has right inverse

$\Rightarrow$ $t_A : F^n \to F^n$ is an epimorphism

$\Rightarrow$ rank $(t_A) = n$

$\Rightarrow$ rank $(A) = n$.

Conversely suppose that,

rank $(A) = n$

$\Rightarrow$ rank $(t_A) = n$

$\Rightarrow$ $t_A : F^n \to F^n$ is an epimorphism

$\Rightarrow$ $t_A$ has a right inverse.

$\Rightarrow$ A has a right inverse

Since $t_A : F^n \to F^n$ is an epimorphism

$\Rightarrow$ $t_A : F^n \to F^n$ is also a monomorphism

$\Rightarrow$ $t_A$ has a left inverse

$\Rightarrow$ $A$ has a left inverse.

Hence the matrix $A$ is invertible.

**Theorem 5.** *Prove that the following statements are equivalent for any matrix* $A \in F^{n \times n}$

***(i)*** *A has a left inverse in* $F^{n \times n}$,

***(ii)*** *nullity* $(A) = 0$

***(iii)*** *rank* $(A) = n$

***(iv)*** *A has a right inverse in* $F^{n \times n}$

***(v)*** *A has two sided inverse in* $F^{n \times n}$.

**Proof :** Let $t_A : F^n \to F^n$ be the linear transformation corresponding to the matrix $A$.

Now, A has a left inverse in $F^{n \times n}$, so $t_A$ has a left inverse.

$\Rightarrow$ $t_A$ is a monomorphism .....*(i)*

$\Rightarrow$ Ker $(t_A) = \{0\}$

$\Rightarrow$ nullity $(t_A) = 0$

$\Rightarrow$ nullity $(A) = 0$ .....*(ii)*

$\Rightarrow$ rank $(A) = n$ .....*(iii)*

$\Rightarrow$ $t_A : F^n \to F^n$ is an epimorphism

$\Rightarrow$ $t_A$ is an isomorphism

$\Rightarrow$ $t_A$ is invertible transformation

$\Rightarrow$ $A$ is invertible matrix

$\Rightarrow$ $A$ has a right inverse in $F^{n \times n}$ .....*(iv)*

$\Rightarrow$ $t_A$ has a right inverse

$\Rightarrow$ $t_A$ is an epimorphism

$\Rightarrow$ $t_A$ is an isomorphism

$\Rightarrow$ $t_A$ has two sided inverse

$\Rightarrow$ $A$ has two sided inverse in $F^{n \times n}$ .....(v)

$\Rightarrow$ $A$ has a left inverse in $F^{n \times n}$ .....(vi)

## 11.4 Change of basis

Let $V$ be a finite dimensional vector space over a field $F$. Let $B = \{b_1, b_2, ..., b_n\}$ and $B' = \{b_1', b_2', ..., b_n'\}$ be any two bases of $V$. Suppose $t : V \to V$ be a linear transformation, and $M_B^B(t)$ be the matrix of $t$ relative to basis $B$. Let $M_{B'}^{B'}(t)$ be the matrix $t$, relative to basis $B'$.

Let $I_V$ be the identity map, *i.e.* $I_V : V \to V$ such that

$$I_V(v) = v, \qquad \forall\, v \in V.$$

Then, for each $j$, we have $I_V(b_j') = b_j'$

$$= \sum_{i=1}^{n} a_{ij}\, b_i, \qquad \text{for} \quad a_{ij} \in F$$

$$= a_{1j}b_1 + a_{2j}b_2 + ... + a_{nj}b_n$$

Putting $j = 1, 2, ..., n$, we get

$$b_1' = a_{11}b_1 + a_{21}b_2 + ... + a_{n1}b_n$$

$$b_2' = a_{12}b_1 + a_{22}b_2 + ... + a_{n2}b_n$$

$$\vdots$$

$$b_n' = a_{1n}b_1 + a_{2n}b_2 + ... + a_{nn}b_n$$

The transpose of the above coefficient matrix is called the transition matrix or change of basis matrix $P$ from basis $B$ to $B'$.

This $P = M_B^{B'}(I_V)$ is the associated matrix of the identity map which expresses the basis vectors of $B'$ in terms of the basis vectors of $B$.

## 11.5 Similar matrices

Matrices $A$ and $B$ of order $n$ over a field $F$ are said to be similar if there exists an invertible matrix $P$ of order $n$ over $F$ such that

$$B = P^{-1}AP.$$

***Ex.1.*** *Let* $B = \{b_1 = (1,0), b_2 = (0,1)\}$ *and* $B' = \{b_1' = (1,3), b_2' = (2,5)\}$ *be any two bases of* $R^2$. *Then*

*(i) Determine the transition matrix P from the basis B to the basis B′ .*

*(ii) Determine the transition matrix Q from the basis B′ to the bases B.*

**Sol.** Let $I_V : R^2 \to R^2$ such that

$$I_V(v) = v, \qquad \forall\, v \in R^2,$$

then $$I_V\left(b_j'\right) = b_j' \qquad j = 1, 2,.$$

Put $j = 1, 2$, we get

$$I_V\left(b_1'\right) = b_1' = (1, 3)$$

$$= 1 \cdot (1, 0) + 3\,(0, 1)$$

$$b_1' = b_1 + 3\,b_2$$

$$I_V\left(b_2'\right) = b_2'$$

$$= (2, 5)$$

$$= 2 \cdot (1, 0) + 5\,(0, 1)$$

$$= 2b_1 + 5\,b_2$$

Thus $$P = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$$

***(ii)*** Now to determine the transition matrix $Q$ from the basis $B'$ to the basis $B$, we have

$$b_1 = (1, 0)$$

$$= -5\,(1, 3) + 3\,(2, 5)$$

$$= -5b_1' + 3b_2'.$$

$$b_2 = (0, 1)$$

$$= 2\,(1, 3) - 1\,(2, 5)$$

$$= 2b_2' - b_2'$$

Thus $$Q = \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix}$$

Now $$PQ = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$= I$$

Similarly, $$QP = I$$

Hence $$PQ = I = QP$$

Thus $$Q = P^{-1}.$$

**Theorem 6.** *Two matrices over a field F are similar iff they correspond to the same linear transformation of a vector space V over F to it self, with respect to two different bases.*

**Proof :** Let $V$ be an n-dimensional vector space over field $F$ and $t \in \text{Hom}(V, V)$. Let $I_V$ be the identity map from $V$ to $V$, and $P = M_B^{B'}(I_V)$ be the transition matrix, where $B$ and $B'$ are any two bases of $V$.

Let $A = M_B^B(t)$ and $A' = M_{B'}^{B'}(t)$ are matrices of $t$ relative to bases $B$ and $B'$ respectively.

To prove that $A$ and $A'$ are similar,

now, $$P^{-1}AP = P^{-1}M_B^B(t)P$$
$$= M_{B'}^B(I_V)\, M_B^B(t)\, M_{B'}^{B'}(I_V)$$
$$= M_{B'}^{B'}(t) = A'$$

Thus $$P^{-1}AP = A'$$

$\Rightarrow$ $A$ and $A'$ are similar matrices.

Conversely suppose that $A$ and $A'$ are similar matrices of order $n \times n$ and $B$ be the basis for $V$ such that

$$M_B^B(t) = A, \qquad \text{where} \qquad B = \{b_1, b_2, ..., b_n\}.$$

Then we have to show that $A'$ is matrix of $t$ relative to some new basis of $V$.

$\because$ $A$ and $A'$ are similar matrices, so that there exists an invertible matrix $P$ such that

$$A' = P^{-1}AP \qquad .....(1)$$

Now, we define a new basis $B' = \{b_1', b_2', ..., b_n'\}$ of $V$, so by definition

$$P = M_B^{B'}(I_V)$$

Now, from equation (1), we have

$$A' = M_{B'}^B(I_V)\, M_B^B(t)\, M_B^{B'}(I_V)$$
$$= M_{B'}^{B'}(t).$$

Thus $A'$ is the matrix of $t$ relative to some new basis $B'$ of $V$.

## Self-learning exercise-1

**1.** True/False statements :

***(i)*** Transition matrix from a basis to the same basis is the zero matrix. (T/F)

***(ii)*** Similar matrix are matrices of the same linear transformation (T/F)

***(ii)*** If a matrix $A$ is similar to a matrix $B$, then it is not necessary that $B$ is similar to $A$. (T/F)

**2.** Fill in the blanks :

***(i)*** Two square matrices $A$ and $B$ of order $n$ are similar iff there exists a ..... matrix $P$ order $n$ such that $B$ = ..... .

*(ii)* Being similar is an ..... relation on the set of all $n \times n$ matrices having entries in the same field.

## 11.6 Eigenvalues and eigenvectors of a linear transformation

Let $V$ be finite dimensional vector space over field $F$. A non-zero vector $v \in V$ is called an eigenvector of a linear transformation $t : V \to V$, if there exists $\lambda \in F$ such that

$$t(v) = v\lambda$$

If $v$ is an eigenvector of $t$, the corresponding $\lambda$ is called the eigenvalue of $t$ corresponding to $v$.

**Eigenspace :** The set $S_\lambda$, of all eigenvectors of $t$, with eigenvalue $\lambda$ is called the eigenspace of $\lambda$.

**Theorem 7.** *Let $V$ be a finite dimensional vector space over a field $F$. Then the set of all eigenvectors corresponding to an eigenvalue $\lambda$ of a linear transformation $t : V \to V'$ by adjoining zero vector to it, is a subspace of $V$.*

**Proof :** Let $S_\lambda$ be the set of all eigenvectors of $t$, corresponding to eigenvalue $\lambda$, so that,

$$S_\lambda = \{v \in V : t(v) = v\lambda\}.$$

Since $\lambda$ is an eigenvalue, $S_\lambda$ is clearly a non-empty subset of $V$.

Let $u, v \in S_\lambda \Rightarrow t(u) = u\lambda$ and $t(v) = v\lambda$.

Also let $\alpha, \beta \in F$, then

$$\begin{aligned} t(\alpha u + \beta b) &= \alpha t(u) + \beta t(v) && [\because\ t \text{ is linear}] \\ &= \alpha(u\lambda) + \beta(v\lambda) \\ &= u(\alpha\lambda) + v(\beta\lambda) \\ &= (u\alpha)\lambda + (v\beta)\lambda \\ &= (u\alpha + v\beta)\lambda \\ &= (\alpha u + \beta v)\lambda \end{aligned}$$

$$\Rightarrow \qquad \alpha u + \beta v \in S_\lambda, \qquad \forall\, u, v \in S_\lambda \quad \text{and} \quad \alpha, \beta \in F$$

Thus $S_\lambda \cup \{\mathbf{0}\}$ is a vector subspace of $V$.

### 11.6.1 Eigenvalues and eigenvectors of a matrix

Let $A$ be an $n \times n$ matrix over a field $F$ and let $t_A : F^n \to F^n$ be the linear transformation corresponding to $A$. Then eigenvectors or eigenvalues of matrix $A$ are the eigenvectors or eigenvalues of the linear transformation $t_A : F^n \to F^n$.

**Theorem 8.** *Let $A$ be an $n \times n$ matrix over a field $F$. Then a non-zero vector $X \in F^n$ (or a column vector $X$) iff there exists a scalar $\lambda \in F$ such that*

$$AX = X\lambda$$

**Proof :** First suppose that $\mathbf{0} \neq X \in F^n$ is an eigenvector of matrix $A$. Then $X$ is an eigenvector of linear transformation $t_A : F^n \to F^n$ and so there exists a scalar $\lambda \in F$ such that

$$t_A(X) = X\lambda$$

$$\Rightarrow \quad AX = X\lambda \qquad [\because\ t_A(X) = AX]$$

Conversely suppose that $\mathbf{0} \neq X \in F^n$ such that

$$AX = X\lambda$$

$$\Rightarrow \quad t_A(X) = X\lambda$$

$\Rightarrow$ $X$ is an eigenvector of $t_A$

$\Rightarrow$ $X$ is an eigenvector of $A$.

**Theorem 9.** *Let $V$ be a finite dimensional vector space over a field $F$ and $t : V \to V$ be a linear transformation. Suppose that $v_1, v_2, ..., v_n$ are distinct eigenvectors of $t$ corresponding to distinct eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$.*

*Then $\{v_1, v_2, ..., v_n\}$ is a linearly independent set.*

**Proof :** Given that $v_1, v_2, ..., v_n$ be the $n$ distinct eigenvectors of $t$ corresponding to the $n$ distinct eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$ respectively.

Then, we have $t(v_i) = v_i \lambda_i, \qquad i = 1, 2, ..., n.$

In order to prove that $v_1, v_2, ..., v_n$ are linearly independent, we shall use induction on $n$.

The result is clearly true for $n = 1$, because $v_1 \neq \mathbf{0}$. Let us assume that $v_1, v_2, ..., v_k$, where $k < n$, are linearly independent.

Now consider the vectors $v_1, v_2, ..., v_{k+1} \in F$. Let there exist scalars $\alpha_1, \alpha_2, ..., \alpha_{k+1} \in F$ such that

$$\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_k v_k + \alpha_{k+1} v_{k+1} = \mathbf{0} \qquad .....(1)$$

$$\Rightarrow \quad t(\alpha_1 v_1 + ... + \alpha_k v_k + \alpha_{k+1} v_{k+1}) = t(0) = \mathbf{0}$$

$$\Rightarrow \quad t(v_1)\alpha_1 + ... + t(v_k)\alpha_k + ... + t(v_{k+1})\alpha_{k+1} = \mathbf{0}$$

$$\Rightarrow \quad (v_1\lambda_1)\alpha_1 + ... + (v_k\lambda_k)\alpha_k + ... + (v_{k+1}\lambda_{k+1})\alpha_{k+1} = \mathbf{0} \qquad .....(2)$$

Multiplying equation (1) by $\lambda_{k+1}$ on right and then substracting it from equation (2), we have

$$v_1(\lambda_1 - \lambda_{k+1})\alpha_1 + ... + v_k(\lambda_k - \lambda_{k+1})\alpha_k = \mathbf{0}$$

$$\Rightarrow \quad (\lambda_1 - \lambda_{k+1})\alpha_1 = ... = (\lambda_k - \lambda_{k+1})\alpha_k = 0 \qquad [\because\ v_1, ..., v_k \text{ are linearly independent}]$$

$$\Rightarrow \quad \alpha_1 = \alpha_2 = ... = \alpha_k = 0 \qquad [\because\ \lambda_i \neq \lambda_{k+1},\ i = 1, 2, ..., k]$$

Using these in equation (1), we get

$$v_{k+1}\lambda_{k+1} = \mathbf{0}$$

$$\Rightarrow \quad \lambda_{k+1} = 0 \qquad [\because\ v_{k+1} \neq \mathbf{0}]$$

Thus, $v_1\alpha_1 + \ldots + v_k\alpha_k + v_{k+1}\,\alpha_{k+1} = \mathbf{0}$

$\Rightarrow \quad \alpha_1 = \ldots = \alpha_k = \alpha_{k+1} = 0$

$\Rightarrow \quad v_1, v_2, \ldots, v_k, v_{k+1}$ are linearly independent.

Hence by induction $\{v_1, v_2, \ldots, v_n\}$ is linearly independent set.

**Theorem 10.** *Let $A$ be an $n \times n$ matrix over a field $F$, has $n$ distinct eigenvalues $\lambda_i$, $i = 1, 2, \ldots, n$, then there exists an invertible matrix $P$ such that*

$$P^{-1}AP = diag.\ (\lambda_1, \lambda_2, \ldots, \lambda_n)$$

**Proof :** Suppose that $X_i \in F^n$, $i = 1, 2, \ldots, n$ be the $n$ distinct eigenvectors of $A$ corresponding to $n$ distinct eigenvalues $\lambda_i$, $i = 1, 2, \ldots, n$ respectively. Then we have

$$AX_i = X_i\,\lambda_i, \qquad i = 1, 2, \ldots, n.$$

Suppose $P = [X_1, X_2, \ldots, X_n]$. Then it is an $n \times n$ matrix over $F$. Since $\{X_1, X_2, \ldots, X_n\}$ is linearly independent set so rank $(P) = n$, and $P$ is an invertible matrix.

Now by definition of matrix multiplication, $i^{\text{th}}$ column of $AP = A$ ($i^{\text{th}}$ column of $P$)

$$= AX_i, \qquad i = 1, 2, \ldots, n$$

Thus
$$AP = [AX_1, AX_2, \ldots, AX_n]$$
$$= [X_1\lambda_1, X_2\lambda_2, \ldots, X_n\lambda_n]$$

$$= [X_1, X_2, \ldots, X_n] \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & & & \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}$$

$$= P \text{ diag } (\lambda_1, \lambda_2, \ldots, \lambda_n)$$

Thus $P^{-1}AP = \text{diag } (\lambda_1, \lambda_2, \ldots, \lambda_n)$.

**Theorem 11.** *Let $V$ be a finite dimensional vector space over a field $F$ and $t : V \to V$ be a linear transformation. Then*

***(i)*** *The matrix $A$ of $t$ is a diagonal matrix having the eigenvalues of $t$ as diagonal entries iff $A$ is corresponding to a basis of $V$ consisting of eigenvectors of linear transformation $t$.*

***(ii)*** *The eigenvalues of $t$ are exactly the diagonal entries of $A$ and each appearing on the diagonal as many times as the dimension of its eigenspace.*

**Proof :** ***(i)*** Let $B = \{b_1, b_2, \ldots, b_n\}$ be a basis of $V$, consisting of eigenvectors of $t : V \to V$.

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of $t$ corresponding to the eigenvectors $b_1, b_2, \ldots, b_n$ respectively. Then

$$t\,(b_j) = b_j\,\lambda_j, \qquad j = 1, 2, \ldots, n.$$

Let $B^* = \{f_1, f_2, \ldots, f_n\}$ be the basis dual to $B$.

Now, for all $i, j = 1, 2, \ldots, n$, we have

$$A = \left[ M_B^B (t) \right] = f_i (t (b_j))$$
$$= f_i (b_j \lambda_j)$$
$$= f_i (b_j) \lambda_j$$
$$= \delta_{ij} \lambda_j$$

Thus
$$A = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & & \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

which is a diagonal matrix of $t$ having diagonal elements as eigenvalues of $t$.

Conversely suppose that matrix $A = \left[ a_{ij} \right]_{n \times n}$ having diagonal elements $\lambda_i$, $i = 1, 2, ..., n$, is a diagonal matrix, of $t$ relative to basis $B = \{b_1, b_2, ..., b_n\}$ of $V$.

We have $a_{ii} = \lambda_i, \qquad i = 1, 2, ..., n$

and $a_{ij} = 0, \quad \text{if} \quad i \neq j, \qquad i, j = 1, 2, ..., n$

and $M_B^B (t) = A$

Now, $t (b_j) = \sum_{s=1}^{n} b_s \, a_{sj}, \qquad j = 1, 2, ..., n$
$$= b_1 a_{1j} + b_2 a_{2j} + ... + b_j a_{jj} + ... + b_n a_{nj}$$
$$= 0 + 0 + ... + b_j a_{jj} + ... + 0$$
$$= b_j \, a_{jj}$$
$$= b_{ij} \lambda_j$$

Thus $b_j, j = 1, 2, ..., n$ are eigenvectors of $t$ corresponding to the eigenvalues $\lambda_j, j = 1, 2, ..., n$ respectively.

***(ii)*** Assume that a non-zero vector $v \in V$ is an eigenvector of $t$ with eigenvalue $\lambda$, then we have

$$t (v) = v\lambda$$

where $v = \alpha_1 b_1 + ... + \alpha_n b_n, \quad \text{for} \quad \alpha_1, ..., \alpha_n \in F.$

So $$t\left( \sum_{i=1}^{n} b_i \, \alpha_i \right) = \left( \sum_{i=1}^{n} b_i \, \alpha_i \right) \lambda$$

$\Rightarrow$ $$\sum_{i=1}^{n} t(b_i) \alpha_i = \sum_{i=1}^{n} b_i (\alpha_i \lambda)$$

$\Rightarrow$ $$\sum_{i=1}^{n} (b_i \, \lambda_i) \alpha_i = \sum_{i=1}^{n} b_i (\alpha_i \, \lambda)$$

$$\Rightarrow \qquad \sum_{i=1}^{n} b_i (\lambda_i - \lambda)\alpha_i = \mathbf{0}$$

$$\Rightarrow \qquad (\lambda_i - \lambda)\,\alpha_i = 0, \qquad i = 1, 2, ..., n$$

[$\because$ $B = \{b_1, b_2, ..., b_n\}$ is a basis for $V$]

But $\alpha_i \neq 0$, $i = 1, 2, ..., n$, because $v \neq \mathbf{0}$.

Thus $\lambda = \lambda_i$

Hence any eigenvalue $\lambda$ of $t$ must be one of the diagonal elements of $A$.

Let us assume that an eigenvalue of $t$ appears $m$ times on the diagonal of matrix $A$. Let us suppose that first $m$ entries on the diagonal are same, that is, $\lambda_1 = \lambda_2 = ... = \lambda_m = \lambda$.

Suppose that $S$ be the subspace of $V$ spans by the corresponding eigenvectors $b_i$, $i = 1, 2,..., m$.

Thus $$S = \left\{ \sum_{i=1}^{m} b_i\, \alpha_i : \alpha_i \in F \right\}$$

Now $\{b_1, b_2, ..., b_m\}$ is a subset of $B$ and so it is linearly independent and thus dim $S = m$.

We have to show that $S = S_\lambda$, where $S_\lambda = \{v \in V : t(v) = v\lambda\}$ be the eigenspace of $\lambda$.

Let $u \in S \Rightarrow$ $$u = \sum_{i=1}^{m} b_i\, \mu_i, \qquad \mu_i \in F$$

$$\Rightarrow \qquad t(u) = \left( t \sum_{i=1}^{m} b_i\, \mu_i \right)$$

$$= \sum_{i=1}^{m} t(b_i)\, \mu_i$$

$$= \sum_{i=1}^{m} (b_i\, \lambda)\, \mu_i$$

$$= \left( \sum_{i=1}^{m} (b_i\, \mu_i) \right) \lambda$$

Hence $u$ is an eigenvector of $t$ corresponding to eigenvalue $\lambda$ and $u \in S_\lambda$.

Thus $S \subseteq S_\lambda$ .....(1)

Now suppose that $u \in S_\lambda \Rightarrow t(u) = u\lambda$

But $u \in S_\lambda \Rightarrow u \in V$

$$\Rightarrow u = \sum_{i=1}^{n} b_i\, \alpha_i, \qquad \alpha_i \in F$$

Thus, $$t\left( \sum_{i=1}^{n} b_i\, \alpha_i \right) = \left( \sum_{i=1}^{n} b_i\, \alpha_i \right) \lambda$$

$$\Rightarrow \quad \sum_{i=1}^{n} t(b_i)\alpha_i = \sum_{i=1}^{n} (b_i\ \alpha_i)\lambda$$

$$\Rightarrow \quad \sum_{i=1}^{n} (b_i\ \lambda_i)\alpha_i = \sum_{i=1}^{n} (b_i\ \alpha_i)\lambda$$

$$\Rightarrow \quad \sum_{i=1}^{n} b_i(\lambda_i - \lambda)\alpha_i = \mathbf{0}$$

$(\lambda_i - \lambda)\ \alpha_i = 0, \qquad i = 1, 2, ..., n \quad [\because\ B = \{b_1, b_2, ..., b_n\}$ is a basis for $V]$

But $\lambda_i \neq \lambda, \quad$ if $\quad i = m + 1, m + 2, ..., n,$

so that $\alpha_i = 0, \quad$ if $\quad i = m + 1, m + 2, ..., n,$

Thus $u = \sum_{i=1}^{m} b_i\ \alpha_i \in S$

Therefore $S_\lambda \subseteq S$ .....(2)

From equations (1) and (2), we have

$$S = S_\lambda \ \Rightarrow\ dim\ S_\lambda = m,$$

which shows that an eigenvalue $\lambda$ of $t$ appears on the diagonal of $A$, as may times as the dimension of its eigenspace.

## 11.7 Summary

In this unit we have studied a general over view of the rank, nullity of linear maps as well as of matrices. We have also studied eigenvalues and eigenvectors of linear maps and of corresponding matrices. Several examples have been solved in this unit to understand the concept introduced, and we have gone through a variety of problems in exercises to grasp the concept of the topic.

## 11.8 Answers to self-learning exercises

### Self-learning exercise-1

**1.** ***(i)*** F ***(ii)*** T ***(iii)*** F

**2.** ***(i)*** non-singular, $P^{-1}AP$ ***(ii)*** equivalence.

## 11.9 Exercises

**1.** Let $A$ be an $n \times n$ matrix over field $F$, and has $n$ distinct eigenvalues in $F$ then $A$ is similar to a diagonal matrix.

**2.** If $A$ is an $n \times n$ invertible matrix. Then prove that rank $(A) = n$.

**3.** Prove that any two similar matrices have the same column rank and the same row rank.

**4.** The matrix $A$ of linear transformation $t : V \to V$ is diagonal if $A$ is relative to a basis of $V$ consisting of eigenvectors of $t$.

□ □ □

# UNIT 12 : Determinants of Matrices

**Structure of the Unit**



## 12.0 Objectives

After reading this unit you will be able to learn the importance of determinants of matrices and their applications, such as, in the solution of system of linear equations and many other in linear algebra. you will also see that determinant is a function and it is a multilinear alternative form through which various properties of determinant of a matrix will be explained.

## 12.1 Introduction

In previous unit we have been discussing linear transformations and their matrices. This unit introduces the concept of determinant of a matrix and their properties. In this unit we shall study determinant function as a multilinear alternating form, existence and uniqueness of determinants, Cramer's rule, characteristic polynomial and eigen values, and Cayley-Hamilton theorem.

## 12.2 Multilinear and alternating function

**Multilinear function :**

A function $t : V_1 \times V_2 \times ... \times V_n \to S$, is called $n$-multilinear if it is linear in each factor, when the other entries are kept constant, *i.e.* for all, $i = 1, 2, ..., n$

$$t(v_1, v_2, ..., v_{i-1}, \alpha v_i + \beta v_i', v_{i+1}, ..., v_n)$$

$= \alpha t\,(v_1, v_2, ..., v_{i-1}, v_i, v_{i+1}, ..., v_n)$

$= \beta t\,(v_1, v_2, ..., v_{i-1}, v_i', v_{i+1}, ..., v_n)$

$\forall\, v_i, v_i' \in V_i$ and $\alpha, \beta \in F$.

Where $V_1, V_2, ..., V_n, S$ are vector spaces over the same field $F$.

If $V_i = V$, $i = 1, 2, ..., n$, then $t$ is called an n-multilinear function on $V$ and if in addition $S = F$, then $t$ is called an $n$-multilinear form on $V$.

An $n$-multilinear function $t$ on $V$ is said to be an **alternating function** if $t(v_1, v_2, ..., v_n) = \mathbf{0}$ when $v_i = v_{i+1}$ for $i = 1, 2, ..., n-1$. The multilinear function $t$ on $V$ is said to be **symmetric** if interchanging $v_i$ and $v_j$ for any $i$ and $j$, the value of $t(v_1, v_2, ..., v_n)$ does not change.

**Alternating form :**

An $n$-multilinear form $t$ on $V$ is said to be an alternating form if

$t(v_1, v_2, ..., v_n) = 0$ when $v_i = v_{i+1}$ for $i = 1, 2, ..., n-1$.

## 12.3 Determinant function

An $n \times n$ determinant function, is a mapping,

$$\det : M_{n \times n}(F) \to F \text{ such that}$$

***(i)*** the det is an $n$-multilinear alternating form on $F^n$, and

***(ii)*** $\det(I) = 1$, where $I$ is the identity matrix of order $n$.

The above mapping is infact, $\det : F^n \times F^n \times ... \times F^n \to F$ with two axioms. We also write

$$\det(A) = \det(A_1, A_2, ..., A_n),$$

where $A_1, A_2, ..., A_n$ are the $n$ columns of square matrix $A$.

## 12.4 Existence and uniqueness of determinants

**Theorem 1.** (*Existence*). *There exists a multilinear function det :* $(F^n)^n \to F$ *such that*

$$\begin{aligned} det\,(A) &= det\,(A_1, A_2, ..., A_n) \\ &= \sum_{\rho \in S_n} \in (\rho)\, a_{\rho(1)1}\, a_{\rho(2)2} \cdots a_{\rho(n)n}, \quad \forall\, A_i \in F^n, \end{aligned}$$

*satisfying the axioms of determinant function.*

**Proof :** Suppose that $A_1, A_2, ..., A_n$ be the $n$ column vector in a $n \times n$ matrix $A = [a_{ij}]$. We shall prove that the above multilinear function satisfy the axioms of a determinant function, that is, det is an alternating form with $\det(I) = 1$, where $I$ is an identity matrix of order $n$.

Given that $\det : (F^n)^n \to F$ such that

$$\begin{aligned} \det(A) &= \det(A_1, A_2, ..., A_n) \\ &= \sum_{\rho \in S_n} \in (\rho)\, a_{\rho(1)1}\, a_{\rho(2)2} \cdots a_{\rho(n)n}, \end{aligned} \qquad .....(1)$$

If $A = I = (e_1, e_2, ..., e_n)$, then each term in equation (1) is zero except the term corresponding to the identity permutation *i.e.*

$$\begin{aligned}\det(I) &= \det(e_1, e_2, ..., e_n)\\ &= a_{11}\, a_{22} \ldots a_{nn}\\ &= 1 \cdot 1 \cdot \ldots \cdot 1 = 1\end{aligned}$$

Thus $$\det(I) = 1 \qquad .....(2)$$

Now to prove that det is an alternating form, let $A_j = A_{j+1}$. Then we shall prove that

$$\det(A_1, A_2, ..., A_n) = 0$$

Since, $$A_j = A_{j+1}$$

$\Rightarrow$ $$a_{ij} = a_{i(j+1)}, \qquad i = 1, 2, ...., n$$

Let $\rho \in S_n$ be any odd permutation, then from equation (1),

$$\det(A) = (-1)\, a_{\rho(1)1}\, a_{\rho(2)2} \ldots a_{\rho(n)n} \qquad .....(3)$$

Suppose that $h = (j, j+1)$ be a transposition, then $\rho = \sigma h$, for some even permutation $\sigma \in S_n$, so that

$$\rho(j) = (\sigma h)(j) = \sigma(h(j)) = \sigma(j+1) \qquad .....(4a)$$

$$\rho(j+1) = (\sigma h)(j+1) = \sigma(h(j+1)) = \sigma(j) \qquad .....(4b)$$

$$\rho(r) = (\sigma h)(r) = \sigma(h(r)) = \sigma(r) \qquad \forall\, r \neq j, j+1 \qquad .....(4c)$$

For even permutation $\sigma$, we have from equation (1),

$$\begin{aligned}\det(A) &= (+1)\, a_{\sigma(1)1}\, a_{\sigma(2)2} \ldots a_{\sigma(n)n}\\ &= a_{\sigma(1)1}\, a_{\sigma(2)2} \ldots a_{\sigma(j)j}\, a_{\sigma(j+1)(j+1)} \ldots a_{\sigma(n)n}\\ &= a_{\sigma(1)1}\, a_{\sigma(2)2} \ldots a_{\sigma(j)(j+1)} a_{\sigma(j+1)j} \ldots a_{\sigma(n)n} \qquad \left[\because a_{ij} = a_{i(j+1)}\right]\\ &= a_{\rho(1)1}\, a_{\rho(2)2} \ldots a_{\rho(j+1)(j+1)}\, a_{\rho(j)j} \ldots a_{\rho(n)n}\end{aligned}$$

$[\because$ $F$ is commutative$]$

$\therefore$ $$\det(A) = a_{\rho(1)1}\, a_{\rho(2)2} \ldots a_{\rho(j)j}\, a_{\rho(j+1)j+1} \ldots a_{\rho(n)n} \qquad .....(5)$$

From equations (3) and (5), we have

$$\det(A) = \det(A_1, A_2, ..., A_n) = 0$$

when $$A_j = A_{j+1}.$$

**Theorem 2.** (*Uniqueness*). *Let det and det′ be two determinant functions, then for all column vectors* $A_1, A_2, ..., A_n, \in F^n$,

$$det(A_1, A_2, ..., A_n) = det'(A_1, A_2, ..., A_n).$$

**Proof :** We define a function $\Delta$ such that

$$\Delta(A_1, A_2, ..., A_n) = \det(A_1, A_2, ..., A_n) - \det'(A_1, A_2, ..., A_n)$$

Since both det and det′ satisfy axioms of determinant function, $\Delta$ is a multilinear function with

$$\Delta(e_1, e_2, ..., e_n) = 0 \text{ and } \Delta(A_1, A_2, ..., A_n) = 0 \text{ if any } A_i = A_j \text{ for } i \neq j$$

and $\Delta(A_1, A_2, ..., A_n)$ changes sign if any two $A_i$ and $A_j$ are interchanged. Here $\{e_1, e_2, ..., e_n\}$ be the standard ordered basis of $F^n$. Now, it is sufficient to show that

$$\Delta(A_1, A_2, ..., A_n) = 0$$

Now, for $\alpha_{ij} \in F$, we have

$$A_j = \alpha_{1j} e_1 + \alpha_{1j} e_2 + ... + \alpha_{nj} e_n$$

$$= \sum_{i=1}^{n} \alpha_{rj} e_r$$

Now, $$\Delta(A_1, A_2, ..., A_n) = \Delta\left(\sum_{r=1}^{n} \alpha_{r_1} e_r, A_2 ..., A_n\right)$$

Now using linearity to the first factor, then to second, etc, we get

$$\Delta(A_1, A_2, ..., A_n) = \sum_{r=1}^{n} \alpha_{r1} \Delta\left(e_r, A_2 ..., A_n\right)$$

$$= \sum_{r=1}^{n} \alpha_{r_1} \Delta\left(e_r, \sum_{r_1^2=1}^{n} \alpha_{r_1^2} e_{r_1}, A_3 ..., A_n\right)$$

$$= \sum_{r=1}^{n} \sum_{r_1=1}^{n} \alpha_{r_1} \alpha_{r_1^2} \Delta\left(e_r, e_{r_1}, A_3 ..., A_n\right)$$

$$\vdots$$

$$= \sum_{r=1}^{n} \sum_{r_1=1}^{n} .... \sum_{r_{n-1}=1}^{n} \alpha_{r_1} \alpha_{r_1^2} ... \alpha_{r_{n-1}n} \Delta\left(e_r, e_{r_1}, ..., e_{r_{n-1}}\right)$$

$$= \sum_{r, r...r_{n-1}=1}^{n} \alpha_{r_1} \alpha_{r_1^2} \alpha_{r_2^3} ... \alpha_{r_{n-1}^n} \Delta\left(e_r, e_{r_1}, e_{r_2}, ..., e_{r_{n-1}}\right)$$

$$= 0 \qquad \left[\because \Delta\left(e_r, e_{r_1}, e_{r_2}, ..., e_{r_{n-1}}\right) = 0\right]$$

Thus $$\Delta(A_1, A_2, ..., A_n) = 0.$$

Hence the uniqueness theorem is proved.

**Determinant of a matrix :**

Let $A$ be an $n \times n$ matrix. The determinant of matrix $A$ is denoted by det $(A)$ or $|A|$ and is defined as follows :

$$\det(A) = |A| = \sum_{\rho \in S_n} \in(\rho)\, a_{\rho(1)1}\, a_{\rho(2)2} ... a_{\rho(n)n}$$

## 12.5 Properties of determinant function

Let $A$ be a matrix of order $n \times n$, then

*(a)* If any column (row) of $A$ is zero, then $|A| = 0$,

*(b)* If $\lambda$ is a scalar,, then $|\lambda A| = \lambda^n |A|$

*(a)* If $A$ is a diagonal matrix, then $|A| = a_{11}\, a_{22} \ldots a_{nn}$.

**Theorem 3.** *Let $A$ be a matrix of order $n \times n$ and let $\phi : n \to n$, then*

***(a)*** $\sum_{\rho \in S_n} \in (\rho)\, a_{\rho(1)\phi(1)}\, a_{\rho(2)\phi(2)} \ldots a_{\rho(n)\phi(n)} = \in (\phi) |A|$

***(b)*** $\sum_{\rho \in S_n} \in (\rho)\, a_{\phi(1)\rho(1)}\, a_{\phi(2)\rho(2)} \ldots a_{\phi(n)\rho(n)} = \in (\phi) |A|$.

**Proof :** There are two cases :

**Case (I)** when $\phi$ is not a permutation *i.e.* $\phi \notin S_n$

Then $\in (\phi) = 0$

So that, R.H.S. $= 0$

To prove that L.H.S. $= 0$

Since $\phi$ is not a permutation so that there exist $i, j \in \{1, 2, ..., n\}$ such that

$$i \neq j \quad \text{but} \quad \phi(i) = \phi(j) \qquad .....(1)$$

Let $h = (i, j) \in S_n$ be the transposition.

For any $\rho \in S_n$, suppose that $\sigma = \rho h$, then

$$\sigma(i) = (\rho h)(i) = \rho(h(i)) = \rho(j) \qquad .....(2a)$$

$$\sigma(j) = (\rho h)(j) = \rho(h(j) = \rho(i) \qquad .....(2b)$$

and $$\sigma(s) = (\rho h)(s) = \rho(h(s)) = \rho(s),\ s \in \{1, 2, ..., n\},\ s \neq i, j \qquad .....(2c)$$

Since $h$ is an odd permutation, so that

$$\in (\sigma) = \in (\rho h) = \in (\rho) \in (h) = - \in (\rho) \qquad .....(3)$$

Now, the contribution of the permutations $\rho$ and $\sigma$ in the L.H.S. of $(a)$ is

$$\in (\rho)\, a_{\rho(1)\phi(1)}\, a_{\rho(2)\phi(2)} \ldots a_{\rho(i)\phi(i)} \ldots a_{\rho(j)\phi(j)} \ldots a_{\rho(n)\phi(n)}$$

$$+ \in (\sigma)\, a_{\sigma(1)\phi(1)}\, a_{\sigma(2)\phi(2)} \ldots a_{\sigma(i)\phi(i)} \ldots a_{\sigma(j)\phi(j)} \ldots a_{\sigma(n)\phi(n)}$$

$$= \in (\rho) \Big\{ a_{\rho(1)\phi(1)}\, a_{\rho(2)\phi(2)} \ldots a_{\rho(i)\phi(i)} \ldots a_{\rho(j)\phi(j)} \ldots a_{\rho(n)\phi(n)}$$

$$- a_{\rho(1)\phi(1)}\, a_{\rho(2)\phi(2)} \ldots a_{\rho(j)\phi(i)} \ldots a_{\rho(i)\phi(j)} \ldots a_{\rho(n)\phi(n)} \Big\}$$

$$= \in (\rho) \Big\{ a_{\rho(1)\phi(1)}\, a_{\rho(2)\phi(2)} \ldots a_{\rho(i)\phi(i)} \ldots a_{\rho(j)\phi(j)} \ldots a_{\rho(n)\phi(n)}$$

$$- a_{\rho(1)\phi(1)}\, a_{\rho(2)\phi(2)} \ldots a_{\rho(i)\phi(i)} \ldots a_{\rho(j)\phi(j)} \ldots a_{\rho(n)\phi(n)} \Big\}.$$

$$\left[\because \quad \text{From (1)}, i \neq j \Rightarrow \phi(i) = \phi(j)\right]$$

$= \in (\rho)\ (0)$

$= 0$

Hence, $\sum_{\rho \in S_n} \in (\sigma)\ a_{\rho(1)\phi(1)}\ a_{\rho(2)\phi(2)} \ldots\ a_{\rho(n)\phi(n)} = 0$

**Case (II)** When $\phi$ is a permutation *i.e.* $\phi \in S_n \Rightarrow \phi^{-1} \in S_n$.

For any permutation $\rho \in S_n$, Let $\sigma = \rho\phi^{-1}$

Then $\in (\phi) = 0$

$\Rightarrow$ $\rho = \sigma\phi$

$\Rightarrow$ $\rho(i) = \sigma(\phi(i)), \qquad i = 1, 2, \ldots, n$

Now, $\sum_{\rho \in S_n} \in (\rho)\, a_{\rho(1)\phi(1)}\ a_{\rho(2)\phi(2)} \ldots a_{\rho(n)\phi(n)}$

$$= \sum_{\sigma \in S_n} \in (\sigma\phi)\, a_{\sigma(\phi(1))\phi(1)}\ a_{\sigma(\phi(2))\phi(2)} \ldots a_{\sigma(\phi(n))\phi(n)}$$

$$= \sum_{\sigma \in S_n} \in (\sigma) \in (\sigma\phi)\, a_{\sigma(1)1}\ a_{\sigma(2)2} \ldots a_{\sigma(n)n}$$

[on changing the order of factors]

$$= \in (\sigma) \left[ \sum_{\sigma \in S_n} a_{\sigma(1)1}\ a_{\sigma(2)2} \ldots a_{\sigma(n)n} \right]$$

$= \in (\phi)\ |A|$.

Similarly part (*b*) can also be proved.

**Theorem 4.** *Let $A$ be a matrix of order $n \times n$ and $A^T$ is the transpose of $A$. Then*

$|A^T| = |A|$, *where* $|A|$ *denotes the determinant of* $A$.

**Proof :** Let $A = [a_{ij}]$ be an $n \times n$ order matrix, then we have

$$|A| = \sum_{\rho \in S_n} \in (\rho)\, a_{\rho(1)1}\ a_{\rho(2)2} \ldots a_{\rho(n)n}$$

So that, $$|A^T| = \sum_{\rho \in S_n} \in (\rho)\, a_{1\rho(1)}\ a_{2\rho(2)2} \ldots a_{n\rho(n)}$$

$$= \sum_{\rho \in S_n} \in (\rho)\, a_{\phi(1)\rho(1)}\ a_{\phi(2)\rho(2)} \ldots a_{\phi(n)\rho(n)}$$

where $\phi : n \to n$ is the identity permutation.

Thus $|A^T| = \in (\phi)\ |A|$ [By theorem (3)]

$= 1 \cdot |A|$ [$\because$ $\phi$ is an even permutation]

$= |A|$.

**Theorem 5.** *If any two rows (columns) of a determinant A are equal, then*

$$|A| = 0$$

**Proof :** Suppose $A$ be an $n \times n$ matrix and its $r^{th}$ and $s^{th}$ rows are same, that is

$$A^r = A^s$$

$$\Rightarrow \qquad a_{rj} = a_{sj}, \qquad j = 1, 2,..., n$$

Now,
$$|A| = \sum_{\rho \in S_n} \in (\rho) a_{\rho(1)1}\, a_{\rho(2)2} \ldots a_{\rho(r)r} \ldots a_{\rho(s)s} \ldots a_{\rho(n)n}$$

Let $\phi : n \to n$ such that $\phi(i) = i$, $i \neq r, s$ and $\phi(r) = \phi(s)$,

then
$$|A| = \sum_{\rho \in S_n} \in (\rho) a_{\rho(1)\phi(1)}\, a_{\rho(2)\phi(2)} \ldots a_{\rho(r)\phi(r)} \ldots a_{\rho(s)\phi(s)} \ldots a_{\rho(n)\phi(n)}$$

$$= \in (\phi)\, |A|$$

$$= 0\, |A| \qquad [\because\ \phi \text{ is not a permutation, } \therefore \in (\phi) = 0]$$

$$|A| = 0.$$

**Theorem 6.** *Let A and B be any two matrices of n × n order, then*

$$|AB| = |A| \cdot |B|.$$

**Proof :** Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be two matrices of order $n \times n$ over the same field $F$, then $AB$ is a matrix of order $n \times n$ in $F$, and

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik}\, b_{kj}$$

By definition of determinant function, we have

$$|AB| = \sum_{\rho \in S_n} \in (\rho) (AB)_{\rho(1)1}\, (AB)_{\rho(2)2} \ldots (AB)_{\rho(n)n}$$

$$= \sum_{\rho \in S_n} \in (\rho) \left( \sum_{k=1}^{n} A_{\rho(1)k}\ B_{k1} \right) \ldots \left( \sum_{k=1}^{n} A_{\rho(n)k}\ B_{kn} \right)$$

In these $n$ summations, the index $k$ be replaced respectively by $\phi(1)$, $\phi(2)$, .... $\phi(n)$, and opening the brackets,

$$|AB| = \sum_{\rho \in S_n} \in (\rho) \sum_{\phi(1)=1}^{n} \ldots\ldots \sum_{\phi(n)=1}^{n} a_{\rho(1)\phi(1)}\, b_{\phi(1)1} \ldots\ldots a_{\rho(n)\phi(n)}\, b_{\phi(n)n}$$

$$= \sum_{\rho \in S_n} \in (\rho) \sum_{\phi(1)=1}^{n} \ldots\ldots \sum_{\phi(n)=1}^{n} a_{\rho(1)\phi(1)} \ldots a_{\rho(n)\,\phi(n)}\, b_{\phi(1)1} \ldots b_{\phi(n)n}$$

$$[\because\ F \text{ is commutative}]$$

$$= \sum_{\rho \in \underline{n}^{\underline{n}}} b_{\phi(1)1} \ldots b_{\phi(n)n} \sum_{\rho \in S_n} \in (\rho)\, a_{\rho(1)\phi(1)} \ldots a_{\rho(n)\phi(n)}$$

$$= \sum_{\rho \in \underline{n}^{\underline{n}}} b_{\phi(1)1} \ldots b_{\phi(n)n} \in (\phi) \mid A \mid .$$

$$= \mid A \mid \sum_{\phi \in \underline{n}^{\underline{n}}} \in (\phi) b_{\phi(1)1} \ldots b_{\phi(n)n}$$

$$= \mid A \mid \sum_{\phi \in S_n} \in (\phi) b_{\phi(1)1} \ldots b_{\phi(n)n}$$

[$\in (\phi) = 0$, when $\phi : \underline{n} \to \underline{n}$ is not a permutation]

$$\mid AB \mid = \mid A \mid \cdot \mid B \mid$$

**Corollary :** An $n \times n$ square matrix $A$ over a field $F$ is invertible iff det $(A) \neq 0$ and

$$\det\left(A^{-1}\right) = \frac{1}{\det(A)} .$$

**Proof** : First suppose that $n \times n$ square matrix $A$ is invertible (non-singular), so that $A^{-1}$ exists and

$$AA^{-1} = I$$

$\Rightarrow$ $\det (AA^{-1}) = \det (I)$

$\Rightarrow$ $\det (A) \det (A^{-1}) = 1$

Thus $\det (A) \neq 0$ and $\det (A^{-1}) = \dfrac{1}{\det(A)}$

Conversely suppose that $\det (A) \neq 0$. To prove that $A$ is invertible. To do this it is sufficient to prove that $\rho (A) = n$ *i.e.* rank $\rho (A) = n$. As contradiction suppose that rank $(A) \neq n$

*i.e.* $\rho(A) \neq n$

$\Rightarrow$ $\rho(A) < n$

$\Rightarrow$ $A_1, A_2, \ldots, A_n$ are linearly dependent columns in $F^n$.

$\Rightarrow$ at least one column say $A_i$ out of these $n$ columns is either zero or linear combination of the previous vector.

**Case (I)** When $A_i = 0$, clearly $\det (A) = 0$ which is a contradiction.

**Case (II)** When $A_i$ is the linear combination of $A_1, A_2, \ldots, A_{i-1}$, so that these exist scalars $\alpha_1, \alpha_2, \ldots, \alpha_{i-1} \in F$, such that

$$A_i = \sum_{k=1}^{i-1} A_k \alpha_r$$

$\Rightarrow$ $$\det(A) = \det\left(A_1, A_2, \ldots A_{i-1}, A_i, \ldots, A_n\right)$$

$$= \det\left(A_1, A_2, \ldots A_{i-1}, \sum_{k=1}^{i-1} A_k \alpha_k, \ldots, A_n\right)$$

$$= \sum_{k=1}^{i-1} \det\left(A_1, A_2, \ldots A_{i-1}, A_k, \ldots, A_n\right)$$

$$\det(A) = 0$$

which is again a contradiction of $\det(A) \neq 0$

Thus the columns of $A$ in $F^n$ are linearly independent

Hence, $\rho(A) = n$

$\Rightarrow$ $A$ is invertible.

**Theorem 7.** *Let $A = (A_1, A_2, \ldots, A_n)$ be an $n \times n$ square matrix over a field F, where F is the $i^{th}$ column of A. Then*

*(i)* $\det(A_1, A_2, \ldots, A_i, \ldots, A_j, \ldots, A_n) = 0$, if $A_i = 0$ for some $i$.

*(ii)* $\det(A_1, \ldots A_{i-1}, A_i + \lambda A_j, \ldots, A_j \ldots, A_n)$

$= \det(A_1, \ldots, A_i, \ldots, A_j, \ldots, A_n)$

*(iii)* $\det(A_1, A_2, \ldots, A_n) = 0$, if the set $\{A_1, A_2, \ldots, A_n\}$ is linearly dependent.

*(iv)* $\det(A\alpha) = \alpha^n \det(A)$, for $\alpha \in F$.

*(v)* multiplying one column of $A$ by a scalar $\alpha$, $\det(A)$ multiplies by $\alpha$.

**Proof :** *(i)* Let $\alpha \in F$, then

$$\det = (A_1, A_2, \ldots, \alpha A_i, \ldots, A_n)$$

$$= \alpha \det(A_1, A_2, \ldots, A_i, \ldots, A_n)$$

$$= 0, \quad \text{by taking } \alpha = 0$$

*(ii)* This part follows easily by multilinearity and alternating law.

*(iii)* Given that $\{A_1, \ldots, A_{i-1}, A_i, \ldots, A_n\}$ is linearity dependent, so that some one $A_i$ can be written as a linear combination of the preceding vectors *i.e.* there exist scalars

$\alpha_1, \alpha_2, \ldots, \alpha_{i-1} \in F$, such that

$$A_i = \sum_{k=1}^{i-1} A_k \alpha_k$$

Now, $\det(A_1, A_2, \ldots, A_{i-1}, A_i, \ldots, A_n)$

$$= \det\left(A_1, A_2, \ldots, A_{i-1}, \sum_{k=1}^{i-1} A_k \alpha_k, \ldots, A_n\right)$$

$$= \sum_{k=1}^{i-1} \det\left(A_1, \ldots, A_{i-1}, A_k, \ldots, A_n\right)\alpha_k$$

$= 0$ [$\because$ Two columns of $A$ are same].

*(iv)* $\det(A\alpha) = \det(A_1\alpha, A_2\alpha, \ldots, A_n\alpha)$

$= \det(A_1, A_2, \ldots, A_n)\,\alpha^n$

$= \det(A)\,\alpha^n$

*(v)* If $A_i$ is multiplied by $\alpha$, then

$$\det (A_1, A_2, ...., A_i\alpha, ..., A_n)$$
$$= \det (A_1, A_2, ..., A_i, ..., A_n)\ \alpha$$
$$= \det (A)\ \alpha$$

**Theorem 8.** *Let A and B be any two matrices of order $n \times n$. If matrix B is obtained by*

*(i) interchanging two columns (rows) of A, then*

$$det\ (B) = -\ det\ (A)$$

or, symbolically,

$$\det [e^{(i,j)}\ (A)] = -\det (A).$$

*(ii)* Adding to a column (row) of $A$ by a scalar multiplier of another column (row) of $A$, then $\det (B) = \det (A)$ or, symbolically,

$$\det [e^{i+(j)\lambda}\ (A)] = \det (A).$$

**Proof :** Let $A = [a_{ij}]$ and $B = e^{(i,j)}\ (A)$,

then $B_i = A_j, B_j = A_i$ and $B_r = A_r, r \neq i,\ j$

Now,

$$\det [e^{(i,j)}\ (A)] = \det (B)$$
$$= \sum_{\rho \in S_n} \in (\rho) b_{\rho(1)1}\ b_{\rho(2)2} ... b_{\rho(i)i} ... b_{\rho(j)j} ... b_{\rho(n)n}$$
$$= \sum_{\rho \in S_n} \in (\rho)\, a_{\rho(1)1} ...\ a_{\rho(2)2} ... a_{\rho(i)j} ... a_{\rho(i)j} ... a_{\rho(n)n} \qquad .....(1)$$

$[\because\ B_i = A_j$ and $B_j = A_i]$

Let $h = (i, j) \in S_n$ be a transposition, then $h\ (i) = j$, $h\ (j) = 1$ and $h\ (r) = r,\ r \in \{1, 2, ...., n\}$ and $r \neq i, j$.

Using these in equation (1), we get

$$\det B = \det\ [e^{(i,j)}\ (A)]$$
$$= \sum_{\rho \in S_n} \in (\rho)\, a_{\rho(1)h(1)} a_{\rho(2)h(i)} ... a_{\rho(i)h(i)} ... a_{\rho(j)h(j)} ... a_{\rho(n)h(n)}$$
$$= \in (h) \det (A)$$
$$= -\det (A) \qquad [h \text{ is an odd permutation}]$$

*(ii)* Let $B = e^{(i)+(j)\lambda}\ (A)$, then $B_r = A_r, r \in \{1, 2, ..., n\}, r \neq i,$

and $B_i = A_i + \lambda\ A_j$

$B_r = A_r$

$\Rightarrow$ $b_{mr} = a_{mr},\ \forall\ m = 1, 2, ...\ n$

$B_i = A_i + \lambda A_j$

$b_{mi} = a_{mi} + \lambda a_{nj},$

$\forall\ m = 1, 2, ...\ n$

Then,

$$\det(B) = \det\left[e^{(i)+(j)\lambda}(A)\right]$$

$$= \sum_{\rho \in S_n} \in(\rho) b_{\rho(1)i} \ldots b_{\rho(i)i} \ldots b_{\rho(j)j} \ldots b_{\rho(n)n}$$

$$= \sum_{\rho \in S_n} \in(\rho) a_{\rho(1)1} \ldots \left(a_{\rho(i)i} + \lambda a_{\rho(j)j}\right) \ldots a_{\rho(j)j} \ldots a_{(n)n}$$

$$= \sum_{\rho \in S_n} \in(\rho) a_{\rho(1)1} \ldots a_{\rho(i)i} \ldots a_{\rho(j)j} \ldots a_{(n)n}$$

$$+\lambda \sum_{\rho \in S_n} \in(\rho) a_{\rho(1)1} \ldots a_{\rho(j)j} \ldots a_{\rho(j)j} \ldots a_{\rho(n)n}$$

$$= \det(A) + 0$$

$$= \det(A).$$

**Theorem 9.** *Let A and B are two similar matrices over the same field F, then*

$$det\,(A) = det\,(B).$$

**Proof :** Given that $A$ and $B$ are similar matrices, so that there exists an invertible matrix $P$ such that

$$B = PAP^{-1}$$

$$\det(B) = \det(PAP^{-1})$$

$$= \det(P) \det(A) \det(P^{-1})$$

$$= \det(P) \det(A) \frac{1}{\det(P)}$$

$$= \det(A)$$

**Theorem 10.** *(Cramer's rule) Let A be an n × n square matrix over a field F and $A_1$, $A_2$, ... $A_n$ are columns of A, and B = $\alpha_1 A_1 + \ldots + \alpha_n A_n$ for some $\alpha_1$, $\alpha_2$, ... $\alpha_n \in F$, then*

$$\alpha_i = \frac{1}{\det(A)} \det\left(A_1, \ldots A_{i-1}, B, A_{i+1}, \ldots, A_n\right)$$

**Proof :** We have

$$\det(A_1, \ldots, A_{i-1}, B, A_{i+1}, \ldots, A_n)$$

$$= \det\left(A_1, \ldots, A_{i-1}, \sum_{i=1}^{n} \alpha_i A_i, A_{i+1}, \ldots, A_n\right)$$

$$= \alpha_1 \det(A_1, \ldots, A_{i-1}, A_1, A_{i+1}, \ldots, A_n) + \ldots$$

$$+ \alpha_2 \det(A_1, \ldots, A_{i-1}, A_2, A_{i+1}, \ldots, A_n) + \ldots$$

$$+ \alpha_n \det(A_1, \ldots, A_{i-1}, A_n, A_{i-1}, \ldots, A_n)$$

$$= \alpha_i \det(A)$$

Thus,

$$\alpha_i = \frac{1}{\det(A)} \det\left(A_1, ..., A_{i-1}, B, A_{i+1}, ..., A_n\right)$$

**Singular matrix :** A square matrix $A$ is said to be singular if det $(A) = 0$, and matrix $A$ is non-singular if det $(A) \neq 0$.

**Determinant rank :** Let $A$ be an $m \times n$ non-zero matrix over a field $F$. A positive integer $r$ is called the rank of matrix $A$, if there exists an $r \times r$ submatrix $M$ of $A$ such that det $(M) \neq 0$ and if $N$ is any $s \times s$ submatrix of $A$ and $s > r$, then det $(N) = 0$.

The rank of a zero matrix is defined to be zero.

## 12.6 Characteristic polynomial and eigenvalue

**Theorem 12.** *Let $A$ be a matrix of order $n \times n$ over a field $F$. Then a scalar $\lambda \in F$ is an eigenvalue of $A$ iff det $(A - \lambda I) = 0$.*

**Proof :** Suppose that $\lambda$ is an eigen value of matrix $A \Leftrightarrow$ There exists a non-zero vector $X \in F^n$ such that $AX = X\lambda$

$\Leftrightarrow$ $AX - X\lambda I = 0$

$\Leftrightarrow$ $(A - \lambda I)\, X = 0$

$\Leftrightarrow$ $t_{A-\lambda I}\,(X) = 0$ $\quad [t_{A-\lambda I} : F^n \rightarrow F^n]$

$\Leftrightarrow$ a non-zero $X \in$ Ker $(t_{A-\lambda I})$

$\Leftrightarrow$ dim ker $(t_{A-\lambda I}) > 0$

$\Leftrightarrow$ nullity $(t_{A-\lambda I}) > 0$

$\Leftrightarrow$ rank $(t_{A-\lambda I}) < n$ $\quad$ [By Sylvester's law of nullity]

$\Leftrightarrow$ rank $(A - \lambda I) < n$

$\Leftrightarrow$ det $(A - \lambda I) = 0$.

**Characteristic Polynomial :**

Let $A = [a_{ij}]$ be an $n \times n$ matrix over a field $F$ and $\lambda \in F$. Then

$$C_A(\lambda) = \det(A - \lambda I)$$

$$= \begin{bmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{bmatrix}$$

$$C_A(\lambda) = C_0 + C_1 \lambda + ... + C_n \lambda^n \qquad .....(1)$$

where $C_0, C_1, \ldots, C_n$ are scalars in $F$. This expression $C_A(\lambda)$ is called the characteristic polynomial of matrix $A$ and the equation $C_A(\lambda) = 0$ is called the characteristic equation of matrix $A$. From theorem (12), it is obvious that eigen values of matrix $A \in M_{n\times n}(F)$ are the roots of the characteristic polynomial of matrix $A$.

**Theorem 13.** *Similar matrices have the same characteristic polynomial and hence the same eigen values.*

**Proof :** Let $A$ and $B$ be similar matrices over the field $F$. Then there exists an invertible matrix $P$ such that

$$B = PA\ P^{-1}$$

The characteristic polynomial of $B$ is

$$C_B(\lambda) = \det(B - \lambda I)$$

Putting $B = PAP^{-1}$, we get

$$\begin{aligned} C_B(\lambda) &= \det(PAP^{-1} - \lambda I) \\ &= \det(PAP^{-1} - P(\lambda I)P^{-1}) \\ &= \det(P(A - \lambda I)P^{-1}) \\ &= \det P \det(A - \lambda I) \det(P^{-1}) \\ &= \det P \det(A - \lambda I) \frac{1}{\det(P)} \\ &= \det(A - \lambda I) \\ &= C_A(\lambda) \end{aligned}$$

Hence the similar matrices $A$ and $B$ have the same characteristic polynomial. Since the eigen values of a matrix are roots of the characteristic polynomial, so $A$ and $B$ have the same eigen values.

**Theorem 14.** *Let $A$ be an $n \times n$ matrix over a field $F$. Then $A$ and $A^T$ have the same eigen values.*

**Proof :** The eigen values of a matrix $A$ are the roots of the characteristic equation of $A$.

The characteristic equations of $A$ and $A^T$ are $\det(A - \lambda I) = 0$ and $(A^T - \lambda I) = 0$.

Now,

$$(A - \lambda I)^T = A^T - \lambda I$$

$$\Rightarrow \quad \det[(A - \lambda I)^T] = \det(A^T - \lambda I)$$

$$\Rightarrow \quad \det(A - \lambda I) = \det(A^T - \lambda I) \qquad [\because \det(A - \lambda I)^T = \det(A - \lambda I)]$$

Thus

$$\det(A^T - \lambda I) = 0 \quad \text{iff} \quad \det(A - \lambda I) = 0$$

Hence every root of $\det(A - \lambda I) = 0$ is also a root of $\det(A^T - \lambda I) = 0$ and vice-versa. Thus $A$ and $A^T$ have the same eigen values.

**Theorem 15.** *Let A and B be similar matrices. If X is an eigen vector of A corresponding to the eigen value λ , then PX is an eigen vector of B, corresponding to the same eigen value , where*

$$B = PAP^{-1}$$

**Proof :** Given that λ is an eigen value of A corresponding to the eigen vector $X$, so that

$$AX = X\lambda$$

Now,

$$\begin{aligned} B(PX) &= (PAP^{-1})(PX) \\ &= PA(P^{-1}P)X \\ &= PAX \\ &= P(AX) \\ &= P(X\lambda) \\ &= (PX)\lambda \end{aligned}$$

Thus $PX$ is eigenvector of $B$ corresponding to eigenvalue λ.

**Theorem 16.** *(Cayley-Hamilton Theorem) Every square matrix A over a field F has $C_A(A) = 0$, where $C_A(\lambda)$ is characteristic polynomial of A. In other words every square matrix A over a field F satisfies its characteristic equation.*

**Proof :** Let $A$ be an $n \times n$ matrix and

$$\begin{aligned} C_A(\lambda) &= \det(A - \lambda I) \\ &= C_0 + C_1\lambda + C_2\lambda^2 + .... + C_n\lambda^n \end{aligned}$$ be its characteristic polynomial,

where $C_0, C_1, ....., C_n$ are scalars in λ. Now consider the adjoint of matrix $(A - \lambda I)$. Since each entry of matrix $(A - \lambda I)$ is linear is λ or a constant, so that each entry of matrix $adj\ (A - \lambda I)$ is a polynomial of degree $(n - 1)$ in λ. It is noted that each entry of $adj\ (A - \lambda I)$ is the determinant of an $(n - 1) \times (n - 1)$ submatrix of $A - \lambda I$.

Hence, $$adj\ (A - \lambda I) = B_0 + B_1\lambda + .... + B_{n-1}\lambda^{n-1},$$

where $B_0, B_1, ...., B_{n-1}$ are $n \times n$ square matrices over field $F$. We have

$$(A - \lambda I)\ adj\ (A - \lambda I) = \det(A - \lambda I)\ I$$

$$\Rightarrow \quad (A - \lambda I)(B_0 + B_1\lambda + .... + B_{n-1}\lambda_{n-1}) = (C_0 + C_1\lambda + ... + C_n\lambda^n)\ I.$$

Equating the coefficients of the various powers of λ on both sides, respectively, we get

$$\begin{aligned} AB_0 &= C_0 I \\ AB_1 - IB &= C_1 I \end{aligned}$$

$$AB_2 - IB_1 = C_2 I$$

$$\vdots$$

$$AB_{n-1} - IB_{n-2} = C_{n-1} I$$

$$-I B_{n-1} = C_n I$$

On multiplying the above $(n + 1)$ equations on the right by $I, A, A^2, ..., A^n$ respectively and adding, we get

$$C_0 I + C_1 A + C_2 A^2 + ..... + C_n A^n = 0$$

$$\Rightarrow \quad C_A (A) = 0$$

Thus each square matrix $A$ over field $F$, satisfies its characteristic equation.

## Self-learning exercise-1

**1. True/False Statements :**

*(i)* A function is bilinear if it is linear in each component. [T/F]

*(ii)* If $\lambda \in F$ is a scalar, then $|\lambda A| = \lambda^n |A|$, where $A$ be a matrix of order $n \times n$. [T/F]

*(iii)* If any two rows of a determinant $A$ are equal then $|A| = 0$ [T/F]

**2. Fill in the blanks :**

*(i)* Let $A$ and $B$ by any two similar matrices over the same field $F$, then

$$\det A = .....$$

*(ii)* A square matrix $A$ is called ..... if $\det (A) = 0$ and matrix $A$ is called ..... if $\det (A) \neq 0$

*(iii)* The equation $C_A (\lambda) = \det (A - \lambda I) = 0$ is called ..... for matrix $A$, where $\lambda \in F$.

## 12.7 Summary

In this unit we studied various properties of determinants, their existence and uniqueness. We have also studied the characteristic polynomial and eigen values of determinants, Cramer's rule and a very important result known as the Cayley-Hamilton theorem.

## 12.8 Answers to self-learning exercises

### Self-learning exercise-1

**1.** *(i)* T *(ii)* T *(iii)* T

**2.** *(i)* det B *(ii)* singular, non-singular *(iii)* characteristic equation

## 12.9 Exercise

**1.** Find the characteristic roots and characteristic spaces of the matrix

$$\begin{bmatrix} 2 & 2 & 0 \\ 2 & 1 & 1 \\ -7 & 2 & -3 \end{bmatrix}$$

**2.** Prove that the eigen vectors corresponding to distinct eigen values of a matrix are linearly independent.

**3.** Prove that det $(adj\ A) = [\det (A)]^{n-1}$ for an $n \times n$ matrix $A$, $n \geq 2$.

**4.** Prove that the characteristic polynomial of an $n \times n$ triangular matrix $A = [a_{ij}]$ is

$$(a_{11} - \lambda)\ (a_{22} - \lambda)\ \ldots\ldots\ (a_{nn} - \lambda).$$

□□□

# UNIT 13 : Real Inner Product Space-I

**Structure of the Unit**



## 13.0 Objectives

In this unit, we consider only real vector spaces, that is vector spaces over the real field. Our main object is to study vector spaces in which it makes sense to speak of 'length' of a vector and of the 'angle' between two vectors. Thus we study a certain type of scalar-valued function on pairs of vector, known as an inner product. One known example of an inner product is the scalar or dot product of vectors in $R^2$. A vector space equipped with an inner product is called an inner-product space. After reading this unit we will be able to understand various properties of these spaces and their usefulness.

## 13.1 Introduction

This unit introduces the concept of an inner-product, inner-product space, length or norm of a vector and orthogonal vectors. In this unit we shall study some important results, such as, Schwartz's inequality, Triangle inequality, Pythagoras theorem and Gram-Schmidt orthogonalization.

## 13.2 Real inner product space

Let $V$ be a vector space over the real field $R$. An **inner-product** on $V$ is a function $< . >$ : $V \times V \to R$ which assigns to each pair of vectors $(u, v)$ of $V$, a real number $< u, v >$ such that the following are satisfied :

*(i)* $< u, u > \geq 0$ and $< u, u > = 0$ if and only if $u = \mathbf{0}$, (Positive definiteness)

*(ii)* $< u, v > = < v, u >$ (symmetry)

*(iii)* $< \alpha u + \beta v, w > = \alpha < u, w > + \beta < v, w >$ (Linearity)

$(V, <, >)$ is called a **real inner-product space.** It is also called a Euclidean space.

In is clear from *(ii)* and *(iii)* that

$< u, \alpha v + \beta w > = \alpha < u, v > + \beta < u, w >$, so an inner product is linear in each of its arguments.

Note that $< u, v > = 0$, if $u = \mathbf{0}$ or $v = \mathbf{0}$ or both $u = v = \mathbf{0}$.

Thus $< \mathbf{0}, v > = < 0\ v, v > = 0 < v, v > = 0$.

***Ex.*1.** *Prove that $R^n$ is a real inner-product space with an inner-product defined by*

$$< u, v > = a_1 b_1 + a_2 b_2 + ... + a_n b_n \qquad .....(1)$$

*where* $u = (a_1, a_2, ..., a_n)$ *and* $v = < b_1, b_2, ..., b_n >$.

**Sol.** We know that $R^n$ is a vector space over $R$. Now we shall show that the product defined in (1) satisfies all the properties of an inner-product.

***(i)* Positive definiteness :** We have for all $u \in R^n$.

$$< u, u > = a_1 a_1 + ... + a_n a_n$$
$$= a_1^2 + ... + a_n^2 > 0 \quad \text{if} \quad u \neq \mathbf{0},$$

and it is zero if and if $a_1 = a_2 = ... = a_n = 0$, *i.e.* $u = \mathbf{0}$.

***(ii)* Symmetry :** By the definition of product given in (1), we have

$$< u, v > = a_1 b_1 + ... + a_n b_n$$
$$= b_1 a_1 + ... + b_n a_n \quad \text{[multiplication is commutative in } R\text{]}$$
$$= < v, u >$$

$\therefore \quad < u, v > = < v, u >$ for all $u, v \in R^n$.

***(iii)* Linearity :** Let $u = (a_1, ..., a_n)$, $v = (b_1, ..., b_n)$ and $w = (c_1, ..., c_n)$ be in $R^n$ and $\alpha, \beta \in R$.

Then $< \alpha u + \beta v, w > = (\alpha a_1 + \beta b_1) c_1 + ... + (\alpha a_n + \beta b_n) c_n$

$$= [(\alpha a_1) c_1 + ... + (\alpha a_n) c_n] + [(\beta b_1) c_1 + ... + (\beta b_n) c_n]$$
$$= \alpha (a_1 c_1 + ... + a_n c_n) + \beta (b_1 c_1 + ... + b_n c_n)$$
$$= \alpha < u, w > + \beta < v, w > .$$

Hence $R^n$ is a real inner-product space for the product defined by (1).

The product defined by (1) is called the usual or standard inner product.

***Ex.2.*** *If* $u = (a_1, a_2)$ $v = (b_1, b_2) \in R^2$, *then*

$$< u, v > = a_1 b_1 - a_2 b_1 - a_1 b_2 + 4 a_2 b_2 \qquad .....(1)$$

*defines an inner-product.*

**Sol.** We know that $R^2$ is a vector space over $R$. We shall now show that the product defined above satisfies all the properties of an inner-product.

***(i)*** **Positive definiteness :** For any $u = (a_1, a_2) \in R^2$, we have

$$\begin{aligned} < u, u > &= a_1 a_1 - a_2 a_1 - a_1 a_2 + 4 a_2 a_2^2 \\ &= a_1^2 - 2a_1 a_2 + 4 a_2^2 \\ &= (a_1 - a_2)^2 + 3 a_2^2 \qquad .....(2) \end{aligned}$$

Now (2) is a sum of two non-negative real numbers. Therefore it is $\geq 0$. Thus $< u, u > \geq 0$

Also $< u, u > = 0$

$$\begin{aligned} \Leftrightarrow \quad & (a_1 - a_2)^2 + 3 a_2^2 = 0 \\ \Leftrightarrow \quad & (a_1 - a_2)^2 = 0, 3 a_2^2 = 0 \\ \Leftrightarrow \quad & a_1 - a_2 = 0, a_2 = 0 \\ \Leftrightarrow \quad & a_1 = 0, a_2 = 0, \Leftrightarrow u = \mathbf{0} \end{aligned}$$

***(ii)*** **Symmetry :** We have

$$\begin{aligned} < u, v > &= a_1 b_1 - a_2 b_1 - a_1 b_2 + 4 a_2 b_2 \\ &= b_1 a_1 - b_2 a_1 - b_1 a_2 + 4 b_2 b_2 \\ &\qquad \text{[By commutativity of multiplication in } R\text{]} \\ &= < u, v > \end{aligned}$$

***(iii)*** **Linearity :** For any $u = (a_1, a_2)$, $v = (b_1, b_2)$, $w = (c_1, c_2)$ in $R^2$ and $\alpha, \beta \in R$, we have

$$\begin{aligned} < \alpha u + \beta v, w > &= (\alpha a_1 + \beta b_1) c_1 - (\alpha a_2 + \beta b_2) c_1 - (\alpha a_1 + \beta b_1) c_2 \\ &\qquad + 4(\alpha a_2 + \beta b_2) c_2 \text{ [From (1)]} \\ &= (\alpha a_1 c_1 - \alpha a_2 c_1 - \alpha a_1 c_2 + 4 \alpha a_2 c_2) \\ &\qquad + (\beta b_1 c_1 - \beta b_2 c_1 - \beta b_1 c_2 + 4 \beta b_2 c_2) \\ &= \alpha (a_1 c_1 - a_2 c_1 - a_1 c_2 + a_2 c_1) + \beta (b_1 c_1 - b_2 c_1 - b_1 c_2 + 4 b_2 c_2) \\ &= \alpha < u, w > + \beta < v, w > . \qquad \text{[From (1)]} \end{aligned}$$

Hence $R^2$ the product defined in (1) is an inner-product on $R^2$. Also with respect to this inner product $R^2$ is an inner-product space.

***Ex.3.*** *Show that* $< u, v > = a_1 b_1 + a_2 b_2 - a_3 b_3$ *is not an inner-product on* $R^3$ *where* $u = (a_1, a_2, a_3)$ *and* $v = (b_1, b_2, b_3)$.

**Sol.** Let $u = (3, 4, 5) \neq \mathbf{0}$. Then

$$< u, v > = 3 \cdot 3 + 4 \cdot 4 - 5 \cdot 5 = 0$$

and so the positive definiteness property of an inner product is not satisfied.

***Ex.4.*** *Show that* $< u, v > = a_1 b_1 a_2 b_2$ *is not an inner-product on* $R^2$ *where* $u = (a_1, a_2)$, $v = (b_1, b_2)$.

**Sol.** Let $\alpha = 2$ and $u = (1, 3)$, $v = (1, 1)$. Then

$$\alpha u = (2, 6)$$

and we have $< u, v > = 1 \cdot 3 \cdot 1 \cdot 1 = 3$

and $< \alpha u, v > = 2 \cdot 6 \cdot 1 \cdot 1 = 12.$

Thus $\alpha < u, v > = 2 \cdot 3 = 6$ is not equal to $< \alpha u, v >$ and so, the linearity property of an inner product is not satisfied.

## 13.3 Norm of a vector

**Norm :** Let $V$ be an inner-product space. Then any $v \in V$, the non-negative square root $\sqrt{< v, v >}$ is called the **norm** or the length of the vector $v$ and is denoted as

$$\| v \| = \sqrt{< v, v >}$$

If $\| v \| = 1$, for $v \in V$, then $v$ is called a **unit vector.**

***Ex.5.*** *Find* $\| v \|$ *using the usual inner product in* $R^2$, *where* $v =$ (3, 4)

**Sol.** $\| v \|^2 = < v, v > = < (3, 4), (3, 4) > = 3 \cdot 3 + 4 \cdot 4 = 9 + 16 = 25$ Hence $\| v \| = 5$.

**Theorem 1.** *If V is an inner-product space and* $v \in V$, *then*

*(i)* $\| v \| \geq 0$ : and $\| v \| = 0$ if and any if $v = \mathbf{0}$

*(ii)* $\| \alpha v \| = | \alpha | \| v \|$

**Proof.** ***(i)*** We have

$$\| v \| = < v, v > \qquad \text{[by def. of norm]}$$

$$\Rightarrow \quad \| v \|^2 = < v, v >$$

$$\Rightarrow \quad \| v \|^2 \geq 0 \qquad [\because < v, v > \geq 0]$$

$$\Rightarrow \quad \| v \| \geq 0$$

Also $< v, v > = 0$ iff $v = \mathbf{0}$

$\therefore \quad \| v \|^2 = 0$ iff $v = \mathbf{0}$ *i.e.* $\| v \| = 0$ iff $v = \mathbf{0}$.

***(ii)*** We have

$$\| \alpha v \|^2 = < \alpha v, \alpha v > \qquad \text{[By def. of norm]}$$

$$= \alpha < v, \alpha v > \qquad \text{[by linearly properly]}$$

$$= \alpha \alpha < v, v >$$

$$= \alpha^2 \| v \|^2$$

Hence $\| \alpha v \| = | \alpha | \| v \|$.

**Note :** If $v$ is any nonzero vector of an inner product space $V$, then $\frac{1}{\| v \|} v$ is a unit in $V$.

We have $\| v \| \neq 0$ because $v \neq \mathbf{0}$ Therefore $\dfrac{1}{\| v \|} v \in V$.

Now $\left\langle \dfrac{1}{\| v \|} v, \dfrac{1}{\| v \|} v \right\rangle = \dfrac{1}{\| v \|} \left\langle v, \dfrac{1}{\| v \|} v \right\rangle$

$$= \frac{1}{\| v \|^2} \langle v, v \rangle = \frac{1}{\| v \|^2} \| v \|^2 = 1$$

Therefore $\dfrac{v}{\| v \|} = 1$ and thus $\dfrac{v}{\| v \|}$ is a unit vector.

For example if $v = (2, 1, -1)$ is a vector in $R^3$ with standard inner-product, then

$$\| v \| = \sqrt{\langle v, v \rangle} = \sqrt{4+1+1} = \sqrt{6}$$

Therefore $\dfrac{1}{\sqrt{6}}(2, 1, -1)$ *i.e.* $\left( \dfrac{2}{\sqrt{6}}, \dfrac{1}{\sqrt{6}}, \dfrac{-1}{\sqrt{6}} \right)$ is a unit vector.

Now we shall prove an elementary but familiar result about real quadratic equations.

**Lemma :** *If $a, b, c \in R$ such that $a > 0$ and $f(t) = at^2 + bt + c \geq 0$* for all real numbers $t$, then $b^2 \leq 4ac$.

**Proof.** We have $f(t) = at^2 + bt + c$

$$= \frac{1}{a} \left( a^2 t^2 + abt + ac \right)$$

$$= \frac{1}{a} \left[ (at)^2 + 2(at) \frac{b}{2} + \frac{b^2}{4} - \frac{b^2}{4} + ac \right]$$

$$= \frac{1}{a} \left[ \left( at + \frac{b}{2} \right)^2 - \left( \frac{b^2 - 4ac}{4} \right) \right]$$

Since $at^2 + bt + c \geq 0$ for all $t \in R$.

Thus $\dfrac{b^2 - 4ac}{-4a} \geq 0$

$\Rightarrow$ $\dfrac{1}{4a} \left( 4ac - b^2 \right) \geq 0 \quad \Rightarrow \quad b^2 \leq 4ac \qquad (\because a > 0)$

## 13.4 The Schwartz inequality

**Theorem 2.** *Let V be an inner-product space. Then for arbitrary vectors $u, v \in V$, we have* $| < u, v > | \leq \| u \| \| v \|$ .

**Proof.** If either $u = \mathbf{0}$ or $v = \mathbf{0}$ *i.e.* either $\| u \|$ or $\| v \|$ is zero, then both $< u, v > = 0$ and $\| u \| \| v \| = 0$ so the theorem is true.

Now let $u \neq \mathbf{0}$, Then $\| u \| > 0$. Then for any scalar $\alpha \in R$, we have

$$< \alpha u + v, \alpha u + v > \geq 0 \qquad \text{[By positive definiteness]}$$

$$\Rightarrow \quad < \alpha u, \alpha u > + < \alpha u, v > + < v, \alpha u > + < v, v > \geq 0$$

$$\Rightarrow \quad \alpha^2 < u, u > + \alpha < u, v > + \alpha < v, u > + < v, v > \geq 0$$

$$\Rightarrow \quad \alpha^2 \| u \|^2 + 2\alpha < u, v > + \| v \|^2 \geq 0$$

Since $\alpha^2 \| u \|^2 + 2\alpha < u, v > + \| v \|^2$ is a quadratic expression in $\alpha$ with $\| u \|^2 > 0$ as coefficient of $\alpha^2$, therefore by lemma, we have

$$4 (< u, v >)^2 \leq 4 \| u \|^2 \| v \|^2$$

$$\Rightarrow \quad (< u, v >)^2 \leq \| u \|^2 \| v \|^2$$

$$\Rightarrow \quad | < u, v > | \leq \| u \| \| v \|$$

**Corollary :** For any real numbers $x_1, x_2, y_1$ and $y_2$

$$| x_1 y_1 + x_2 y_2 | \leq \left( x_1^2 + x_2^2 \right)^{1/2} \left( y_1^2 + y_2^2 \right)^{1/2}.$$

**Proof.** We know that $R^2$ is an inner-product space with respect to standard inner-product. Therefore by Schwarz inequality $| < u, v > | \leq \| u \| \| v \|$ for all $u, v \in R^2$

Taking $u = (x_1, x_2)$ and $v = (y_1, y_2)$, we get $| x_1 y_1 + x_2 y_2 | \leq \left( x_1^2 + x_2^2 \right)^{1/2} \left( y_1^2 + y_2^2 \right)^{1/2}$

## 13.5 Triangle inequality

**Theorem 3.** *For any two vector u and v in an inner-product space V,*

$$\| u + v \| \leq \| u \| \| v \|.$$

**Proof.** We have

$$\| u + v \|^2 = < u + v, u + v >$$

$$= < u, u > + 2 < u, v > + < v, v >$$

[By using linearity and symmetry]

$$= \| u \|^2 + 2 < u, v > + \| v \|^2 \leq \| u \|^2 + 2 \| u \| \| v \| + \| v \|^2$$

[By Schwartz inequality]

$$\therefore \quad \| u + v \|^2 = ( \| u \| \| v \| )^2$$

Taking sequence root of both the sides, we get $\| u + v \| \leq \| u \| \| v \|$.

**Theorem 4.** *For any two vectors u and v in an inner-product space V,*

$$\| u \| - \| v \| \leq \| u - v \|.$$

**Proof.** Let $\| u \| = \|(u - v) + v \| \leq \| u - v \| + \| v \|$, by triangle inequality.

$$\Rightarrow \quad \| u \| - \| v \| \leq \| u - v \| \qquad \text{.....(1)}$$

Interchanging $u$ and $v$ in the above inequality, we get

$$\| v \| - \| u \| \leq \| v - u \| = \| (-1) (u - v) \|$$

$$= |- 1| \| u - v \| = \| u - v \|$$

$\Rightarrow$ $-(\| u \| - \| v \|) \le \| u - v \|$ .....(2)

From (1) and (2), we have

$$\pm(\| u \| - \| v \|) \le \| u - v \|$$

Hence $$\big|\, \| u \| - \| v \| \,\big| \le \| u - v \|$$

**Theorem 5.** *Let V be an inner-product space. Prove that for any two vector v, u ∈ V,*

*(i)* $\| u + v \|^2 - \| u - v \|^2 = 4 < u, v >$ it is known as **polarization identity,** and

*(ii)* $\| u + v \|^2 + \| u - v \|^2 = 2 (\| u \|^2 + \| v \|^2)$, it is **parallelogram law.**

**Proof.** We have $\| u + v \|^2 = < u + v, u + v >$

$= < u, u > + < u, v > + < v, u > + < v, v >$

$= < u, u > + 2 < u, v > + < v, v >$ $[\because < u, v > = < v, u >]$

$= \| u \|^2 + 2 < u, v > + \| v \|^2$ .....(1)

and $\| u \pm v \|^2 = < u - v, u - v >$

$= < u, u > + < u, - v > + < - v, u > + < - v, - v >$

$= < u, u > - < u, v > - < v, u > + (-1)^2 < v, v >$

$= \| u \|^2 - 2 < u, v > + \| v \|^2$ .....(2)

From (1) and (2), we get

$\| u + v \|^2 - \| u - v \|^2 = 4 < u, v >$ and

$\| u + v \|^2 + \| u - v \|^2 = 2 (\| u \|^2 + \| v \|^2)$

Hence the theorem.

**Self-learning exercise–1**

**1.** Expand :

*(i)* $< u, v_1 + v_2 >$ *(ii)* $< u, kv >$ *(iii)* $< u + v, u - v >$

**2.** If $u = (1, 2, 4)$, $v = (2, -3, 5)$, $w = (4, 2, -3) \in R^3$ Then with respect to the usual inner-product in $R^3$, find

*(i)* $u \cdot v$ *(ii)* $(u + v) \cdot w$ *(iii)* $\| u \|$ *(iv)* $\| u + v \|$

## 13.6 Orthogonality

In this section, we shall define the angle between two nonzero vectors of an inner-product space and orthogonal vectors.

By the Schwartz inequality, we have $$-1 \le \frac{\langle u, v \rangle}{\| u \| \| v \|} \le 1.$$

It follows that there exists a unique $\theta \in [0, \pi]$ such that $\cos \theta = \dfrac{\langle u, v \rangle}{\| u \| \| v \|}$. Thus $\theta$ is called the **angle** between two nonzero vectors $u$ and $v$, and thus

$$\theta = \cos^{-1}\frac{\langle u, v\rangle}{\| u \| \| v \|}, \quad 0 \le \theta \le \pi$$

***Ex.6.*** *Find* $\cos\theta$ *for the angle* $\theta$ *between* $u = (5, 1)$ *and* $v = (-2, 3)$ *in Euclidean space* $R^2$. *In which quadrant does* $\theta$ *lie* ?

**Sol.** We have

$$< u, v > = -10 + 3 = -7, \| u \|^2 = 25 + 1 = 26, \| v \|^2 = 4 + 9 = 13. \text{ Thus}$$

$$\cos\theta = \frac{\langle u, v\rangle}{\| u \| \| v \|} = \frac{-7}{\sqrt{13}\sqrt{26}} = \frac{-7}{13\sqrt{2}}$$

Since $\theta$ is negative, $\theta$ lies in the second quadrant.

**Orthogonal vectors :**

Let $V$ be an inner-product space and $u, v \in V$. Then u is said to be orthogonal to $v$ if

$$< u, v > = 0$$

If $u$ is orthogonal to $v$, then we write $u \perp v$. The symmetry of the inner-product implies that whenever $u$ is orthogonal to $v$, then $v$ is orthogonal to $u$, *i.e.,* if $u \perp v$, then $v \perp u$. Further every vector $u \in V$ is orthogonal to zero vector, since $< u, \mathbf{0} > = 0$. And, obviously zero vector is the only vector which is orthogonal to itself.

A vector $u$ is perpendicular to a subspace $S$ of an inner-product space $V$ if for all $v \in S$

$$< u, v > = 0.$$

**Orthogonal subspace :**

Two subspaces $S_1$ and $S_2$ of an inner-product space $V$ are said to be orthogonal if for all $u \in S_1$, $v \in S_2$, $< u, v > = 0$.

**Orthogonal set :**

A set of vectors $S$ in an inner-product space $V$ is said to be an **orthogonal set**, if any two distinct vectors in it are orthogonal.

If $S$ is also a basis of $V$, then it is called an **orthogonal basis** of the inner-product space $V$.

**Theorem 6.** *Any orthogonal set of nonzero vectors in an inner-product space is linearly independent.*

**Proof.** Let $V$ be an inner product space and $S$ be an orthogonal set in $V$. Let $S_1 = \{v_1, ..., v_n\}$ be a finite subset of $S$. Suppose for scalars $\alpha_1, \alpha_2 \ldots \alpha_n$ in $R$,

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = \mathbf{0}$$

Taking the inner-product of both sides with $v_1$ and using the bilinear properties of the inner-product, we obtain.

$$\alpha_1 < v_1, v_1 > + \alpha_2 < v_2, v_1 > + \ldots + \alpha_n < v_n, v_1 > = < \mathbf{0}, v_1 > = 0$$

$$\Rightarrow \quad \alpha_1 < v_1, v_1 > = 0$$

$\Rightarrow$ $\alpha_1 = 0$, since $< v_1, v_1 > \neq 0$ because $v_1 \neq \mathbf{0}$.

Taking the inner-product with $v_2,..., v_n$ in turn, we have $\alpha_2 = \alpha_3 = ... \alpha_n = 0$, which implies that the set $S_1$ is linearly independent. Thus every finite subset of $S$ is linearly independent. Hence $S$ is linearly independent.

**Orthogonal complement of a vector :**

Let $V$ be an inner-product space and $u \in V$, then the set of all vectors orthogonal to $u$ in $V$ is called the orthogonal complement of $u$ and is denoted by $u^{\perp}$.

Thus $u^{\perp} = \{v \in V \mid < u, v > = 0\}$

The symbol "$u^{\perp}$" *is* usually read as "$u$ perpendicular.

It is easy to see that $u^{\perp}$ is a subspace of $V$.

**Orthogonal complement of a set :**

Let $S^{\perp}$ be a subspace of an inner-product space $V$. Then orthogonal complement of $S$ is denoted by $S^{\perp}$ and is defined to be is the set of all vectors in $V$ which are orthogonal to each vector in $S$.

Thus $S^{\perp\perp} = \{v \in V \mid < u, v > = 0$ for all $u \in S\}$

**Theorem 7.** *If $S$ is a subspace of an inner product space $V$, then so is $S^{\perp}$.*

**Proof.** Since zero vector is orthogonal to every vector in $S$, therefore $\mathbf{0} \in S^{\perp}$. Hence $S^{\perp}$ is a non-empty subset of $V$.

Let $v_1, v_2 \in S^{\perp}$. Then $< u, v_1 > = 0, < u, v_2 > = 0$ for all $u \in S$.

for any $\alpha, \beta \in R$, we have

$$< \alpha v_1 + \beta v_2, u > = \alpha < v_1, u > + \beta < v_2, u >$$

$$= \alpha \cdot 0 + \beta \cdot 0 = 0$$

Therefore $\alpha v_1 + \beta v_2 \in S^{\perp}$ for all $v_1, v_2 \in S^{\perp}$ and all $\alpha, \beta \in R$.

Hence $S^{\perp}$ is a subspace of $V$.

Note that $S \cap S^{\perp} = \{\mathbf{0}\}$, for if $u \in S \cap S^{\perp}$, it must be self orthogonal, *i.e.* $< u, u > = 0$, which implies that $u = \mathbf{0}$.

**Orthogonal complement of an orthogonal complement :** If $S$ is a subset of an inner product space $V$, then $S^{\perp}$ is subspace of $V$. The set of all vectors in $V$ which are orthogonal to each vector of $S^{\perp}$ is called orthogonal complement of an orthogonal complement and is denoted by $S^{\perp\perp}$. Thus

$$S^{\perp\perp} = \left\{ v \in V \mid \langle v, u \rangle = 0 \text{ for all } u \in S^{\perp} \right\} .$$

Obviously $S^{\perp\perp}$ is a subspace of $V$ and $S \subset S^{\perp\perp}$.

## 13.7 Pythagoras theorem

**Theorem 8.** *Let V be an inner-product space, and u, v, ∈ V be orthogonal to each other. Then*

$\| u + v \|^2 = \| u \|^2 + \| v \|^2$. *More generally, Let* $\{v_1, v_2 \ldots v_n\}$ *be a set of vector in V such that they are pairwise orthogonal, then*

$$\left\| \sum_{i=1}^{n} v_i \right\|^2 = \sum_{i=1}^{n} \| v_i \|^2 .$$

**Proof.** We have

$$\begin{aligned} \| u + v \|^2 &= < u + v, u + v > \\ &= < u, u > + <u, v > + < v, u > + < v, v > \\ &= \| u \|^2 + 2 <u, v > + \| v \|^2 \qquad [\because < u, v > = < v, u >] \\ &= \| u \|^2 + \| v \|^2, \quad \text{since} \quad < u, v > = 0 \text{ as } u \perp v. \end{aligned}$$

Further,
$$\left\| \sum_{i=1}^{n} v_i \right\|^2 = \left\langle \sum_{i=1}^{n} v_i, \sum_{i=1}^{n} v_i, \right\rangle$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \langle v_i, v_j \rangle = \sum_{i=1}^{n} \langle v_i, v_i \rangle, \qquad \text{since } < v_i, v_j > = 0, \; i \neq j$$

$$= \sum_{i=1}^{n} \| v_i \|^2 .$$

Note that if $B = \{v_1, v_2, \ldots, v_n\}$ is an orthogonal basis of an inner product space $V$, then any vector $u \in V$ can be expressed as

$$u = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n \qquad \ldots\ldots(1)$$

for some $\alpha_1, \alpha_2, \ldots, \alpha_n \in R$, where each $\alpha_i$ is obtained as

$$\begin{aligned} < u, v_1 > &= < \alpha_1 v_1 + \ldots + \alpha_n v_n, v_1 > \\ &= \alpha_1 < v_1, v_1 > + \alpha_2 < v_2, v_1 > + \ldots + \alpha_n < v_n, v_1 > \\ &= \alpha_1 < v_1, v_1 >, \text{ since } = < v_i, v_j > = 0, \; i \neq j \\ &= \alpha_1 \| v_1 \|^2 \end{aligned}$$

Hence
$$\alpha_1 = \frac{\langle u, v_1 \rangle}{\| v_1 \|^2}$$

Similarly, we obtain
$$\alpha_2 = \frac{\langle u, v_2 \rangle}{\| v_2 \|^2}, \ldots, \alpha_n = \frac{\langle u, v_n \rangle}{\| v_n \|^2}$$

Hence expression (1) can also be written as $u$

$$= \sum_{i=1}^{n} v_i \frac{\langle u, v_i \rangle}{\| v_i \|^2}.$$

## 13.8 Orthonormal set

**Orthonormal set :**

*Let V be an inner-product space. An* ***orthonormal set*** *of vectors in V is any set* $\{v_1,...,v_m\}$ *satisfying conditions* :

***(i)*** $\| v_i \| = 1,\ 1 \le i \le m,$ *i.e.* $< v_i, v_i > = 1$

***(ii)*** $< v_i, v_j > = 0,\ i \ne j$

*i.e.* an orthogonal set of vectors in $V$ is an orthonormal set if each vector is of unit length.

**Orthonormal basis :**

A basis $\{v_1, v_2,..., v_n\}$ of $V$ is said to be **orthonormal basis** if we have

$$< v_i, v_j > = \delta_{ij} \text{ for } 1 \le i, j \le n.$$

**Theorem 9.** *Every orthonormal set of vectors is a linearly independent set in an inner-product space.*

**Proof :** Let $\{v_1, v_2,..., v_m\}$ be an orthonormal set in an inner product space $V$.

Suppose $$\sum_{i=1}^{n} \alpha_i v_i = \mathbf{0}, \quad \text{for} \quad \alpha_i \in R$$

Taking the inner product of both sides with each $v_j$, $1 \le j \le m$, we obtain

$$\left\langle \sum_{i=1}^{m} \alpha_i v_i, v_j \right\rangle = < \mathbf{0}, v_j > = 0$$

$$\Rightarrow \quad \sum_{i=1}^{m} \alpha_i < v_i, v_j > = 0 \qquad \text{[By linearity property]}$$

$$\Rightarrow \quad \sum_{i=1}^{m} \alpha_i \delta_{ij} = 0$$

$$\Rightarrow \quad \alpha_j = 0, \text{ for each } j = 1, 2, ..., m$$

Hence the orthonormal set $\{v_1, v_2,..., v_n\}$ is linearly independent.

***Ex.*****7.** *The set of unit vectors* $\{e_1, e_2,..., e_n\}$ *in* $R^n$ *is an orthonormal set, where the inner product is given by*

$$< u, v > = \sum_{i=1}^{n} a_i b_i$$

*for vector* $u = (a_1, a_2,..., a_n),\ v = (b_1, b_2,..., b_n)$ *in* $R^n$.

This set of unit vectors also form an orthonormal basis of $R^n$, with respect to the above inner product defined in $R^n$.

***Ex.*****8.** *Let* $\{v_1, v_2,..., v_n\}$ *be an orthogonal basis of an inner product space V. Then the set* $\{e_i\}_{i=1}^{n}$ *is an orthonormal basis of V, where for each i,* $1 \le i \le n$

$$e_i = \frac{v_i}{\| v_i \|}$$

**Theorem 10.** *If $S = \{v_1, v_2, ..., v_n\}$ is an orthonormal set in an inner product space V. Then for any vector $v \in V$, the vector*

$$u = v - \sum_{i=1}^{n} v_i < v, v_i >$$

*is orthogonal to each of the vectors $v_1, ..., v_n$ and consequently, to the subspace spanned by S.*

**Proof :** For any $v_j$, $1 \le j \le n$, we have

$$< u, v_j > = \left\langle v - \sum_{i=1}^{n} v_i < v, v_i >, v_j \right\rangle$$

$$= < v, v_j > - \sum_{i=1}^{n} < v, v_i > < v_i, v_j >$$

$$= < v, v_j > - \sum_{i=1}^{n} v_i < v, v_i > \delta_{ij}$$

$$= < v, v_j > - < v, v_j > = 0 \qquad .....(1)$$

Hence $u$ is orthogonal to each $v_j$, $1 \le j \le n$,.

To show that $u$ is orthogonal to the subspace $[S]$, let $w \in [S]$. Then there exist $\alpha_1, \alpha_2, ..., \alpha_n$ in $R$ such that

$$w = \sum_{i=1}^{n} \alpha_i v_i$$

Now $$< u, w > = \left\langle u, \sum_{i=1}^{n} \alpha_i v_i \right\rangle = \sum_{i=1}^{n} \alpha_i < u, v_i > = 0 \qquad \text{[From (11)]}$$

Hence $u$ is orthogonal to the subspace $[S]$, spanned *b*y $S$.

The proves the theorem.

**Theorem 11.** *If $B = \{u_1, u_2, ..., u_n\}$ be an orthonormal basis of an inner product space V and $v \in V$ be any arbitrary vector. Then the coordinates of v relative to the basis B of V are $< v, u_i >$, $i = 1, 2, ..., n$ and*

$$\| v \|^2 = \sum_{i=1}^{n} | \langle v, v_i \rangle |^2$$

**Proof :** Let $$v = \sum_{j=1}^{n} \alpha_i u_j,$$

where $\alpha_j \in F$ are coordinates of $v \in V$ relative to the basis $B = \{u_1, u_2, ..., u_n,\}$ of $V$.

Now for $i$, $1 \le i \le n$

$$< v, u_i > = \left\langle \sum_{j=1}^{n} \alpha_j u_j, u_i \right\rangle$$

$$= \sum_{j=1}^{n} \alpha_j < u_j, u_i >$$

$$= \sum_{j=1}^{n} \alpha_j \, \delta_{ji} = \alpha_i$$

Thus coordinates of $v$ relative to $B$ are,

$\alpha_1 = < v, u_1 >, \alpha_2 = < v, u_2 >, ..., \alpha_n = < v, u_n >$.

Further, $\| v \|^2 = < v, v >$

$$= \left\langle \sum_{i=1}^{n} \alpha_i u_i \sum_{i=1}^{n} \alpha_j u_j \right\rangle$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j < u_i, u_j >$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j \delta_{ij} \qquad [\because \ B \text{ is an orthonormal basis}]$$

$$= \sum_{i=1}^{n} \alpha_i^2 \quad = \sum_{i=1}^{n} | < v, u_i > |^2$$

### Self-learning exercise-2

1. If $u$ is orthogonal to every $v \in V$, then $u$ must be ..... .
2. Suppose $u$ and $v$ are non-zero in $V$. Then $u$ and $v$ are orthogonal if ad only if they are ..... .
3. Whether the following set $S$ of vectors in $R^3$ is orthogonal ?

   $S = \{u_1 = (1, 2, 1), u_2 = (2, 1, -4), u_3 = (3, -2, 1)\}$.
4. The ..... vector is the only vector which is orthogonal to itself.
5. If $\alpha$ and $\beta$ are orthogonal unit vectors (that is, $(\alpha, \beta)$ is an orthonormal set), what is the distance between $\alpha$ and $\beta$ ?
6. Find the norm of the vector $v = (1, -2, 5)$. Also normalise this vector.

## 13.9 Gram-Schmidt orthogonalization process

We now proceed to give an inductive procedure for constructing an orthonormal basis from a given set of basis vectors of an inner product space. This procedure is known as the **Gram-Schmidt orthogonalization process.**

**Theorem 12.** *Every finite dimensional inner product space has an orthonormal basis.*

**Proof :** Let $V$ be an $n$-dimensional inner product space and let $B = \{v_1, ....., v_n\}$ be a basis of $V$. From this basis we shall construct an orthonormal set of $n$ vectors. This set is linearly independent. Therefore the basis which we shall construct would be an orthonormal basis of $V$.

We seek $n$ vectors $u_1$,....., $u_n$ each of length 1 such that $i \neq j$. $< u_i, u_j > = 0$. In fact we shall finally produce them in the following form : $u_1$ will be a multiple of $v_1$; $u_2$ will be in the linear span of $v_1$ and $v_2$ ; $u_3$ in the linear span of $v_1$, $v_2$ and $v_3$ and more generally ; $u_i$ in the linear span of $v_1$, $v_2$,..., $v_i$.

Since $B$ is a basis of $V$, it is linearly independent, and so $v_i \neq \mathbf{0}$ for $1 \leq i \leq n$.

Let $u_1 = \dfrac{v_1}{\| v_1 \|}$, so that $u_1 \neq \mathbf{0}$ and $\| u_1 \| = 1$.

$u_1$ is in the linear span of $v_1$ and $\{u_1\}$ forms an orthonormal set in $V$.

Let $w_2 = v_2 - u_1 < v_2, u_1 >$, $w_2 \neq \mathbf{0}$ because if $w_2 = \mathbf{0}$ then $v_2$ must be a linear combination of $v_1$, which is a contradiction since these are basis vectors.

So let $u_2 = \dfrac{w_2}{\| w_2 \|}$, so that $\| u_2 \| = 1$ and it is in the linear span of $v_1$ and $v_2$.

Also
$$< u_2, u_1 > = < \frac{w_2}{\| w_2 \|}, u_1 >$$
$$= \frac{1}{\| w_2 \|} < w_2, u_1 >$$
$$= \frac{1}{\| w_2 \|} < v_2 - u_1 < v_2, u_1 >, u_1 >$$
$$= \frac{1}{\| w_2 \|} \left[ < v_2, u_1 > - < v_2, u_1 > < u_1, u_1 > \right]$$
$$= 0 \quad \text{since} \quad \| u \| = 1$$

Hence $\{u_1, u_2\}$ is an orthonormal set in $V$.

We continue this process and suppose we have constructed an orthonormal set $\{u_1, u_2, ..., u_r\}$, $r < n$ of vectors such that these are in the linear span of $v_1, v_2, ..., v_r$.

Let
$$w_{r+1} = v_{r+1} - \sum_{i=1} u_i < v_{r+1}, u_i > .$$

By theorem 10, we can show that $w_{r+1}$ is orthogonal to each $u_i$, $i = 1, 2, ...., r$.

Also $w_{r+1} \neq \mathbf{0}$ for $w_{r+1} = \mathbf{0}$ implies that $\{v_1, v_2, ..., v_{r+1}\}$ is linearly independent which is a contradiction.

So let
$$u_{r+1} = \frac{w_{r+1}}{\| w_r + 1 \|}, \text{ so that } \| u_{r+1} \| = 1.$$

Further $u_{r+1} \neq u_i$, $i = 1, 2,..., r$. For otherwise $v_{r+1}$ would be a linear combination of $v_1, v_2,..., v_r$ .

Thus inductively, we can construct an orthonormal set $\{u_1, u_2,...., u_n\}$ in $V$. Since an orthonormal set is always linearly independent and dim $V = n$, hence $B = \{u_1, u_2,..., u_n\}$ forms an orthonormal basis of the inner product space $V$.

We will now illustrate the construction used in this theorem by the following example.

***Ex.2.*** *Apply the Gram-Schmidt process to the vectors* $v_1 = (1, 0, 1)$, $v_2 = (1, 0, -1)$, $v_3 = (0, 3, 4)$ *to obtain an orthonormal basis for* $R^3$ *with the standard inner product.*

**Sol.** We have $\| v_1 \|^2 = < v_1, v_1 > = 1 \cdot 1 + 0 \cdot 0 + 1 \cdot 1 = 2$

Let $$v_1 = \frac{v_1}{\| v_1 \|} = \frac{1}{\sqrt{2}}(1,0,1) = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$$

Now let $$w_2 = v_2 - u_1 < v_2, u_1>$$

Since $$< v_2, u_1> = 1 \cdot \frac{1}{\sqrt{2}} + 0 \cdot 0 + \left(\frac{-1}{\sqrt{2}}\right) = 0$$

Therefore $$w_2 = v_2 - u_1 \cdot 0 = v_2 = (1, 0, -1)$$

$\Rightarrow$ $$\| w_2 \| = \sqrt{< w_2, w_2 >} = \sqrt{1 \cdot 1 + 0 \cdot 0 + (-1)(-1)} = \sqrt{2}$$

Let $$u_2 = \frac{w_2}{\| w_2 \|} = \left(\frac{1}{\sqrt{2}}, 0 \frac{-1}{\sqrt{2}}\right)$$

Further let $$w_3 = v_3 - u_1 < v_3, u_1> - u_2 < v_3, u_2 >$$

we have, $$< v_3, u_1> = 0 \cdot \frac{1}{\sqrt{2}} + 3 \cdot 0 + 4 \cdot \frac{1}{\sqrt{2}} = 2\sqrt{2}$$

and $$< v_3, u_2> = 0 \cdot \frac{1}{\sqrt{2}} + 3 \cdot 0 + 4 \cdot \left(\frac{-1}{\sqrt{2}}\right) = -2\sqrt{2}$$

Therefore $$w_3 = (0,3,4) - \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) 2\sqrt{2} + \left(\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}\right) 2\sqrt{2}$$

$$= (0, 3, 4) - (2, 0, 2) + (2, 0, -2)$$

$$= (0, 3, 0)$$

$\Rightarrow$ $$\| w_3 \| = \sqrt{< w_3, w_3 >} = \sqrt{0 \cdot 0 + 3 \cdot 3 + 0 \cdot 0} = 3$$

Let $$u_3 = \frac{w_3}{\| w_3 \|} = (0,1,0)$$

Thus $\{u_1, u_2, u_3,\}$

where $$u_1 = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), u_2 = \left(\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}\right)$$ and

$$u_3 = (0, 1, 0)$$

is the required orthonormal basis for $R^3$.

## 13.10 Summary

In this unit we have studied the concept of length or norm of a vector, scalar valued function known as inner product and various properties of real inner product space. We also studied some important theorems, *i.e.* Schwartz inequality and Gram-Schmidt orthogonalization process.

## 13.11 Answers self-learning exercises

### Self learning exercise-1

**1.** ***(i)*** $< u, v_1 > + < u, v_2 >$ ***(ii)*** $k < u, v >$ ***(iii)*** $\| u \|^2 - \| v \|^2$

**2.** ***(i)*** 16 ***(ii)*** $-17$ ***(iii)*** $\sqrt{21}$ ***(iv)*** $\sqrt{91}$

### Self learning exercise-2

**1.** zero vector **2.** perpendicular **3.** yes as $u_2 \cdot u_2 = 0, u_1 \cdot u_3 = 0, u_2 \cdot u_3 = 0$

**4. zero** **5.** $d(\alpha, \beta) = \| \alpha - \beta \| = \sqrt{2}$

**6.** $\| v \| = \sqrt{30}$, normalised vector $= \left( \frac{1}{\sqrt{30}}, \frac{-2}{\sqrt{30}}, \frac{5}{\sqrt{30}} \right)$

## 13.12 Exercises

**1.** Define an inner product and an inner product space.

**2.** Let $V$ be a vector space of real continuous functions on the interval $a \leq t \leq b$. Show that the following is an inner product on $V$;

$$< f, g > = \int_a^b f(t)\, g(t)\, dt .$$

**3.** Let $V(R)$ be a vector space of polynomials with inner product defined by

$$< f, g > = \int_0^1 f(t)\, g(t)\, dt .$$

In $f(x) = x^2 + x - 4$, $g(x) = x - 1$, then find $< f, g >$ and $\| g \|$.

**[Ans.** $< f, g > = \frac{7}{4}, \| g \| = \frac{1}{3}$ **]**

**4.** If $\alpha, \beta$ be vectors in a real inner product space such that

$$\| \alpha \| = \| \beta \|.$$

Then prove that $(\alpha + \beta, \alpha - \beta) = 0$

**5.** Prove that $u$ and $v$ are orthogonal in an inner product space $V$ if and only if

$$\| u - v \|^2 = \| u \|^2 + \| v \|^2$$

**6.** Which of the following defines inner products in $V_2(R)$ ?

*(i)* $< u, v > = x_1 y_1 - x_2 y_1 - x_1 y_2 + 5 x_2 y_2$

*(ii)* $< u, v > = 2x_1 y_1 + x_2 y_1 + x_1 y_2 + x_2 y_2$

*(iii)* $< u, v > = x_1 y_1 - 2x_1 y_2 - 2x_2 y_1 + 5 x_2 y_2$ **[Assume** $u = < x_1, x_2 >, v = < y_1, y_2 >$]

[**Ans.** *(i)*, *(ii)* are innerproducts and *(iii)* is not]

**7.** If in an inner product space $\| u + v \| = \| u \| + \| v \|$, then prove that the vectors $u$ and $v$ are linearly dependent.

**8.** It $u$ and $v$ are vectors in a real inner product space and if $u + v$ is orthogonal to $u - v$, then prove that $\| u \| = \| v \|$.

**9.** Prove that two vectors $u$ and $v$ in a real inner product space are orthogonal if and only if $\| u + v \|^2 = \| u \|^2 + \| v \|^2$

**10.** Let $V$ be an inner product space, and $u$, $v$ be vectors in $V$. Show that $u = v$ if and only if $< u, w > = < v, w >$ for every $w$ in $V$.

□ □ □

# UNIT 14 : Real Inner Product Space-II

**Structure of the Unit**



## 14.0 Objectives

After reading this unit we will be able to understand results on real inner-product spaces related to orthonormal sets, orthonormal complete sets, adjoint of a linear transformation, self-adjoint linear transformation and the corresponding matrix

## 14.1 Introduction

This unit is a continuation of the previous unit on real inner product spaces. It introduces some important theorems, *e.g.* Bessel's inequality, Parseval's identity and their applications. In this unit we shall study direct sum and their properties, adjoint of a linear transformation, self-adjoint linear transformation and their matrices.

## 14.2 Bessel's inequality

**Theorem 1.** *If* $\{u_1, u_2, ..., u_n\}$ *is any finite orthonormal set in an inner product space V and v is any vector in V, then*

$$\sum_{i=1}^{n} |\langle v, u_i \rangle|^2 \leq \|v\|^2 .$$

*And equality holds if and only if v is in the subspace generated by* $\{u_1, u_2, ..., u_n\}$.

**Proof :** Given $\{u_1, u_2, ..., u_n\}$ is a finite orthonormal set.

$\Rightarrow$ $< u_i, u_j > = \delta_{ij}$, where $\delta_{ij}$ is kronecker's dalta

*i.e.* $$< u_i, u_j > = \begin{cases} 0 & i \neq j \\ 1 & i \neq j \end{cases} \quad \text{.....(1)}$$

Now, consider the vector

$$w = v - \sum_{i=1}^{n} u_i < v, u_i >$$

Evidently $w \in V$.

So $$\| w \|^2 = < w, w >$$

$$= \left\langle v - \sum_{i=1}^{n} u_i < v, u_i >, v - \sum_{j=1}^{n} u_j < v, u_j > \right\rangle$$

$$= < v, v > - \sum_{i=1}^{n} < v, u_i > < u_i, v >$$

$$- \sum_{j=1}^{n} < v, u_j > < v, u_j >$$

$$+ \sum_{i=1}^{n} \sum_{j=1}^{n} < v, u_i > < v, u_j > < u_i, u_j >$$

$$= \| v \|^2 - \sum_{i=1}^{n} | < v, u_i > |^2 - \sum_{j=1}^{n} | < v, u_j > |^2$$

$$+ \sum_{i=1}^{n} \sum_{j=1}^{n} < v, u_i > < v, u_j > \delta_{ij}$$

$$= \| v \|^2 - \sum_{i=1}^{n} | < v, u_i > |^2 - \sum_{j=1}^{n} | < v, u_j > |^2$$

$$+ \sum_{j=1}^{n} | < v, u_j > |^2 \quad \text{[on summing over } i \text{ and using (1)]}$$

$$= \| v \|^2 - \sum_{i=1}^{n} | < v, u_i > |^2$$

Since $\| w \|^2 \geq, 0$, for all $w \in V$, we have

$$\| v \|^2 - \sum_{i=1}^{n} | < v, u_i > |^2 \geq 0$$

Hence $$\sum_{i=1}^{n} | < v, u_i > |^2 \leq \| v \|^2$$

Further, if the equality holds, then

$$\sum_{i=1}^{n} |< v, u_i >|^2 = \| v \|^2,$$

the equality holds iff $\| w \|^2 = 0$

$\Leftrightarrow$ $w = \mathbf{0}$

$\Leftrightarrow$ $v = \sum_{i=1}^{n} u_i < v, u_i >$

$\Leftrightarrow$ $v$ is a linear combination of vector $u_1, u_2, ..., u_n$

$\Leftrightarrow$ $v$ is in the subspace generated by $u_1, u_2, ..., u_n$.

**Cor.** *If $\{u_1, u_2, ..., u_n\}$ is any finite orthogonal set of nonzero vectors in an inner product space $V$ and $v \in V$, then*

$$\sum_{i=1}^{n} \frac{|< v, u_i >|^2}{\| u_i \|^2} \leq \| v \|^2$$

**Proof :** Let $v_i = \frac{u_i}{\| u_i \|}$, $i = 1, 2, ..., n$. Then $\{v_1, ..., v_n\}$ is a finite orthonormal set of nonzero vectors in an inner product space $V$. So by theorem 1, we have

$$\sum_{i=1}^{n} |< v, u_i >|^2 \leq \| v \|^2 \qquad .....(1)$$

Now $$< v, v_i > = \left\langle v, \frac{u_i}{\| u_i \|} \right\rangle$$

$$= \frac{1}{\| u_i \|} < v, u_i > \qquad .....(2)$$

Using (2) in (1), we have

$$\sum_{i=1}^{n} \left| \frac{1}{\| u_i \|} < v, u_i > \right|^2 \leq \| v \|^2$$

Hence $$\sum_{i=1}^{n} \frac{|< v, u_i >|^2}{\| u_i \|^2} \leq \| v \|^2$$

## 14.3 Complete orthonormal set

An orthonormal set is said to be **complete** if it is not contained in any larger orthonormal set.

Thus an orthonormal set $\{v_i\}$ in a inner product space $V$ is complete if it is not possible to adjoin a vector $v$ to $\{v_i\}$ in such a way that $\{v_i, v\}$ is an orthonormal set which properly contains $\{v_i\}$.

**Theorem 2.** (*Parseval's identity*) *If* $A = \{u_1, u_2, ..., u_n\}$ *is any orthonormal set in any finite dimensional inner product space V, then the following are equivalent :*

***(i)*** *orthonormal set A is complete.*

***(ii)*** *if* $u \in V$ *and* $< u, u_i > = 0$ *for* $1 \leq i \leq n$, *then* $u = \mathbf{0}$

***(iii)*** $< A > = V$, *that is A generates V.*

***(iv)*** if $u \in V$ *then* $u = \sum_{i=1}^{n} < u, u_i > u_i$

***(v)*** if $u, v \in V$, *then*

$$< u, v > = \sum_{i=1}^{n} < u, u_i > < v, u_i >$$ (*Parseval's identity*)

***(vi)*** *if* $u \in V$, *then* $\| u \|^2 = \sum_{i=1}^{n} | < u, u_i > |^2$

**Proof :** We shall prove the implications in following way :

*(i)* $\Rightarrow$ *(ii)* $\Rightarrow$ *(iii)* $\Rightarrow$ *(iv)* $\Rightarrow$ *(v)* $\Rightarrow$ *(vi)* $\Rightarrow$ *(i)*

*(i)* $\Rightarrow$ *(ii)* Let orthonormal set $A = \{u_1, u_2, ..., u_n\}$ be complete in $V$. So that there is no non-zero $u \in V$ such that $u \perp u_i$, $1 \leq i \leq n$ and $< u, u > = 1$ *i.e.* $\| u \| = 1$. More over $\{u_1, u_2, ..., u_n\}$ is not an orthonormal set.

Let $u \in V$ be arbitrary vector such that $< u, u_i > = 0$, $1 \leq i \leq n$. We have to show that $u = \mathbf{0}$.

Let if possible $u \neq \mathbf{0}$, then $\frac{u}{\| u \|}$ exists and we have

$$< u, u_i > = 0 \Rightarrow u \perp u_i \Rightarrow \frac{u}{\| u \|} \perp u_i, \quad 1 \leq i \leq n$$

Also $$\left\| \frac{u}{\| u \|} \right\| = 1$$

$\Rightarrow$ $\{u_1, u_2, ..., u_n, \frac{u}{\| u \|}\}$ is an orthonormal set.

$\Rightarrow$ $A$ is not complete, which is a contradiction.

So $$u = \mathbf{0}.$$

*(ii)* $\Rightarrow$ *(iii)* Let $u \in V$ and $< v, u_i > = 0$, $1 \leq i \leq n$, then $u = \mathbf{0}$. We shall show that $A$ generates $V$, *i.e.* each vector in $V$ can be expressed as a linear combination of vectors of $A$. Now take any arbitrary vector $v \in V$ then the vector

$$w = v - \sum_{i=1}^{n} u_i < v, u_i >$$

is orthogonal to each $u_i$, *i.e.* $< w, u_i > = 0$, $1 \leq i \leq n$.

So $w = \mathbf{0}$ by *(ii)*

Consequently $v = \sum_{i=1}^{n} u_i < v, u_i >,$ and $< A > = V.$

*(iii)* $\Rightarrow$ *(iv)* Let $u \in V$, then by *(iii)*, $u$ can be expressed as a linear combination of vector of $A$ *i.e.*

$$u = \sum_{i=1}^{n} \alpha_i u_i, \quad \alpha_i \in R$$

Now $$< u, u_j > = \left\langle \sum_{i=1}^{n} \alpha_i u_i, u_j \right\rangle$$

$$= \sum_{i=1}^{n} \alpha_i < u_i, u_j >$$

$$= \sum_{i=1}^{n} \alpha_i \delta_{ij} = \alpha_j, \qquad 1 \le i \le n$$

for $\{u_1, u_2, ..., u_n\}$ is orthonormal

$\Rightarrow$ $$< u_i, u_j > = \delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases}$$

Hence $$u = \sum_{i=1}^{n} < u, u_i > u_i$$

*(iv)* $\Rightarrow$ *(v)* Suppose $u, v \in V$ are expressible as

$$u = \sum_{i=1}^{n} < u, u_i > u_i$$

$$v = \sum_{j=1}^{n} < v, u_j > u_j$$

Thus $$< u, v > = \left\langle \sum_{i=1}^{n} u_i < u, u_i >, \sum_{j=1}^{n} u_j < v, u_j > \right\rangle$$

$$= \sum_{i} \sum_{j} < u, u_i > < v, u_j > < u_i, u_j >$$

$$= \sum_{i} \sum_{j} < u, u_i > < v, u_j > \delta_{ij}$$

$$= \sum_{i=1}^{n} < u, u_i > < v, u_i >$$

*(v)* $\Rightarrow$ *(vi)* Suppose $u, v \in V \Rightarrow < u, v > = \sum_{i=1}^{n} < u, u_i >< v, u_i >$

Let $u = v$ in *(v)*, we set

$$< u, u > = \| u \|^2 = \sum_{i=1}^{n} < u, u_i >< u, u_i >$$

Hence $$\| u \|^2 = \sum_{i=1}^{n} |< u, u_i >|^2$$

*(vi)* $\Rightarrow$ *(i)* Let $u \in V \Rightarrow \| u \|^2 = \sum_{i=1}^{n} |< u, u_i >|^2$ .....(1)

We shall show that $A$ is complete.

Suppose $A$ is not complete, then there is some $v \in V$ such that $\| v \| = 1$ and $v \perp u_i$, $1 \leq i \leq n$. That is, $< v, u_i > = 0$ for $1 \leq i \leq n$ with these values (1) becomes

$$1^2 = \sum_{i=1}^{n} |0|^2 = 0$$

*i.e.* $1 = 0$, a contradiction.

Thus $A = \{u_1, u_2, ..., u_n\}$ is a complete orthonormal set in $V$.

## 14.4 Direct sum

Let $V$ be a finite dimensional inner product space and $w_1, w_2, ..., w_n$ are subspaces of $V$, then $V$ is called **direct sum** of $w_1, w_2, ..., w_n$ if every vector $v \in V$ can be uniquely expressed as

$$v = w_1 + w_2 + ... + w_n$$

where $w_1 \in W_1, w_2 \in W_2, ..., w_n \in W_n$

The direct sum is denoted by

$$v = W_1 \oplus W_2 \oplus ... \oplus W_n.$$

**Theorem 3.** *Let V be a finite dimensional inner product space and W be its subspace. Then V is direct sum of W and* $W^{\perp}$.

*Symbolically* $V = W \oplus W^{\perp}$.

*OR*

*Let V be a finite dimensional inner product space and W be its subspace. Then prove that every vector* $v \in V$ *can be uniquely expressed as* $v = v_1 + v_2$ *where* $v_1 \in W$ *and* $v_2 \in W^{\perp}$.

**Proof :** Since $W$ is a subspace of finite dimensional inner product space $V$, therefore $W$ itself is a finite dimensional inner product space (its inner product being that of $V$ restricted to $W$). Thus it has an orthonormal basis $\{u_1, u_2, ..., u_m\}$

Let $v \in V$. Then the vector $w = v - \sum_{i=1}^{n} < v, u_i > $ is orthogonal to each $u_i$, for $1 \leq i \leq m$ and therefore it is orthogonal to the subspace $W$.

Clearly $\sum_{i=1}^{m} u_i < v, u_i > \in W$ [being linear combination of basis vector]

Thus for each $v \in V$, we have

$$v = w + \sum_{i=1}^{m} u_i < v, u_i >$$

*i.e.* $v$ is the sum of a vector of $W$ and a vector of $W^{\perp}$.

Thus $V = W + W^{\perp}$ .....(1)

Further, if $u \in W \cap W^{\perp}$, then

$$u \in W \quad \text{and} \quad u \in W^{\perp}$$

$$\Rightarrow \quad < u, u > = 0 \Rightarrow u = \mathbf{0}$$

Hence $W \cap W^{\perp} = \{\mathbf{0}\}$ .....(2)

(1) and (2) implies that

$$V = W \oplus W^{\perp}$$

**Corollary.** If $W$ is any subspace of a finite dimensional inner product space. Then

$$\dim W^{\perp} = \dim V - \dim W.$$

When we have $V = W \oplus W^{\perp}$, then each vector $v \in V$ has a unique representation as $v = v_1 + v_2$ for same $v_1 \in W$ and $v_2 \in W^{\perp}$. In such a case $v_1$ and $v_2$ are called the orthogonal projections of $v$ on the subspaces $W$ and $W^{\perp}$.

**Theorem 4.** *If $W$ is any subspace of a finite dimensional inner product space $V$, then*

$$(W^{\perp})^{\perp} = W$$

**Proof :** Since $W$ is any subspace of $V$, therefore so is $W^{\perp}$. We have

$$V = W \oplus W^{\perp}$$

$$\Rightarrow \quad \dim V = \dim W + \dim W^{\perp} \qquad ...(1)$$

Since $W^{\perp}$ is a subspace of $V$, therefore again $(W^{\perp})^{\perp}$ is also a subspace of $V$ and

$$V = W^{\perp} \oplus (W^{\perp})^{\perp}$$

$$\Rightarrow \quad \dim V = \dim W^{\perp} + \dim (W^{\perp})^{\perp} \qquad .....(2)$$

(1) and (2) implies that

$$\dim W = \dim (W^{\perp})^{\perp} \qquad .....(3)$$

Let $v \in W$. Then $< v, u > = 0$ for all $u \in W^{\perp}$

$$\Rightarrow \quad v \in (W^{\perp})^{\perp} \qquad [\because \ (W^{\perp})^{\perp} = \{v \in V : < v, u > = 0 \ \forall \ u \in W^{\perp}\}]$$

Thus $v \in W \Rightarrow v \ (W^{\perp})^{\perp}$

$\therefore$ $W \subset (W^{\perp})^{\perp}$ .....(4)

Since $W$ and $(W^{\perp})^{\perp}$ are subspaces of $V$ and $W \subset (W^{\perp})^{\perp}$.

Consequently $W$ is a subspace of $(W^{\perp})^{\perp}$ and dim $W$ = dim $(W^{\perp})^{\perp}$, therefore

$$W = (W^{\perp})^{\perp}$$

***Ex.*1.** *If W is a subspace of an inner product space* $R^3$ *spanned by*

$$B_1 = \{(1, 0, 1), (1, 2, -2)\},$$

*then find a basis of orthogonal complement* $W^{\perp}$.

**Sol.** Let $u_1 = (1, 0, 1)$, $u_2 = (1, 2, -2)$. Clearly $B_1$ is a basis for $W$. Thus dim $W = 2$.

Since $\dim W + \dim W^{\perp} = \dim V$,

therefore $\dim W^{\perp} = 1$.

Let $B_1' = \{v_1\} = \{(a, b, c)\}$ be basis for $W^{\perp}$.

Then it should be orthogonal to $B_1$. Therefore we have

$$< u_1, v_1 > = a + c = 0$$

and $$< u_2, v_1 > = a + 2b - 2c = 0$$

Solving there two equations, we set

$$a = -c \quad \text{and} \quad b = \frac{3}{2}c$$

Thus $B_1' = \left\{\left(-c, \frac{3}{2}c, c\right)\right\}$ is the basis of $W^{\perp}$.

**Theorem 5.** *Let V be a finite dimensional inner product space. If* $A = \{u_1, u_2, ..., u_n\}$ *is an orthonormal basis of a subspace W of V, and* $B = \{v_1, v_2, ..., v_m\}$ *is an orthonormal basis of* $W^{\perp}$. *Then* $\{u_1, ..., u_n, v_1, ..., v_m\}$ *is an orthonormal basis of V.*

**Proof :** Since $A = \{u_1, u_2, ..., u_n\}$ is an orthonormal basis of $W$,

$$< u_i, u_j > = \delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases} \quad \text{for } 1 \leq i \leq n.$$

Also $B = \{v_1, v_2, ..., v_m\}$ is an orthonormal basis of $W^{\perp}$, so

$$< v_r, r_s > = \delta_{rs} = \begin{cases} 1 & \text{if} \quad r = s \\ 0 & \text{if} \quad r \neq s \end{cases} \quad \text{for } 1 \leq r, s \leq m.$$

$$W, W^{\perp} \subset V \Rightarrow u_i, v_r \in V \quad \text{for} \quad 1 \leq i \leq n, 1 \leq r \leq n,$$

since $W^{\perp} = \{v \in V : < u, v > = 0 \ \forall \ u \in W\}$

$\therefore$ any $v \in W^{\perp} \Rightarrow < u, v > = 0 \ \forall \ u \in W$

$\Rightarrow$ $< u_i, v > = 0 \quad \text{for} \quad 1 \leq i \leq n.$

This is true for every $v \in W^{\perp}$.

Thus $< u_i, v_r > = 0 \quad \text{for} \quad 1 \leq i \leq n, 1 \leq r \leq n.$

Thus $\{u_1, u_2, ..., u_n, v_1, v_2, ..., v_m\}$ is an orthonormal set of $(n + m)$ vectors in $V$. It is linearly

independent because it is orthonormal set of vectors. We have, now

$$\dim V = \dim W + \dim W^{\perp}$$

$$= n + m.$$

Therefore $\{u_1,...,u_n, v_1,...,v_m\}$ is an orthonormal basis of $V$.

## 14.5 Adjoint of a linear transformation

Let $V$ and $V'$ be finite dimensional inner product spaces and $t : V \to V'$ be a linear transformation. Then a linear transformation $t^* : V' \to V$ is called **adjoint** of $t$ if

$$< t^* (v') > = < v', t (v) > \text{ for all } v \in V \text{ and } v' \to V'.$$

**Theorem 6.** *If $t_1$ and $t_2$ are linear transformations of finite dimensional inner product spaces $V$ to $V'$,*

*then* $$(t_1 + t_2)^* = t_1^* + t_2^*.$$

**Proof :** Since $t_1 : V \to V'$ and $t_2 : V \to V'$ are linear transformations. Therefore so is $t_1 + t_2 : V \to V'$.

For any $v \in V$ and $v' \to V'$, we have

$$< v', (t_1 + t_2)\, v > = < v', t_1 (v) + t_2 (v) >$$

$$= < v', t_1 (v) > + < v', t_2 (v) >$$

$$= < t_1^* (v'), v > + < t_2^* (v'), v >$$

$$= < t_1^* (v') + t_2^* (v'), v >$$

$$= < (t_1^* + t_2^*) (v'), v > \quad .....(1)$$

But $$< v', (t_1 + t_2) (v) > = < (t_1 + t_2)^* (v'), v > \quad .....(2)$$

Thus from (1) and (2), we have

$$(t_1 + t_2)^* = t_1^* + t_2^*$$

**Theorem 7.** *If $t_1$ and $t_2$ are linear transformations of a finite dimensional inner product spaces $V$. Then*

$$(t_1\, t_2)^* = t_2^*\, t_1^*.$$

**Proof :** Since $t_1 : V \to V$ and $t_2 : V \to V$ are linear transformations. Therefore so is $t_1\, t_2 : V \to V$.

For any $u, v \in V$, we have

$$< (t_1\, t_2) (u), v > = < t_1 (t_2 (u)), v >$$

$$= < t_2 (u), t_1^* (v) >$$

$$= < u, t_2^* (t_1^* (v) >$$

$$= < u, (t_2^*\, t_1^*) (v) > \quad .....(1)$$

But $$< (t_1\, t_2) (u), v > = < u, (t_1\, t_2)^* (v) > \quad .....(2)$$

So from (1) and (2), we have

$$(t_1\ t_2)^* = t_2^*\ t_1^*.$$

**Adjoint operator :**

Let $V$ be a finite dimensional inner product space and $t$ be a linear operator on $V$. Then a linear operator $t^*$ on $V$ is said to be an **adjoint operator** of $t$ if

$$< t(u),\ v > = < u,\ t^*(v) > \quad \text{for all} \quad u,\ v \in V.$$

Recall that a linear transformation $t : V \to V$ is called a **linear operator**, and we say $t$ is a linear operator on $V$.

**Theorem 8.** *Let $V$ be a finite dimensional inner product space. Let $t : V \to V$* be a linear transformation then there exists a unique linear transformation

*then* $t^* : V \to V$ *such that*

$$< t(u),\ v > = < u,\ t^*(v) >, \quad \text{for all} \quad u,\ v \in V.$$

**Proof :** Let $\{b_1, b_2, ..., b_n\}$ be an orthonormal basis of $V$. For $v \in V$, choose

$$t^*(v) = \sum_{i=1}^{n} < t(b_i),\ v > b_i$$

then
$$< b_k,\ t^*(v) > = \left\langle b_k, \sum_{i=1}^{n} < t(b_i),\ v > b_i \right\rangle$$

$$= \sum_{i=1}^{n} < t(b_i),\ v > < b_k,\ b_i >$$

$$= \sum_{i=1}^{n} < t(b_i),\ v > \delta_{ki}, \quad \text{where} \quad \delta_{ki} = \begin{cases} 1 & k = i \\ 0 & k \neq i \end{cases}$$

$$= < t(b_k),\ v >, \qquad \text{.....(1)}$$

Now for any $u \in V$, where

$$u = \sum_{i=1}^{n} \alpha_i\ b_i, \quad \alpha_i \in R, \text{ we have}$$

$$< t(u),\ v > = \left\langle t\left( \sum_{i=1}^{n} \alpha_i\ b_i \right), v \right\rangle$$

$$= \left\langle \sum_{i=1}^{n} \alpha_i\ t(b_i), v \right\rangle$$

$$= \sum_{i=1}^{n} \alpha_i < t(b_i),\ v >$$

$$= \sum_{i=1}^{n} \alpha_i < b_i,\ t^*(v) > \qquad \text{[by (1)]}$$

$$= \left\langle \sum_{i=1}^{n} \alpha_i b_i,\ t^*(v) \right\rangle$$

$$= < u, t^* (v) >$$

Thus $t^*$ exists.

To show that $t^*$ is unique, if possible let there be $t_1^*$ and $t_2^*$ such that

$$< t (u), v > = < u, t_1^* (v) >$$

and $$< t (u), v > = < u, t_2^* (v) > \quad \text{for all} \quad u, v \in V.$$

This implies that

$$< u, t_1^* (v) - < u, t_2^* (v) > = 0$$

$$\Rightarrow \quad < u, t_1^* (v) - t_2^* (v) > = 0, \quad \text{for all} \quad u \in V.$$

$$\Rightarrow \quad < t_1^* (v) - t_2^* (v), t_1^* (v) - t_2^* (v) > = 0 \quad \text{for all} \quad u = t_1^* (v) - t_2^* (v)$$

$$\Rightarrow \quad t_1^* (v) - t_2^* (v) = 0, \quad \text{for all} \quad v \in V.$$

$$\Rightarrow \quad t_1^* = t_2^*$$

Finally, to show that $t^*$ is linear, let $u, v, w \in V$ and $\alpha, \beta \in R$.

$$< u, t^* (\alpha v + \beta w) = < t (u), \alpha v + \beta w >$$

$$= \alpha < t (u), v > + \beta < t (u), w >$$

$$= \alpha < u, t^* (v) > + \beta < u, t^* (w) >$$

$$= < u, \alpha t^* (v) > + < u, \beta t^* (w) >$$

$$= < u, \alpha t^* (v) + \beta t^* (w) >, \quad \text{for all} \quad u \in V.$$

Thus $$t^* (\alpha v + \beta w) = \alpha t^* (v) + \beta t^* (w), \quad \text{for all} \quad u, v \in V.$$

It follows that $t^*$ is linear.

**Theorem 9.** *It t is a linear map on an inner product space V, then* $t = \hat{0}$ *(zero map) if and only if* $< t (u), v > = 0$, *for all* $u, v \in V$.

**Proof :** Let $t = \hat{0}$ (zero map), then

$$< (t (u), v) > = < \hat{0} (u), v > = < \mathbf{0}, v > = 0$$

conversely, let $< t (u), v > = 0$ for all $u, v \in V$.

Then $< t (u), t (u) > = 0$ for $v = t (u)$

$$\Rightarrow \quad t (u) = 0 \quad \text{for all} \quad u \in V.$$

$$\Rightarrow \quad t = \hat{0}$$

## 14.6 Self-adjoint linear transformation and matrices

Let $V$ and $V'$ be finite dimensional inner product spaces. A linear transformation $t : V \to V$ (linear operator on $V$) is said to be **self-adjoint** if

$$t = t^*$$

*i.e.* $\quad < t(u), v > = < u, t(v) > \quad$ for all $\quad u, v \in V.$

A self-adjoint operator is called **symmetric** or Hermitian according as the space is real inner product space $V(R)$ or unitary space $V(C)$.

The identity transformation and the null transformation are symmetric transformations.

A square matrix $A = [a_{ij}]$ over $R$ is said to be symmetric if $A^T = A$. A linear transformation $t : V \to V$, where $V$ is a finite dimensional real inner product space is called **skew-symmetric** if $t^* = -t$. It is called **skew-Hermitian** if the space is unitary $V(C)$.

A square matrix $A = [a_{ij}]$ over $R$ is said to be **skew-symmetric** if $A^T = -A$.

If $t$ is a linear operator on a unitary space $V(C)$, such that $t\, t^* = t^*\, t$, then $t$ is called **normal operator.**

**Theorem 10.** *A linear transformation t from a finite dimensional inner product space V into itself is symmetric if and only if its matrix* $A = [a_{ij}]$ *relative to some orthonormal basis B of V is symmetric.*

**Proof :** Let $B = \{u_1, u_2, ..., u_n\}$ be an orthonormal basis of $V$. Since $A = [a_{ij}]$ is the matrix of $t$ relative to basis $B$, therefore

$$t(u_i) = \sum_{k=1}^{n} a_{ki}\, u_k, \qquad 1 \le i \le n.$$

First suppose that $t : V \to V$ is a symmetric transformation. Then we have

$$< t(u), v > = < u, t(u) > \quad \text{for all} \quad u, v \in V.$$

$$\Rightarrow \qquad < t(u_i), u_j > = < u_i, t(u_j) > \quad \text{for} \quad i, j = 1, 2, ..., n.$$

Since $B$ is the basis of $V$.

$$\Rightarrow \qquad \left\langle \sum_{k=1}^{n} a_{ki}\, u_k, u_j \right\rangle = \left\langle u_i, \sum_{k=1}^{n} a_{kj}\, u_k \right\rangle$$

$$\Rightarrow \qquad \sum_{k=1}^{n} a_{ki} < u_k, u_j > = \sum_{k=1}^{n} a_{kj} < u_i\, u_k >$$

$$\Rightarrow \qquad \sum_{k=1}^{n} a_{ki}\, \delta_{kj} = \sum_{k=1}^{n} a_{kj}\, \delta_{ik}$$

$$\Rightarrow \qquad a_{ji} = a_{ij} \quad \text{for all} \quad i, j$$

$$\Rightarrow \qquad A^T = A$$

$\Rightarrow$ $A$ is symmetric.

Conversely, let $A = [a_{ij}]$ be a symmetric matrix. Then

$$A^T = A$$

$$\Rightarrow \qquad a_{ji} = a_{ij} \quad \text{for all} \quad i, j = 1, 2, ..., n.$$

Now, since $A$ is the matrix corresponding to the linear transformation $t : V \to V$ relative to the basis $B$, therefore, we have

$$< t(u_i), u_j > = \left\langle \sum_{k=1}^{n} a_{ki}\, u_k, u_j \right\rangle$$

$$= \sum_{k=1}^{n} a_{ki} < u_k, u_j >$$

$$= \sum_{k=1}^{n} a_{ki}\, \delta_{kj}$$

$$= a_{ji}$$

and $$< u_i, t(u_j) > = \left\langle u_i, \sum_{k=1}^{n} a_{kj}\, u_k \right\rangle$$

$$= \sum_{k=1}^{n} a_{kj} < u_i, u_k >$$

$$= \sum_{k=1}^{n} a_{kj}\, \delta_{ik}$$

$$= a_{ij}$$

Thus, $< t(u_i), u_j > = < u_i, t(u_j) >$ for all $i, j = 1, 2, ..., n$.

Since $B = \{u_1, u_2, ..., u_n\}$ is the basis of $V$, hence

$$< t(u), v > = < u, t(v) > \quad \text{for all} \quad u, v \in V$$

It follows that $t$ is symmetric in $V$.

**Theorem 11.** *A linear transformation t from a finite dimensional inner-product space V to itself is skew symmetric iff they commute.*

**Proof :** Let $t_1$ and $t_2$ be two self-adjoint transformations on an inner-product space $V$.

Let the product $t_1\, t_2$ be a self-adjoint transformation. Then

$$(t_1\, t_2)^* = t_1\, t_2$$

$$\Rightarrow \quad t_2^*\, t_1^* = t_1\, t_2$$

$$\Rightarrow \quad t_2\, t_1 = t_1\, t_2$$

$\Rightarrow$ $t_1$ and $t_2$ commute.

Conversely suppose that $t_1$ and $t_2$ commute with each other *i.e.*

$$t_1\, t_2 = t_2\, t_1$$

$$\Rightarrow \quad (t_1\, t_2)^* = (t_2\, t_1)^*$$

$$\Rightarrow \quad (t_1\, t_2)^* = t_1^*\, t_2^*$$

$$\Rightarrow \quad (t_1\, t_2)^* = t_1\, t_2$$

$\Rightarrow$ $t_1\, t_2$ is a self adjoint transformation.

## Self-learning exercise-1

**1.** In an inner product space $V$, a set of orthogonal vectors is always linearly dependent. (T/F)

**2.** In an inner product space $V$, a set of orthonormal vector is always linearly independent. (T/F)

**3.** Orthonormal vector in an inner product space are unit vectors which are not orthogonal to each other. (T/F)

**4.** Let $A = \{v_1, v_2, ..., v_n\}$ be an orthonormal basis of an inner product space $V$, then

$$< v_i, v_j > = .... \quad \text{for} \quad i, j = 1, 2, ..., n.$$

**5.** If an inner product space $V$ is direct sum of its subspaces $S$ and $S^{\perp}$, then

$$S \cap S^{\perp} = ....$$

**6.** Let $t$ be a linear transformation on an inner product space $V$.

If $< t(u), v > = 0$ for all $u, v \in V$

then $t = ...$ .

## More illustrative results

**Theorem 12.** *If $M$ and $N$ are subspaces of a finite dimensional inner product space $V$. Then*

$$(M + N)^{\perp} = M^{\perp} \cap N^{\perp}.$$

**Proof :** Any $z \in (M + N)^{\perp} \Rightarrow < z, u > = 0 \quad \forall \ u \in M + N$

$\Rightarrow \quad < z, x + y > = 0 \quad \forall \ x + y \in M + N$

$\Rightarrow \quad < z, x > + < z, y > = 0 \quad$ where $\quad x \in M, y \in N$

$\Rightarrow \quad < z, x > = 0 = < z, y > \quad \forall \ x \in M \quad$ and $\quad \forall \ y \in N$ for $< u, v > \geq 0 \ \forall \ u, v \in V$

$\Rightarrow \quad z$ is orthogonal to $M$ and $N$ both

$\Rightarrow \quad z \perp M \quad$ and $\quad z \perp N$

$\Rightarrow \quad z \in M^{\perp} \quad$ and $\quad z \in N^{\perp}$

$\Rightarrow \quad z \in M^{\perp} \cap N^{\perp}$

Thus any $z \in (M + N)^{\perp} \Rightarrow z \in M^{\perp} \cap N^{\perp}$

$\Rightarrow \quad (M + N)^{\perp} \subset M^{\perp} \cap N^{\perp}$ .....(1)

Conversely, let $z \in M^{\perp} \cap N^{\perp} \Rightarrow z \in M^{\perp}$ and $z \in N^{\perp}$

$\Rightarrow \quad z$ is orthogonal to $M$ and $N$ both

$\Rightarrow \quad (z, x) = 0 = (z, y) \quad \forall \ x \in M \quad$ and $\quad y \in N$

$\Rightarrow \quad < z, x > + < z, y > = 0 + 0 = 0 = < z, x + y > = 0 \quad \forall \ x + y \in M + N$

$\Rightarrow \quad Z$ is orthogonal to $M + N$

$\Rightarrow \quad z \in (M + N)^{\perp}$

$\therefore \quad z \in M^{\perp} \cap N^{\perp} \Rightarrow z \in (M + N)^{\perp}$

$\Rightarrow \quad M^{\perp} \cap N^{\perp} \subset (M + N)^{\perp}$ .....(2)

Combining (1) and (2), we get

$$(M + N)^{\perp} = M^{\perp} \cap N^{\perp}.$$

**Theorem 13.** *Let M and N are subspaces of a finite dimensional inner product space V, then*

$$(M + N)^{\perp} = M^{\perp} + N^{\perp}$$

**Proof :** We know that

$$(M + N)^{\perp} = M^{\perp} \cap N^{\perp} \quad \text{.....(1)}$$

Since $V$ is finite dimensional inner product space and $M$, $N$ be its subspace. Therefore we have

$$M^{\perp\perp} = M$$

$$N^{\perp\perp} = N \quad \text{.....(2)}$$

Replacing $M$ and $N$ in (1) by $M^{\perp}$ and $N^{\perp}$ respectively, we get

$$(M^{\perp} + N^{\perp})^{\perp} = M^{\perp\perp} \cap N^{\perp\perp}$$

$\Rightarrow \quad (M^{\perp} + N^{\perp})^{\perp} = M \cap N$ using (2)

$\Rightarrow \quad ((M^{\perp} + N^{\perp})^{\perp})^{\perp} = (M \cap N)^{\perp}$

$\Rightarrow \quad M^{\perp} + N^{\perp} = (M \cap N)^{\perp}$ using (2)

**Theorem 14.** *Let t be a linear transformation on inner product space V. Show that*

$$t = 0 \Leftrightarrow < t(u), v > = 0 \quad \forall \ u, v \in V.$$

**Proof :** Suppose $t$ is a linear transformation on an inner product space $V$. Suppose $u, v \in V$.

Suppose $t = 0$

$\Rightarrow \quad < t(u), v) = < 0(u), v > = < 0, v > = 0$

Conversely let $< t(u), v > = 0 \quad \forall \ u, v \in V$

$\Rightarrow \quad < t(u), t(u) > = 0, \quad$ where $\quad v = t(u)$

$\Rightarrow \quad t(u) = 0 \quad \forall \ u \in V$

$\Rightarrow \quad t = 0$

**Theorem 15.** *Let V and V′ be finite dimensional inner product spaces. If t be a linear transformation from V to V′ and* $\alpha$ *is a scalar. Then*

***(i)*** $(at)^* = a\, t^*$

***(ii)*** $(t^*)^* = t.$

**Proof :** for any $u, v \in V$, we have

***(i)*** $< (at)(u), v > = < at(u), v >$

$= a < t(u), v >$

$$= a < u, t^*(v) >$$

$$= < u, at^*(v) >$$

$$= < u, (at^*)\, v > \quad \text{.....(1)}$$

But $< (at)(u), v > = < u, (at)^*\, v > \quad$ .....(2)

Thus from (1) and (2), we have

$$< u, (at^*)\, v > = < u, (at)^*\, v >$$

Uniqueness of adjoint

$$\Rightarrow \quad at^* = (at)^*$$

***(ii)*** $\because \quad < u, t(v) > = < t^*(u), v >$

Also $\quad < t^*(u), v > = < u, (t^*)^*\, v >$

$$\Rightarrow \quad < u, t(v) > = < u, (t^*)^*\, v >$$

$$\Rightarrow \quad t = (t^*)^*.$$

**Theorem 16.** *Show that if $t : V \to V$ is a self adjoint linear transformation on an inner product space V, then $s^*\, t\, s$ is self-adjoint for every linear transformation $s : V \to V$. Further if s is invertible and $s^*\, t\, s$ is self adjoint, then t is self adjoint.*

**Proof :** Let $t$ be self adjoint so that

$$t = t^* \quad \text{.....(1)}$$

$\because \quad (s^*\, t\, s) = s^*\, t^*\, s^{**}$

$$= s^*\, t^*\, s$$

$$= s^*\, t\, s \quad \text{using (1)}$$

Hence $s^*\, t\, s$ is self adjoint. Again let $s$ be invertible, so that $s^{-1}$ and $(s^*)^{-1}$ exist.

Since $s^*\, t\, s$ is self adjoint, therefore

$$(s^*\, ts)^* = s^*\, ts$$

$$\Rightarrow \quad s^*\, t^*\, s^{**} = s^*\, ts$$

$$\Rightarrow \quad s^*\, t^*\, s = s^*\, ts$$

$$\Rightarrow \quad s^*\, t^*\, (ss^{-1}) = s^*\, t\, (ss^{-1})$$

$$\Rightarrow \quad s^*\, t^*\, I = s^*\, t\, I$$

$$\Rightarrow \quad (s^*)^{-1}\, s^*\, t^* = (s^*)^{-1}\, s^*\, t$$

$$\Rightarrow \quad I\, t^* = I\, t$$

$$\Rightarrow \quad t^* = t$$

$\Rightarrow$ $t$ is self adjoint.

**Theorem 17.** *If both t and s are self adjoint linear transformation on an inner product space V, then ts + st is self adjoint. It both t and s are skew-adjoint, then ts – st is skew-adjoint.*

**Proof :** Let $t$ and $s$ both the self adjoint so that

$$t^* = t, \quad s^* = s. \qquad \text{.....(1)}$$

$$\therefore \quad (ts + st)^* = (ts)^* + (st)^*$$

$$= s^* t^* + t^* s^*$$

$$= st + ts$$

$$\therefore \quad (ts + st)^* = ts + st$$

$\Rightarrow$ $ts + st$ is self adjoint. Again let $t$ and $s$ both be skew-adjoint so that

$$t^* = -t, \quad s^* = -s$$

$$\Rightarrow \quad (ts - st)^* = (ts)^* - (st)^*$$

$$= s^* t^* - t^* s^*$$

$$= st - ts$$

$$= -(ts - st)$$

$$\therefore \quad (ts - st)^* = -(ts - st)$$

$$\Rightarrow \quad s^* t^* s = s^* t s$$

$\Rightarrow$ $(ts - st)$ is skew-adjoint.

## 14.7 Summary

In this unit we have studied some important results, such as Bessel's inequality, Parseval's identity, and that a finite dimensional inner product is the direct sum of its subspace and its orthogonal complement. We also studied self-adjoint linear transformation between finite dimensional inner product spaces and the results on the corresponding matrix.

## 14.8 Answers to self-learning exercises

### Self learning exercise-1

**1.** false **2.** true **3.** false

**4.** $\delta_{ij}$ **5.** $\{\mathbf{0}\}$ **6.** zero map

## 14.9 Exercises

**1.** Let $T$ be a symmetric linear transformation on an inner product space $V$. Show that

$$< T(\alpha), \alpha > = 0 \Leftrightarrow T = 0 .$$

**2.** If $T$ and $S$ are self-adjoint operators on an inner product space $V$, then $TS$ is self-adjoint

$$\Leftrightarrow TS = ST.$$

**3.** If $T$ is self-adjoint operator on an inner product space $V(K)$, then $\alpha T$ is self-adjoint, $\alpha \in K \Leftrightarrow$ $\alpha$ real

**4.** If $T$ is a self-adjoint linear transformation on an inner product space $V$ and if $T^2(u) = 0$, then $T(u) = 0$.

**5.** Suppose $< T(u), v > = 0$ for every $u, v \in V$. Show that $T = 0$

**6.** Show that $T^* T$ and $TT^*$ are self-adjoint for any operator $T$ on $V$.

**7.** Show that $T + T^*$ is self-adjoint for any operator $T$ on $V$.

**8.** Show that any operator $T$ is the sum of a self-adjoint operator and a skew-adjoint operator.

□ □ □

# UNIT 15 : Real Inner Product Space-III

**Structure of the Unit**



## 15.0 Objectives

This unit is in continuation of the earlier units on real inner-product spaces. After reading this unit we will be able to understand the importance of orthogonal linear transformation which preserves the length and angle between two vectors, the corresponding orthogonal matrix and a very important result, known as the principal axis theorem.

## 15.1 Introduction

This unit introduces the concept of orthogonal linear transformation from an inner product-space to an inner-product space, and many results based on this concept. We shall also study orthogonal matrices and the theorems on orthogonal linear transformations and the orthogonal matrices. We shall also study the results on eigenvalues of a self-adjoint linear transformations and the principal axis theorem.

## 15.2 Orthogonal linear transformation

A linear transformation $t : V \to V'$ from an inner product space $V$ into an inner product space $V'$ is said to **orthogonal if**

$$< t(u), t(v) > = < u, v > \text{ for all } u, v \in V.$$

**Theorem 1.** *Let $V$ and $V'$ be inner product spaces. Then every orthonormal linear transformation $t : V \to V'$ preserves the length and angle between two vectors.*

**Proof.** For any $u, v \in V$ we have

$$< t(u), t(v) > = < u, v >$$

In particular, $$< t(u), t(u) > = < u, u >$$

$\Rightarrow$ $\| t(u) \|^2 = \| u \|^2$

$\Rightarrow$ $\| t(u) \| = \| u \|$ for all $u \in V$.

Thus an orthogonal linear transformation preserves the length.

If f is the angle between two vectors $u$ and $v$ of $V$, then

$$\cos \theta = \frac{\langle u, v \rangle}{\| u \| \| v \|}$$

$$\text{or } \cos \theta = \frac{\langle t(u), t(v) \rangle}{\| t(u) \| \| t(v) \|}$$

Therefore an orthogonal linear transformation preserves the angle also.

**Theorem 2.** *Let $V$ and $V'$ be inner product spaces. Then every orthogonal linear transformation $t : V \to V'$ is a monomorphism of vector spaces.*

**Proof.** In order to prove that $t : V \to V'$ is a monomorphism, it is sufficient to prove that $Ker\ (t) = \{\mathbf{0}\}$.

Let $v \in Ker\ (t)$. Then $t\ (v) = \mathbf{0} \in V'$.

Now for any $u \in V$, we have

$$< u, v > = < t\ (u), t\ (v) > \qquad (\because\ t \text{ is orthogonal})$$
$$= < t\ (u), \mathbf{0} >$$
$$= 0$$

Thus if $v \in Ker\ (t) \Rightarrow < u, v > = 0$ for all $u \in V$.

In particular, $< v, v > = 0 \Rightarrow v = \mathbf{0}$

Hence $v \in Ker\ (t) \Rightarrow v = \mathbf{0}$

Therefore $Ker\ (t) = \{\mathbf{0}\}$.

Hence $t$ is a monomorphism.

The following theorem proves that an orthogonal linear transformation carries an orthogonal list of vectors to an orthogonal list.

**Theorem 3.** *Let $V$ and $V'$ be inner-product spaces. If $(u_1, u_2,...,u_n)$ is an orthonormal list of vectors in $V$. Then the list $(t\ (u_1), t\ (u_2),..., t\ (u_n))$ is orthonormal in $V'$, where $t : V \to V'$ be an orthogonal linear transformation.*

**Proof.** Since the list $(u_1, u_2,...,u_n)$ is an orthonormal list in $V$, therefore

$$< u_i, u_j > = \delta_{ij} = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases} \qquad \text{for } i, j = 1, 2,...,n.$$

Since $t$ is orthogonal, therefore

$$< t\ (u_i), t\ (u_j) > = < u_i, u_j > \qquad \text{for } i, j = 1, 2,...,n.$$
$$= \delta_{ij}$$

Therefore $(t\ (u_1),..., t(u_n))$ is an orthonormal list of vectors in $V'$.

**Theorem 4.** *Let $V$ and $V'$ be inner-product spaces. Then a linear transformation $t : V \to V'$ is orthogonal if and only if*

$$|| t(u) || = || u || \quad \text{for all} \quad u \in V.$$

**Proof.** Let $t : V \to V'$ be an orthogonal linear transformation. Then

$$< t(u), t(v) > = < u, v > \text{ for all } u, v \in V.$$

Thus $< t(u), t(u) > = < u, u >$

$$|| t(u) ||^2 = || (u) ||^2$$

Hence $|| t(u) || = || u ||$, for all $u \in V$.

Conversely, let $t : V \to V'$ be a linear transformation such that $|| t(u) || = || u ||$ for all $u \in V$. Then we have to show that $t$ is an orthogonal linear transformation.

Let $u, v \in V$. Then

$$|| t(u + v) || = || u + v ||$$

$$\Rightarrow \quad || t(u + v) ||^2 = || u + v ||^2$$

$$\Rightarrow \quad < t(u + v), t(u + v) > = < u + v, u + v >$$

$$\Rightarrow \quad < t(u) + t(v), t(u) + t(v) > = < u + v, u + v >, \text{ since } t \text{ is linear.}$$

$$\Rightarrow \quad < t(u), t(u) > + < t(v), t(v) > + 2 < t(u), t(v) > = < u, u > + < v, v > + 2 < u, v >$$

$$\Rightarrow \quad || t(u) ||^2 + || t(v) ||^2 + 2 < t(u), t(v) > = || u ||^2 + || v ||^2 + 2 < u, v >$$

$$\Rightarrow \quad < t(u), t(v) > = < u, v > \text{ for all } u, v \in V.$$

Hence $t$ is an orthogonal transformation.

**Theorem 5.** *The composite of two orthogonal transformations, when defined, is an orthogonal transformation.*

**Proof.** Let $V$, $V'$ and $V''$ be finite dimensional inner product spaces and $t : V \to V'$ and $s' : V' \to V''$ be orthogonal linear transformations;

Then for any $u, v \in V$, we have

$$< (sot)(u), (sot)(v) > = < st(u), s\,t(v) >$$

$$= < t(u), t(v) > \qquad (\because s \text{ is orthogonal})$$

$$= < u, v > \qquad (\because t \text{ is orthogonal})$$

Hence sot is an orthogonal transformation.

**Theorem 6.** *The inverse of an orthogonal linear transformation, when defined, is an orthogonal transformation.*

**Proof.** Let $V$ be a finite dimensional inner product space and $t : V \to V$ be an orthogonal linear transformation. By theorem 2, $t$ is a monomorphism and hence it is an isomorphism. Therefore $t$ is invertible.

Let $u, v \in V$, we have

$\Rightarrow \quad < tot^{-1}(u), tot^{-1}(v) > = < t(t^{-1}(u), t(t^{-1}(v)) >$

$\Rightarrow \quad < u, v > = < t^{-1}(u), t^{-1}(v) >$ $(\because\ t$ is orthogonal)

Hence $t^{-1}$ is an orthogonal transformation.

**Theorem 7.** *Let V be a finite dimensional inner product space. Then the set of all orthogonal transformation on V* (*all automorphism of the inner-product space*) *is a group.*

**Proof.** Let $A(V)$ be the set of all automorphism of the inner-product space $V$. By theorem 5, for any $t_1, t_2 \in A(V)$, $(t_1 o t_2) \in A(V)$. Therefore $A(V)$ is closed.

Since orthogonal transformation are functions and composition of functions is always associative. Therefore, for any

$$t_1, t_2, t_3 \in A(V) \qquad (t_1 o t_2) o t_3 = t_1 o (t_2 o t_3) \qquad .....(1)$$

Since

$$< I_u(u). I_u(u) > = < u, u > \qquad \text{for all } u \in V.$$

Hence the identify map $I_u$ is an orthogonal transformation. Therefore

$$I_u \in A(V) \qquad .....(2)$$

$t \in A(V) \Rightarrow t$ is an orthogonal transformation on $V$

$\Rightarrow t$ is a monomorphism from $V$ into itself [By theorem 2]

$\Rightarrow\ t$ is an isomorphism $[\because\ t : V \to V]$

$\Rightarrow\ t^{-1} : V \to V$ exists.

For any $u, v \in V$

$< u, v > = < t\, t^{-1}(u), t\, t^{-1}(v) >$

$= < t(t^{-1}(u), t(t^{-1}(v)) >$

$= < t^{-1}(u), t^{-1}(v) >$ $(\because\ t$ is orthogonal)

$\therefore \quad t^{-1}$ is orthogonal transformation.

$\Rightarrow \quad t^{-1} \in A(V)$ .....(3)

Hence $A(V)$ is a group. $A(V)$ is called orthogonal group of $V$.

**Theorem 8.** *Let* $B = \{u_1, u_2, ..., u_n\}$ *be an orthonormal basis of an inner product space V. Then a linear transformation t from V to an inner-product space V′ is orthogonal if and only if the set* $\{t(u_1), t(u_2), ..., t(u_n)\}$ *is orthogonal in V′.*

**Proof.** First suppose that $t : V \to V'$ is an orthogonal transformation. Then we have to prove that $\{t(u_1), t(u_2), ..., t(u_n)\}$ is an orthogonal set.

Let $u, v \in V$, then

$$< t(u), t(v) > = < u, v >$$

Thus for all $i, j = 1, 2,...,n$.

$$< t(u_i), t(u_j) > = < u_i, u_j > = \delta_{ij}.$$

Since $B$ is an orthogonal basis of $V$.

Hence $\{t(u_1), t(u_2),..., t(u_n)\}$ is an orthogonal set in $V'$.

Conversely, let $t$ be a linear transformation such that $\{t(u_1), t(u_2),..., t(u_n)\}$ is an orthogonal set in $V'$. Then we have to show that $t$ is an orthogonal transformation. In order to prove this it is sufficient to show that $\| t(u) \| = \| u \|$ for all $u \in V$. Let $u \in V$ be any vector, so for some $\alpha_i \in R$, we have

$$u = \sum_{i=1}^{n} \alpha_i \, u_i$$

$$\Rightarrow \qquad t(u) = \sum_{i=1}^{n} \alpha_i \, t(u_i)$$

Now, $\| u \|^2 = < u, u >$

$$= \left\langle \sum_{i=1}^{n} \alpha_i \, u_i, \sum_{j=1}^{n} \alpha_j \, u_j \right\rangle$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \, u_j \left\langle u_i, u_j \right\rangle$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \, \alpha_j \delta_{ij}$$

$$= \sum_{i=1}^{n} \alpha_i^{\,2} \qquad \text{(on summing over } j\text{)} \qquad .....(1)$$

and $\| t(u) \|^2 = < t(u), t(u) >$

$$= \left\langle \sum_{i=1}^{n} \alpha_i \, t(u_i), \sum_{j=1}^{n} \alpha_j \, t(u_j) \right\rangle$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \, \alpha_j \left\langle t(u_i), t(u_j) \right\rangle$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j \, \delta_{ij}$$

$$= \sum_{i=1}^{n} \alpha_i^2 \qquad \text{(on summing over } j\text{)} \qquad .....(2)$$

So from (1) and (2), we have $\| t(u) \| = \| u \|$ for all $u \in V$.

Hence $t$ is orthogonal.

**Theorem 9.** *If $t : V \to V$ is any map from an inner-product space $V$ to itself such that*

*(i)* $t(\mathbf{0}) = \mathbf{0}$

*(ii)* $\| t(u) - t(v) \| = \| u - v \|$

Then $t$ is an orthonormal linear transformation.

**Proof.** Using *(i)* and *(ii)*, we have

$$\| t(u) \| = \| t(u) - \mathbf{0} \| = \| t(u) - t(\mathbf{0}) \|$$
$$= \| u - \mathbf{0} \| = \| u \|$$

That is, $\| t(u) \| = \| u \|$ for all $u \in V$.

Also, using *(ii)*, we get

$$\| t(u) - t(v) \|^2 = \| u - v \|^2 = \langle u - v, u - v \rangle$$
$$= \langle u, u \rangle + \langle v, v \rangle - 2 \langle u, v \rangle \quad \ldots(1)$$

and $$\| t(u) - t(v) \|^2 = \langle t(u) - t(v), t(u) - t(v) \rangle$$
$$= \langle t(u), t(u) \rangle + \langle t(v), t(v) \rangle - 2 \langle t(u), t(v) \rangle$$
$$= \| t(u) \|^2 + \| t(v) \|^2 - 2 \langle t(u), t(v) \rangle$$
$$= \| u \|^2 + \| v \|^2 - 2 \langle t(u), t(v) \rangle \quad \ldots..(2)$$

Thus form (1) and (2), we obtain

$$\langle t(u), t(v) \rangle = \langle u, v \rangle \text{ for all } u, v \in V.$$

So, $t$ preserves the inner product.

Let $\{u_1, u_2, \ldots, u_n\}$ be an orthonormal basis of $V$, then by theorem 8, $\{t(u_1), t(u_2), \ldots, t(u_n)\}$ is an orthonormal basis too.

Let, $$u = \sum_{i=1}^{n} \alpha_i u_i, \qquad \alpha_i \in R$$

Then coordinates of $t(u)$ are $\langle t(u), t(u_i) \rangle$,

*i.e.* $$t(u) = \sum_{i=1}^{n} \langle t(u), t(u_i) \rangle t(u_i)$$

$$= \sum_{i=1}^{n} \langle u, u_i \rangle t(u_i)$$

$$= \sum_{i=1}^{n} \alpha_i t(u_i)$$

Thus $$t(u) = \sum_{i=1}^{n} \alpha_i t(u_i)$$

*i.e* $$t\left( \sum_{i=1}^{n} \alpha_i u_i \right) = \sum_{i=1}^{n} \alpha_i t(u_i)$$

Hence $t$ is linear.

## 15.3 Orthogonal matrix

A square matrix $A$ over $R$ is said to be orthogonal if its columns are orthonormal in the standard inner product of $R^n$.

Thus a square matrix $A$ is orthogonal if $C_i\, C_j = \delta_{ij}$ for all $i, j = 1, 2,...,n$, where $C_i$ stands for the $i^{th}$ column vector in $R^n$.

or $\quad \sum_{k=1}^{n} a_{ki}\, a_{kj} = \delta_{ij} \qquad$ for all $i, j = 1, 2,...,n$.

**Theorem 10.** *Let $V$ be a finite dimensional inner product space. Then a linear transformation $t : V \to V$ is orthogonal if and only if its matrix relative to an orthonormal basis is orthogonal.*

**Proof.** Let $B = \{u_1, u_2,..., u_n\}$ be an orthonormal basis of $V$ and $A = [a_{ij}]$ over $R$ be the matrix of t relative to the basis $B$. So that

$$t(u_i) = \sum_{k=1}^{n} a_{ki}\, u_k \qquad 1 \le i \le n.$$

Now, $t$ is orthogonal,

$\Leftrightarrow \quad < t(u), t(v) > = < u, v > \qquad$ for all $u,v \in V$,

$\Leftrightarrow \quad < t(u_i), t(u_j) > = < u_i, u_j >, \qquad i, j = 1, 2,...,n.$

$$\Leftrightarrow \quad \left\langle \sum_{k=1}^{n} a_{ki}\, u_k, \sum_{r=1}^{n} a_{rj} u_r \right\rangle = \delta_{ij}$$

$$\Leftrightarrow \quad \sum_{k=1}^{n}\sum_{r=1}^{n} a_{ki}\, a_{rj} \langle u_k, u_r \rangle = \delta_{ij}$$

$$\Leftrightarrow \quad \sum_{k=1}^{n}\sum_{r=1}^{n} a_{ki}\, a_{rj}\, \delta_{kr} = \delta_{ij} \qquad (\because\ B \text{ is orthonormal basis})$$

$$\Leftrightarrow \quad \sum_{k=1}^{n} a_{ki}\, a_{kj} = \delta_{ij} \qquad \text{(on summing over } r\text{)}$$

$\Leftrightarrow \quad C_i\, C_j = \delta_{ij} \qquad$ for all $\ i,\ j = 1, 2,..., n.$

$\Leftrightarrow \quad A$ is orthogonal matrix.

**Theorem 11.** *A orthogonal matrix is always non singular.*

**Proof.** Let $A$ be an orthogonal matrix. Therefore columns of $A$ are orthonormal in standard product of $R^n$. We know that a list of orthonormal vectors in an inner product space is linearly independent. Consequently rank $(A) = n$, which implies that $|A| \neq 0$. Hence $A$ is non-singular.

**Theorem 12.** *Let V be a finite dimensional inner product space and $t \in Hom(V, V)$. Then linear transformation t is orthogonal if and only if the matrix A of t with respect to an orthonormal basis satisfies the condition $A^TA = I$ and $AA^T = I$.*

**Proof.** Let $A\ [a_{ij}]$ be an orthogonal matrix. Then

$$t(u_i) = \sum_{j=1}^{n} a_{ji} u_j, \qquad 1 \le i \le n$$

Now,
$$< t(u_i), t(u_i) > = \left\langle \sum_{j=1}^{n} a_{ji}\, u_j, \sum_{k=1}^{n} a_{ki}\, u_k \right\rangle$$

$$= \sum_{j=1}^{n} \sum_{k=1}^{n} a_{ji}\, a_{ki} \left\langle u_k, u_j \right\rangle$$

$$= \sum_{j=1}^{n} \sum_{k=1}^{n} a_{ji}\, a_{ki}\, \delta_{jk} = \sum_{j=1}^{n} a_{ji}^2 = 1 \quad \text{(on summing over } k\text{)} \qquad .....(1)$$

$$< t(u_i), t(u_j) > = \left\langle \sum_{k=1}^{n} a_{ki}\, u_k, \sum_{r=1}^{n} a_{rj}\, u_r \right\rangle$$

$$= \sum_{k=1}^{n} \sum_{r=1}^{n} a_{ki}\, a_{rj} \left\langle u_k, u_r \right\rangle$$

$$= \sum_{k=1}^{n} \sum_{r=1}^{n} a_{ki}\, a_{rj}\, \delta_{kr} = \sum_{k=1}^{n} a_{ki}\, a_{kj} = 0 \quad \text{(on summing over } r\text{)} \qquad .....(2)$$

Equation (1) and (2) imply that $A^TA = I$, since the $(ik)^{th}$ entry of $A^T$ is $a_{ki}$. Similarly, $AA^T = I$.

Conversely, $A^TA = I$, implies that the conditions in (1) and (2) are satisfied and hence $\{ t(u_1), t(u_2),..., t(u_n)\}$ is an orthonormal set which is basis of $V$. Thus $t$ takes an orthonormal basis to an orthonormal basis, so $t$ is an orthogonal linear transformation.

**Theorem 13.** *The determinant of an orthogonal matrix is* $\pm 1$.

**Proof :** Let $A$ be an orthogonal matrix. Then

$$AA^T = \mathrm{I}$$

$\Rightarrow$ $\quad \mathrm{Det}\,(AA^T) = \mathrm{Det}\,(I)$

$\Rightarrow$ $\quad \mathrm{Det}\,(A)\ \mathrm{Det}\,(A^T) = 1$

$\Rightarrow$ $\quad$ Hence $\mathrm{Det}\,(A) = \pm 1 \qquad [\because\ \mathrm{Det}\,(A) = \mathrm{Det}\,(A^T)]$

## 15.4 Principal axis theorem

**Lemma :** The eigenvalues of a self-adjoint linear transformation are real.

**Proof :** Let $V$ be a finite dimensional real inner-product space. Let $t : V \to V$ be a self-adjoint linear transformation. Then matrix $A$ of $t$ with respect to an orthonormal basis of $V$, is real symmetric

matrix. Let $\lambda$ be an eigenvalue of $t$, then it is an eigenvalue of $A$. Also let $X$ be an eigenvector of $A$ corresponding to the eigenvalue $\lambda$ , then we have

$$AX = X\lambda$$

$$\Rightarrow \quad \left(\overline{AX}\right)^T = \left(\overline{X\lambda}\right)^T$$

$$\Rightarrow \quad \bar{X}^T \bar{A}^T = \bar{X}^T \bar{\lambda}^T$$

$$\Rightarrow \quad \bar{X}^T A = \bar{X}^T \bar{\lambda} \qquad (A \text{ is symmetric and real})$$

Now $$\bar{X}^T AX = \left(\bar{X}^T \bar{\lambda}\right) X = \left(\bar{X}^T X\right)\bar{\lambda} \qquad .....(1)$$

Also $$\bar{X}^T AX = \bar{X}^T (AX) = \bar{X}^T (X\lambda) = \left(\bar{X}^T X\right)\lambda \qquad .....(2)$$

So, from *(i)* & *(ii)*, we have

$$\left(\bar{X}^T X\right)\bar{\lambda} = \left(\bar{X}^T X\right)\lambda$$

$$\Rightarrow \quad \left(\bar{X}^T X\right)\left(\bar{\lambda} - \lambda\right) = \mathbf{0}$$

$$\Rightarrow \quad \lambda = \bar{\lambda}$$

Hence $\lambda$ is real, because $X \neq \mathbf{0}$ and so $\bar{X}^T X \neq \mathbf{0}$.

## Self-learning exercise-1

**1.** An orthogonal linear transformation between inner product spaces preserves the lengths, but not the angle between two vectors. [T/F]

**2.** If $C_i$ and $C_j$ are columns of an orthogonal matrix, then $C_i\, C_j$ = ..... .

**3.** Matrix $A$ is orthogonal then $A^T = A^{-1}$. [T/F]

**4.** Matrix $A$ is orthogonal then $AA^T$ = ..... .

**Theorem 14.** ***(Principal axis theorem).*** *Let t be a self-adjoint linear transformation on a finite dimensional inner product space V. Then the matrix of t with respect to some orthonormal basis is a diagonal matrix whose diagonal elements are eigenvalues of t these are real, and each appears on the diagonal as many times as its multiplicity.*

**Proof :** We shall prove the theorem by induction on the dimension $n$ of inner product space $V$.

For $n = 1$, let $B = \{v_1\}$ be an orthonormal basis.

So $$t(v_1) = v_1 \lambda_1, \quad \text{for some} \quad \lambda_1 \in R.$$

And therefore, $v_1$ is an eigenvector of $t$ and $\lambda_1$ is the corresponding eigenvalue which is real and $A = \{\lambda_1\}$. Hence it is true for $n = 1$.

Now assume that it is true for a vector space whose dimension is $(n - 1)$. Then we shall show that it is true for $V$ whose dimension is n.

Now let $W = v_1^{\perp} = \{u \in V \mid < u, v_1 > = 0\}$ be the orthogonal complement of $v_1$. It is sub-space of $V$. We shall now show that $W$ is invarient under $t$. To do so, let $v \in W$. Then

$$< v, v_1 > = 0 \qquad \text{.....(1)}$$

We have $< t(v), v_1 > = < v, t(v_1) >$, since t is self-adjoint

$$= < v, v_1 \lambda_1 >$$

$$= \lambda_1 < v, v_1 >$$

$$= 0 \qquad \text{[from (1)]}$$

Thus $t(v) \in W$.

Hence $v \in W \Rightarrow t(v) \in W$.

Thus $t$ maps vectors of $W$ to vectors of $W$ only. Therefore the restriction of $t$ to $W$ is $t_w : W \to W$ is also a self adjoint linear transformation.

Now, $\dim W + \dim W^{\perp} = n$

$\Rightarrow$ $\dim W = n - \dim W^{\perp}$

$\Rightarrow$ $\dim W = n - 1$, because $W^{\perp} = < \{v_1\} >$ and $\dim W^{\perp} = 1$

Hence $t_w : W \to W$ is a self adjoint linear transformation from an $(n-1)$ dimensional inner-product space $W$ to itself, therefore by induction assumption there is an orthonormal basis of eigen-vectors, say $\{v_2, v_3, ...., v_n\}$ of $W$ for which matrix of $t_w$ is a diagonal matrix. Combining basis $\{v_1\}$ of $W^{\perp}$ to basis $\{v_2, v_3, ...., v_n\}$ of $W$ so that we have an orthonormal basis of eigenvectors $\{v_1, v_2, ...., v_n\}$ of $V$, for which matrix of self adjoint $t$ is a diagonal matrix and each diagonal element is an eigenvalue of t, appears as many times as its multiplicity and the eigenvalues of $t$ are also real.

## 15.5 Summary

In this unit we have studied orthogonal linear transformation between inner product spaces, orthogonal matrices related to orthogonal linear transformations and their properties, and a very useful and fundamental result known as the principal axis theorem.

## 15.6 Answers to self learning exercises

### Self learning exercise-1

**1.** false **2.** $\delta_{ij}$ **3.** true **4.** $I$

## 15.7 Exercises

1. Prove that if $W_1$ and $W_2$ are subspaces of $V$ such that $\dim W_1 = \dim W_2$, then there exists an orthogonal transformation $t$ such that $t(W_1) = W_2$.
2. If $A$ is an orthogonal matrix, show that $A^T$ and $A^{-1}$ are orthogonal matrices.

□□□

## Reference Books


1. **Contemporary Abstract Algebra**

   Joseph A. Gallian

   Narosa Publishing House, New Delhi, 1998

2. **A First Course in Abstract Algebra**

   John B. Fraleigh

   Addison-Wesley Publishing Company, 1970

3. **Topics in Algebra**

   I.N. Herstein

   Vani Educational Books, New Delhi, 1975

4. **Studies in Algebra**

   Dileep S. Chauhan and K.N. Singh

   Jaipur Publishing House, Jaipur, 2009