M.Sc. Previous Year

Mathematics, MM-01

ADVANCED ABSTRACT ALGEBRA

मध्यप्रदेश भोज (मुक्त) विश्वविद्यालय – भोपाल

MADHYA PRADESH BHOJ (OPEN) UNIVERSITY - BHOPAL

Reviewer Committee

1. Dr. Piyush Bhatnagar
   Professor
   Govt. MLB College, Bhopal
2. Dr Anil Rajput
   Professor
   Govt. C. S. Azad (PG) College, Sehore
3. Dr. Rajkumar Bhimte
   Assistant Professor
   Govt. College, Vidisha, MP

Advisory Committee

1. Dr. Jayant Sonwalkar
   Hon'ble Vice Chancellor
   Madhya Pradesh Bhoj (Open) University, Bhopal (MP)
2. Dr. L. S. Solanki
   Registrar
   Madhya Pradesh Bhoj (Open) University, Bhopal (MP)
3. Dr. Neelam Wasnik
   Dy. Director Printing
   Madhya Pradesh Bhoj (Open) University, Bhopal (MP)
4. Dr. Piyush Bhatnagar
   Professor
   Govt. MLB College, Bhopal
5. Dr. Anil Rajput
   Professor
   Govt. C. S. Azad (PG) College, Sehore
6. Dr. Rajkumar Bhimte
   Assistant Professor
   Govt. College, Vidisha, MP

COURSE WRITERS

V K Khanna, Formerly Associate Professor, Department of Mathematics, Kirori Mal College, University of Delhi

SK Bhambri, Formerly Associate Professor, Department of Mathematics, Kirori Mal College, University of Delhi

Units: (1-5)

Published by Registrar, MP Bhoj (open) University, Bhopal in 2020

VIKAS® PUBLISHING HOUSE PVT. LTD.
E-28, Sector-8, Noida - 201301 (UP)
Phone: 0120-4078900 • Fax: 0120-4078999
Regd. Office: A-27, 2nd Floor, Mohan Co-operative Industrial Estate, New Delhi 1100 44
• Website: www.vikaspublishing.com • Email: helpline@vikaspublishing.com

SYLLABI-BOOK MAPPING TABLE

Advanced Abstract Algebra

Syllabi	Mapping in Book
UNIT - 1: **Groups:** Normal and Subnormal Series, Composition Series, Jordan-Hölder Theorem, Solvable Groups, Nilpotent Groups.	**Unit-1:** Groups **(Pages 3-71)**
UNIT - 2: **Canonical Forms:** Similarity of Linear Transformations, Invariant Subspaces, Reduction to Triangular Forms, Nilpotent Transformation, Index of Nilpotency, Invariants of a Nilpotent Transformation, The Primary Decomposition Theorem, Jordan Block and Jordan Forms. Cyclic Modules: Simple Modules, Semi-Simple Modules, Schur's Lemma. Free Modules.	**Unit-2:** Canonical Forms **(Pages 73-138)**
UNIT - 3: **Field Theory:** Extension Fields, Algebraic and Transcendental Extensions, Separable and Inseparable Extensions, Normal Extensions, Perfect Fields, Finite Fields, Primitive Elements, Algebraically Closed Fields, Automorphisms of Extensions, Galois Extensions, Fundamental Theorem of Galois Theory, Solution of Polynomial Equations by Radicals, Lnsolvcibility of the General Equation of Degree 5 by Radicals.	**Unit-3:** Field Theory **(Pages 139-217)**
UNIT - 4: **Noetherian and Artinian Modules and Rings:** Hilbert Basis Theorem, Wedderburn-Artin Theorem, Uniform Modules, Primary Modules, and Noether Lasker Theorem, Smith Normal Form Over a Principal Ideal Domain and Rank.	**Unit-4:** Noetherian and Artinian Modules and Rings **(Pages 219-248)**
UNIT - 5: Fundamental Structure Theorem For Finitely Generated Modules Over a Principal Ideal Domain and its Applications to Finitely Generated Abelian Groups, Rational Canonical Form, Generalised Jordan form Over any Field.	**Unit-5:** Abelian Groups and Jordan Form **(Pages 249-276)**

CONTENTS

INTRODUCTION

NOTES

In algebra, which is a broad division of mathematics, abstract algebra, also occasionally called modern algebra, is the study of algebraic structures. Algebraic structures include groups, rings, fields, modules, vector spaces, lattices, and algebras. The term abstract algebra was coined in the early 20th century to distinguish this area of study from older parts of algebra, and more specifically from elementary algebra, the use of variables to represent numbers in computation and reasoning.

Algebraic structures, with their associated homomorphisms, form mathematical categories. Category theory is a formalism that allows a unified way for expressing properties and constructions that are similar for various structures. Universal algebra is a related subject that study the different types of algebraic structures as single objects. For example, the structure of groups is a single object in universal algebra, which is called the variety of groups.

Major themes in algebraic equations include, solving of systems of linear equations which led to linear algebra, attempts to find formulas for solutions of general polynomial equations of higher degree that resulted in discovery of groups as abstract manifestations of symmetry and arithmetical investigations of quadratic and higher-degree forms and diophantine equations that directly produced the notions of a ring and ideal.

In mathematics, a canonical, normal or standard form of a mathematical object is a standard way of presenting that object as a mathematical expression. It provides the simplest representation of an object which allows it to be identified in a unique way. The distinction between 'Canonical' and 'Normal' forms varies from subfield to subfield. In most fields, a canonical form specifies a unique representation for every object, while a normal form simply specifies its form, without the requirement of uniqueness. The canonical form of a positive integer in decimal representation is a finite sequence of digits that does not begin with zero.

A 'Ring' is a set equipped with two operations, called addition and multiplication. Fundamentally, the 'Ring' is a 'Group' under addition and satisfies some of the properties of a group for multiplication. A 'Field' is a 'Group' under both addition and multiplication.

Artinian and Noetherian rings have some measure of finiteness associated with them. In fact, the conditions for Artinian and Noetherian rings, called respectively, the descending and ascending chain conditions, are often termed the minimum and maximum conditions.

Leonhard Euler considered algebraic operations on numbers modulo an integer—modular arithmetic—in his generalization of Fermat's little theorem. These investigations were further analysed by Carl Friedrich Gauss, who considered the

NOTES

structure of multiplicative groups of residues mod n and established many properties of cyclic and more general abelian groups.

This book is divided into five units which explains groups, canonical forms, cyclic modules, field theory, Galois theory and extensions, Noetherian and Artinian modules and rings, and fundamental structure theorem.

The book follows the self-instruction mode or the SIM format where in each unit begins with an 'Introduction' to the topic followed by an outline of the 'Objectives'. The content is presented in a simple and structured form interspersed with Answers to 'Check Your Progress' for better understanding. A list of 'Summary' along with a 'Key Terms' and a set of 'Self-Assessment Questions and Exercises' is provided at the end of each unit for effective recapitulation.

UNIT 1 GROUPS

NOTES

Structure

1.0 INTRODUCTION

In mathematics, a group is a set equipped with an operation that combines any two elements to form a third element while being associative as well as having an identity element and inverse elements. These three conditions, called group axioms, hold for number systems and many other mathematical structures.

A group is an algebraic structure consisting of a set together with a binary operation known as the group operation that combines any two of its elements to form a third element. A subgroup series is a chain of subgroups which simplifies the study of a group to the study of simpler subgroups and their relations. A Sylow subgroup is a subgroup having order which is a power of a prime number and which is not strictly contained in any other subgroup with the same property. The Sylow theorems concern subgroups with maximal prime power size.

A composition series provides a way to break up an algebraic structure, such as, a group or a module, into simple pieces. The need for considering composition series in the context of modules arises from the fact that many naturally occurring modules are not semisimple, hence cannot be decomposed into a direct sum of simple modules. A composition series of a module M is a finite increasing filtration of M by submodules such that the successive quotients are simple and serves as a replacement of the direct sum decomposition of M into its simple constituents.

In this unit, you will learn about the groups, normal and subnormal, composition series, Jordan-Holder Series, solvable groups, nilpotent groups, conjugate elements, Sylow p-subgroups and Sylow's theorems and their simple applications.

NOTES

1.1 OBJECTIVES

After going through this unit, you will be able to:

- Define groups, and normal, subnormal and composition series
- Know about the Jordan-Holder theorem and solvable groups
- Describe about the nilpotent groups
- Explain class equation for a finite group
- Know the properties of finite groups up to order 15

1.2 GROUPS

Definition: A non-empty set G, together with a binary compsition $*$ (star) is said to form a group, if it satisfies the following postulates

(*i*) ***Associativity*****:** $a * (b * c) = (a * b) * c$, for all $a, b, c \in G$

(*ii*) ***Existence of Identity*****:** $\exists$ an element $e \in G$, such that,

$$a * e = e * a = a \quad \text{for all } a \in G$$

(e is then called *identity*)

(*iii*) ***Existence of Inverse*****:** For every $a \in G$, $\exists\ a' \in G$ (depending upon a) such that,

$$a * a' = a' * a = e$$

(a' is then called inverse of a)

Notes:

1. Since $*$ is a binary composition on G, it is understood that for all $a, b \in G$, $a * b$ is a unique member of G. This property is called closure property.
2. If, in addition to the above postulates, G also satisfies the *commutative law*

   $a * b = b * a$ for all $a, b \in G$

   then G is called an *abelian group* or *a commutative group*.
3. Generally, the binary composition for a group is denoted by '.' (dot) which is so convenient to write (and makes the axioms look so natural too).

This binary composition '.' is called product or multiplication (although it may have nothing to do with the usual multiplication, that we are so familiar with). In fact, we even drop '.' and simply write ab in place of $a \,.\, b$.

In future, whenever we say that G is a group it will be understood that there exists a binary composition '.' on G and it satisfies all the axioms in the definition of the group.

If the set G is finite (i.e., has finite number of elements) it is called a *finite group* otherwise, it is called an *infinite group*.

Definition: By order of a group, we will mean the number of elements in the group and shall denote it by $o(G)$ or $| G |$.

NOTES

We now consider a few cases of systems that form groups (or do not form groups).

Case 1: The set **Z** of integers forms an abelian group with respect to the usual addition of integers.

It is easy to verify the postulates in the definition of a group as sum of two integers is a unique integer (thus closure holds). Associativity of addition is known to us. 0 (zero) will be identity and negatives will be the respective inverse elements. Commutativity again being obvious.

Case 2: One can easily check, as in the previous case, that sets **Q** of rationals, **R** of real numbers would also form abelian groups with respect to addition.

Case 3: Set of integers, with respect to usual multiplication does not form a group, although closure, associativity, identity conditions hold.

Note 2 has no inverse with respect to multiplication as there does not exist any integer a such that, $2 \,.\, a = a \,.\, 2 = 1$.

Case 4: The set G of all +ve irrational numbers together with 1 under multiplication does not form a group as closure does not hold. Indeed $\sqrt{3} \,.\, \sqrt{3} = 3 \notin G$, although one would notice that other conditions in the definition of a group are satisfied here.

Case 5: Let G be the set $\{1, -1\}$. Then it forms an abelian group under multiplication. It is again easy to check the properties.

1 would be identity and each element is its own inverse.

Case 6: Set of all 2×2 matrices over integers under matrix addition would be another example of an abelian group.

Case 7: Set of all non zero complex numbers forms a group under multiplication defined by

$$(a + ib)\,(c + id) = (ac - bd) + i\,(ad + bc)$$

$1 = 1 + i.0$ will be identity,

$$\frac{a}{a^2 + b^2} - i\frac{b}{a^2 + b^2} \text{ will be inverse of } a + ib.$$

Note: $a + ib$ non zero means that not both a and b are zero. Thus, $a^2 + b^2 \neq 0$.

Case 8: The set G of all nth roots of unity, where n is a fixed positive integer forms an abelian group under usual multiplication of complex numbers.

We know that complex number z is an nth root of unity if $z^n = 1$ and also that there exist exactly n distinct roots of unity.

NOTES

In fact the roots are given by,

$$e^{\frac{2\pi ir}{n}}$$

Where, $r = 1, 2, ..., n$ and $e^{ix} = \cos x + i \sin x$.

If $a, b \in G$ be any two members, then $a^n = 1$, $b^n = 1$ thus $(ab)^n = a^n b^n = 1$.

$\Rightarrow$ ab is an nth root of unity

$\Rightarrow$ $ab \in G \Rightarrow$ closure holds.

Associativity of multiplication is true in complex numbers.

Again, since $1.\ a = a\ .\ 1 = a$, 1 will be identity.

Also for any $a \in G$, $\frac{1}{a}$ will be its inverse as $\left(\frac{1}{a}\right)^n = \frac{1}{a^n} = 1$.

So, inverse of $e^{2\pi ir/n}$ is $e^{2\pi i(n-r)/n}$ and identity is $e^{2\pi i0/n} = 1$

Commutativity being obvious, we find G is an abelian group.

As a particular case, if $n = 4$ then G is $\{1, -1, i, -i\}$

Case 9: (*i*) Let $G = \{\pm 1, \pm i, \pm j, \pm k\}$. Define product on G by usual multiplication together with

$$i^2 = j^2 = k^2 = -1,\ ij = -ji = k$$
$$jk = -kj = i$$
$$ki = -ik = j$$

then G forms a group. G is not abelian as $ij \neq ji$.

This is called the ***Quaternion Group***.

(*ii*) If set G consists of the eight matrices

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

$$\begin{bmatrix}0 & i\\ i & 0\end{bmatrix}, \begin{bmatrix}0 & -i\\ -i & 0\end{bmatrix}, \text{ where } i = \sqrt{-1}$$

then G forms a non abelian group under matrix multiplication. (Compare with part (*i*)).

Case 10: Let $G = \{(a, b) \mid a, b$ rationals, $a \neq 0\}$. Define $*$ on G by

$$(a, b) * (c, d) = (ac, ad + b)$$

Closure follows as $a, c \neq 0 \Rightarrow ac \neq 0$

$$[(a, b) * (c, d)] * (e, f) = (ac, ad + b) * (e, f)$$
$$= (ace, acf + ad + b)$$
$$(a, b) * [(c, d) * (e, f)] = (a, b) * (ce, cf + d)$$
$$= (ace, acf + ad + b)$$

proves associativity.

(1, 0) will be identity and $(1/a, -b/a)$ will be inverse of any element (a, b).

G is not abelian as

$$(1, 2) * (3, 4) = (3, 4 + 2) = (3, 6)$$
$$(3, 4) * (1, 2) = (3, 6 + 4) = (3, 10).$$

NOTES

Case 11 (a): The set G of all 2×2 matrices of the form $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ over reals, where $ad - bc \neq 0$, forms a non abelian group under matrix multiplication.

It is called the general linear group of 2×2 matrices over reals and is denoted by $GL(2, \mathbf{R})$.

The matrix $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ will act as identity and

the matrix $\begin{bmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix}$ will be inverse of $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

One can generalize and prove.

(b) If G be the set of all $n \times n$ invertible matrices over reals, then G forms a group under matrix multiplication.

Case 12: Let $G = \{2^r \mid r = 0, \pm 1, \pm 2, ...\}$

We show G forms a group under usual multiplication.

For any $2^r, 2^s \in G$, $2^r . 2^s = 2^{r+s} \in G$

Thus closure holds.

Associativity is obvious.

Again as $1 \in G$, and $x . 1 = 1 . x = x$ for all $x \in G$

1 is identity.

For any $2^r \in G$, as $2^{-r} \in G$ and $2^r . 2^{-r} = 2^0 = 1$,

We find each element of G has inverse. Commutativity is evidently true.

Case 13: *Group of Residues* : Let $G = \{0, 1, 2, 3, 4\}$. Define a composition $\oplus_5$ on G by $a \oplus_5 b = c$ where c is the least non negative integer obtained as remainder when $a + b$ is divided by 5. For example, $3 \oplus_5 4 = 2$, $3 \oplus_5 1 = 4$, etc. Then $\oplus_5$ is a binary composition on G (called addition modulo 5). It is easy to verify that G forms a group under this.

One can generalize this result to

$$G = \{0, 1, 2, ..., n - 1\}$$

under addition modulo n where n is any positive integer.

We thus notice

$$a \oplus_n b = \begin{cases} a+b & \text{if } a+b<n \\ a+b-n & \text{if } a+b \geq n \end{cases}$$

Also, in case there is no scope of confusion we drop the sub suffix n and simply write $\oplus$. This group is generally denoted by $\mathbf{Z}_n$.

NOTES

Case 14: Let $G = \{x \in \mathbf{Z} \mid 1 \leq x < n, x, n \text{ being co-prime}\}$ where $\mathbf{Z}$ = set of integers and x, n being co-prime means H.C.F of x and n is 1.

We define a binary composition $\otimes$ on G by $a \otimes b = c$ where c is the least +ve remainder obtained when $a \,.\, b$ is divided by n. This composition $\otimes$ is called multiplication modulo n.

We show G forms a group under $\otimes$.

Closure: For $a, b \in G$, let $a \otimes b = c$. Then $c \neq 0$, because otherwise $n \mid ab$ which is not possible as a, n and b, n are co-prime.

Thus $c \neq 0$ and also then $1 \leq c < n$.

Now if c, n are not co-prime then $\exists$ some prime number p such that, $p \mid c$ and $p \mid n$.

Again as $ab = nq + c$ for some q

We get $p \mid ab$ $\quad [p \mid n \Rightarrow p \mid nq, p \mid c \Rightarrow p \mid nq + c]$

$\Rightarrow p \mid a$ or $p \mid b$ (as p is prime)

If $p \mid a$ then as $p \mid n$ it means a, n are not co-prime.

But a, n are co-prime.

Similarly $p \mid b$ leads to a contradiction.

Hence c, n are co-prime and thus $c \in G$, showing that closure holds.

Associativity: Let $a, b, c \in G$ be any elements.

Let $a \otimes b = r_1$, $(a \otimes b) \otimes c = r_1 \otimes c = r_2$

then r_2 is given by $r_1 c = nq_2 + r_2$

Also $\quad a \otimes b = r_1$ means

$$ab = q_1 n + r_1$$

thus $\quad ab - q_1 n = r_1$

$$\Rightarrow (ab - q_1 n)c = r_1 c = nq_2 + r_2$$

$$\Rightarrow \quad (ab)c = r_2 + nq_2 + nq_1 c = n(q_1 c + q_2) + r_2$$

or that r_2 is the least non-negative remainder got by dividing $(ab)c$ by n.

Similarly, if $a \otimes (b \otimes c) = r_3$ then we can show that r_3 is the least non-negative remainder got by dividing $a(bc)$ by n.

But since $a(bc) = (ab)c$, $r_2 = r_3$

Hence $a \otimes (b \otimes c) = (a \otimes b) \otimes c$.

Existence of Identity: It is easy to see that

$$a \otimes 1 = 1 \otimes a = a \quad \text{for all } a \in G$$

or that 1 will act as identity.

Existence of Inverse: Let $a \in G$ be any element then a and n are co-prime and thus we can find integers x and y such that, $ax + ny = 1$

By division algorithm, we can write

$$x = qn + r, \quad \text{where } 0 \le r < n$$
$$\Rightarrow \quad ax = aqn + ar$$
$$\Rightarrow \quad ax + ny = aqn + ar + ny$$
$$\Rightarrow \quad 1 = aqn + ar + ny$$

or that $ar = 1 + (-aq - y)n$

i.e., $a \otimes r = 1$. Similarly $r \otimes a = 1$. If r, n are co-prime, r will be inverse of a.

If r, n are not co-prime, we can find a prime number p such that, $p \mid r$, $p \mid n$

$$\Rightarrow p \mid qn \text{ and } p \mid r$$
$$\Rightarrow p \mid qn + r$$
$$\Rightarrow p \mid x$$
$$\Rightarrow p \mid ax \text{ also } p \mid ny$$
$$\Rightarrow p \mid ax + ny = 1$$

which is not possible. Thus r, n are co-prime and so $r \in G$ and is the required inverse of a.

It is easy to see that G will be abelian. We denote this group by U_n or $U(n)$ and call it the group of integers under multiplication modulo n.

Note: Suppose $n = p$, a prime, then since all the integers 1, 2, 3, ..., $p - 1$ are co-prime to p, these will all be members of G. One can show that

$$G = \{2, 4, 6, ..., 2(p - 1)\}$$

where $p > 2$ is a prime forms an abelian group under multiplication modulo $2p$.

Case 15: Let $G = \{0, 1, 2\}$ and define $*$ on G by

$$a * b = | a - b |$$

Then closure is established by taking a look at the composition table

$*$	0	1	2
0	0	1	2
1	1	0	1
2	2	1	0

Since $a * 0 = | a - 0 | = a = 0 * a$, 0 is identity

and $a * a = | a - a | = 0$ shows each element will be its own inverse.

But the system $(G, *)$ fails to be a group as associativity does not hold.

Indeed $1 * (1 * 2) = 1 * 1 = 0$

but $(1 * 1) * 2 = 0 * 2 = 2$

NOTES

NOTES

Case 16: Let $S = \{1, 2, 3\}$ and let $S_3 = A(S)$ = set all permutations of S. This set satisfies associativity, existence of identity and existence of inverse conditions in the definition of a group. Also clearly, since f, g permutations on S imply that fog is a permutation on S the closure property is ensured. Hence S_3 forms a group. That it is not abelian follows by the fact that $fog \neq gof$. This would, in fact, be the smallest non abelian group.

Note: Let X be a non-empty set and let $M(X)$ = set of all maps from X to X, then $A(X) \subseteq M(X)$. $M(X)$ forms a semi group under composition of maps. Identity map also lies in $M(X)$ and as a map is invertible iff it is 1-1, onto, i.e., a permutation, we find $A(X)$ the subset of all permutations forms a group, denoted by S_X and is called symmetric group of X. If X is finite with say, n elements then $o(M(X)) = n^n$ and $o(S_X) = \underline{|n}$ and in that case we use the notation S_n for S_X.

In the definition of a group, we only talked about the existence of identity and inverse of each element. We now show that these elements would also be unique, an elementary but exceedingly useful result. We prove it along with some other results in

Lemma: *In a group G,*

(1) *Identity element is unique.*

(2) *Inverse of each $a \in G$ is unique.*

(3) $(a^{-1})^{-1} = a$, *for all* $a \in G$, *where* a^{-1} *stands for inverse of* a.

(4) $(ab)^{-1} = b^{-1} a^{-1}$ *for all* $a, b \in G$

(5) $ab = ac \Rightarrow b = c$

$ba = ca \Rightarrow b = c$ *for all* $a, b, c \in G$

(called the cancellation laws).

Proof: (1) Suppose e and e' are two elements of G which act as identity.

Then, since $e \in G$ and e' is identity,

$$e'e = ee' = e$$

and as $e' \in G$ and e is identity

$$e'e = ee' = e'$$

The two $\Rightarrow e = e'$

which establishes the uniqueness of identity in a group.

(2) Let $a \in G$ be any element and let a' and a'' be two inverse elements of a, then

$$aa' = a'a = e$$

$$aa'' = a''a = e$$

Now, $a' = a'e = a'(aa'') = (a'a)a'' = ea'' = a''$.

Showing thereby that inverse of an element is unique. We shall denote inverse of a by a^{-1}.

NOTES

(3) Since a^{-1} is inverse of a

$$aa^{-1} = a^{-1}a = e$$

Which also implies a is inverse of a^{-1}

Thus, $(a^{-1})^{-1} = a$.

(4) We have to prove that ab is inverse of $b^{-1}a^{-1}$ for which we show

$$(ab)\,(b^{-1}a^{-1}) = (b^{-1}a^{-1})\,(ab) = e.$$

Now, $(ab)\,(b^{-1}a^{-1}) = [(ab)\,b^{-1}]\,a^{-1}$

$$= [(a(bb^{-1})]\,a^{-1}$$

$$= (ae)\,a^{-1} = aa^{-1} = e$$

Similarly $(b^{-1}a^{-1})\,(ab) = e$

and thus the result follows.

(5) Let, $ab = ac$, then

$$b = eb = (a^{-1}a)b$$

$$= a^{-1}(ab) = a^{-1}\,(ac)$$

$$= (a^{-1}\,a)c = ec = c$$

Thus $ab = ac \Rightarrow b = c$

Which is called the left cancellation law.

One can similarly, prove the right cancellation law.

Case 17 (*a*): Let $X = \{1, 2, 3\}$ and let $S_3 = A(X)$ be the group of all permutations on X. Consider $f, g, h \in A(X)$, defined by

$$f(1) = 2, \quad f(2) = 3, \quad f(3) = 1$$
$$g(1) = 2, \quad g(2) = 1, \quad g(3) = 3$$
$$h(1) = 3, \quad h(2) = 1, \quad h(3) = 2$$

It is easy then to verify that $fog = goh$

But $f \neq h$.

(*b*) If we consider the group in case 10, we find

$$(1, 2) * (3, 4) = (3, 6) = (3, 0) * (1, 2)$$

But $(3, 4) \neq (3, 0)$

Hence we notice, cross cancellations *may not* hold in a group.

Theorem 1.1: *For elements a, b in a group G, the equations ax = b and ya = b have unique solutions for x and y in G.*

Proof: Now, $ax = b$

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

$$\Rightarrow \quad ex = a^{-1}b$$

or $x = a^{-1}b$

which is the required solution of the equation $ax = b$.

NOTES

Suppose $x = x_1$ and $x = x_2$ are two solutions of this equation, then

$$ax_1 = b \text{ and } ax_2 = b$$

$$\Rightarrow \quad ax_1 = ax_2$$

$$\Rightarrow \quad x_1 = x_2 \quad \text{by left cancellation}$$

Showing that the solution is unique.

Similarly $y = ba^{-1}$ will be unique solution of the equation $ya = b$.

Theorem 1.2: *A non-empty set G together with a binary composition '.' is a group if and only if*

(1) $a(bc) = (ab)c$ *for all* $a, b, c \in G$

(2) *For any* $a, b \in G$, *the equations* $ax = b$ *and* $ya = b$ *have solutions in G.*

Proof: If G is a group, then (1) and (2) follow by definition and previous theorem. *Conversely*, let (1) and (2) hold. To show G is a group, we need prove existence of identity and inverse (for each element).

Let $a \in G$ be any element.

By (2) the equations $ax = a$

$ya = a$

have solutions in G.

Let $x = e$ and $y = f$ be the solutions.

Thus $\exists\ e, f \in G$, such that, $ae = a$

$fa = a$

Let now $b \in G$ be *any* element then again by (2) $\exists$ some x, y in G such that,

$ax = b$

$ya = b.$

Now, $ax = b \quad \Rightarrow f.(a\,.\,x) = f\,.\,b$

$\Rightarrow (f\,.\,a)\,.\,x = f\,.\,b$

$\Rightarrow a\,.\,x = f\,.\,b$

$\Rightarrow b = f\,.\,b$

Again, $y\,.\,a = b \quad \Rightarrow (y\,.\,a)\,.\,e = b\,.\,e$

$\Rightarrow y\,.\,(a\,.\,e) = b\,.\,e$

$\Rightarrow y\,.\,a = be$

$\Rightarrow b = be$

thus we have $b = fb$ ...(1.1)

$b = be$ ...(1.2)

for any $b \in G$

Putting $b = e$ in Equation (1.1) and $b = f$ in Equation (1.2) we get

$e = fe$

$f = fe$

$\Rightarrow \quad e = f.$

Hence, $ae = a = fa = ea$

i.e., $\exists\ e \in G$, such that, $ae = ea = a$

$\Rightarrow$ e is identity.

Again, for any $a \in G$, and (the identity) $e \in G$, the equations $ax = e$ and $ya = e$ have solutions.

Let the solutions be $x = a_1$, and $y = a_2$

then $aa_1 = e, \quad a_2a = e$

Now, $a_1 = ea_1 = (a_2a)a_1 = a_2(aa_1) = a_2e = a_2.$

Hence, $aa_1 = e = a_1a$ for any $a \in G$

i.e., for any $a \in G$, $\exists$ some $a_1 \in G$ satisfying the above relations $\Rightarrow a$ has an inverse. Thus each element has inverse and, by definition, G forms a group.

Note: While proving the above theorem we have assumed that equations of the type $ax = b$ and $ya = b$ have solutions in G. The result may fail, if only one type of the above equations has solution. Consider for example:

G to be a set with at least two elements. Define '.' on G by $a \,.\, b = b$ for all $a, b \in G$.

then $a \,.\, (b \,.\, c) = a \,.\, c = c$

$(a \,.\, b) \,.\, c = b \,.\, c = c$

shows associativity holds.

Again as $ab = b$, the equation $ax = b$ has a solution for any $a, b \in G$.

We notice that G is not a group, as cancellation laws do not hold in G.

As let $a, b \in G$ be any two distinct members, then

$ab = b$

$bb = b \Rightarrow ab = bb$

But, $a \neq b.$

Definition: A non empty set G together with a binary composition '.' is called a *semi-group* if

$a \,.\, (b \,.\, c) = (a \,.\, b) \,.\, c$ for all $a, b, c \in G$

Obviously then every group is a semi-group. That the converse is not true follows by considering the set **N** of natural numbers under addition.

The set G in Case 15 is not a semi group.

Theorem 1.3: *Cancellation laws may not hold in a semi-group*

Proof: Consider M the set of all 2×2 matrices over integers under matrix multiplication, which forms a semi-group.

NOTES

NOTES

If we take $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$, $C = \begin{bmatrix} 0 & 0 \\ 3 & 0 \end{bmatrix}$

then clearly $AB = AC = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

But, $B \neq C$.

Set of natural numbers under addition is an example of a semi-group in which cancellation laws hold.

Theorem 1.4: *A finite semi-group in which cancellation laws hold is a group.*

Proof: Let $G = \{a_1, a_2, ..., a_n\}$ be a finite semi-group in which cancellation laws hold.

Let $a \in G$ be any element, then by closure property

$aa_1, aa_2, ..., aa_n$

are all in G.

Suppose any two of these elements are equal

say, $aa_i = aa_j$ for some $i \neq j$

then $a_i = a_j$ by cancellation

But, $a_i \neq a_j$ as $i \neq j$

Hence no two of $aa_1, aa_2, ..., aa_n$ can be equal.

These being n in number, will be distinct members of G (Note $o(G) = n$).

Thus, if $b \in G$ be any element then

$b = aa_i$ for some i

i.e., for $a, b \in G$ the equation $ax = b$ has a solution ($x = a_i$) in G.

Similarly, the equation $ya = b$ will have a solution in G.

G being a semi-group, associativity holds in G.

Hence G is a group (by Theorem 1.2).

Note: The above theorem holds only in finite groups. The semi-group of natural numbers under addition being an example where cancellation laws hold but which is not a group.

Theorem 1.5: *A finite semi-group is a group if and only if it satisfies cancellation laws.*

Proof: Follows by previous theorem.

Definition: A non-empty set G together with a binary composition '.' is said to form a *monoid* if

(*i*) $a(bc) = (ab)c \quad \forall\ a, b, c \in G$

(*ii*) $\exists$ an element $e \in G$, such that, $ae = ea = a \quad \forall\ a \in G$

e is then called identity of G. It is easy to see that e is unique.

So all groups are monoids and all monoids are semi groups.

When we defined a group, we insisted that $\exists$ an element e which acts both as a right as well as a left identity and each element has both sided inverse. We show now that it is not really essential and only one sided identity and the *same* sided inverse for each element could also make the system a group.

NOTES

Theorem 1.6: *A system* $< G, . >$ *forms a group if and only if*

(*i*) $a(bc) = (ab)c$ *for all* $a, b, c \in G$

(*ii*) $\exists\ e \in G$, *such that*, $ae = a$ *for all* $a \in G$

(*iii*) *for all* $a \in G$, $\exists\ a' \in G$, *such that*, $aa' = e$.

Proof: If G is a group, we have nothing to prove as the result follows by definition.

Conversely, let the given conditions hold.

All we need show is that $ea = a$ for all $a \in G$

and $a'a = a$ for any $a \in G$

Let $a \in G$ be any element.

By (*iii*) $\exists\ a' \in G$, such that, $aa' = e$

$\therefore$ For $a' \in G$, $\exists\ a'' \in G$, such that, $a'a'' = e$ (using (*iii*))

Now $a'a = a'(ae) = (a'a)e = (a'a)(a'a'')$

$= a'(aa')a'' = a'(e)a'' = (a'e)a'' = a'a'' = e.$

Thus for any $a \in G$, $\exists\ a' \in G$, such that, $aa' = a'a = e$

Again $ea = (aa')a = a(a'a) = ae = a$

$\therefore$ $ae = ea = a$ for all $a \in G$

i.e., e is identity of G.

Hence G is a group.

(Refer Example 1.6 for another proof).

Theorem 1.7: *A system* $< G, . >$ *forms a group if and only if*

(*i*) $a(bc) = (ab)c$ *for all* $a, b, c \in G$

(*ii*) $\exists\ e \in G$, *such that,* $ea = a$ *for all* $a \in G$

(*iii*) *for all* $a \in G$, $\exists$ *some* $a' \in G$, *such that*, $a'a = e$.

Proof: A natural question to crop up at this stage would be what happens, when one sided identity and the other sided inverse exists. Would such a system also form a group? The answer to which is provided by the following illustration.

Let G be a finite set having at least two elements. Define '.' on G by

$ab = b$ for all $a, b \in G$

then clearly associativity holds in G.

Let $e \in G$ by any fixed element.

NOTES

Then as $ea = a$ for all $a \in G$

e will act as left identity.

Again $a \cdot e = e$ for all $a \in G$

$\Rightarrow$ e is right inverse for any element $a \in G$.

But we know G is not a group (cancellation laws do not hold in it).

Hence for a system $< G, . >$ to form a group it is essential that the same sided identity and inverse exist.

A Notation: Let G be a group with binary composition '.'. If $a \in G$ be any element then by closure property $a \cdot a \in G$. Similarly $(a \cdot a) \cdot a \in G$ and so on.

It would be very convenient (and natural!) to denote $a \cdot a$ by a^2 and $a \cdot (a \cdot a)$ or $(a \cdot a) \cdot a$ by a^3 and so on. Again $a^{-1} \cdot a^{-1}$ would be denoted by a^{-2}. And since $a \cdot a^{-1} = e$, it would not be wrong to denote $e = a^0$. It is now a simple matter to understand that under our notation

$$a^m \cdot a^n = a^{m+n}$$

$$(a^m)^n = a^{mn}$$

where m, n are integers.

In case the binary composition of the group is denoted by +, we will talk of sums and multiples in place of products and powers. Thus here $2a = a + a$, and $na = a + a + ... + a$ (n times), if n is a +ve integer. In case n is negative integer then $n = -m$, where m is positive and we define $na = -ma = (-a) + (-a) + ... + (-a)$ m times.

Example 1.1: *If G is a finite group of order n then show that for any $a \in G$, $\exists$ some positive integer r, $1 \le r \le n$, such that, $a^r = e$.*

Solution: Since $o(G) = n$, G has n elements.

Let $a \in G$ be any element. By closure property $a^2, a^3, ...$ all belong to G.

Consider $e, a, a^2, ..., a^n$

These are $n + 1$ elements (all in G). But G contains only n elements.

$\Rightarrow$ at least two of these elements are equal. If any of $a, a^2, ..., a^n$ equals e, our result is proved. If not, then $a^i = a^j$ for some i, j, $1 \le i, j \le n$. Without any loss of generality, we can take $i > j$

then $a^i = a^j$

$\Rightarrow$ $a^i \cdot a^{-j} = a^j \cdot a^{-j}$

$\Rightarrow$ $a^{i-j} = e$ where $1 \le i - j \le n$.

Putting, $i - j = r$ gives us the required result.

Example 1.2: *Show that a finite semi-group in which cross cancellation holds is an abelian group.*

NOTES

Solution: Let G be the given finite semi-group. Let $a, b \in G$ be any elements. Since G is a semi-group, by associativity

$$a(ba) = (ab)a$$

By cross cancellation then $ba = ab \Rightarrow G$ is abelian.

Since G is abelian, cross cancellation laws become the cancellation laws. Hence G is a finite semi-group in which cancellation laws hold.

Thus G is a group.

Example 1.3: *If G is a group in which* $(ab)^i = a^i b^i$ *for three consecutive integers i and any a, b in G, then show that G is abelian.*

Solution: Let $n, n+1, n+2$ be three consecutive integers for which the given condition holds. Then for any $a, b \in G$,

$$(ab)^n = a^n b^n \quad ...(1)$$

$$(ab)^{n+1} = a^{n+1} b^{n+1} \quad ...(2)$$

$$(ab)^{n+2} = a^{n+2} b^{n+2} \quad ...(3)$$

Now, $(ab)^{n+2} = a^{n+2} b^{n+2}$

$\Rightarrow (ab)(ab)^{n+1} = a^{n+2}\ b^{n+2}$

$\Rightarrow (ab)(a^{n+1} b^{n+1}) = a^{n+2}\ b^{n+2}$

$\Rightarrow ba^{n+1} = a^{n+1} b$ (using cancellation) ...(4)

Similarly, $(ab)^{n+1} = a^{n+1} b^{n+1}$

gives $(ab)(ab)^n = a^{n+1} b^{n+1}$

i.e., $(ab)(a^n b^n) = a^{n+1} b^{n+1}$

$\Rightarrow ba^n = a^n b$

$\Rightarrow ba^{n+1} = a^n ba$

$\Rightarrow a^{n+1} b = a^n ba$ using Equation (4)

$\Rightarrow ab = ba.$

Hence G is abelian.

Note: Conclusion of the above result may not follow if the given result holds only for two consecutive integers.

Consider, for example, the Quaternion group. One can check that $(ab)^i = a^i b^i$ for $i = 4, 5$ but the group is not abelian.

Example 1.4: *Suppose* $(ab)^n = a^n b^n$ *for all* $a, b \in G$ *where* $n > 1$ *is a fixed integer.*

Show that, (*a*) $(ab)^{n-1} = b^{n-1} a^{n-1}$

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

(*c*) $(aba^{-1} b^{-1})^{n(n-1)} = e$ *for all* $a, b \in G$

NOTES

Solution: (a) We have,

$$[b^{-1}(ba)b]^n = b^{-1}(ba)^n b$$

and $$[b^{-1}(ba)b]^n = (ab)^n$$

$$(ab)^n = b^{-1}(ba)^n b$$

$$\Rightarrow \quad (ab)^{n-1}ab = b^{-1}(b^n a^n)b$$

$$\Rightarrow \quad (ab)^{n-1} = b^{n-1}a^{n-1} \quad \text{for all } a, b \in G$$

(b) Now, $(a^{-1}b^{-1}ab)^n = a^{-n}b^{-n}a^n b^n$

and $(a^{-1}b^{-1}ab)^n = a^{-n}(b^{-1}ab)^n$

$$= a^{-n}b^{-1}a^n b$$

$$\therefore \quad a^{-n}b^{-n}a^n b^n = a^{-n}b^{-1}a^n b$$

$$\Rightarrow \quad a^n b^{n-1} = b^{n-1}a^n \quad \text{for all } a, b \in G$$

(c) Consider $(aba^{-1}b^{-1})^{n(n-1)}$

$$= [(aba^{-1}b^{-1})^{n-1}]^n$$

$$= [(ba^{-1}b^{-1})^{n-1}a^{n-1}]^n \text{ by } (i)$$

$$= [ba^{-(n-1)}b^{-1}a^{n-1}]^n = [b(a^{-(n-1)}b^{-1}a^{n-1})]^n$$

$$= b^n(a^{-(n-1)}b^{-1}a^{n-1})^n = b^n a^{-(n-1)}b^{-n}a^{n-1}$$

$$= a^{-(n-1)}b^n b^{-n}a^{n-1} \quad \text{by } (ii)$$

$$= e \quad \text{for all } a, b \in G.$$

Example 1.5: *Let G be a group and suppose there exist two relatively prime positive integers m and n such that* $a^m b^m = b^m a^m$ *and* $a^n b^n = b^n a^n$ *for all a, b* $\in G$*. Show that G is abelian.*

Solution: Since m, n are relatively prime, there exist integers x and y such that, $mx + ny = 1$.

For any a, b we have

$$(a^m b^n)^{mx} = (a^m b^n)(a^m b^n).....(a^m b^n) \quad mx \text{ times}$$

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

$$= a^m(b^n a^m)^{mx-1}b^n$$

$$= a^m(b^n a^m)^{mx}(b^n a^m)^{-1}b^n$$

$$= a^m c^m(b^n a^m)^{-1}b^n \text{ where } c = (b^n a^m)^x$$

$$= c^m a^m(b^n a^m)^{-1}b^n$$

$$= c^m \; a^m a^{-m}b^{-n}b^n = c^m = (b^n a^m)^{mx}$$

Similarly, $(a^m b^n)^{ny} = (b^n a^m)^{ny}$

giving $(a^m b^n)^{mx+ny} = (b^n a^m)^{mx+ny}$

$$\Rightarrow \quad a^m b^n = b^n a^m \quad \text{for all } a, b \in G \qquad ...(1)$$

Now, $$ab = a^{mx+ny} \; b^{mx+ny}$$

$$= a^{mx} \,.\, (a^{ny} \; b^{mx})b^{ny}$$

NOTES

$= a^{mx}(a^m\ k^m)b^{ny}$ where $d = a^y$, $k = b^x$

$= a^{mx}(k^m\ d^m)b^{ny}$ by (1)

$= a^{mx} \cdot b^{mx} \cdot a^{ny} \cdot b^{ny}$

$= (a^x)^m \cdot (b^x)^m \cdot (a^y)^n \cdot (b^y)^n$

$= (b^x)^m \cdot (a^x)^m \cdot (b^y)^n \cdot (a^y)^n$

$= b^{mx}(a^{mx} \cdot b^{ny}) \cdot a^{ny} = b^{mx}\ (b^{ny} \cdot a^{mx}) \cdot a^{ny}$

$= b^{mx+ny} \cdot a^{mx+ny} = ba.$

Hence G is abelian.

Note: In the following Theorem, we give another proof to Theorem 1.6 done earlier.

Subgroups

We have seen that **R**, the set of real numbers, forms a group under addition, and **Z**, the set of integers, also forms a group under addition. Also **Z** is a subset of **R**. It is one of the many situations which prompts us to make

Definition: A non empty subset H of a group G is said to be a subgroup of G, if H forms a group under the binary composition of G.

Obviously, if H is a subgroup of G and K is a subgroup of H, then K is subgroup of G.

If G is a group with identity element e then the subset $\{e\}$ and G are trivially subgroups of G and we call them the *trivial* subgroups. All other subgroups will be called non-trivial (or proper subgroups).

Notice that $\mathbf{Z}_5 = \{0, 1, 2, 3, 4\}$ mod 5 is not a subgroup of **Z** under addition as addition modulo 5 is not the composition of **Z**. Similarly, $\mathbf{Z}_5$ is not a subgroup of $\mathbf{Z}_6$, etc.

We sometimes use the notation $H \leq G$ to signify that H is a subgroup of G and $H < G$ to mean that H is a proper subgroup of G.

It may be a little cumbersome at times to check whether a given subset H of a group G is a subgroup or not by having to check all the axioms in the definition of a group. The following two theorems (especially the second one) go a long way in simplifying this exercise.

Theorem 1.8: *A non-empty subset H of a group G is a subgroup of G iff*

(i) $a, b \in H \Rightarrow ab \in H$

(ii) $a \in H \Rightarrow a^{-1} \in H$.

Proof: Let H be a subgroup of G then by definition it follows that (i) and (ii) hold.

Conversely, let the given conditions hold in H.

Closure holds in H by (i).

Again, $a, b, c \in H \Rightarrow a, b, c \in G \quad \Rightarrow \quad a(bc) = (ab)c$

NOTES

Hence associativity holds in H.

Also for any $a \in H$, $a^{-1} \in H$ and so by (i)

$$aa^{-1} \in H \Rightarrow e \in H$$

thus H has identity.

Inverse of each element of H is in H by (ii).

Hence H satisfies all conditions in the definition of a group and thus it forms a group and therefore a subgroup of G.

Theorem 1.9: *A non-void subset H of a group G is a subgroup of G iff* $a, b \in H \Rightarrow ab^{-1} \in H$.

Proof: If H is a subgroup of G then, $a, b \in H \Rightarrow ab^{-1} \in H$ (follows easily by using definition).

Conversely, let the given condition hold in H.

That associativity holds in H follows as in previous theorem.

Let $a \in H$ be any element $(H \neq \varphi)$

then $a, a \in H \Rightarrow aa^{-1} \in H \Rightarrow e \in H$.

So H has identity.

Again, for any $a \in H$, as $e \in H$

$$ea^{-1} \in H \Rightarrow a^{-1} \in H$$

i.e., H has inverse of each element.

Finally, for any $a, b \in H$,

$$a, b^{-1} \in H$$

$$\Rightarrow a(b^{-1})^{-1} \in H \Rightarrow ab \in H$$

i.e., H is closed under multiplication.

Hence H forms a group and therefore a subgroup of G.

Note: If the binary composition of the group is denoted by +, the above condition would read as $a, b \in H \Rightarrow a - b \in H$. Note also that e is always in H.

The following theorem may not prove to be very useful in as much as it confines itself to finite subsets only but nevertheless it has its importance.

Theorem 1.10: *A non empty finite subset H of a group G is a subgroup of G iff H is closed under multiplication.*

Proof: If H is a subgroup of G then it is closed under multiplication by definition, so there is nothing to prove.

Conversely, let H be a finite subset such that,

$$a, b \in H \Rightarrow ab \in H$$

Now, $a, b, c \in H \Rightarrow a, b, c \in G$

$\Rightarrow a(bc) = (ab)c$

$\therefore$ Associativity holds in H.

$\Rightarrow H$ is a semi-group.

Again, trivially the cancellation laws hold in H (as they hold in G) and thus H is a finite semi-group in which cancellation laws hold. Hence H forms a group.

Aliter: Let H be a finite subset such that, $a, b \in H \Rightarrow ab \in H$

We show $a \in H \Rightarrow a^{-1} \in H$.

If $a = e$ then $a^{-1} = a \in H$

Let $a \neq e$, then by closure $a, a^2, a^3 \ldots \in H$

Since H is finite, for some n, m, $a^n = a^m$, $n > m$

i.e., $a^{n-m} = e$, $n - m > 1$ as $a \neq e$

i.e., $a^{n-m-1}.\ a = e$

$\Rightarrow a^{n-m-1} = a^{-1}$

where $n-m-1 \geq 1$ and therefore,

$a^{n-m-1} \in H$. Hence $a \in H \Rightarrow a^{-1} \in H$ and thus H is a subgroup of G (Theorem 1.8).

Definition: Let G be a group. Let

$$Z(G) = \{x \in G \mid xg = gx \text{ for all } g \in G\}$$

then $Z(G)$ is called *centre* of the group G.

Theorem 1.11: *Centre of a group G is a subgroup of G.*

Proof: Let $Z(G)$ be the centre of the group G.

Then $Z(G) \neq \varphi$ as $e \in Z(G)$

Again, $x, y \in Z(G) \Rightarrow xg = gx$

$yg = gy$ for all $g \in G$

$\Rightarrow g^{-1} x^{-1} = x^{-1} g^{-1}$

$g^{-1} y^{-1} = y^{-1} g^{-1}$ for all $g \in G$

Now, $g(xy^{-1}) = (gx)y^{-1} = (xg)y^{-1}$

$= (xg)y^{-1} (g^{-1}g)$

$= xg(y^{-1} g^{-1})g = xg(g^{-1} y^{-1})g$

$= x(gg^{-1})y^{-1}g$

$= (xy^{-1})g$ for all $g \in G$

$\Rightarrow xy^{-1} \in Z(G)$

Hence $Z(G)$ is a subgroup.

Note: Obviously, G is abelian iff $Z(G) = G$.

Definition: Let G be a group. $a \in G$ be any element. The subset

$N(a) = \{x \in G \mid xa = ax\}$ is called *normalizer* or *centralizer* of a in G.

It is easy to see that normalizer is a subgroup of G.

NOTES

NOTES

Example 1.6: *Find centre of* S_3.

Solution: We have, $S_3 = \{I, (12), (13), (23), (123), (132)\}$

Centre of S_3, $Z(S_3) = \{\sigma \in S_3 \mid \sigma\theta = \theta\sigma \text{ for all } \theta \in S_3\}$

Since $(12)(13) = (132)$

$(13)(12) = (123)$

We find (12), (13) do not commute.

$\Rightarrow$ (12) and (13) do not belong to $Z(S_3)$

Again, $(23)(132) = (12)$

$(132)(23) = (13)$

$\Rightarrow$ (23), (132) do not belong to $Z(S_3)$

Also, $(123)(12) = (13)$

$(12)(123) = (23)$

Shows $(123) \notin Z(S_3)$

Hence $Z(S_3)$ contains only I.

Example 1.7: *Let G be the group of all 2×2 non singular matrices over the reals. Find centre of G.*

Solution: If $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be any element of the centre $Z(G)$ of G then it should commute with all members of G. In particular we should have,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$\Rightarrow b = c,\ a = d$$

Also, $\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ gives

$$\begin{bmatrix} a+b & b \\ c+d & d \end{bmatrix} = \begin{bmatrix} a & b \\ a+c & b+d \end{bmatrix}$$

$$\Rightarrow a + b = a,\ b = c = 0$$

Hence any member $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ of $Z(G)$ turns out to be of the type $\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$.

In other words, members of the centre $Z(G)$ are the 2×2 scalar matrices of G.

Example 1.8: *Let G be a group in which*

$$(ab)^3 = a^3b^3$$

$$(ab)^5 = a^5b^5, \quad \textit{for all } a, b \in G$$

Show that G is abelian.

Solution: We first show that $b^2 \in Z(G)$ for all $b \in G$.

We know $(a^{-1}ba)^3 = a^{-1}\ b^3a$

By given condition $(a^{-1}ba)^3 = a^{-3}\ (ba)^3 = a^{-3}\ b^3a^3$

$$\Rightarrow \quad a^{-1} b^3 a = a^{-3} b^3 a^3$$

$$\Rightarrow \quad a^2 b^3 = b^3 a^2 \quad \text{for all } a, b \in G$$

Similarly, $(a^{-1} ba)^5 = a^{-1} b^5 a$

$$(a^{-1} ba)^5 = a^{-5} b^5 a^5$$

$$\Rightarrow \quad a^{-1} b^5 a = a^{-5} b^5 a^5$$

$$\Rightarrow \quad a^4 b^5 = b^5 a^4 \Rightarrow a^4 b^3 b^2 = b^5 a^4$$

$$\Rightarrow \quad (a^2)^2 b^3 b^2 = b^5 a^4 \Rightarrow b^3 a^4 b^2 = b^5 a^4$$

$$\Rightarrow \quad a^4 b^2 = b^2 a^4 \Rightarrow a a^3 b^2 = b^2 a^4$$

$$\Rightarrow \quad ab^2 a^3 = b^2 a^4$$

$$\Rightarrow \quad ab^2 = b^2 a \quad \text{for all } a, b \in G$$

$$\therefore \quad b^2 \in Z(G) \quad \text{for all } b \in G$$

Now, $(ab)^4 = (ab)^5 (ab)^{-1} = a^5 b^5 b^{-1} a^{-1}$

$$= a^5 b^4 a^{-1} = a^5 a^{-1} b^4, \quad \text{as } b^2 \in Z(G)$$

$$= a^4 b^4$$

$\therefore$ $(ab)^i = a^i b^i$ for three consecutive integers $i = 3, 4, 5$

So, $ab = ba$ for all $a, b \in G$, by example done earlier.

Hence G is abelian.

Example 1.9: *Show that union of two subgroups may not be a subgroup.*

Solution: Let, $H_2 = \{2n \mid n \in \mathbf{Z}\}$

$H_3 = \{3n \mid n \in \mathbf{Z}\}$

where $(\mathbf{Z}, +)$ is the group of integers. H_2 and H_3 will be subgroups of $\mathbf{Z}$. Indeed

$$2n - 2m = 2(n - m) \in H_2$$

Now $H_2 \cup H_3$ is not a subgroup as $2, 3 \in H_2 \cup H_3$

but $2 - 3 = -1 \notin H_2 \cup H_3$

Theorem 1.12: *Union of two subgroups is a subgroup iff one of them is contained in the other.*

Proof: Let H, K be two subgroups of a group G and suppose $H \subseteq K$ then $H \cup K = K$ which is a subgroup of G.

Conversely, let H, K be two subgroups of G such that, $H \cup K$ is also a subgroup of G. We show one of them must be contained in the other. Suppose it is not true i.e.,

$$H \not\subseteq K, K \not\subseteq H$$

Then, $\exists\, x \in H$ such that, $x \notin K$

$\exists\, y \in K$ such that, $y \notin H$

Also then $x, y \in H \cup K$ and since $H \cup K$ is a subgroup, $xy \in H \cup K$

$$\Rightarrow xy \in H \text{ or } xy \in K$$

NOTES

If $xy \in H$, then as $x \in H$, $x^{-1}(xy) \in H \Rightarrow y \in H$, which is not true.

Again, if $xy \in K$, then as $y \in K$, $(xy)y^{-1} \in K \Rightarrow x \in K$ which is not true.

i.e., either way we land up with a contradiction.

Hence our supposition that $H \not\subseteq K$ and $K \not\subseteq H$ is wrong.

Thus one of the two is contained in the other.

Definition 1: Let H be a subgroup of a group G. For $a, b \in G$, we say a is congruent to b mod H if $ab^{-1} \in H$.

In notational form, we write $a \equiv b \bmod H$.

It is easy to prove that this relation is an equivalence relation. Corresponding to this equivalence relation, we get equivalence classes. For any $a \in G$, the equivalence class of a, we know will be given by

$$cl(a) = \{x \in G \mid x \equiv a \bmod H\}.$$

Definition 2: Let H be a subgroup of G and let $a \in G$ be any element. Then $Ha = \{ha \mid h \in H\}$ is called a right coset of H in G.

We show in the following theorem that any right coset of H in G is an equivalence class. To be exact we state:

Theorem 1.13: $Ha = \{x \in G \mid x \equiv a \bmod H\} = cl(a)$ *for any* $a \in G$.

Proof: Let, $x \in Ha$

Then, $x = ha$ for some $h \in H$

$\Rightarrow xa^{-1} = h$

$\Rightarrow xa^{-1} \in H$

$\Rightarrow x \equiv a \bmod H$

$\Rightarrow x \in cl(a)$

Thus, $Ha \subseteq cl(a)$.

Again let $x \in cl(a)$ be any element.

Then, $x \equiv a \bmod H$

$\Rightarrow xa^{-1} \in H$

$\Rightarrow xa^{-1} = h$ for some $h \in H$

$\Rightarrow x = ha \in Ha$

thus $cl(a) \subseteq Ha$

and hence $Ha = cl(a)$.

Having established that right cosets are equivalence classes, we are free to use the results that we know about equivalence classes. We can, therefore, say now that *any two right cosets are either equal or have no element in common* and also that *union of all the right cosets of H in G will equal G*.

Note: Note that a coset is not essentially a subgroup. If G be the Quaternion group then $H = \{1, -1\}$ is a subgroup of G. Take $a = i$, then $Ha = \{i, -i\}$ which is not a subgroup of G. (it doesn't contain identity). Refer Theorem 1.15 ahead.

Lemma: *There is always a* 1 – 1 *onto mapping between any two right cosets of H in G.*

Proof: Let *Ha*, *Hb* be any two right cosets of *H* in *G*.

Define a mapping $f : Ha \to Hb$, such that,

$$f(ha) = hb$$

Then $h_1a = h_2a \Rightarrow h_1 = h_2 \Rightarrow h_1b = h_2b$

$$\Rightarrow f(h_1a) = f(h_2a)$$

i.e., *f* is well defined.

$$f(h_1a) = f(h_2a) \Rightarrow h_1b = h_2b \Rightarrow h_1 = h_2 \Rightarrow h_1a = h_2a$$

Showing *f* is 1–1.

That *f* is onto, is easily seen, as for any $hb \in Hb$, *ha* would be its pre image.

The immediate utility of this lemma is seen, if the group *G* happens to be finite, because in that case the lemma asserts that any two right cosets of *H* in *G* have the same number of elements. Since *H* = *He* is also a right coset of *H* in *G*, this leads us to state that all right cosets of *H* in *G* have the *same* number of elements as are in *H* (*G*, being, of course, finite). We are now ready to prove

Theorem 1.14 (Lagrange's): *If G is a finite group and H is a subgroup of G then o*(*H*) *divides o*(*G*).

Proof: Let $o(G) = n$.

Since corresponding to each element in *G*, we can define a right coset of *H* in *G*, the number of distinct right cosets of *H* in *G* is less than or equal to *n*.

Using the properties of equivalence classes we know

$$G = Ha_1 \cup Ha_2 \cup \ldots \cup Ha_t$$

where, *t* = Number of distinct right cosets of *H* in *G*.

$$\Rightarrow \quad o(G) = o(Ha_1) + o(Ha_2) + \ldots + o(Ha_t)$$

(Reminding ourselves that two right cosets are either equal or have no element in common).

$$\Rightarrow \quad o(G) = o(H) + o(Ha) \underset{t \text{ times}}{+ \ldots +} o(H) \quad \text{using the above lemma}$$

$$\Rightarrow \quad o(G) = t.\ o(H)$$

or that $o(H) \mid o(G)$

and we have proved a very important theorem.

But a word of caution here. Converse of Lagrange's theorem does not hold.

Note: If *G* is a group of prime order, it will have only two subgroups *G* and $\{e\}$. Refer Theorem 1.25 also.

We have been talking about *right cosets* of *H* in *G* all this time. Are there left cosets also? The answer should be an obvious yes. After all we can similarly

NOTES

NOTES

talk of

$$aH = \{ah \mid h \in H\}, \quad \text{for any } a \in G$$

which would be called a *left coset*. One can by defining similarly an equivalence relation ($a \equiv b \bmod H \Leftrightarrow a^{-1} b \in H$) prove all similar results for left cosets. It would indeed be an interesting 'brushing up' for the reader, by proving these results independently.

We now come to a simple but very important

Theorem 1.15: *Let H be a subgroup of G then,*

(*i*) $Ha = H \Leftrightarrow a \in H$; $aH = H \Leftrightarrow a \in H$

(*ii*) $Ha = Hb \Leftrightarrow ab^{-1} \in H$; $aH = bH \Leftrightarrow a^{-1} b \in H$

(*iii*) Ha (*or* aH) *is a subgroup of G iff* $a \in H$.

Proof: (*i*) Let $Ha = H$

Since $e \in H$, $ea \in Ha \Rightarrow ea \in H \Rightarrow a \in H$.

Let $a \in H$, we show $Ha = H$.

Let $x \in Ha \Rightarrow x = ha$ for some $h \in H$

Now $h \in H$, $a \in H \Rightarrow ha \in H \Rightarrow x \in H \Rightarrow Ha \subseteq H$

Again, let $y \in H$, since $a \in H$

$$ya^{-1} \in H$$

$$\Rightarrow \quad ya^{-1} = h \quad \text{for some } h \in H$$

$$\Rightarrow \quad y = ha \in Ha$$

$$\Rightarrow \quad H \subseteq Ha$$

Hence $Ha = H$.

(*ii*) $Ha = Hb$

$$\Leftrightarrow \quad (Ha)b^{-1} = (Hb)b^{-1}$$

$$\Leftrightarrow \quad Hab^{-1} = He$$

$$\Leftrightarrow \quad Hab^{-1} = H$$

$$\Leftrightarrow \quad ab^{-1} \in H \quad \text{using } (i)$$

(*iii*) If $a \in H$ then $Ha = H$ which is a subgroup. Conversely, if Ha is a subgroup of G then $e \in Ha$ and thus the right cosets Ha and He have one element e in common and hence $Ha = He = H \Rightarrow a \in H$ by (*i*).

Corresponding results for left cosets can be tackled similarly.

Definition: Let G be a group and H, a subgroup of G. Then *index* of H in G is the number of distinct right (left) cosets of H in G. It is denoted by $i_G(H)$ or $[G:H]$.

A look at the proof of Lagrange's theorem suggests that if G is a finite group, then $i_G(H) = \dfrac{o(G)}{o(H)}$.

It is, of course, possible for an infinite group G to have a subgroup $H (\neq G)$ with finite index.

NOTES

Case 18: Let $< \mathbf{Z}, + >$ be the group of integers under addition.

Let $H = \{3n \mid n \in \mathbf{Z}\}$ then H is a subgroup of $\mathbf{Z}$. We show H has only three right cosets in $\mathbf{Z}$ namely H, $H + 1$, $H + 2$.

If $a \in \mathbf{Z}$ be any element $(\neq 0, 1, 2)$ then we can write (by division algorithm),

$$a = 3n + r, \qquad 0 \leq r < 3$$

which gives

$$H + a = H + (3n + r) = (H + 3n) + r = H + r$$

where $0 \leq r < 3$

Hence H has only 3 right cosets in $\mathbf{Z}$ and thus has index 3.

Notice, $H - 1 = (H + 3) - 1 = H + (3 - 1) = H + 2$, etc.

Case 19: Let $G = < \mathbf{R} - \{0\}, . >$, i.e., let G be the group of non zero real numbers under multiplication. Let $H = \{1, -1\}$. Then H is a subgroup of G where $i_G(H)$ is infinite. Notice H has infinite number of right cosets in G, these being $\{2, -2\}$, $\{3, -3\}$, $\{4, -4\}$, ..., etc.

Definition: Let H be a subgroup of a group G, we define

$C(H) = \{x \in G \mid xh = hx \text{ for all } h \in H\}$ then $C(H)$ is called *centralizer* of H in G.

Also the set

$$N(H) = \{x \in G \mid xH = Hx\}$$
$$= \{x \in G \mid xHx^{-1} = H\}$$

is called *normalizer* of H in G.

It is an easy exercise to see that both $C(H)$ and $N(H)$ are subgroups of G.

Again as, $x \in C(H) \Rightarrow xh = hx$ for all $h \in H$

$$\Rightarrow xH = Hx$$
$$\Rightarrow x \in N(H)$$

we notice $C(H) \subseteq N(H)$.

However, $C(H)$ need not be equal to $N(H)$ as consider the Quaternion group $G = \{\pm 1, \pm i, \pm j, \pm k\}$ and let $H = \{\pm 1, \pm i\}$.

Then $N(H) = G$ and $C(H) = \{\pm 1, \pm i\}$.

Showing that $C(H) \neq N(H)$

Note: One can define $C(H)$ or $N(H)$ in the same way even if H happens to be only a non-empty *subset* of G.

Example 1.10: *Show that* $C(H) = G \Leftrightarrow H \subseteq Z(G)$.

Solution: Let $C(H) = G$. Let $h \in H$ be any element. Then, $x \in G \Rightarrow x \in C(H)$ $\Rightarrow xh = hx \Rightarrow$ any element h in H commutes with all elements of $G \Rightarrow h \in Z(G) \Rightarrow H \subseteq Z(G)$.

Conversely, let $H \subseteq Z(G)$. Let $x \in G$. Since $H \subseteq Z(G)$ each element of H commutes with every element of G.

$\Rightarrow xh = hx$ for all $h \in H$

$\Rightarrow x \in C(H) \Rightarrow G \subseteq C(H) \Rightarrow G = C(H)$.

NOTES

Example 1.11: *Show that there exists a one-one onto map between the set of all left cosets of H in G and the set of all right cosets of H in G where H is a subgroup of a group G.*

Solution: Let, $\Im$ = Set of all left cosets of H in G.

$\Re$ = Set of all right cosets of H in G.

Define a mapping $\theta : \Im \to \Re$, such that,

$$\theta(aH) = Ha^{-1} \quad a \in G$$

θ is well defined as $aH = bH$

$$\Rightarrow a^{-1} b \in H$$
$$\Rightarrow Ha^{-1} = Hb^{-1}$$
$$\Rightarrow \theta(aH) = \theta(bH)$$

Taking the steps backwards, we find θ is 1–1. Again, for any $Ha \in \Re$, $a^{-1}H$ is the required pre-image under θ proving that θ is onto.

If G is finite, then the above result reduces to saying that number of left cosets of H in G is same as the number of right cosets of H in G.

Example 1.12: *Let H be a subgroup of a group G and* $N(H) = \{a \in G \mid aHa^{-1} = H\}$. *Prove that N(H) is a subgroup of G which contains H.*

Solution: $N(H) \neq \phi$ subset of G as

$$eHe^{-1} = H \Rightarrow e \in N(H)$$

Let now $a, b \in N(H)$ be any two elements, then

$$aHa^{-1} = H$$
$$bHb^{-1} = H$$

then, $bHb^{-1} = H \Rightarrow b^{-1}(bHb^{-1})b = b^{-1}Hb$

$$\Rightarrow (b^{-1}b)Hb^{-1}b = b^{-1}Hb$$
$$\Rightarrow H = b^{-1}Hb$$
$$\Rightarrow aHa^{-1} = a(b^{-1}Hb)a^{-1}$$
$$\Rightarrow aHa^{-1} = ab^{-1}Hba^{-1}$$
$$\Rightarrow H = (ab^{-1})H(ab^{-1})^{-1}$$
$$\Rightarrow ab^{-1} \in N(H)$$ i.e., $N(H)$ is a subgroup of G.

Since $h \in H \Rightarrow hHh^{-1} = H$ ($Ha = H \Leftrightarrow a \in H$, etc.)

we find $h \in N(H)$ showing that $H \subseteq N(H)$.

Example 1.13: *Suppose that H is a subgroup of a group G such that whenever* $Ha \neq Hb$ *then* $aH \neq bH$. *Prove that* $gHg^{-1} \subseteq H$ *for all* $g \in G$.

Solution: It is given that if $Ha \neq Hb$ then $aH \neq bH$

thus if $aH = bH$ then $Ha = Hb$. ...(1)

Let now $g \in G, h \in H$ be any elements, then

$$(g^{-1}h)H = g^{-1}(hH) = g^{-1}H \quad (h \in H)$$

$\therefore$ By Equation (1) $H(g^{-1}h) = Hg^{-1}$

$\Rightarrow (g^{-1}h)(g^{-1})^{-1} \in H \quad (Ha = Hb \Rightarrow ab^{-1} \in H)$

$\Rightarrow g^{-1}hg \in H \quad$ for all $h \in H$

$\Rightarrow g^{-1}Hg \subseteq H.$

NOTES

Example 1.14: *If $G = S_3$ and $H = \{I, (13)\}$, write all the left cosets of H in G.*

Solution: $_{(12)}H = \{(12)I, (12)(13)\} = \{(12), (132)\}$

$= {}_{(123)}H$ (Show!)

$_{(23)}H = \{(23)I, (23)(13)\} = \{(23), (132)\} = {}_{(132)}H$

$_{(13)}H = H$ as $(13) \in H$

$IH = H$

are all the left cosets of H in G.

Definition: Let H and K be two subgroups of a group G. We define $HK = \{hk \mid h \in H, k \in K\}$ then HK will be a non-empty subset of G (Sometimes, called the *complex* of H and K). Will it form a subgroup? The answer is provided by

Theorem 1.16: *HK is a subgroup of G iff $HK = KH$.*

Proof: Let HK be a subgroup of G. We show $HK = KH$

Let, $x \in HK$ be any element

Then, $x^{-1} \in HK$ (as HK is a subgroup)

$\Rightarrow x^{-1} = hk$ for some $h \in H, k \in K$

$\Rightarrow x = (hk)^{-1} = k^{-1}h^{-1} \in KH$

Thus, $HK \subseteq KH$

Again let $y \in KH$ be any element

Then, $y = kh$ for some $k \in K, h \in H$

$\Rightarrow y^{-1} = h^{-1}k^{-1} \in HK$

$\Rightarrow y \in HK$ (as HK is a subgroup)

$\Rightarrow KH \subseteq HK$

Hence, $HK = KH.$

Conversely, let $HK = KH$.

Let, $a, b \in HK$ be any two elements, we show $ab^{-1} \in HK$

$a, b \in HK \Rightarrow a = h_1k_1$ for some $h_1, h_2 \in H$

$b = h_2k_2 \quad k_1, k_2 \in K$

Then, $ab^{-1} = (h_1k_1)(h_2k_2)^{-1} = (h_1k_1)(k_2^{-1}h_2^{-1})$

$= h_1(k_1k_2^{-1})h_2^{-1}$

Now, $(k_1k_2^{-1})h_2^{-1} \in KH = HK$

Thus, $(k_1k_2^{-1})h_2^{-1} = hk$ for some $h \in H, k \in K$

NOTES

Then, $ab^{-1} = h_1(hk) = (h_1h)k \in HK$

Hence, HK is a subgroup.

Notes: 1. $HK = KH$ does not mean that each element of H commutes with every element of K. It only means that for each $h \in H$, $k \in K$, $hk = k_1h_1$ for some $k_1 \in K$ and $h_1 \in H$.

2. If G has binary composition +, we define

$$H + K = \{h + k \mid h \in H, k \in K\}.$$

Theorem 1.17: *If H and K are finite subgroups of a group G then*

$$o(HK) = \frac{o(H) \cdot o(K)}{o(H \cap K)}.$$

Proof: Let $D = H \cap K$ then D is a subgroup of K and as in the proof of Lagrange's theorem, $\exists$ a decomposition of K into disjoint right cosets of D in K and

$$K = Dk_1 \cup Dk_2 \cup ... \cup Dk_t$$

and also $t = \frac{o(K)}{o(D)}$

Again, $HK = H(\bigcup_{i=1}^{t} Dk_i)$ and since $D \subseteq H$, $HD = H$

Thus, $HK = \bigcup_{i=1}^{t} Hk_i = Hk_1 \cup Hk_2 \cup ... \cup Hk_t$

Now no two of $Hk_1, Hk_2, ..., Hk_t$ can be equal as if $Hk_i = Hk_j$ for i, j then $k_ik_j^{-1} \in H \Rightarrow k_ik_j^{-1} \in H \cap K \Rightarrow k_ik_j^{-1} \in D \Rightarrow Dk_i = Dk_j$

which is not true.

Hence, $o(HK) = o(Hk_1) + (Hk_2) + ... + o(Hk_t)$

$= o(H) + o(H) + ... + o(H)$

$= t \cdot o(H)$

$= \frac{o(H) \cdot o(K)}{o(H \cap K)}$

which proves the result.

Aliter: We have $HK = \{hk \mid h \in H, k \in K\}$.

Let $H \cap K = \{x_1, x_2, ..., x_n\}$ and suppose $o(H) = r$, $o(K) = s$

Now $hk = (hx_i)(x_i^{-1}k) \in HK \quad \forall\ i = 1, 2, ..., n$

Also, $hx_i \in H,\ x_i^{-1}\ k \in K$ as $x_i \in H$ and K

Thus, $hk = (hx_i)\left(x_i^{-1}\ k\right) \in HK \quad \forall\ i = 1, 2, ..., n$

or that hk can be written in at least n different ways. We show these are the only n ways that hk can be expressed as an element of HK.

Suppose $hk = h_1k_1$

$\Rightarrow\ h^{-1}h_1 = kk_1^{-1} \in H \cap K$

$\Rightarrow\ h^{-1}h_1 = x_i$

and $kk_1^{-1} = x_i$ for some $i = 1, 2, ..., n$

or that $h_1 = hx_i$

$k_1 = x_i^{-1}k$

and thus $hk = h_1k_1 = (hx_i)(x_i^{-1}k)$

Hence each hk can be written in exactly n different ways.

Since h can be chosen in r ways, k can be chosen in s ways, we find hk can be chosen in $\frac{rs}{n}$ ways.

Thus, $o(HK) = \frac{rs}{n} = \frac{o(H).o(K)}{o(H \cap K)}$.

Note $o(H \cap K) \geq 1$ as $H \cap K \neq \varphi$ as $e \in H \cap K$.

Corollary: If H and K are subgroups of a finite group G such that $o(H) > \sqrt{o(G)}$, $o(K) > \sqrt{o(G)}$ then $o(H \cap K) > 1$.

Proof: We have,

$$o(G) \geq o(HK) = \frac{o(H)\,o(K)}{o(H \cap K)} > \frac{\sqrt{o(G)}.\sqrt{o(G)}}{o(H \cap K)} = \frac{o(G)}{o(H \cap K)}$$

$$\Rightarrow\ o(H \cap K) > 1.$$

Example 1.15: *Suppose G is a finite group of order pq, where p, q are primes and p > q. Show that G has at most one subgroup of order p.*

Solution: Suppose H, K are two subgroups of order p.

Then, as $o(H \cap K) \mid o(H) = p$, we find

$o(H \cap K) = 1$ or p

If, $o(H \cap K) = 1$, then

$$o(HK) = \frac{o(H).o(K)}{o(H \cap K)} = \frac{p.p}{1} = p^2 > pq = o(G)$$

$[p > q \ \Rightarrow\ p^2 > pq]$

which is not possible. Hence $o(H \cap K) = p = o(H)$

and as $H \cap K \subseteq H$, we find $H \cap K = H$

Similarly, $H \cap K = K$ and hence $K = H$.

There exists at least one subgroup of order p. A group of order 15 will have only one subgroup of order 5.

Note: We have defined the product HK of two subgroups H and K. The same definition can be used for the product, even if H, K happen to be subsets of G.

Example 1.16: *Let H, K be subgroups of G. Show that HK is a subgroup of G if and only if HK = KH.*

NOTES

NOTES

Solution: Suppose HK is a subgroup of G.

Then, $HK = (HK)^{-1} = K^{-1}H^{-1} = KH$

Conversely, let $HK = KH$

$$\begin{aligned}(HK)(HK)^{-1} &= (HK)(K^{-1}H^{-1}) \\ &= (HK)\,(KH) = H(KK)H \\ &= H(KH) = H(HK) \\ &= (HH)K = HK\end{aligned}$$

Then, HK is a subgroup of G.

Cyclic Groups

Definition: Order of an element : Let G be a group and $a \in G$ be any element. We say a is of order (or period) n if n is the least +ve integer such that, $a^n = e$. If binary composition of G is denoted by +, this would read as $na = 0$, where 0 is identity of G.

If it is not possible to find such n, we say a has infinite order. Order of a will be denoted by $o(a)$. It is obvious that $o(a) = 1$ iff $a = e$.

***Cyclic Group*:** A group G is called a *cyclic* group if $\exists$ an element $a \in G$, such that every element of G can be expressed as a power of a. In that case a is called *generator* of G. We express this fact by writing $G = < a >$ or $G = (a)$.

Thus G is called cyclic if $\exists$ an element $a \in G$ such that, $G = \{a^n \mid n \in \mathbf{Z}\}$. Again, if binary composition of G is denoted by +, the words 'power of a' would mean multiple of a.

Note we are not saying that generator is unique. Indeed if a is generator so would be a^{-1}. A simple example of a cyclic group is the group of integers under addition, 1 being its generator.

Again the group $G = \{1, -1, i, -i\}$ under multiplication is cyclic as we can express its members as i, i^2, i^3, i^4. Thus i (or $-i$) is a generator of this group.

Case 20: The group $\mathbf{Z}_n = \{0, 1, 2, \ldots, n-1\}$ addition modulo $n (n \geq 1)$ is a cyclic group. 1 and $-1 = n - 1$ will be its generators. But it can have more generators besides these. (Refer Theorem 1.30 ahead).

Consider, $\mathbf{Z}_8 = \{0, 1, 2, \ldots 7\}$ addition modulo 8

Then we can check that 1, 3, 5, 7 will be generators of $\mathbf{Z}_8$

Notice that,

$3^1 = 3$, $3^2 = 3 \oplus 3 = 6$, $3^3 = 3 \oplus 3 \oplus 3 = 1$

$3^4 = 3 \oplus 3 \oplus 3 \oplus 3 = 4$ and so on

i.e., $< 3 > = \{3, 6, 1, 4, 7, 2, 5, 0\}$ or that 3

is a generator of $\mathbf{Z}_8$. Observe also that 1, 7 and 3, 5 are each others inverses.

On the other hand, U_n, the group under multiplication modulo n is not cyclic for every n. For instance U_5 is cyclic. But U_8 is not cyclic.

Theorem 1.18: *Order of a cyclic group is equal to the order of its generator.*

Proof: Let $G = < a >$ i.e., G is a cyclic group generated by a.

Case (*i*): $o(a)$ is finite, say n, then n is the least +ve integer such that, $a^n = e$.

Consider the elements $a^0 = e, a, a^2, ..., a^{n-1}$

These are all elements of G and are n in number.

Suppose any two of the above elements are equal

say $a^i = a^j$ with $i > j$

Then, $a^i \cdot a^{-j} = e \Rightarrow a^{i-j} = e$

But $0 < i - j \leq n - 1 < n$, thus $\exists$ a positive integer $i - j$, such that, $a^{i-j} = e$ and $i - j < n$, which is a contradiction to the fact that $o(a) = n$.

Thus no two of the above n elements can be equal, i.e., G contains at least n elements. We show it does not contain any other element. Let $x \in G$ be any element. Since G is cyclic, generated by a, x will be some power of a.

Let $x = a^m$

By division algorithm, we can write

$m = nq + r$ where $0 \leq r < n$

Now, $a^m = a^{nq+r} = (a^n)^q \cdot a^r = e^q . a^r = a^r$

$\Rightarrow$ $x = a^r$ where $0 \leq r < n$

i.e., x is one of $a^0 = e, a, a^2, ..., a^{n-1}$

or G contains precisely n elements

$\Rightarrow$ $o(G) = n = o(a)$

Case (*ii*): $o(a)$ is infinite.

In this case no two powers of a can be equal as if $a^n = a^m$ $(n > m)$ then $a^{n-m} = e$, i.e., it is possible to find a positive integer $n - m$ such that, $a^{n-m} = e$ meaning thereby that a has finite order.

Hence no two powers of a can be equal. In other words G would contain infinite number of elements.

Example 1.17: *If $a \in G$ be of finite order n and also $a^m = e$ then show that $n \mid m$.*

Solution: Let $o(a) = n$, then by definition n is the least positive integer such that, $a^n = e$.

Suppose $a^m = e$ for some m

By division algorithm, $m = nq + r$, where $0 \leq r < n$

$a^m = a^{nq + r}$

$\Rightarrow$ $e = a^{nq} . a^r = (a^n)^q . a^r = e^q . a^r = a^r$

where $0 \leq r < n$

Since n is such least positive integer, we must have $r = 0$

i.e., $m = nq$ or that $n \mid m$.

NOTES

NOTES

Example 1.18: *If G is a finite abelian group then show that o(ab) is a divisor of l.c.m. of o(a), o(b).*

Solution: Let $o(a) = n$, $o(b) = m$, $o(ab) = k$.

Let l = l.c.m. (m, n)

then $m \mid l$, $n \mid l$, $\Rightarrow l = mr_1$, $l = nr_2$

Now, $(ab)^l = a^l b^l$ (G is abelian)

$= a^{nr2} b^{mr1} = e \cdot e = e$

$\Rightarrow o(ab) \mid l$

$\Rightarrow k \mid l$.

Example 1.19: *If in a group G, $a^5 = e$, $aba^{-1} = b^2$ for a, $b \in G$ then show that $o(b) = 31$.*

Solution: We have $b^2 = aba^{-1}$

$\Rightarrow b^4 = (aba^{-1})\ (aba^{-1})$

$= ab(a^{-1}a)^{nr}b^2\ ab^{mr1} = ab^2a^{-1}$

$= a(aba^{-1})\ a^{-1}$

$\Rightarrow b^4 = a^2ba^{-2}$

$\Rightarrow b^8 = (a^2ba^{-2})\ (a^2ba^{-2}) = a^2b^2a^{-2}$

$= a^2(aba^{-1})a^{-2} = a^3ba^{-3}$

$\Rightarrow b^{16} = a^4ba^{-4}$ (as above)

$\Rightarrow b^{32} = a^5ba^{-5} = b$ as $a^5 = e$

$\Rightarrow b^{31} = e \Rightarrow 31$ is a multiple of $o(b)$

Since 31 is a prime number, it is the least positive integer such that $b^{31} = e$

$\Rightarrow o(b) = 31$.

We are, of course, taking $b \neq e$.

Theorem 1.19: *A subgroup of a cylic group is cyclic.*

Proof: Let $G = < a >$ and let H be a subgroup of G. If $H = \{e\}$, there is nothing to prove. Let $H \neq \{e\}$. Members of H will be powers of a. Let m be the least positive integer such that, $a^m \in H$. We claim $H = < a^m >$.

Let $x \in H$ be any element. Then $x = a^k$ for some k. By division algorithm, $k = mq + r$ where $0 \leq r < m$

$\Rightarrow r = k - mq$

$\Rightarrow a^r = a^k \cdot a^{-mq} = x \cdot (a^m)^{-q} \in H$

But m is the least positive integer such that, $a^m \in H$, meaning thereby that $r = 0$.

Thus, $k = mq$

or that $x = a^k = (a^m)^q$

i.e., any member of H is a power of a^m.

or that H is cyclic, generated by a^m.

Note: Any subgroup of $< \mathbf{Z}, + >$ will therefore, be of the type $n\mathbf{Z}$ = set of multiples of n, where n is an integer (≥ 0). We write $n\mathbf{Z} = < n >$.

NOTES

Also $m\mathbf{Z} \subseteq n\mathbf{Z}$ if and only if $n \mid m$. So $m\mathbf{Z} = n\mathbf{Z}$ if and only if $m = \pm n$.

Case 21: Let $H = < a > = \{an \mid n \in \mathbf{Z}\} = a\mathbf{Z}$

$K = < b > = \{bm \mid m \in \mathbf{Z}\} = b\mathbf{Z}$

be two subgroups of $< \mathbf{Z}, + >$, then $\mathbf{Z}$ being abelian, $H + K = K + H$

$\Rightarrow H + K$ is a subgroup of $\mathbf{Z}$.

[Note here $HK = H + K$].

We show $H + K = < d > = d\mathbf{Z}$, where d = g.c.d.(a, b)

Now, $x \in H + K$

$\Rightarrow$ $x \in < a > + < b >$

$\Rightarrow$ $x = an + bm, \quad n, m \in \mathbf{Z}$

$\Rightarrow$ $x \in < d >$ [as $d \mid a, d \mid b \Rightarrow d \mid an + bm \Rightarrow d \mid x$]

Thus $H + K \subseteq < d >$.

Again, $y \in < d > \Rightarrow y = td$

$\Rightarrow y = t(ax + by) = atx + bty \in H + K$

Hence $H + K = < d >$

i.e., $a\mathbf{Z} + b\mathbf{Z} = (a, b)\mathbf{Z}$.

Theorem 1.20: *A cyclic group is abelian.*

Proof: Let $G = < a >$. If $x, y \in G$ be any elements then $x = a^n$, $y = a^m$ for some integers m, n.

Now $xy = a^n \cdot a^m = a^{n+m} = a^{m+n} = a^m \cdot a^n = y \cdot x$

Hence G is abelian.

Note: In view of the above result, all non abelian groups are non-cyclic. $< \mathbf{Q}, + >$ the group of rationals under addition serves as an example of an abelian group which is not cyclic. For, suppose $\frac{m}{n} \in \mathbf{Q}$ is a generator of $\mathbf{Q}$, then any element of $\mathbf{Q}$ should be a multiple of $\frac{m}{n}$. Now $\frac{1}{3n} \in \mathbf{Q}$, and if $\frac{m}{n}$ is a generator, we should be able to write $\frac{1}{3n} = k \frac{m}{n}$, for some k

$$\Rightarrow \frac{1}{3} = km$$

Which is not possible as k, m are integers, whereas $\frac{1}{3}$ is not. Hence no element can act as generator of $\mathbf{Q}$.

NOTES

Klein's four group would be an example of a finite abelian group which is not cyclic. It is the group of matrices $\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}, \begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}$ and $\begin{bmatrix}-1 & 0\\ 0 & -1\end{bmatrix}$ under matrix multiplication.

Theorem 1.21: *If G is a finite group, then order of any element of G divides order of G.*

Proof: Let $a \in G$ be any element.

Let $H = \{a^n \mid n \text{ an integer}\}$

then H is a cyclic subgroup of G, generated by a, as

$$x, y \in H \Rightarrow x = a^n, y = a^m$$

$$\therefore \quad xy^{-1} = a^n \cdot a^{-m} = a^{n-m} \in H$$

By Lagrange's theorem $o(H) \mid o(G)$. But $o(H) = o(a)$

$$\therefore \quad o(a) \mid o(G).$$

Corollary: If G is a finite group then for any $a \in G$

$$a^{o(G)} = e$$

Proof: $o(a) \mid o(G) \Rightarrow o(G) = o(a)k$ For some k

Now, $a^{o(G)} = a^{o(a)k} = (a^{o(a)})^k = e^k = e$

Thus any element of a finite group, has finite order (which is less than or equal to the order of the group). Converse is, however, not true.

Case 22: The group $< \mathbf{Z}, + >$ of integers is an example of a group in which each non identity element is of infinite order.

As another example consider $G = \{2^r : r = 0, \pm 1, ...\}$

then we know G forms a group under multiplication. No non-identity element in G has finite order as

$$(2^r)^n = 1 \quad \text{iff } 2^{rn} = 1$$

iff $r = 0$ or $n = 0$.

Note: If G is a finite group of order n and $\exists$ an element $a \in G$, such that, $o(a) = n$ then G is cyclic, generated by a. Clearly $o(a) = n$ gives $a^n = e$, and lesser powers not equal to e and thus $G = \{a, a^2, ..., a^n = e\}$.

Example 1.20: *Let G be a finite group whose order is not divisible by* 3. *Suppose* $(ab)^3 = a^3b^3$ *for all* $a, b \in G$, *then show that G is abelian.*

Solution: Let $a, b \in G$ be any elements.

Then as, $(ab)^3 = a^3b^3$

we get $ababab = a^3b^3$

$\Rightarrow \quad baba = a^2b^2$ (cancellation)

$\Rightarrow \quad (ba)^2 = a^2b^2$ ...(1)

Again as, $(ba)^3 = b^3a^3$

we get $(ba)(ba)^2 = b^3a^3$

$\Rightarrow (ba)a^2b^2 = b^3a^3$ using Equation (1)

$\Rightarrow a^3b^2 = b^2a^3$ ...(2)

Consider now, $(a^{-1}b^{-2}ab^2)^3 = (a^{-1})^3 (b^{-2}ab^2)^3 = a^{-3} (b^{-2}ab^2)^3$

$= a^{-3} (b^{-2}a^3b^2)$

$= a^{-3} (b^{-2}b^2a^3)$ from Equation (2)

$= a^{-3}a^3 = e$

$\Rightarrow o(a^{-1}b^{-2}ab^2) \mid 3$

$\Rightarrow o(a^{-1}b^{-2}ab^2) = 1$ or 3

If $o(a^{-1}b^{-2}ab^2) = 3$ then $3 \mid o(G)$ which is not true.

Hence $o(a^{-1}b^{-2}ab^2) = 1$

$\Rightarrow a^{-1}b^{-2}ab^2 = e$

$\Rightarrow ab^2 = b^2a$ ...(3)

Again from, (1) $(ba)^2 = a^2b^2 = a(ab^2) = a(b^2a)$ using Equation (3)

$\Rightarrow (ba)(ba) = ab^2a$

$\Rightarrow bab = ab^2 \Rightarrow ba = ab$

or that G is abelian.

Theorem 1.22: *Converse of Lagrange's theorem holds in finite cyclic groups.*

Proof: Let $G = < a >$ be a finite cyclic group of order n.

Then, $o(G) = o(a) = n$

Suppose $m \mid n$. We show $\exists$ a subgroup of G having order m.

Since $m \mid n$, $\exists$ k such that, $n = mk$

Let H be the cyclic group generated by a^k

then H is a subgroup of G and $o(H) = o(a^k)$

We show $o(a^k) = m$

Now, $(a^k)^m = a^{km} = a^n = e$, as $o(a) = n$

Suppose now, that $(a^k)^t = e$

Then, $a^{kt} = e$

$\Rightarrow o(a) \mid kt \Rightarrow n \mid kt$

$\Rightarrow km \mid kt \Rightarrow m \mid t$

thus $o(a^k) = m$

which proves the result.

Note: One can go a step further here and show that such a subgroup (as taken above) would also be unique. Suppose H' is another subgroup of G such that, $o(H') = m$. Since H' is a subgroup of a cyclic group $G = < a >$, H' will be generated by some power of a.

NOTES

NOTES

Let $H' = < a^p >$

By division algorithm,

$$p = kq + r \qquad 0 \le r < k$$

$$\Rightarrow mp = mkq + mr \qquad 0 \le mr < mk$$

$$\Rightarrow a^{mp} = a^{mkq + mr} = (a^{mk})^q \cdot a^{mr}$$

$$= a^{mr} \quad (o(a) = n = mk)$$

Now, $o(H') = o(a^p) = m$

$\Rightarrow (a^p)^m = e$

thus $a^{mr} = e$ where $0 \le mr < n$

But this $\Rightarrow mr = 0$ (as $o(a) = n$)

$\Rightarrow r = 0$ as $m \neq 0$

hence $p = kq$

Thus $H' = < a^p > = < a^{kq} > \subseteq < a^k > = H$

But $o(H') = o(H)$

$\Rightarrow H = H'$.

We thus conclude:

Theorem 1.23: *If G is a finite cyclic group of order n then the number of distinct subgroups of G is the number of distinct divisors of n, and there is at most one subgroup of G of any given order.*

Proof: So subgroups of G are of the type $< a^k >$ where k is a divisor of n and $< a^{n/m} >$ is the unique subgroup of order m. As a particular case, suppose $G = < a >$ has order 30. Since divisors of 30 are 1, 2, 3, 5, 6, 10, 15, 30, $\exists$ eight subgroups of G, namely

$$< a > = \{e, a, a^2, ..., a^{29}\} = G$$

$$< a^2 > = \{e, a^2, a^4, ..., a^{28}\}$$

$$< a^3 > = \{e, a^3, a^6, ..., a^{27}\}$$

$< a^5 >, < a^6 >, < a^{10} >, < a^{15} >$ and $< a^{30} > = \{e\}$ having order 30, 15, 10, 6, 5, 3, 2, 1.

Consider again, the cyclic group $\mathbf{Z}_{30} = \{0, 1, 2, ..., 29\}$ under addition modulo 30 . $o(\mathbf{Z}_{30}) = 30$ and as 30 has 8 divisors 1, 2, 3, 5, 6, 10, 15, 30, $\mathbf{Z}_{30}$ will have eight subgroups namely

$$< 1 > = \{0, 1, 2, ..., 29\} = \mathbf{Z}_{30}$$

$$< 2 > = \{0, 2, 4, ..., 28\}$$

$$< 3 > = \{0, 3, 6, ..., 27\}$$

$$< 5 >, < 6 >, < 10 >, < 15 >, < 30 > = \{0\}$$

having order 30, 15, 10, 6, 5, 3, 2, 1.

In view of the above theorem these would be the only subgroups of $\mathbf{Z}_{30}$.

NOTES

Theorem 1.24: *A group G of prime order must be cyclic and every element of G other than identity can be taken as its generator.*

Proof: Let $o(G) = p$, a prime

Take any $a \in G$, $a \neq e$

and let $H = \{a^n \mid n \text{ an integer}\}$ then H is a cyclic subgroup of G.

$\therefore \quad o(H) \mid o(G) \Rightarrow o(H) = 1$ or p

But, $\quad o(H) \neq 1$ as $a \in H$, $a \neq e$,

Thus $o(H) = p \Rightarrow H = G$, i.e., G is a cyclic group generated by a. Since a was taken as any element (other than e), any element of G can act as its generator.

Corollary: A group of prime order is abelian.

Theorem 1.25: *A group G of prime order cannot have any non-trivial subgroups.*

Proof: If H is any subgroup of G then as $o(H) \mid o(G) = p$, a prime

We find $o(H) = 1$ or p

i.e., $\quad H = \{e\}$ or $H = G$.

Theorem 1.26: *A group of finite composite order has at least one non-trivial subgroup.*

Proof: Let $o(G) = n = rs$ where $1 < r, s < n$

Since $n > 1$, $\exists\ e \neq a \in G$. Consider a^r.

Case (*i*): $a^r = e$

then $\quad o(a) \leq r$, let $o(a) = k$

then $\quad 1 < k \leq r < n \quad (k > 1, \text{ as } a \neq e)$

Let, $\quad H = \{a, a^2, a^3, ..., a^k = e\}$

then H is a non-empty finite subset of G and it is closed under multiplication, thus H is a subgroup of G. Since $o(H) = k < n$, we have proved the result.

Case (*ii*): $a^r \neq e$, then since $(a^r)^s = a^{rs} = a^n = a^{o(G)} = e$

$o(a^r) \leq s$. Let $o(a^r) = t$ then $1 < t \leq s < n$.

If we take $K = \{a^r, a^{2r}, ..., a^{tr} = e\}$ then K is a non empty finite subset of G, closed under multiplication and is therefore a subgroup of G. Its order being less than n, it is the required subgroup.

Theorem 1.27: *If G is a group having no non-trivial subgroups then G must be finite having prime order.*

Proof: Suppose G has infinite order.

Then we can find $a \in G$, such that, $a \neq e$.

Let $H = < a >$, then H is a cyclic subgroup of G and $H \neq \{e\}$. But G has no non-trivial subgroups.

Thus, $\quad H = G$

$\Rightarrow \quad G = < a >$

NOTES

Consider now the subgroup $K = < a^2 >$

Now $a \notin < a^2 >$, because if $a \in < a^2 >$ then $a = a^{2t}$ for some integer t

$$\Rightarrow a^{2t-1} = e \Rightarrow o(a) \leq 2t - 1$$

meaning thereby that $o(a)$ if finite, which is not true. Thus $a \notin < a^2 >$.

Again $< a^2 > \neq \{e\}$, because then $a^2 = e$ would again mean that $o(a)$ is finite (≤ 2).

Thus $< a^2 >$ is a non-trivial subgroup of G which is not possible. Hence $o(G)$ cannot be infinite.

So $o(G)$ is finite and as it cannot be composite by previous theorem, it must be prime.

Summing up, what we have done above proves

Theorem 1.28: *The only groups which have no non-trivial subgroups are the cyclic groups of prime order and the group* $\{e\}$.

All this time we have been talking about cyclic groups and their generators without being very sure as to how many generators a cyclic group could have. To resolve this, we consider

Theorem 1.29: *An infinite cyclic group has precisely two generators.*

Proof: Let $G = < a >$ be an infinite cyclic group.

As mentioned earlier, if a is a generator of G then so would be a^{-1}.

Let now b be any generator of G,

Then as $b \in G$, a generates G, we get $b = a^n$ for some integer n

Again as $a \in G$, b generates G, we get $a = b^m$ for some integer m

$$\Rightarrow a = b^m = (a^n)^m = a^{nm}$$

$$\Rightarrow a^{nm-1} = e \Rightarrow o(a) \text{ is finite and } \leq nm - 1$$

Since $o(G) = o(a)$ is infinite, the above can hold only if

$$nm - 1 = 0 \Rightarrow nm = 1$$

$$\Rightarrow m = \frac{1}{n} \text{ or } n = \pm 1 \text{ as } m, n \text{ are integers.}$$

i.e., $b = a$ or a^{-1}

In other words, a and a^{-1} are precisely the generators of G.

Question to be answered now is how many generators a finite cyclic group would have. Before we come to the answer we first define what is popularly known as the ***Euler's φ function*** (or Euler's totient function).

For any integer n, we define $\varphi(1) = 1$ and for $n > 1$, $\varphi(n)$ to be the number of positive integers less then n and relatively prime to n. As an example $\varphi(6) = 2$, $\varphi(10) = 4$, etc.

Note 1, 5 are less than 6 and relatively prime to 6 and 1, 3, 7, 9 (four in number) are less than 10 and relatively prime to 10, etc. Obviously, $\varphi(p) = p - 1$, if p is a prime. The following two results can be helpful at times.

(*i*) If $p_1, p_2, ..., p_n$ are distinct prime factors of n (>1), then

$$\varphi(n) = n\left(1-\frac{1}{p_1}\right)\left(1-\frac{1}{p_2}\right) ... \left(1-\frac{1}{p_k}\right)$$

(*ii*) If m, n are co-prime then

$\varphi(mn) = \varphi(m)\ \varphi(n)$, $(m, n \geq 1)$

We are now ready to prove

NOTES

Theorem 1.30: *Number of generators of a finite cyclic group of order n is* $\varphi(n)$.

Proof: Let $G = < a >$ be a cyclic group of order n

then $o(a) = o(G) = n$

We claim a^m is generator of G iff $(m, n) = 1$, i.e., m, n are relatively prime.

[For instance, if $n = 8$, then $\varphi(8) = 4$ will be number of generators as we will show a, a^3, a^5, a^7 will generate G and no other element can generate G. So here m can have values 1, 3, 5, 7].

Let now a^m be a generator of G for some m

Since $a \in G$, $a = (a^m)^i$ for some i

$$\Rightarrow a^{mi-1} = e \Rightarrow o(a) \mid mi - 1$$

$$\Rightarrow n \mid mi - 1$$

$$\Rightarrow mi - 1 = nj \quad \text{for some integer } j$$

i.e., $mi - nj = 1$

$$\Rightarrow (m, n) = 1.$$

Conversely, let $(m, n) = 1$

Then $\exists$ integers x and y such that,

$$mx + ny = 1$$

$$\Rightarrow a^{mx+ny} = a$$

$$\Rightarrow a^{mx} \cdot a^{ny} = a$$

$$\Rightarrow a^{mx} (a^n)^y = a$$

$$\Rightarrow a^{mx} = a \quad \text{as } o(a) = n$$

$$\Rightarrow a = (a^m)^x$$

Since every element of G is a power of a and a itself is a power of a^m, we find a^m generates G, which proves our result.

Note: We thus realize that if a is a generator of a finite cyclic group G of order n, then other generators of G are of the type a^m where m and n are coprime.

In fact an integer k will be a generator of $\mathbf{Z}_n$ if and only if k and n are coprime, and thus generators of $\mathbf{Z}_n$ would indeed be the elements of U_n.

Theorem (Euler's) 1.31: *Let a, n* $(n \geq 1)$ *be any integers such that* g.c.d. $(a, n) = 1$. *Then*

$$a^{\varphi(n)} \equiv 1 \ (mod\ n).$$

NOTES

Proof: Let $U_n = \{x \mid x \text{ is an integer, } (x, n) = 1, 1 \leq x < n\}$

Then U_n is a group under multiplication modulo n.

By definition of Euler's φ-function,

$$o(U_n) = \varphi(n).$$

If $n = 1$, then $\varphi(n) = \varphi(1) = 1$ and $a^{\varphi(n)} = a^1 \equiv 1 \pmod 1$ (as 1 divides $a - 1$)

Let $n > 1$

Now by Euclid's algorithm

$a = nq + r$, for some integers q, r where $0 \leq r < n$.

If $r = 0$ then $a = nq \Rightarrow n \mid a \Rightarrow (a, n) = n > 1$, a contradiction

$\therefore \quad 1 \leq r < n$

Also $(r, n) = d \Rightarrow d \mid r, d \mid n \Rightarrow d \mid a–nq, d \mid nq$

$\Rightarrow d \mid a, d \mid n$

$\Rightarrow d \mid (a, n) = 1$

$\Rightarrow d = 1$

$\therefore \quad (r, n) = 1$ and $1 \leq r < n$

$\Rightarrow r \in U_n$

Also $a = nq + r \Rightarrow a \equiv r \pmod n$

It follows from Lagrange's theorem that,

$$r \otimes r \otimes \ldots \otimes r = \text{identity of } U_n = 1 \quad [a^{o(G)} = e]$$

where $\otimes$ is composition multiplication modulo n in U_n and $\varphi(n)$ is order of group U_n.

$\therefore \quad r^{\varphi(n)} – nq_1 = 1$, for some integer q_1

$\Rightarrow r^{\varphi(n)} \equiv 1 \pmod n$

$\Rightarrow a^{\varphi(n)} \equiv 1 \pmod n$

so, $a \equiv r \pmod n \Rightarrow a^{\varphi(n)} \equiv r^{\varphi(n)} \pmod n$.

Theorem (Fermat's) 1.32: *For any integer a and prime p,*

$$a^p \equiv a \ (mod\ p).$$

Proof: If $(a, p) = 1$, then by Euler's theorem

$a^{\varphi(p)} \equiv 1 \pmod p$

$\Rightarrow a^{p-1} \equiv 1 \pmod p$ as $\varphi(p) = p - 1$

$\Rightarrow a^{\,p} \equiv a \pmod p$

If $(a, p) = p$, then $p \mid a \Rightarrow p \mid a^p$

$\therefore \quad p \mid a^p - a$

$\Rightarrow a^p \equiv a \pmod{p}$.

(Note $(a, p) = 1$ or p as 1 and p are only divisors of p).

Example 1.21: *Show that if G is a group of order* 10 *then it must have a subgroup of order* 5.

NOTES

Solution: By Lagrange's theorem such a subgroup can exist.

We first claim that all elements of G cannot be of order 2. Suppose it is so.

Let $a, b \in G$ be two different elements with order 2.

Let $H = < a >$, $K = < b >$ be the cyclic subgroups generated by a and b

then $o(H) = 2$, $o(K) = 2$

Since all elements of G are of order 2, it must be abelian.

$\therefore$ $HK = KH \Rightarrow HK$ is a subgroup of G

and as $$o(HK) = \frac{o(H) \cdot o(K)}{o(H \cap K)} = \frac{2 \times 2}{1} = 4$$

[Note $H \cap K = \{e\}$ as $a \neq b$]

By Lagrange's theorem $o(HK)$ would divide $o(G)$

i.e., 4 | 10 which is not true hence our assumption is wrong and thus all elements of G cannot have order 2.

Again, since G is finite, $o(a) \mid o(G)$ for all $a \in G$

$\Rightarrow \exists$ at least one element $a \in G$, such that, $o(a) = 5$ or 10.

If $o(a) = 5$, then $H = < a >$ is a subgroup of order 5.

If $o(a) = 10$, then $H = < a^2 >$ is a subgroup of order 5.

In any case our result is proved.

Example 1.22: *Let G be a group such that intersection of all its subgroups which are different from* $\{e\}$ *is a subgroup different from* $\{e\}$. *Prove that every element of G has finite order.*

Solution: Let $a \in G$ be any element.

If $a = e$, $o(a) = 1$

Let $a \neq e$ and suppose $o(a)$ is not finite.

Consider the cyclic subgroups $< a >, < a^2 >, < a^3 >, \ldots$

Since each $< a^i > \neq \{e\}$ as $o(a)$ is not finite

$< a > \cap < a^2 > \cap < a^3 > \cap \ldots \neq \{e\}$ by given condition.

As intersection of cyclic subgroups is cyclic subgroup

$$\bigcap_i < a^i > = < a^m > \text{ for some integer } m$$

Again, $< a^m > \subseteq < a^i >$ for all i

In particular, $< a^m > \subseteq < a^{2m} >$

But $< a^{2m} > \subseteq < a^m >$

(multiples of $2m$ are multiples of m)

NOTES

$$\Rightarrow < a^m > = < a^{2m} >$$

Thus, $a^m \in < a^m > \Rightarrow a^m \in < a^{2m} >$

$$\Rightarrow a^m = (a^{2m})^k$$

$$\Rightarrow a^{m(2k-1)} = e$$

$\Rightarrow o(a)$ is finite, a contradiction.

Hence the result follows.

Theorem 1.33: *If G is a finite group of order n and for every divisor d of n ∃ unique subgroup of order d, then G is cyclic.*

Proof: Let $d \mid n$.

Define $A(d) = \{x \in G \mid o(x) = d\}$

Suppose $A(d) \neq \varphi$. Then $\exists\ x \in G$ such that, $o(x) = d$.

Let $H = < x >$. Then $o(x) = o(H) = d$. This gives $\varphi(d)$ generators of H or $\varphi(d)$ elements of order d in H. If $\exists\ y \in G, y \notin H$ such that, $o(y) = d$, then $K = < y >$ is a subgroup of order d. It is given that G has unique subgroup of order d. So, $K = H \Rightarrow y \in H$, a contradiction. Thus, the number of elements in G of order d is $\varphi(d)$.

So, $o(A(d)) = \varphi(d)$ if $A(d) \neq \varphi$

and $o(A(d)) = 0$ if $A(d) = \varphi$ for all $d \mid n$

Clearly, $G = \underset{d|n}{\cup} A(d)$

Let $d_1, ..., d_s$ be all divisors of n.

Suppose $A(d_1) = \varphi, ..., A(d_i) = \varphi$

and $A(d_{i+1}) \neq \varphi, ..., A(d_s) \neq \varphi$

(Note, if $A(d) = \varphi$ for all $d \mid n$, then $o(G) = 0$, a contradiction. So, $A(d) \neq \varphi$ for some $d \mid n$)

$\therefore$ $o(A(d_1)) = ... = o(A(d_i)) = 0$

and $o(A(d_{i+1})) = \varphi(d_{i+1}) ..., o(A(d_s)) = \varphi(d_s)$

Now $G = \underset{d|n}{\cup} A(d) \Rightarrow o(G) = \sum_{d|n} o(A(d))$

$$\Rightarrow n = \varphi(d_{i+1}) + ... + \varphi(d_s)$$

By Example 1.21, $n = \sum_{d|n} \varphi(d)$

$\Rightarrow \varphi(d_1) + ... + \varphi(d_i) + \varphi(d_{i+1}) + ... + \varphi(d_s) = \varphi(d_{i+1}) + ... + \varphi(d_s)$

$\Rightarrow \varphi(d_1) + ... + \varphi(d_i) = 0$, a contradiction

So, $A(d) \neq \varphi$ for all $d \mid n$. In particular

$A(n) \neq \varphi \Rightarrow \exists\ x \in A(n) \Rightarrow \exists\ x \in G$ such that, $o(x) = n = o(G) \Rightarrow G$ is a cyclic group.

Example 1.23: *Show that in a cyclic group of order n, ∃ $\varphi(m)$ elements of order m for every divisor m of n. Deduce that* $n = \sum_{d|n} \varphi(d)$.

NOTES

Solution: Let m divide n. Then $\exists$ a unique subgroup H of G such that $o(H) = m$.

Let $H = < b >$

Then $m = o(H) = o(b)$

The number of elements of order m in H equals the number of generators of H. But the number of generators of H is $\varphi(m)$. So, the number of elements of order m in H is $\varphi(m)$. If $k \in G$ such that, $o(k) = m$, then $K = < k >$ has order m. Since G, has unique subgroup of order m, $K = H$.

$\therefore$ $k \in H$. So, all elements of order m belong to H.

This gives total number of elements of order m in G to be $\varphi(m)$.

Let $a \in G$ such that, $o(a) = d$. Then $d \mid o(G) = n$.

From above $\exists$ $\varphi(d)$ elements of order d in G. In this way, count all elements of G to get $n = \sum_{d|n} \varphi(d)$.

Example 1.24: *Let G be a group.*

Show that $o(a^n) = \dfrac{o(a)}{(o(a), n)}$ *for all* $a \in G$

where n is an integer and $(o(a), n) = g.c.d.\ (o(a), n)$.

Solution: Let $o(a) = m$.

Let $d = (m, n) \Rightarrow \dfrac{m}{d}, \dfrac{n}{d}$ are integers

$\therefore$ $(a^n)^{m/d} = (a^m)^{n/d} = e^{n/d} = e$

Let $(a^n)^r = e \Rightarrow a^{nr} = e$

$\Rightarrow o(a) \mid nr$

$\Rightarrow m \mid nr$

$\Rightarrow \dfrac{m}{d} \Big| \dfrac{n}{d} r$

$\Rightarrow \dfrac{m}{d} \Big| r$ as $\left(\dfrac{m}{d}, \dfrac{n}{d}\right) = 1$

$\Rightarrow r \geq \dfrac{m}{d}$

$\therefore$ $o(a^n) = \dfrac{m}{d} = \dfrac{o(a)}{(o(a), n)}$.

Example 1.25: *Let G be a group. Suppose a, b* $\in$ *G, such that,*

(*a*) $ab = ba$

(*b*) $(o(a), o(b)) = 1$.

Show that $o(ab) = o(a)\ o(b)$.

Solution: Let $o(a) = m$, $o(b) = n$

NOTES

Then, $(ab)^{mn} = a^{mn}b^{mn}$ as $ab = ba$

$= (a^m)^n (b^n)^m$

$= e$

Let, $(ab)^r = e \Rightarrow a^r b^r = e$

$\Rightarrow ar = b^{-r}$

$\Rightarrow (a^r)^n = (b^{-r})^n = (b^n)^{-r} = e$

$\Rightarrow o(a) \mid rn$

$\Rightarrow m \mid rn$

$\Rightarrow m \mid r$ as $(m, n) = 1$

Similarly, $n \mid r$

$\Rightarrow$ l.c.m. of $(m \& n) \mid r$

$\Rightarrow mn \mid r \Rightarrow mn \leq r$

$\therefore \quad o(ab) = mn.$

1.2.1 Normal and Subnormal Series

Definition: A normal subgroup H of a group G is called a *maximal normal* subgroup of G if $H \neq G$ and there exists no normal subgroup K of G such that, $H \subset K \subset G$.

Thus $H \neq G$ is a maximal normal subgroup of G if whenever $K \triangleleft G$ such that, $H \subseteq K \subseteq G$ then either $K = H$ or $K = G$.

In fact, a subgroup $H \neq G$ is called maximal subgroup of G if whenever $H \leq K \leq G$ then either $K = G$ or $K = H$.

Case 23: A_3 is a maximal normal subgroup of S_3. $o(A_3) = 3$ whereas $o(S_3) = 6$. Clearly there cannot be any subgroups of order 4 or 5 in S_3. We also notice that $o\left(\frac{S_3}{A_3}\right) = 2$, a prime and thus $\frac{S_3}{A_3}$ is a simple group.

Case 24: If G is a simple group then it has no non-trivial normal subgroups and so $\{e\}$ will be a (and only) maximal normal subgroup in G.

Theorem 1.34: *H is a maximal normal subgroup of G iff G/H is simple.*

Proof: Let H be maximal normal in G. Any subgroup of G/H is of the form K/H where $K \leq G$ and $H \subseteq K$ and also K/H is normal in $G/H \Leftrightarrow K \trianglelefteq G$.

Thus any subgroup K/H will be non trivial normal subgroup of G/H if $H \triangleleft K \triangleleft G$, which is not true as H is maximal normal. So G/H has no non trivial normal subgroup and is, therefore, simple.

Conversely: Let G/H be simple. Suppose H is not maximal normal, then $\exists$ a normal subgroup K of G such that,

$H \subset K \subset G$ and thus K/H will be normal subgroup of G/H where $K/H \subset G/H$, a contradiction as G/H is simple.

NOTES

Example 1.26: *Any finite group G* (*with at least two elements*) *has a maximal normal subgroup.*

Solution: If G is simple then it has no proper normal subgroup except $\{e\}$ and thus $\{e\}$ is a maximal normal subgroup of G.

Suppose G is not simple. Then it has at last one normal subgroup $N \neq G$, $N \neq \{e\}$. If N is maximal normal, we are done. If not, then $\exists$ at least one normal subgroup M where $N \subsetneq M \subsetneq G$. If M is maximal normal, we are done. If not, we continue like this. Since G is finite, it can have finite number of subgroups and hence the above process must end after a finite number of steps. Hence G will have a maximal normal subgroup.

Example 1.27: *Let H, K be two distinct maximal normal subgroups of G then G = HK and H ∩ K is a maximal normal subgroup of H as well as K.*

Solution: Since H, K are normal, HK is normal in G.

Since $H \subseteq HK \subseteq G$ and HK is maximal normal.

We must have $HK = H$ or $HK = G$

Similarly, $HK = K$ or $HK = G$

Hence $HK = G$ (as $HK \neq G \Rightarrow HK = H$, $HK = K \Rightarrow H = K$).

Again by isomorphism theorem

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

Thus, $$\frac{K}{H \cap K} \cong \frac{G}{H}$$

Since H is maximal normal, $\frac{G}{H}$ is simple

i.e., $\frac{K}{H \cap K}$ is simple

$\Rightarrow H \cap K$ is maximal normal in K

Similarly, it is maximal normal in H.

Example 1.28: *Show that* $(\mathbf{Q}, +)$ *has no maximal normal subgroup.*

Solution: Suppose H is a maximal normal subgroup of $(\mathbf{Q}, +)$, then $\frac{\mathbf{Q}}{H}$ is simple and so $\frac{\mathbf{Q}}{H}$ has no non trivial normal subgroup i.e., it will have no non trivial subgroup ($\mathbf{Q}$ being abelian, all subgroups are normal). Thus $\frac{\mathbf{Q}}{H}$ is a cyclic group of prime order p.

Let $H + x \in \frac{\mathbf{Q}}{H}$ be any element

Then $p(H + x) = H$

NOTES

i.e., $H + px = H$ or that $px \in H \quad \forall\, x \in \mathbf{Q}$

Let now $y \in \mathbf{Q}$ be any element, then $\frac{y}{p} \in \mathbf{Q}$

If $\frac{y}{p} = x$ then $y = px \Rightarrow y \in H$ or that

$\mathbf{Q} \subseteq H \subseteq \mathbf{Q} \Rightarrow H = \mathbf{Q}$, a contradiction.

Hence the result follows.

Definition: Let G be a group. A sequence of subgroups

$$\{e\} = G_0 \subseteq G_1 \subseteq G_2 \subseteq \ldots\ldots\ldots \subseteq G_n = G \qquad \ldots(1.3)$$

is called a *normal series* of G if G_i is a normal subgroup of G_{i+1},

$$\forall\ i = 0,\ 1\ \ 2,\ldots,n-1.$$

The factor (quotient) groups $\frac{G_{i+1}}{G_i}$ $(\forall\, i)$ are called the *factors* of the normal series.

Here each G_i is normal in G_{i+1}, although it may not be normal in G. Also it is possible that $G_i = G_{i+1}$ for some i. The number of distinct members of Equation (1.3) excluding G is called the length of the normal series.

The above is expressed in short by saying that $N = (G_0, G_1, \ldots\ldots\ldots, G_n)$ is a normal series of G. If N and M are two normal series of G such that, $N \subseteq M$ then M is called a *refinement* of N (a proper refinement if $N \subsetneqq M$).

Note: Some authors prefer to call the above a subnormal series. It is then called a normal series if G_i is normal in $G\ \forall\ i$.

If G is any group then

$$\{e\} = G_0 \subseteq G_1 = G$$

is an obvious example of a normal series.

Case 25: $\{I\} \subseteq A_3 \subseteq S_3$ is a normal series of S_3.

$\{I\} \subseteq E \subseteq K_4 \subseteq A_4 \subseteq S_4$ is a normal series of S_4, where

$E = \{I, (12)(34)\}$, $K_4 = \{I, (12)(34), (13)(24), (14)(23)\}$

1.2.2 Composition Series

Definition: Let G a group. A sequence of subgroups

$$\{e\} = G_0 \subset G_1 \subset G_2 \subset \ldots\ldots\ldots \subset G_n = G$$

of G is called a *composition series* of G if

(*i*) Each G_i is normal subgroup of G_{i+1} $(i = 0, 1, \ldots\ldots\ldots n-1)$,

(*ii*) $G_i \neq G_{i+1}$ for any i and

(*iii*) $\frac{G_{i+1}}{G_i}$ is a simple group $\forall\ i$.

The factor (quotient) groups $\frac{G_{i+1}}{G_i}$ are called factors of the series.

The condition (*iii*) can be replaced by 'G_i is a maximal normal subgroup of G_{i+1}' $\forall$ *i*.

NOTES

We notice that a composition series is a normal series (converse being not true) and that a composition series has no 'Gaps'.

A group can have more then one composition series.

Case 26: $\{0\} \subset <8> \subset <4> \subset \mathbf{Z}$
is a normal series of the group (**Z**, +), but it is not a composition series as $<4>$ is not maximal normal in **Z**. Notice $<4> \subset <2> \subset \mathbf{Z}$.

Case 27: Consider the quaternion group *G*. Then

$$\{1\} \subset \{1, -1\} \subset \{1, -1, i, -i\} \subset G$$
$$\{1\} \subset \{1, -1\} \subset \{1, -1, j, -j\} \subset G$$
$$\{1\} \subset \{1, -1\} \subset \{1, -1, k, -k\} \subset G$$

are all composition series of *G*. If we write the first series as $G_0 \subset G_1 \subset G_2 \subset G$ then

$$o\left(\frac{G}{G_2}\right) = \frac{8}{4} = 2, \quad o\left(\frac{G_2}{G_1}\right) = \frac{4}{2} = 2, \quad o\left(\frac{G_1}{G_0}\right) = 2$$

i.e., all the factor groups are of prime order and thus have no trivial normal subgroups and hence are simple.

The existence of a composition series is ensured by

Theorem 1.35: *Every finite group G* (*with more than one element*) *has a composition series.*

Proof: We use induction on $o(G)$.

If $o(G) = 2$ then $\{e\} = G_0 \subset G_1 = G$ is (only) composition series of *G*.

Notice $\frac{G_1}{G_0} = \frac{G}{\{e\}} \cong G$ and as $o(G) = 2$, a prime it is simple group and, therefore, $\frac{G_1}{G_0}$ is simple.

Suppose now that the result holds for groups with order less than $o(G)$. We show result holds for *G*. If *G* is a simple group then $\{e\} \subset G$ is the composition series for *G*. Suppose *G* is not simple.

Since *G* is finite, it has a maximal normal subgroup $N \neq G$ and as $o(N) < o(G)$, result holds for *N* which then has a composition series, say,

$$\{e\} \subset N_1 \subset N_2 \subset \ldots \subset N$$

Then the series

$\{e\} \subset N_1 \subset N_2 \subset \ldots \subset N \subset G$ will be a composition series for *G*.

Hence the result holds.

NOTES

Note: If $o(G) = 1$, we sometimes say that the result holds trivially as then (G) is a composition series of G (without factors).

Definition: Two composition series.

$$C_1 : \{e\} = N_0 \subset N_1 \subset ..... \subset N_t = G \quad ...(1.4)$$

$$C_2 : \{e\} = H_0 \subset H_1 \subset ..... \subset H_m = G \quad ...(1.5)$$

of a group G are said to be equivalent if $\exists$ a 1-1 onto mapping between the factors of Equation (1.4) and factors of Equation (1.5) such that the corresponding factor groups are isomorphic. In other words (1.4) and (1.5) will be equivalent if $t = m$ and each factor group of Equation (1.4) is isomorphic to some factor group of Equation (1.5).

Also in this case, we write $C_1 \sim C_2$. It is easy to see that $\sim$ is an equivalence relation.

We have seen that a finite group can have more than one composition series. The next theorem shows the equivalence of any two such composition series.

1.2.3 Jordan-Holder Theorem

Theorem 1.36 (Jordan-Hölder): *Let G be a finite group. Let*

$$C_1 : \{e\} = N_0 \subset N_1 \subset ... \subset N_{t-1} \subset N_t = G \quad ...(1.6)$$

$$C_2 : \{e\} = H_0 \subset H_1 \subset ... \subset H_{m-1} \subset H_m = G \quad ...(1.7)$$

be two composition series of G. Then m = t and there exists a permutation $i \to i'$ *of* 0, 1, 2, ..., t–1 such that, $\frac{N_{i+1}}{N_i} \cong \frac{H_{i'+1}}{H_{i'}}$, $0 \le i \le t-1$

i.e., C_1 *and* C_2 *are equivalent.*

Proof: Let $o(G) = n$. We use induction on n.

If $n = 2$, we have seen (Theorem 1.35) G has only one composition series. Hence result holds in this case.

Let now the result hold for groups with order less than $o(G)$.

Case (*i*) $N_{t-1} = H_{m-1}$. Consider the series

$$\{e\} = N_0 \subset N_1 \subset ... \subset N_{t-1} \quad ...(1.8)$$

$$\{e\} = H_0 \subset H_1 \subset ... \subset H_{m-1} = N_{t-1} \quad ...(1.9)$$

Then these are composition series for finite group N_{t-1} and as $o(N_{t-1}) < o(G)$, the result holds for Equation (1.8) Equation and Equation (1.9), i.e., Equation (1.8) and Equation (1.9) are equivalent.

Thus, $t - 1 = m - 1 \Rightarrow t = m$

and also factors of Equation (1.8) and Equation (1.9) are isomorphic under some permutation.

Now, $$\frac{N_t}{N_{t-1}} = \frac{G}{N_{t-1}} = \frac{G}{H_{m-1}} = \frac{H_m}{H_{m-1}}$$

Thus, Equation (1.6) and Equation (1.7) will be equivalent (as $t = m$ and factors of Equation (1.6) and Equation (1.7) are isomorphic). Hence result holds in this case.

Case (*ii*) $N_{t-1} \neq H_{m-1}$. Let $K = N_{t-1} \cap H_{m-1}$

Then K is a finite group and has a composition series. Let

$\{e\} = K_0 \subset K_1 \subset ... \subset K_s = K$ be a composition series of K.

Since N_{t-1}, H_{m-1} are normal in G, $K = N_{t-1} \cap H_{m-1}$ will be normal subgroup of G

Again, as N_{t-1}, H_{m-1} are maximal normal subgroups of G

$$N_{t-1}.\ H_{m-1} = G$$

and $N_{t-1} \cap H_{m-1} = K$ is maximal normal subgroup of N_{t-1} and H_{m-1}. (Refer Example 1.27)

So, $K \subset N_{t-1}$, $K \subset H_{m-1}$

Consider now the series,

$$\{e\} = K_0 \subset K_1 \subset ..... \subset K_s = K \subset N_{t-1} \subset N_t = G \quad ...(1.10)$$

$$\{e\} = K_0 \subset K_1 \subset ..... \subset K_s = K \subset H_{m-1} \subset H_m = G \quad ...(1.11)$$

We show these are composition series of G. For this we need show that $\frac{N_{t-1}}{K}$ and $\frac{H_{m-1}}{K}$ are simple.

By isomorphism theorem

$$\frac{N_{t-1}}{N_{t-1} \cap H_{m-1}} \cong \frac{N_{t-1}\, H_{m-1}}{H_{m-1}} = \frac{G}{H_{m-1}}$$

So, $\frac{N_{t-1}}{N_{t-1} \cap H_{m-1}} \cong \frac{G}{H_{m-1}}$ and similarly $\frac{H_{m-1}}{N_{t-1} \cap H_{m-1}} \cong \frac{G}{N_{t-1}}$ ...(1.12)

Now, $\frac{G}{H_{m-1}} = \frac{H_m}{H_{m-1}}$ is simple as Equation (1.7) is a composition series of G

$\Rightarrow$ $\frac{N_{t-1}}{N_{t-1} \cap H_{m-1}}$ is simple

i.e., $\frac{N_{t-1}}{K}$ is simple.

Similarly, $\frac{H_{m-1}}{K}$ is simple.

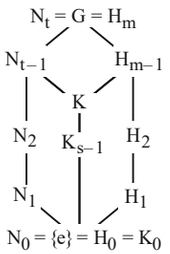
Now Equation (1.10) and Equation (1.11) would be equivalent as

NOTES

$$\frac{N_{t-1}}{K} \cong \frac{H_m}{H_{m-1}} \quad \text{and} \quad \frac{N_t}{N_{t-1}} \cong \frac{H_{m-1}}{K} \qquad \text{from Equation (1.12)}$$

Also lengths of Equation (1.10) and Equation (1.11) are equal both being $s + 2$

NOTES



Now Equation (1.6) and Equation (1.10) are two composition series of $N_t = G$ and applying case (i) to these (second last terms are equal = N_{t-1}) we find they are equivalent. Hence they have same length, i.e., $t = s + 2$

Similarly, (1.7) and (1.11) give $m = s + 2$

$$\Rightarrow \; t = m$$

Now Equation (1.6) ~ Equation (1.10), Equation (1.10) ~ Equation (1.11) $\Rightarrow$ (1.6) ~ Equation (1.11)

Also Equation (1.7) ~ (1.11) thus Equation (1.6) ~ Equation (1.7) as ~ is an equivalence relation.

Hence the theorem is proved.

Example 1.29: *Find all the composition series of $G = < a >$, a cyclic group of order 6 and show they are equivalent.*

Solution: $G = \{e, a, a^2 a^3, a^4, a^5\}$. Since $o(G) = 6$ has four divisors 1, 2, 3, 6, G will have four subgroups, namely $\{e\}$, G and $< a^2 > = \{e, a^2, a^4\}, < a^3 > = \{e, a^3\}$

Composition series of G will be

$$\{e\} \subset < a^3 > \subset G$$
$$\{e\} \subset < a^2 > \subset G$$

Notice $o\left(\frac{G}{< a^3 >}\right) = \frac{6}{2} = 3,\ o\left(\frac{< a^3 >}{\{e\}}\right) = o(< a^3 >) = 2$ which are primes and so the factors are simple groups.

Again, $\quad \frac{G}{< a^3 >} \cong \mathbf{Z}_3, \qquad \frac{< a^3 >}{\{e\}} \cong < a^3 > \cong \mathbf{Z}_2$

$$\frac{G}{< a^2 >} \cong \mathbf{Z}_2, \qquad \frac{< a^2 >}{\{e\}} \cong < a^2 > \cong \mathbf{Z}_3$$

$$\Rightarrow \frac{G}{< a^3 >} \cong \frac{< a^2 >}{\{e\}}, \quad \frac{< a^3 >}{\{e\}} \cong \frac{G}{< a^2 >}$$

Hence the two composition series are equivalent.

NOTES

Example 1.30: *Find all the composition series of* $\mathbf{Z}_{30}$ *and show they are equivalent.*

Solution: $\mathbf{Z}_{30} = \{0, 1, 2, ..., 29\}$ addition modulo 30. Besides $\{0\}$ and $\mathbf{Z}_{30}$, the other subgroups of $\mathbf{Z}_{30}$ are

$< 2 > = \{0, 2, 4, 6, ... 28\}$

$< 3 > = \{0, 3, 6, ..., 27\}$

and $< 5 >, < 6 >, < 10 >, < 15 >$

Composition series will be

$\{0\} \subset < 15 > \subset < 5 > \subset G \qquad \{0\} \subset < 15 > \subset < 3 > \subset G$

$\{0\} \subset < 10 > \subset < 5 > \subset G \qquad \{0\} \subset < 10 > \subset < 2 > \subset G$

$\{0\} \subset < 6 > \subset < 3 > \subset G \qquad \{0\} \subset < 6 > \subset < 2 > \subset G$

Here each $\frac{G_{i+1}}{G_i}$, factor group is simple.

For instance, $o\left(\frac{< 5 >}{< 15 >}\right) = \frac{o(< 5 >)}{o(< 15 >)} = \frac{6}{2} = 3$, a prime and so $\frac{< 5 >}{< 15 >}$ is simple.

Equivalence of any two composition series can be shown as in the previous example.

Theorem 1.37: *An abelian group G has a composition series iff G is finite.*

Proof: If G is finite, we have already shown that (Theorem 1.35) G has a composition series.

Conversely, let G be an abelian group and suppose it has a composition series

$$\{e\} = G_0 \subset G_1 \subset G_2 \subset ..... \subset G_k = G$$

then since $\frac{G_i}{G_{i-1}}$ is an abelian simple group $\forall\ i = 1, 2, ..., k$

it will be a group of prime order, say, p_i

Thus, $o\left(\frac{G_i}{G_{i-1}}\right) = p_i$

and by above problem then $o(G) = p_1 p_2 \ldots p_k$

Hence G is a finite group.

Corollary: An infinite abelian group has no composition series.

1.3 SOLVABLE GROUPS

NOTES

Definition: A group G is said to be *solvable* (or soluble) if $\exists$ a chain of subgroups

$$\{e\} = H_0 \subseteq H_1 \subseteq H_2 \subseteq \ldots\ldots\ldots \subseteq H_n = G \qquad \ldots(1.13)$$

such that, each H_i is a normal subgroup of H_{i+1} and $\frac{H_{i+1}}{H_i}$ is abelian

$$\forall\ i = 0, 1, 2, \ldots, n-1.$$

Also then, the series Equation (1.13) is referred to as *solvable series* of G.

Thus G is solvable if it has a normal series $(H_0, H_1, \ldots, H_n)$ such that, its factor groups are abelian.

Case 28: Any abelian group G is solvable. Since $\{e\} = G_0 \subset G_1 = G$ is a normal series for G where, $\frac{G}{\{e\}} \cong G$ is abelian.

Case 29: Every cyclic group is solvable.

Case 30: S_3 and S_4 are solvable. Since $\{I\} \subseteq A_3 \subseteq S_3$ is a normal series for S_3 where its factor groups $\frac{S_3}{A_3}$ and $\frac{A_3}{\{I\}}$ are abelian as these are of prime order.

So S_3 is an example of a non abelian group that is solvable.

$\{I\} \subseteq K_4 \subseteq A_4 \subseteq S_4$ will serve as the required normal series for S_4. Notice that $\frac{K_4}{\{I\}} \cong K_4 \Rightarrow o\left(\frac{K_4}{\{I\}}\right) = o(K_4) = 4$ and we know a group of order 4 is abelian.

Note: Any non abelian simple group is not solvable. If G is simple, it has no proper normal subgroup except $\{e\}$. So $\{e\} \subset G$ is the only normal series of G and as $\frac{G}{\{e\}} \cong G$, $\frac{G}{\{e\}}$ is not abelian as G is non abelian. Hence G is not solvable.

We have defined commutator subgroup G' of a group G.

Now let G' be commutator subgroup of a group G.

And let $(G')' = G'' = G^{(2)}$ be commutator subgroup of G' and $G^{(3)}$ be commutator subgroup of $G^{(2)}$ and so on then $G^{(n)}$ is called the nth commutator subgroup of G. We use this to provide us with an equivalent definition of a solvable group.

Theorem 1.38: *A group G is solvable iff $G^{(n)} = \{e\}$ for some positive integer n.*

Proof: Let G be solvable. Then there exists a normal series

$$\{e\} = G_0 \subseteq G_1 \subseteq G_2 \subseteq \ldots\ldots\ldots \subseteq G_n = G$$

such that, $\frac{G_{i+1}}{G_i}$ is abelian $\quad \forall\ i = 0, 1, 2, ..., n-1$

Since $\frac{G_n}{G_{n-1}} = \frac{G}{G_{n-1}}$ is abelian, we get $G' \subseteq G_{n-1}$

$\Rightarrow (G')' \subseteq G'_{n-1}$

i.e., $G^{(2)} \subseteq G'_{n-1}$

Again as $\frac{G_{n-1}}{G_{n-2}}$ is abelian, we get $G'_{n-1} \subseteq G_{n-2} \Rightarrow G^{(2)} \subseteq G_{n-2}$

Continuing like this, we will get $G^{(n)} \subseteq G_0 = \{e\}$

which gives $G^{(n)} = \{e\}$.

Conversely, let $G^{(n)} = \{e\}$. Consider the series

$\{e\} = G^{(n)} \subseteq G^{(n-1)} \subseteq G^{(n-2)} \subseteq ... \subseteq G^{(2)} \subseteq G^{(1)} \subseteq G^{(0)} = G$

which will be a normal series for G, where

$\frac{G^{(i)}}{G^{(i+1)}} = \frac{G^{(i)}}{(G^{(i)})'}$ is abelian $\forall\ i$

and, of course, $G^{(i)} \trianglelefteq G^{(i-1)} \quad \forall\ i$

$\Rightarrow G$ is solvable

That solvability is hereditary follows by.

Theorem 1.39: *A subgroup of a solvable group is solvable.*

Proof: Let H be any subgroup of a solvable group G.

Since, G is solvable, $G^{(n)} = \{e\}$ for some positive integer n.

Now, $H \subseteq G \Rightarrow H' \subseteq G' \Rightarrow (H')' \subseteq (G')'$, i.e., $H^{(2)} \subseteq G^{(2)}$

Continuing like this, we get $H^{(n)} \subseteq G^{(n)} = \{e\}$

$\Rightarrow H^{(n)} = \{e\}$

$\Rightarrow H$ is solvable.

Theorem 1.40: *Homomorphic image of a solvable group is solvable.*

Proof: Let $f: G \to H$ be an onto homomorphism, where G is solvable. Then $\exists$ a positive integer n such that, $G^{(n)} = \{e\}$

Let $a, b \in G$ be any elements, then $f(a), f(b) \in H$

$\Rightarrow f(a)\, f(b)\, (f(a))^{-1}\, (f(b))^{-1} \in H'$

Also, $a, b \in G \Rightarrow aba^{-1}b^{-1} \in G'$ and as

$f(aba^{-1}\, b^{-1}) = f(a)\, f(b)\, ((f(a))^{-1}\, (f(b))^{-1} \in H'$, we find

$f(G') \subseteq H'$ as $aba^{-1}b^{-1} \in G'$

NOTES

NOTES

Since f is onto, we find $f(G') = H'$

Again $f: G \to H$ onto means $f(G) = H$

and, therefore, $(f(G))' = H'$

i.e., $(f(G))' = f(G')$

So $H' = f(G')$

$\Rightarrow$ $(H')' = (f(G'))' = [f(G')'] = f(G'') = f(G^{(2)})$

or that $H^{(2)} = f(G^{(2)})$

Continuing like this we get

$H^{(n)} = f(G^{(n)}) = f(\{e\}) = \{e_1\}$ where e_1 is identity of H

i.e., H is solvable.

Theorem 1.41: *Quotient group of a solvable group is solvable.*

Proof: Follows from above as a quotient group is a homomorphic image of the group under the natural homomorphism.

Example 1.31: *Let H be a subgroup of a solvable group G. If*

$\{e\} = N_0 \subseteq N_1 \subseteq ... \subseteq N_{n-1} \subseteq N_n = G$ *be a solvable series of G then show that*

$\{e\} = N_0 \cap H \subseteq N_1 \cap H \subseteq ... \subseteq N_{n-1} \cap H \subseteq N_n \cap H = H$ *is a solvable series of H. Hence show that H is solvable.*

Solution: Let us put $H_i = N_i \cap H, \quad i = 0, 1, 2, ..., n.$

Then we show that

$$\{e\} = H_0 \subseteq H_1 \subseteq H_2 \subseteq ..... \subseteq H_{n-1} \subseteq H_n = H \qquad ...(1)$$

is a solvable series for H.

Since $N_i \trianglelefteq N_{i+1}$ we find $N_i \cap H \trianglelefteq N_{i+1} \cap H$

i.e., $H_i \trianglelefteq H_{i+1} \quad i = 0, 1, 2, ..., n-1$

We show now $\frac{H_i}{H_{i+1}}$ is abelian $\quad \forall\ i = 0, 1, 2, ..., n-1$

Define a map $\theta: H_{i+1} \to \frac{N_{i+1}}{N_i}$, such that,

$$\theta(x) = xN_i \quad (i = 0, 1, 2, ..., n-1)$$

$x \in H_{i+1} = N_{i+1} \cap H \Rightarrow x \in N_{i+1}, x \in H$

Thus $xN_i \in \frac{N_{i+1}}{N_i}$ and θ is well defined

Now $\theta(xy) = xyN_i = xN_i\, yN_i = \theta(x)\,\theta(y)$ shows θ is a homomorphism

Again, $x \in \text{Ker}\theta \quad \Leftrightarrow \theta(x) = N_i$

$\Leftrightarrow xN_i = N_i$

$\Leftrightarrow x \in N_i \Leftrightarrow x \in N_i \cap H$

Hence Ker $\theta = N_i \cap H = H_i$

By Fundamental theorem,

$$\theta(H_{i+1}) \cong \frac{H_{i+1}}{\text{Ker}\,\theta}$$

i.e. $$\frac{H_{i+1}}{H_i} \cong \theta(H_{i+1})$$

where $\theta(H_{i+1})$ is a subgroup of $\frac{N_{i+1}}{N_i}$, which is abelian and so $\theta(H_{i+1})$ is abelian and hence because of the above isomorphism $\frac{H_{i+1}}{H_i}$ is abelian.

Thus series Equation (1) is a solvable series of H.

Example 1.32: *Let G be a solvable group and suppose $H \neq \{e\}$ is a subgroup of G then show that $H' \neq H$.*

Solution: Suppose $H' = H$, then

$$H^{(2)} = (H')' = H' = H \neq \{e\}$$

If $H^{(n)} = H$, then $H^{(n+1)} = H' = H \neq \{e\}$

Thus by induction $H^{(r)} \neq \{e\} \quad \forall\ r \geq 1$

But G solvable $\Rightarrow H$ is solvable $\Rightarrow H^{(r)} = \{e\}$ for some $r \geq 1$, a contradiction. Hence $H' \neq H$.

Example 1.33: *Show that a simple group is solvable if and only if it is abelian.*

Solution: Let G be a simple group. Since $G' \trianglelefteq G$ we find either $G' = \{e\}$ or $G' = G$. If G is solvable then $G' \neq G$ so $G' = \{e\}$. Thus G is abelian.

Conversely, if G is abelian then $G' = \{e\}$ and so G is solvable.

Example 1.34: *Show that S_n $(n \geq 5)$ is not solvable.*

Solution: If S_n is solvable then A_n is solvable. But A_n $(n \geq 5)$ is simple. Thus by above problem A_n is abelian which is not true. [Notice $(123)(234) \neq (234)(123)$].

Hence S_n is not solvable for $n \geq 5$.

Theorem 1.42: *Let N be a normal subgroup of G such that, N and $\frac{G}{N}$ are solvable then G is solvable.*

Proof: Let $\{e\} = N_0 \subseteq N_1 \subseteq \ldots \subseteq N_k = N$ ...(1.14)

and $$\{N\} = \frac{G_0}{N} \subseteq \frac{G_1}{N} \subseteq \frac{G_2}{N} \subseteq \ldots\ldots \subseteq \frac{G_{n-1}}{N} \subseteq \frac{G_n}{N} = \frac{G}{N} \qquad \ldots(1.15)$$

NOTES

NOTES

be solvable series of N and $\frac{G}{N}$. By definition of solvable series then $\frac{G_i}{N} \trianglelefteq \frac{G_{i+1}}{N}$ and $\frac{G_{i+1}/N}{G_i/N}$ is abelian $\forall\ i = 0, 1, 2, ..., n-1$

which gives $G_i \trianglelefteq G_{i+1}\ \forall\ i$

Again by Third theorem of Isomorphism we have

$$\frac{G_{i+1}}{G_i} \cong \frac{G_{i+1}/N}{G_i/N}$$

Since $\frac{G_{i+1}/N}{G_i/N}$ is abelian, we find $\frac{G_{i+1}}{G_i}$ is abelian $\forall\ i$. Consider now the series

$$\{e\} = N_0 \subseteq N_1 \subseteq ... \subseteq N_k = N = G_0 \subseteq G_1 \subseteq ... \subseteq G_n = G$$

then it satisfies all conditions in the definition of a solvable series and hence it is required solvable Equation of G showing thereby that G is solvable.

When we consider the series Equation (1.15), it is clear that $G_0, G_1, ...,$ are all subgroups of G containing H.

Note: We thus conclude that a group G with a normal subgroup N is solvable if both N and G/N are solvable.

Example 1.35: *Show that a finite p-group is solvable, where p is prime.*

Solution: Let G be the given finite p-group, then $o(G) = p^n$ for some $n \geq 0$.

If $n = 1$, then G is a group of prime order and thus it is abelian and so G is solvable.

Suppose now $n > 1$. We use induction on n. Suppose that the result holds for all groups with order p^m where $m < n$, then $o(Z(G)) > 1$.

Let $o(Z(G)) = p^t$, $t \geq 1$ (Notice $o(Z(G)) \mid o(G) = p^n$)

Thus, $$o\left(\frac{G}{Z(G)}\right) = \frac{p^n}{p^t} = p^{n-t} = p^s \text{ where } s < n$$

Since result holds for groups with order p^m where $m < n$ we find $\frac{G}{Z(G)}$ is solvable.

Also $Z(G)$ is solvable as it is abelian.

Hence by above theorem G is solvable.

Example 1.36: *Show that a solvable group contains at least one normal abelian subgroup H.*

NOTES

Solution: Let G be a solvable group. If G is abelian then $H = G$ is the required subgroup.

Let now G be non abelian. Since G is solvable $G^{(n)} = \{e\}$ for some positive integer n.

Now $G' \neq \{e\}$ as if $G' = \{e\}$ then G is abelian, which is not true. Hence $G^{(n)} = \{e\}$, $n \neq 1$

Let $H = G^{(n-1)}$ then H is a subgroup of G.

and as $H' = G^{(n)} = \{e\}$, we find H is abelian and also as $G^{(n-1)}$ is normal subgroup of G, we find H is the required subgroup.

Example 1.37: *Show that a group of order pq is solvable, where p, q are primes.*

Solution: Let $o(G) = pq$. If $p = q$ then $o(G) = p^2$ and thus G is an abelian group. Hence G is solvable. Let now $p > q$. Then number of Sylow p-subgroups of G is $1 + kp$ where $(1 + kp) \mid q$, i.e., $1 + kp = 1$ or q.

If $1 + kp = q$ then $kp = q - 1 \Rightarrow p \mid (q - 1)$ which is not true, as $p > q$.

Hence $1 + kp = 1$ and there exists a unique normal Sylow p-subgroup, say H, of order p.

Since p is prime, H will be cyclic and so abelian and hence solvable.

Again $o\left(\frac{G}{H}\right) = q \quad \Rightarrow \quad \frac{G}{H}$ is abelian $\Rightarrow \frac{G}{H}$ is solvable $\Rightarrow G$ is solvable.

Example 1.38: *Show that the following two statements are equivalent*:

(*a*) *Every group of order $p^m q^n$, where p, q are primes, is solvable.*

(*b*) *Simple groups of order $p^\alpha q^\beta$ are cyclic groups of order p or q.*

Solution: (*a*) $\Rightarrow$ (*b*)

Let G be a simple group of order $p^\alpha q^\beta$. Since G' is normal in G, we find either $G' = \{e\}$ or $G' = G$.

Since G is solvable, by (*a*) $G' = \{e\}$ and so G is abelian.

Let H be a Sylow p-subgroup of G. Then H will be normal as G is abelian and $o(H) = p^\alpha$

Again, G simple means either $H = G$ or $H = \{e\}$

If $H = G$, there $\alpha = 1$, $\beta = 0$ and so G is cyclic of order p

If $H = \{e\}$ then if K is sylow q-subgroup of G, it will be normal and as before, either, $K = G$ or $K = \{e\}$

If $K = G$, then $\alpha = 1$, $\beta = 0$ and so G is cyclic of order q.

If $K = \{e\}$, we get the case where $\alpha = 0$, $\beta = 0$ forcing $G = \{e\}$ which is not true as G is simple. Hence the result follows.

NOTES

We now show that $(b) \Rightarrow (a)$.

Let G be a group of order $p^m q^n$.

Consider a composition series of G (which exists as G is finite) then every composition factor of this series will be a simple group of order $p^\alpha q^\beta$ for some α, β. By (b), each factor would, therefore, be cyclic and so abelian. Hence G is solvable.

Note: There is a famous theorem of Burnside in which it is proved that every group of order $p^m q^n$ where p, q are primes, is solvable.

1.3.1 Nilpotent Groups

Definition I: A group G is called *nilpotent* if it has a normal series

$$\{e\} = G_0 \subseteq G_1 \subseteq G_2 \subseteq \ldots\ldots \subseteq G_n = G$$

such that, $\dfrac{G_i}{G_{i-1}} \subseteq Z\left(\dfrac{G}{G_{i-1}}\right) \quad \forall\ i = 1, 2, \ldots, n$

Definition II: We first define what we mean by nth centre of a group. Let G be a group and $Z(G)$ be its centre. We call $Z(G)$ the first centre of G and put $Z(G) = Z_1(G)$. Consider now the group $\dfrac{G}{Z(G)}$, then centre $Z\left(\dfrac{G}{Z(G)}\right)$ of $\dfrac{G}{Z(G)}$ is a normal subgroup of $\dfrac{G}{Z(G)}$

So, $\quad Z\left(\dfrac{G}{Z_1(G)}\right) \trianglelefteq \dfrac{G}{Z_1(G)}$

Since any normal subgroup of $\dfrac{G}{K}$ is of the form $\dfrac{H}{K}$ for a unique normal subgroup H of G, we find any normal subgroup of $\dfrac{G}{Z_1(G)}$ is of the type $\dfrac{H}{Z_1(G)}$ where $H \trianglelefteq G$

We write $H = Z_2(G)$ (Called second centre of G)

Then $Z_2(G) \trianglelefteq G$ such that, $Z\left(\dfrac{G}{Z_1(G)}\right) = \dfrac{Z_2(G)}{Z_1(G)}$

Continuing like this we get $Z_n(G) \trianglelefteq G$, (called nth centre)

such that, $\quad \dfrac{Z_n(G)}{Z_{n-1}(G)} = Z\left(\dfrac{G}{Z_{n-1}(G)}\right) \quad n > 1$

Let us write $Z_0(G) = \{e\}$, and thus

$$\frac{Z_n(G)}{Z_{n-1}(G)} = Z\left(\frac{G}{Z_{n-1}(G)}\right) \quad \forall\ n = 1, 2, \ldots$$

Also then $Z_0(G) \subseteq Z_1(G) \subseteq Z_2(G) \subseteq \ldots$ are normal subgroups of G. This is called the *upper central series* or *ascending central series* of G.

NOTES

We say a group G is *nilpotent* if $Z_m(G) = G$ for some m. Also in that case the smallest m such that, $Z_m(G) = G$ is called the class of nilpotency of G.

We first show the equivalence of the two definitions.

Definition I ⇒ Definition II

Let G be nilpotent according to Definition I. Then G has a normal series

$$\{e\} = G_0 \subseteq G_1 \subseteq G_2 \subseteq \ldots \subseteq G_n = G$$

such that, $\frac{G_i}{G_{i-1}} \subseteq Z\left(\frac{G}{G_{i-1}}\right) \quad \forall\ i = 1, 2, \ldots, n$

Let $i = 1$, then

$$\frac{G_1}{G_0} \subseteq Z\left(\frac{G}{G_0}\right)$$

If $x \in G_1$ be any element, then

$$G_0 x \in \frac{G_1}{G_0} \Rightarrow G_0 x \in Z\left(\frac{G}{G_0}\right)$$

$$\Rightarrow G_0 x \,.\, G_0 y = G_0 y G_0 x \quad \forall\ G_0 y \in \frac{G}{G_0}$$

$$\Rightarrow G_0 xy = G_0 yx$$
$$\Rightarrow xy\, x^{-1} y^{-1} \in G_0 = \{e\}$$
$$\Rightarrow xy = yx \quad \forall\ y \in G$$
$$\Rightarrow x \in Z(G) = Z_1(G)$$

Hence, $G_1 \subseteq Z_1(G)$

Let, $i = 2$, then

$$\frac{G_2}{G_1} \subseteq Z\left(\frac{G}{G_1}\right)$$

If $x \in G_2$ be any element then proceeding as above we get $xy\, x^{-1}y^{-1} \in G_1$

and as $G_1 \subseteq Z_1(G)$

$$xy\, x^{-1}y^{-1} \in Z_1(G) \quad \forall\ y \in G$$

NOTES

$$\Rightarrow \quad Z_1(G)\ xy = Z_1(G)yx \quad \Rightarrow Z_1(G)x\ Z_1(G)y = Z_1(G)y\ Z_1(G)x$$

$$\Rightarrow Z_1(G)x \in Z\left(\frac{G}{Z_1(G)}\right) = \frac{Z_2(G)}{Z_1(G)}$$

$$\Rightarrow x \in Z_2(G)$$

Hence, $G_2 \subseteq Z_2(G)$

Continuing like this, we get

$$G_i \subseteq Z_i(G) \quad \forall\ i = 1, 2, ..., n$$

Hence, $G = G_n \subseteq Z_n(G)$

or that G is nilpotent according to Definition II.

Definition II $\Rightarrow$ Definition I

Suppose G is nilpotent of class n then $Z_n(G) = G$. Consider the series

$$\{e\} = Z_0(G) \subseteq Z_1(G) \subseteq Z_2(G) \subseteq ..... \subseteq Z_n(G) = G$$

which is a normal series and $\frac{Z_i(G)}{Z_{i-1}(G)} = Z\left(\frac{G}{Z_{i-1}(G)}\right)$

i.e., G is nilpotent according to Definition I.

Case 31: An abelian group is nilpotent. Since G abelian

$$\Rightarrow\ G = Z(G), \text{ i.e., } Z_1(G) = G.$$

Also then all cyclic groups will be nilpotent.

However, a nilpotent group need not be abelian and thus cyclic. Consider G, the quaternion group. Then

$$G_0 = \{1\} \subseteq G_1 = \{1, -1\} \subseteq G_2 = \{1, -1, i, -i\} \subseteq G$$

and $\quad o\left(\frac{G}{G_2}\right) = 2,\ o\left(\frac{G}{G_1}\right) = 4 \Rightarrow \frac{G}{G_1}, \frac{G}{G_2}$ are abelian

$$\Rightarrow \quad Z\left(\frac{G}{G_1}\right) = \frac{G}{G_1},\ Z\left(\frac{G}{G_2}\right) = \frac{G}{G_2}. \text{ Also } Z\left(\frac{G}{G_0}\right) = \{G_{0(1)}, G_{0(-1)}\} \text{ is}$$

abelian

Thus $\quad Z\left(\frac{G}{G_0}\right) = \frac{G}{G_0}$ and so G is nilpotent but not abelian.

Case 32: A finite p-group is nilpotent.

Theorem 1.43: *Every nilpotent group is solvable. Converse is not true.*

Proof: Let G be a nilpotent group, then G has a normal series

$$\{e\} = G_0 \subseteq G_1 \subseteq G_2 \subseteq ... \subseteq G_n = G$$

where $\frac{G_i}{G_{i-1}} \subseteq Z\left(\frac{G}{G_{i-1}}\right)$. $\forall\ i = 1, 2, ..., n$

Which implies that $\frac{G_i}{G_{i-1}}$ is abelian $\forall\ i$

Hence G is solvable.

S_3 is solvable but not nilpotent. Notice that $Z(S_3) = \{I\}$ and so $Z_m(G) = G$ holds for no m.

(In fact S_n is not nilpotent, for $n \geq 3$).

Theorem 1.44: *Any subgroup of a nilpotent group is nilpotent.*

Proof: Let H be a subgroup of a nilpotent group G. Since G is nilpotent, there exists a normal series

$$\{e\} = G_0 \subseteq G_1 \subseteq G_2 \subseteq ..... \subseteq G_n = G$$

such that, $\frac{G_i}{G_{i-1}} \subseteq Z\left(\frac{G}{G_{i-1}}\right)$, $i = 1, 2, ..., n$

Consider the series

$$\{e\} = G_0 \cap H \subseteq G_1 \cap H \subseteq G_2 \cap H \subseteq ... \subseteq G_n \cap H = G \cap H = H$$

It is easy to see that $G_{i-1} \cap H \trianglelefteq G_i \cap H \quad \forall\ i$. We show

$\frac{G_i \cap H}{G_{i-1} \cap H} \subseteq Z\left(\frac{G \cap H}{G_{i-1} \cap H}\right), \forall\ i = 1, 2, ..., n$ which would establish that H is nilpotent.

Let $(G_{i-1} \cap H)x \in \frac{G_i \cap H}{G_{i-1} \cap H}$ be any element

then $x \in G_i \cap H \Rightarrow x \in G_i$ and $x \in H$.

Now, $(G_{i-1} \cap H)x \in Z\left(\frac{G \cap H}{G_{i-1} \cap H}\right)$

if $(G_{i-1} \cap H)x$ commutes with all elements of $\frac{G \cap H}{G_{i-1} \cap H}$

i.e., $(G_{i-1} \cap H)x\ (G_{i-1} \cap H)y = (G_{i-1} \cap H)y\ (G_{i-1} \cap H)x \quad \forall\ y \in G \cap H$

i.e., $(G_{i-1} \cap H)xy = (G_{i-1} \cap H)yx$

i.e., $xy\ x^{-1}y^{-1} \in G_{i-1} \cap H \quad \forall\ y \in G \cap H$

i.e., $xy\ x^{-1}y^{-1} \in G_{i-1}$ and $xy\ x^{-1}y^{-1} \in H \quad \forall\ y \in G \cap H$

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

NOTES

NOTES

Again, since $\frac{G_i}{G_{i-1}} \subseteq Z\left(\frac{G}{G_{i-1}}\right)$ and $x \in G_i$, we find that

$$G_{i-1}\, x \in \frac{G_i}{G_{i-1}} \Rightarrow G_{i-1}\, x \in Z\left(\frac{G}{G_{i-1}}\right)$$

$$\Leftrightarrow G_{i-1}\, x\, G_{i-1}\, y = G_{i-1}\, y\, G_{i-1}\, x \quad \forall\ y \in G$$

$$\Leftrightarrow G_{i-1} xy = G_{i-1} yx$$

$$\Leftrightarrow xy\, x^{-1} y^{-1} \in G_{i-1} \quad \forall\ y \in G$$

and hence over assertion is proved.

Theorem 1.45: *Homomorphic image of a nilpotent group is nilpotent.*

Proof: Let $\theta : G \to H$ be an onto homomorphism and suppose G is nilpotent. Then there exists a normal series

$$\{e\} = G_0 \subseteq G_1 \subseteq G_2 \subseteq \ldots\ldots \subseteq G_n = G$$

such that, $\frac{G_i}{G_{i-1}} \subseteq Z\left(\frac{G}{G_{i-1}}\right) \quad \forall\ i = 1, 2, \ldots, n$

We claim

$\theta(e) = \theta(G_0) \subseteq \theta(G_1) \subseteq \theta(G_2) \subseteq \ldots \subseteq \theta(G_n) = \theta(G) = H$ is the required normal series for H where $\frac{\theta(G_i)}{\theta(G_{i-1})} \subseteq Z\left(\frac{\theta(G)}{\theta(G_{i-1})}\right)$

It is easy to see that $\theta(G_{i-1}) \trianglelefteq \theta(G_i) \quad \forall\ i$ and we leave it for the reader to try and prove it.

Let $\theta(G_i) = H_i \quad i = 1, 2, \ldots, n$

we show $\frac{H_i}{H_{i-1}} \subseteq Z\left(\frac{H}{H_{i-1}}\right)$

Let $H_{i-1}\, x \in \frac{H_i}{H_{i-1}}$ be any element,

we have to show that $H_{i-1}\, x \in Z\left(\frac{H}{H_{i-1}}\right)$

i.e., $(H_{i-1}\, x)\ (H_{i-1}\, y) = (H_{i-1}\, y)\ (H_{i-1}\, x) \quad \forall\ H_{i-1}\, y \in \frac{H}{H_{i-1}}$

i.e., $H_{i-1}\, xy = H_{i-1}\, yx$

i.e., $xy\, x^{-1} y^{-1} \in H_{i-1} \qquad \forall\ y \in H$

Now, $x \in H_i \Rightarrow x \in \theta(G_i) \Rightarrow \exists\ a \in G_i$ such that, $\theta(a) = x$

$y \in H \Rightarrow y = \theta(G) \Rightarrow \exists\ b \in G,$ such that, $\theta(b) = y$

NOTES

Thus, $xy\, x^{-1}y^{-1} = \theta(a)\ \theta(b)\ (\theta(a))^{-1}\ (\theta(b))^{-1} = \theta(ab\ a^{-1}b^{-1}) \in \theta(G_{i-1})$

Since $a \in G_i$, $G_{i-1}\ a \in \dfrac{G_i}{G_{i-1}} \subseteq Z\left(\dfrac{G}{G_{i-1}}\right)$

and so $G_{i-1}\ a\ .\ G_{i-1}\ b = G_{i-1}\ b\ G_{i-1}\ a$

i.e., $G_{i-1}\ ab = G_{i-1}\ ba$

i.e., $ab\ a^{-1}b^{-1} \in G_{i-1}$

i.e., $\theta(ab\ a^{-1}b^{-1}) \in \theta(G_{i-1}) = H.$

Hence the result follows.

Theorem 1.46: *Any quotient group of a nilpotent group is nilpotent.*

Proof: Follows from above theorem as any quotient group of a group is its homomorphic image.

Converse is, however, not true as $\dfrac{S_3}{A_3}$ is abelian and so nilpotent, but S_3 is not nilpotent.

Example 1.39: *If H and K are nilpotent groups then show that $H \times K$ is also nilpotent.*

Solution: Let H and K be nilpotent. Then $\exists$ normal series

$\{e_1\} = H_0 \subseteq H_1 \subseteq H_2 \subseteq \ldots\ldots \subseteq H_n = H$ such that, $\dfrac{H_i}{H_{i-1}} \subseteq Z\left(\dfrac{H}{H_{i-1}}\right)$

$i = 1, 2, \ldots, n$

$\{e_2\} = K_0 \subseteq K_1 \subseteq K_2 \subseteq \ldots\ldots \subseteq K_n = K$ such that, $\dfrac{K_i}{K_{i-1}} \subseteq Z\left(\dfrac{K}{K_{i-1}}\right)$

We can repeat terms in the series with lesser terms.

Consider the series

$\{e_1\} \times \{e_2\} = H_0 \times K_0 \subseteq H_1 \times K_1 \subseteq H_2 \times K_2 \subseteq \ldots\ldots \subseteq H_n \times K_n = H \times K$

Then one can check that this is a normal series in which

$$\frac{H_i \times K_i}{H_{i-1} \times K_{i-1}} \subseteq Z\left(\frac{H \times K}{H_{i-1} \times K_{i-1}}\right)$$

Let $(H_{i-1} \times K_{i-1})\ (h, k) \in \dfrac{H_i \times K_i}{H_{i-1} \times K_{i-1}}$ be any element

then $(H_{i-1} \times K_{i-1})\ (h, k)$ will belong to $Z\left(\dfrac{H \times K}{H_{i-1} \times K_{i-1}}\right)$

NOTES

if $(H_{i-1} \times K_{i-1})\ (h,\ k).(H_{i-1} \times K_{i-1})\ (x,\ y) = (H_{i-1} \times K_{i-1})\ (x\ .\ y).(H_{i-1} \times K_{i-1})\ (h,\ k)$

i.e., if $(h,\ k)\ (x,\ y)\ (h,\ k)^{-1}\ (x,\ y^{-1}) \in H_{i-1} \times K_{i-1}$

i.e., if $(hx\ h^{-1}x^{-1},\ hy\ k^{-1}y^{-1}) \in H_{i-1} \times K_{i-1}$

i.e., if $h\ x\ h^{-1}x^{-1} \in H_{i-1}$

$ky\ k^{-1}y^{-1} \in K_{i-1}$

which is true.

We leave the first part (that $H_i \times K_i \trianglelefteq H_{i+1} \times K_{i+1}$) for the reader to try as an exercise.

Example 1.40: *If H is a proper subgroup of a nilpotent group G then show that H is a proper subgroup of N(H).*

Solution: Since G is nilpotent, it has upper central series

$\{e\} = Z_0(G) \subseteq Z_1(G) \subseteq Z_2(G) \subseteq \ldots\ldots \subseteq Z_n(G) = G$

Now $H \subsetneqq G$, let i be the largest integer such that, $Z_i(G) \subseteq H$

Then we get

$$Z_i(G) \subseteq H \subseteq Z_{i+1}\ (G) \subseteq \ldots$$

Again since $\dfrac{Z_{i+1}(G)}{Z_i(G)} = Z\left(\dfrac{G}{Z_i(G)}\right)$

$\dfrac{Z_{i+1}(G)}{Z_i(G)}$ is abelian.

Let $g \in Z_{i+1}(G)$ and $h \in H$ be any elements, then

$h \in H \subseteq Z_{i+1}(G)$ and so $Z_i(G)g,\ Z_i(G)h \in \dfrac{Z_{i+1}(G)}{Z_i(G)}$ and thus

$$Z_i(G)g\ Z_i(G)h = Z_i(G)h\ Z_i(G)g$$

$$\Rightarrow\ Z_i(G)gh = Z_i(G)hg$$

$$\Rightarrow\ gh\ g^{-1}h^{-1} \in Z_i(G) \subseteq H$$

$$\Rightarrow\ gh\ g^{-1} \in H \quad \forall\ g \in Z_{i+1}(G),\ h \in H$$

$$\Rightarrow\ gH\ g^{-1} \subseteq H \quad \forall\ g \in Z_{i+1}(G)$$

i.e., $gH\ g^{-1} = H \qquad \forall\ g \in Z_{i+1}(G)$

$\Rightarrow$ any $g \in Z_{i+1}(G)$ is such that $g \in N(H)$

or that $Z_{i+1}(G) \subseteq N(H)$

But $H \subsetneqq Z_{i+1}(G)$ and hence H is a proper subgroup of $N(H)$.

Check Your Progress

1. Give the postulates for a group.
2. When set G is called finite or infinite?
3. Define subgroup.
4. When a group G is called cyclic?
5. When maximal normal subgroup of G no exists normal subgroup of K?
6. When can you say that a group of order pq is solvable?
7. What is a nilpotent group?

NOTES

1.4 ANSWERS TO 'CHECK YOUR PROGRESS'

1. A group satisfies the following postulates.
   (*i*) *Associativity:* $a * (b * c) = (a * b) * c$, for all $a, b, c \in G$
   (*ii*) *Existence of Identity:* $\exists$ an element $e \in G$, such that,
   $a * e = e * a = a$ for all $a \in G$
   (e is then called identity)
   (*iii*) Existence of Inverse : For every $a \in G, \exists\, a' \in G$ (depending upon a) such that,
   $a * a' = a' * a = e$
   (a' is then called inverse of a)
2. If the set G is finite (i.e., has finite number of elements) it is called a finite group otherwise, it is called an infinite group.
3. A non empty subset H of a group G is said to be a subgroup of G if H forms a group under the binary composition of G.
4. A group G is called a cyclic group if $\exists$ an element a G such that every element of G can be expressed as a power of a.
5. A normal subgroup H of a group G is called a *maximal normal* subgroup of G if $H \neq G$ and there exists no normal subgroup K of G such that, $H \subset K \subset G$.
6. A group of order pq is solvable if p and q are primes.
7. A group G is called nilpotent if it has a normal series
   $\{e\} = G_0 \subseteq G_1 \subseteq ..... \subseteq G_n = G$ such that,
   $$\frac{G_i}{G_{i-1}} \subseteq Z\left(\frac{G}{G_{i-1}}\right) \forall\, i = 1,2,....,n$$

NOTES

1.5 SUMMARY

- A non-empty set G, together with a binary compsition * (star) is said to form a group, if it satisfies the postulates of associativity, existence of identity and existence of inverse.
- Let G be a group. A sequence of subgroups

  $\{e\} = G_0 \subseteq G_1 \subseteq G_2 \subseteq ..... \subseteq G_n = G$

  is called a normal series of G if G_i is a normal subgroup of G_{i+1},
- A group G is called a cyclic group if $\exists$ an element $a \in G$, such that every element of G can be expressed as a power of a. In that case a is called *generator* of G. We express this fact by writing $G = < a >$ or $G = (a)$.
- Let G a group. A sequence of subgroups

  $\{e\} = G_0 \subset G_1 \subset G_2 \subset ..... \subset G_n = G$

  of G is called a composition series of G if

  (*i*) each G_i is normal subgroup of G_{i+1} ($i = 0, 1, ... n - 1$),

  (*ii*) $G_i \neq G_{i+1}$ for any i and

  (*iii*) is a simple group $\forall i$.
- If G is a finite group, then order of any element of G divides order of G.
- A normal subgroup H of a group G is called a maximal normal subgroup of G if $H \neq G$ and there exists no normal subgroup K of G such that, $H \subset K \subset G$.
- A group G is called nilpotent if it has a normal series

  $\{e\} = G_0 \subseteq G_1 \subseteq G_2 \subseteq ..... \subseteq G_n = G$

  such that, $\forall i = 1, 2, ..., n$
- A group G is called nilpotent if it has a normal series

  $\{e\} = G_0 \subseteq G_1 \subseteq G_2 \subseteq ..... \subseteq G_n = G$

1.6 KEY TERMS

- **Finite group**: If a group has finite number of elements then it is called finite group.
- **Subgroup**: A non-empty subset of a group is said to be a subgroup if it forms a group under the binary composition of the group.
- **Cyclic group:** A group G is called a cyclic group if $\exists$ an element $a \in G$, such that every element of G can be expressed as a power of a. In that case a is called generator of G. We express this fact by writing $G = < a >$ or $G = (a)$.

- **Nilpotent groups:** A group G is called nilpotent if it has a normal series $\{e\} = G_0 \subseteq G_1 \subseteq G_2 \subseteq ..... \subseteq G_n = G$

1.7 SELF-ASSESSMENT QUESTIONS AND EXERCISES

NOTES

Short-Answer Questions

1. Define a group.
2. Specify the term cycle group.
3. What is the difference between normal and subnormal series?
4. Define the term composition series?
5. State the Jordan-Holder theorem.
6. Define the term solvable group.
7. What can you say about the nilpotency is an abelian group?

Long-Answer Questions

1. Check whether the following systems form a group (a semi-group) or not
   (*a*) G = Set of rational numbers under composition $*$ defined by $a * b = \frac{ab}{2}$, $a, b \in G$
   (*b*) $G = \{\pm 1, \pm i\}$, where $i = \sqrt{-1}$ under multiplication.
   (*c*) $G = \{1, w, w^2\}$, where w is cube root of unity under multiplication.
   (*d*) Set of all 2×2 matrices over integers under matrix multiplication.
   (*e*) Set of all matrices of the form $\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$, $\theta \in \mathbf{R}$, under matrix multiplication.
   (*f*) Q = Set of all rational numbers under $*$ where $a * b = a + b - ab$.
   (*g*) $G = \{2, 4, 6, 8\}$ under multiplication modulo 10.
   (*h*) $G = \{1, 2, 3\}$ under multiplication modulo 4.
   (*i*) $G = \{(a, b) \mid a, b \in Z\}$ under $*$ defined by
   $$(a, b) * (c, d) = (ac + bd, ad + bc).$$
2. Let G be the set $\{\pm e, \pm a, \pm b, \pm c\}$ where
   $$e = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, a = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, b = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, c = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$
   Show that G forms a group under matrix multiplication.

NOTES

3. Show that a group G is abelian iff $(ab)^2 = a^2b^2$.
4. In a group G, an element a is called *Idempotent* if $a^2 = a$. Show that a is idempotent iff $a = e$.
5. Show that if G be a group of even order then it has at least one element $(\neq e)$ which is its own inverse.
6. (a) Show that the power set of a finite set X is a finite semi group under intersection, has identity and all elements are idempotent.

   (b) Show that a finite semi-group G with identity is a group iff G contains only one idempotent.
7. Show that a monoid is a group if and only if cancellation laws hold in it.
8. Let G be the Quaternion group. Find centre of G. Find also the normalizer of i in G.
9. If H is a subgroup of G, show that

   $g^{-1}Hg = \{g^{-1}hg \mid h \in H\}$ is a subgroup of G.

   Show further that $g^{-1}Hg$ is abelian if H is abelian.
10. Let G be the group of all 3×3 invertible matrices over reals. Show that

$$H = \left\{ \begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix} a, b, c \in \mathbf{R} \right\} \text{ is a subgroup of } G.$$

11. If $N(H)$ be the normalizer of H in a group G then show that $Z(G) \subseteq N(H)$, where $H \leq G$.
12. If $o(G) = 6$ and $H \neq K$ are subgroups of G each of order 2 then show that HK cannot be a subgroup of G. Show also that G cannot have two subgroups of order 3.
13. If a finite group possesses an element of order 2, show that it possesses an odd number of such elements.
14. Show that every element in U_8 is its own inverse and hence U_8 is not cyclic and let G be a finite group. Let $a \in G$ be such that $o(a) = o(G)$. Show that G is cyclic, generated by a. Hence show that a group of order n is cyclic iff it has an element of order n.
15. Show that a subgroup $(\neq \{e\})$ of an infinite cyclic group is infinite.
16. If G is a cyclic group of order p, a prime then show that any non identity element of G is of order p.
17. Find all the subgroups of the quaternion group G and show that $\exists$ no two non-trivial subgroups H, K of G such that, $H \cap K$ is identity only.
18. Show that a finite cyclic group with three or more elements has even member of generators.
19. Write down all the 12 subgroups of Z_{60}. How many generators it has?

20. Let G be a finite group acting on a finite set S. For any $g \in G$, define $S^g = \{s \in S \mid g * s = s\}$. Prove (Burnside's formula)

$$o(G) \times \text{Number of orbits} = \sum_{g \in G} o(S^g)$$

21. Show that A_n is maximal normal in S_n and write all the maximal normal and maximal sbgroups of S_3.

22. Let G be a finite p-group of order p^n. Show that it has a normal series

$$\{e\} = G_0 \subseteq G_1 \subseteq ... \subseteq G_n = G$$

where $o(G_i) = p^i \quad i = 0, 1, 2, ..., n$

23. Show that a simple group is solvable iff it is cyclic.

24. If all proper subgroups of a non solvable group G are solvable, show that $G = G'$. (A group G such that $G = G'$ is called a perfect group).

25. Show that a finite p-group is nilpotent.

26. Suppose that in a non abelian simple group, $\{e\}$ is the only conjugate class whose order is prime power. Show that a group of order $p^m q^n$ (p, q primes) is a solvable group.

27. Show that every sylow subgroup of a nilpotent group G is normal in G.

NOTES

1.8 FURTHER READING

Herstein, I.N. 1975. *Topics in Algebra*, 3rd Edition. New Delhi: Wiley Eastern Ltd.

Khanna, V.K. and S.K. Bhambri. 2008. *A Course in Abstract Algebra*, 3rd Edition. New Delhi: Vikas Publishing Hous Pvt. Ltd.

Bhattacharya, P.B., S.K. Jain and S.R. Nagpaul. 1997. *Basic Abstract Algebra*, 2nd Edition. New Delhi: Cambridge University Press.

Artin, M.1991. *Algebra*. New Delhi: Prentice-Hall of India.

Lang, S. 1993. *Algebra*, 3rd Edition. New York: Addison-Wesley.

Datta, K.B. 2000. *Matrix and Linear Algebra*. New Delhi: Prentice-Hall of India.

UNIT 2 CANONICAL FORMS

NOTES

Structure

2.0 INTRODUCTION

In mathematics, a module is one of the fundamental algebraic structures used in abstract algebra. A module is an additive abelian group. In a simple module the submodules are the module itself and the module that consists of the element zero.

A canonical, normal, or standard form of a mathematical object is a standard way of presenting that object as a mathematical expression. Often, it is one which provides the simplest representation of an object and which allows it to be identified in a unique way. The distinction between 'Canonical' and 'Normal' forms varies from subfield to subfield. In most fields, a canonical form specifies a unique representation for every object, while a normal form simply specifies its form, without the requirement of uniqueness.

Linear transformation is a function between two vector spaces that preserves the operations of vector addition and scalar multiplication. Transformations satisfying these two conditions simultaneously are called similarity transformations. A nilpotent transformation is one with a power that is the zero map. A Jordan block is a matrix having zeros everywhere except along the diagonal and superdiagonal with each element of the diagonal consisting of a single number and each element of the superdiagonal consisting of a 1. A Jordan form consists of one or more Jordan blocks.

In this unit, you will study about the similarity of linear transformations, invariant subspaces, nilpotent transformation, the primary decomposition theorem and Jordan block's and Jordan forms.

NOTES

2.1 OBJECTIVES

After going through this unit, you will be able to:

- Know about the similarity of linear transformations
- Define invariant subspaces and reduction to triangular form
- Describe the nilpotent transformations, index of nilpotency and invariants of nilpotent transformations
- Analyse the primary decomposition theorem
- Explain about the Jordan blocks and Jordan forms
- Elaborate on the cyclic modules
- Understand the simple modules and schur's lemma
- State and prove fundamental structure theorem for modules

2.2 SIMILARITY OF LINEAR TRANSFORMATIONS

In mathematics and computer science, a '**Canonical, Normal**', or '**Standard Form'** of a mathematical object is a standard way of presenting that object as a mathematical expression. Often, it is one which provides the simplest representation of an object and which allows it to be identified in a unique way. The distinction between 'Canonical' and 'Normal' forms varies from subfield to subfield. In most fields, a canonical form specifies a unique representation for every object, while a normal form simply specifies its form, without the requirement of uniqueness. The canonical form of a positive integer in decimal representation is a finite sequence of digits that does not begin with zero. More generally, for a class of objects on which an equivalence relation is defined, a canonical form consists in the choice of a specific object in each class. For example:

- Jordan normal form is a canonical form for matrix similarity.
- The row echelon form is a canonical form, when one considers as equivalent a matrix and its left product by an invertible matrix.

In computer science, and more specifically in computer algebra, when representing mathematical objects in a computer, there are usually many different ways to represent the same object. A canonical form is a representation such that every object has a unique representation (with canonicalization being the process through which a representation is put into its canonical form). Thus, the equality of two objects can easily be tested by testing the equality of their canonical forms. Canonical forms frequently depend on arbitrary choices (like ordering the variables), which

NOTES

introduce difficulties for testing the equality of two objects resulting on independent computations. Therefore, in computer algebra, normal form is a weaker notion: A normal form is a representation such that zero is uniquely represented. This allows testing for equality by putting the difference of two objects in normal form. Canonical form can also mean a differential form that is defined in a natural (canonical) way.

Given a set S of objects with an equivalence relation R on S, a canonical form is given by designating some objects of S to be canonical form, such that every object under consideration is equivalent to exactly one object in canonical form. In other words, the canonical forms in S represent the equivalence classes, once and only once. To test whether two objects are equivalent, it then suffices to test equality on their canonical forms. A canonical form thus provides a classification theorem and more, in that it not only classifies every class, but also gives a distinguished (canonical) representative for each object in the class.

Formally, a canonicalization with respect to an equivalence relation R on a set S is a mapping $c: S \to S$ such that for all $s, s_1, s_2 \in S$:

1. $c(s) = c(c(s))$ (Idempotence),
2. $s_1 \, R \, s_2$ if and only if $c(s_1) = c(s_2)$ (Decisiveness), and
3. $s \, R \, c(s)$ (Representativeness).

In practical terms, it is often helpful to be able to recognize the canonical forms. There is also a practical, algorithmic question to consider: how to pass from a given object s in S to its canonical form s^*? Canonical forms are generally used to make operating with equivalence classes more effective. For example, in modular arithmetic, the canonical form for a residue class is usually taken as the least non-negative integer in it. Operations on classes are carried out by combining these representatives, and then reducing the result to its least non-negative residue. The uniqueness requirement is sometimes relaxed, allowing the forms to be unique up to some finer equivalence relation, such as allowing for reordering of terms (if there is no natural ordering on terms).

A canonical form may simply be a convention, or a deep theorem. For example, polynomials are conventionally written with the terms in descending powers: it is more usual to write $x^2 + x + 30$ than $x + 30 + x^2$, although the two forms define the same polynomial.

Definition: Let V and U be two vector spaces over the same field F, then a mapping $T: V \to U$ is called a homomorphism or a linear transformation if

$$T(x + y) = T(x) + T(y) \quad \text{for all } x, y \in V$$

$$T(\alpha x) = \alpha T(x) \quad \alpha \in F$$

One can combine the two conditions to get a single condition

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y) \quad x, y \in V; \alpha, \beta \in F$$

It is easy to see that both are equivalent. If a homomorphism happens to be one-one onto also we call it an *isomorphism*, and say the two spaces are isomorphic. (Notation $V \cong U$).

NOTES

Case 1: Identity map $I : V \to V$, such that,

$$I(v) = v$$

and the zero map $O : V \to V$, such that,

$$O(v) = 0$$

are clearly linear transformations.

Case 2: For a field F, consider the vector spaces F^2 and F^3. Define a map $T : F^3 \to F^2$, by

$$T(\alpha, \beta, \gamma) = (\alpha, \beta)$$

then T is a linear transformation as

for any $x, y \in F^3$, if $x = (\alpha_1, \beta_1, \gamma_1)$

$$y = (\alpha_2, \beta_2, \gamma_2)$$

then
$$T(x + y) = T(\alpha_1 + \alpha_2, \beta_1 + \beta_2, \gamma_1 + \gamma_2) = (\alpha_1 + \alpha_2, \beta_1 + \beta_2)$$
$$= (\alpha_1, \beta_1) + (\alpha_2, \beta_2) = T(x) + T(y)$$

and
$$T(\alpha x) = T(\alpha (\alpha_1, \beta_1, \gamma_1)) = T((\alpha\alpha_1, \alpha\beta_1, \alpha\gamma_1)$$
$$= (\alpha\alpha_1, \alpha\beta_1) = \alpha(\alpha_1, \beta_1) = \alpha T(x)$$

Case 3: Let V be the vector space of all polynomials in x over a field F. Define

$T : V \to V$, such that,

$$T(f(x)) = \frac{d}{dx} f(x)$$

then
$$T(f + g) = \frac{d}{dx}(f + g) = \frac{d}{dx} f + \frac{d}{dx} g = T(f) + T(g)$$

$$T(\alpha f) = \frac{d}{dx}(\alpha f) = \alpha \frac{d}{dx} f = \alpha T(f)$$

shows that T is a linear transformation.

In fact, if $\theta : V \to V$ be defined such that

$$\theta(f) = \int_0^x f(t)\ dt$$

then θ will also be a linear transformation.

Case 4: Consider the mapping

$T : \mathbf{R}^3 \to \mathbf{R}$, such that,

$$T(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$$

then T is not a linear transformation.

Consider, for instance,

$$T((1, 0, 0) + (1, 0, 0)) = T(2, 0, 0) = 4$$
$$T(1, 0, 0) + T(1, 0, 0) = 1 + 1 = 2.$$

NOTES

The term similarity transformation is used either to refer to a geometric similarity or to a matrix transformation that results in a similarity.

A similarity transformation is a conformal mapping whose transformation matrix A' can be written in the form

$$A' \equiv BAB^{-1},$$

where A and A' are called similar matrices. Similarity transformations transform objects in space to similar objects.

Hypercompanion Matrix: Let $\{p(\lambda)\}^q$ be one of the elementary divisors of the characteristic matrix of some λ-matrix and $C(p)$ be the companion matrix of $p(\lambda)$. The hypercompanion matrix H associated with the elementary divisor $\{p(\lambda)\}^q$ is given by

$$H = c(p) \qquad If\ q = 1 \quad H = \begin{bmatrix} C(p) & M & 0 & \ldots & 0 & 0 \\ 0 & C(p) & M & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & C(p) & M \\ 0 & 0 & 0 & \ldots & 0 & C(p) \end{bmatrix} \ if\ q > 1$$

where M is a matrix of the same order as $C(p)$ having the element 1 in the lower left-hand corner and zeros elsewhere. The diagonal of the hypercompanion matrix H consists of q identical $C(p)$ matrices. There is a continuous line of 1s just above the diagonal.

Note: Every square matrix A over F is similar to the direct sum of the hypercompanion matrices of the elementary divisors over F of $\lambda I - A$.

Jacobson Canonical Form: The Jacobson canonical form of a square matrix A consists of the direct sum of the hypercompanion matrices of the elementary divisors over F of $\lambda I - A$, i.e., the matrix J,

$$J = \begin{bmatrix} H_1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & H_2 & 0 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & H_{k-1} & 0 \\ 0 & 0 & 0 & \ldots & 0 & H_k \end{bmatrix}$$

where H_i is the hypercompanion matrix associated with the i-th elementary divisor.

Jordan Canonical Form: Let the elementary divisors of the characteristic matrix of a matrix A be powers of linear polynomials. Then the canonical form is the direct sum of hypercompanion matrices of the form

NOTES

$$H = \begin{bmatrix} a_i & 1 & 0 & \dots & 0 & 0 \\ 0 & a_i & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_1 & 0 \\ 0 & 0 & 0 & \dots & 0 & a_1 \end{bmatrix} \quad \dots(2.1)$$

corresponding to the elementary divisor $\{p(\lambda)\}^q = (\lambda - a_i)^q$. The diagonal contains q identical a_i's. This special case of the Jacobson canonical form is known as the **Jordan or classical canonical form.**

- Let F be the field in which the characteristic polynomial of a matrix A factors into linear polynomials. Then A is similar over F to the direct sum of hypercompanion matrices of the form in Equation (2.1), each matrix corresponding to an elementary divisor $(\lambda - a_i)^q$.
- An n-square matrix A is similar to a diagonal matrix if and only if the elementary divisors of $\lambda I - A$ are linear polynomials, i.e., if and only if the minimum polynomial of A is the product of distinct linear polynomials.

Rational Canonical Form: Let A be an $n \times n$ matrix A and let C_i, C_i, C_{i+1}, ..., C_n be the companion matrices of the non-trivial invariant factors of $\lambda I - A$. Then the rational canonical form for all matrices similar to A is

$$S = \begin{bmatrix} C_i & 0 & \dots & 0 \\ 0 & C_{i+1} & \dots & 0 \\ \dots & \dots & \dots & .. \\ 0 & 0 & \dots & C_n \end{bmatrix}$$

In other words, the rational form is the direct sum of the companion matrices C_i, C_i, C_{i+1}, ..., C_n :

$$S = diag\ (C_i, C_i, C_{i+1}, \dots, C_n)$$

- Every square matrix A is similar to the direct sum of the companion matrices of the non-trivial invariant factors of $\lambda I - A$.

Second Canonical Form: Given an $n \times n$ matrix A, let be the companion matrices of the elementary divisors of $\lambda I - A$. Then a canonical form for all matrices similar to A is

$$S = \begin{bmatrix} C_i & 0 & \dots & 0 \\ 0 & C_{i+1} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & C_n \end{bmatrix}$$

We can say that, the form is the direct sum of the companion matrices C_i, C_i, C_{i+1}, ..., C_n ,

$$S = diag\ (C_i, C_i, C_{i+1}, \dots, C_n)$$

- Every square matrix A over F is similar to the direct sum of the companion matrices of the elementary divisors over F of $\lambda I - A$.

NOTES

2.3 INVARIANT SUBSPACES AND REDUCTION TO TRIANGULAR FORM

Definition: Let T be a linear operator on a vector space V. If W is a subspace of V such that, $T(W) \subseteq W$, we say W is *invariant under T* or is *T-invariant*.

Case 5: Since $T(0) = 0$ and $T(V) = V$, both zero subspace and V are invariant subspaces of V.

Case 6: Let $v \in \text{Ker } T$ then $T(v) = 0 \in \text{Ker } T \Rightarrow \text{Ker } T$ is invariant subspace of V. Also $w \in ImT \Rightarrow w = T(v) \Rightarrow Tw = T(Tv)$, $Tv \in V \Rightarrow Tw \in ImT$.

$\therefore$ ImT is an invariant subspace of V.

Case 7: Let $f(t)$ be any polynomial. Let $v \in \text{Ker } (f(T))$ then $f(T)\, v = 0$

Since $f(t)\ .\ t = tf(t)$

$f(T)T = Tf(T)$

Thus, $f(T)\ Tv = Tf(T)\ v = 0$

$\Rightarrow$ $Tv \in \text{Ker } f(T)$

$\Rightarrow$ $\text{Ker } f(T)$ is invariant under T.

Example 2.1: *Let T be a linear operator on* $\mathbf{R}^2$, *the matrix of which in the standard ordered basis is*

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

Prove that the only subspaces of $\mathbf{R}^2$ *invariant under T are* $\mathbf{R}^2$ *and zero subspaces.*

Solution: Characteristic polynomial of A (or T) is $\begin{vmatrix} x-1 & 1 \\ -2 & x-2 \end{vmatrix} = x^2 - 3x + 4$,

whose roots are not real. Thus eigen values of A (or T) do not exist in $\mathbf{R}$. If W is an invariant subspace of $\mathbf{R}^2$ such that, $W \neq 0$, $\mathbf{R}$ then dim $W = 1$. Let W be spanned by v. Then $Tv \in W \Rightarrow Tv = \alpha v$, $v \neq 0 \Rightarrow \alpha$ is an eigen value of T ($\alpha \in \mathbf{R}$), a contradiction. Hence O and $\mathbf{R}^2$ are only invariant subspaces of $\mathbf{R}^2$.

Theorem 2.1: *Let W be an invariant subspace of linear operator T on V. Then T has a matrix representation* $\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$, *where A is matrix of restriction* T_w *of T on W.*

Proof: Let $\{w_1, ..., w_r\}$ be a basis of W. Let $\beta = \{w_1, ..., w_r, v_1, ..., v_s\}$ be a basis of V, obtained by extending basis of W.

Since $T(w) \in W$ for all $w \in W$, we define $T_w : W \to W$ by $T_w(x) = T(x)$ for all $x \in W$.

NOTES

Then T_w is operator in W.

$$T_w(w_1) = T(w_1) = a_{11}w_1 + ... + a_{r1}w_r$$
$$.....$$
$$T_w(w_r) = T(w_r) = a_{1r}w_1 + ... + a_{rr}w_r$$
$$T(v_1) = b_{11}w_1 + ... + b_{r1}w_r + c_{11}v_1 + ... + c_{s1}v_s$$
$$.....$$
$$T(v_s) = b_{1s}w_1 + ... + b_{rs}w_r + c_{1s}v_1 + ... + c_{ss}v_s$$

Thus matrix of T with respect to basis β is $\begin{bmatrix} a_{11} & \cdots & a_{1r} & b_{11} & \cdots & b_{1s} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ a_{r1} & \cdots & a_{rr} & b_{r1} & \cdots & b_{rs} \\ 0 & \cdots & 0 & c_{11} & \cdots & c_{1s} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & c_{s1} & \cdots & c_{ss} \end{bmatrix}$

$$= \begin{bmatrix} A & B \\ 0 & C \end{bmatrix} \text{ where } A = (a_{ij}), B = (b_{ij}), C = (c_{ij})$$

are of order $r \times r$, $r \times s$, $s \times s$ respectively

Clearly, A is matrix of T_w with respect to $\{w_1, ..., w_r\}$ = basis of W. T_w is called restriction of T on W.

We now show that the matrix C obtained in Theorem 2.1 is the matrix of some linear operator on $\frac{V}{W}$ induced by T.

Define $\hat{T} : \frac{V}{W} \to \frac{V}{W}$ such that,

$$\hat{T}(W + v) = W + T(v), \quad v \in V$$

Then $\hat{T}$ is well defined as $W + v = W + v'$

$$\Rightarrow \quad v - v' \in W$$
$$\Rightarrow \quad T(v - v') \in W$$
$$\Rightarrow T(v) - T(v') \in W$$
$$\Rightarrow \quad W + T(v) = W + T(v')$$

Since T is linear transformation, so is $\hat{T}$. Let $\{w_1, ..., w_r\}$ be a basis of W.

Then it can be extended to form a basis of V. Let $\{w_1, ..., w_r, v_1, ..., v_s\}$ be a basis of V. Then $\{W + v_1, ..., W + v_s\}$ is a basis of $\frac{V}{W}$.

NOTES

$$\text{Now, } \hat{T}(W + v_1) = W + T(v_1)$$
$$= W + b_{11}w_1 + ... + b_{r1}w_r + c_{11}v_1 + ... + c_{s1}v_s$$
$$= W + c_{11}v_1 + ... + c_{s1}v_s$$
$$.....$$

$$\hat{T}(W + v_s) = W + T(v_s) = W + b_{1s}w_1 + ... + b_{rs}w_r + c_{1s}w_1 + ... + c_{ss}v_s.$$
$$= W + c_{1s}v_1 + ... + c_{ss}v_s \qquad \text{(as in Theorem 2.1)}$$

$\therefore$ matrix of $\hat{T}$ with respect to basis $\{W + v_1, ..., W + v_s\}$ of $\frac{V}{W}$ is

$$\begin{bmatrix} c_{11} & \cdots & \cdots & c_{1s} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ c_{s1} & \cdots & \cdots & c_{ss} \end{bmatrix} = C$$

A special situation where $B = 0$ in theorem is obtained when V is a direct sum of two invariant subspaces under T.

Example 2.2: *If W and U are invariant subspaces of a linear operator on a* Finite Dimensional Vector Space (*F.D.V.S*). *V over F and* $V = U \oplus W$, *then* $\exists$ *a basis* β *of V such that the matrix of T with respect to* β *is* $\begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix}$, *where A is the matrix of* T_w *on W and C is the matrix of* T_u *on U.*

Solution: Let $\{w_1, ..., w_r\}$ be a basis of W and $\{u_1, ..., u_s\}$ be a basis of U. Then $\{w_1, ..., w_r, u_1, ..., u_s\}$ is a basis of $W \oplus U = V$.

Now $T_w(w_1) = T(w_1) = a_{11}w_1 + ... + a_{r_1}w_r$
$T_w(w_2) = T(w_2) = a_{12}w_1 + ... + a_{r_2}w_2$
.....
$T_w(w_r) = T(w_r) = a_{1r}w_1 + ... + a_{rr}w_r$

as $T(w_i) \in W$ for all $i = 1, ..., r$

Similarly, $T_u(u_1) = T(u_1) = c_{11}u_1 + ... + c_{s1}u_s$
$T_u(u_2) = T(u_2) = c_{12}u_1 + ... + c_{s2}u_s$
.....
$T_u(u_s) = T(u_s) = c_{1s}u_1 + ... + c_{ss}u_s$

as $T(u_j) \in U$ for all $j = 1, ..., s$

So matrix of T with respect to $\beta = \{w_1, ..., w_r, u_1, ..., u_r\}$ of V is given by

NOTES

$$\begin{bmatrix} a_{11} & \cdots & a_{1r} & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ a_{r1} & \cdots & a_{rr} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & c_{11} & \cdots & c_{1s} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & c_{s1} & \cdots & c_{ss} \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix}$$

where $A = (a_{ij})$, $C = (c_{ij})$ are $r \times r$ and $s \times s$ matrices, respectively. Clearly A is the matrix of T_w on W and C is the matrix of T_u on U.

Example 2.3: *Let V be the vector space of all polynomials in x over F, of degree ≤ 5. Let $T : V \to V$ be defined by $T(1) = x^2 + x^4$, $T(x) = x + 1$, $T(x^2) = 1$, $T(x^3) = x^3 + x^2 + 1$, $T(x^4) = x^4$, $T(x^5) = 0$. If W is the linear span of $\{1, x^2, x^4\}$,*

(*a*) Show that W is invariant under T.

(*b*) Find the matrix of T_w in a suitable basis of W.

(*c*) Find the matrix of $\hat{T}$ in a suitable basis of $\frac{V}{W}$.

(*d*) Find the matrix of T in a suitable basis of V.

Solution (*a*): Let $w \in W$. Then $w = a + bx^2 + cx^4$ where $a, b, c \in F$.

$$\begin{aligned} T(w) &= a \,.\, T(1) + bt(x^2) + cT(x^4) \\ &= a(x^2 + x^4) + b + cx^4 \\ &= b + ax^2 + (a + c)\, x^4 \\ &\in W \text{ for all } w \in W \end{aligned}$$

$\therefore$ W is invariant under T.

(*b*): Notice that $\{1, x^2, x^4\}$ is linearly independent set over F and so forms a basis of W, and it can be extended to form a basis, namely $\{1, x^2, x^4, x, x^3, x^5\}$ of V.

Now, $T_w(1) = T(1) = x^2 + x^4 = 0 \,.\, 1 + 1 \,.\, x^2 + 1 \,.\, x^4$

$T_w(x^2) = T(x^2) = 1 = 1 \,.\, 1 + 0 \,.\, x^2 + 0 \,.\, x^4$

$T_w(x^4) = T(x^4) = x^4 = 0 \,.\, 1 + x^2 + 1 \,.\, x^4$

$\therefore$ matrix of T_w with respect to basis $\{1, x^2, x^4\}$ of W is given by

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

(*c*): Now $\{W + x, W + x^3, W + x^5\}$ is basis of $\frac{V}{W}$.

$$\begin{aligned} \therefore \quad \hat{T}(W + x) &= W + T(x) = W + x + 1 \\ &= W + x = 1 \,.\, (W + x) + 0(W + x^3) + 0(W + x^5) \end{aligned}$$

NOTES

$$\hat{T}(W + x^3) = W + T(x^3)$$
$$= W + x^3 + x^2 + 1$$
$$= W + x^3$$
$$= 0(W + x) + 1(W + x^3) + 0(W + x^5)$$
$$\hat{T}(W + x^5) = W + T(x^5)$$
$$= W + 0 = W = \text{zero of } \frac{V}{W}$$
$$= 0(W + x) + 0(W + x^3) + 0(W + x^5)$$

$\therefore$ matrix of $\hat{T}$ with respect to basis $\{W + x, W + x^3, W + x^5\}$ of $\frac{V}{W}$ is given by

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$(d): \quad T(x) = x + 1 = 1 \,.\, 1 + 0 \,.\, x^2 + 0 \,.\, x^4 + 1 \,.\, x + 0 \,.\, x^3 + 0 \,.\, x^5$$
$$T(x^3) = x^3 + x^2 + 1 = 1 \,.\, 1 + 1 \,.\, x^2 + 0 \,.\, x^4 + 0 \,.\, x + 1 \,.\, x^3 + 0 \,.\, x^5$$
$$T(x^5) = 0 = 0 \,.\, 1 + 0 \,.\, x^2 + 0 \,.\, x^4 + 0 \,.\, x + 0 \,.\, x^3 + 0 \,.\, x^5$$

$\therefore$ matrix of T with respect to basis $\{1, x^2, x^4, x, x^3, x^5\}$ of V is given by

$$\begin{bmatrix} 0 & 1 & 0 & \vdots & 1 & 1 & 0 \\ 1 & 0 & 0 & \vdots & 0 & 1 & 0 \\ 1 & 0 & 1 & \vdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \vdots & 1 & 0 & 0 \\ 0 & 0 & 0 & \vdots & 0 & 1 & 0 \\ 0 & 0 & 0 & \vdots & 0 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}, \text{ where } B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Example 2.4: *Let T be a linear operator on a F.D.V.S. V over F. Let W be an invariant subspace of T. Show that the characteristic polynomial $p_T(x)$ of T is given by*

$p_T(x) = p_{T_w}(x)\, p_{\hat{T}_w}(x)$, *where* $p_{T_w}(x)$, $p_{\hat{T}_w}(x)$ *are the characteristic polynomials of* T_w *and* $\hat{T}_w$ *respectively.*

NOTES

Solution: Characteristic polynomial $p_T(x)$ of T is given by

$$\left|\begin{bmatrix} A & B \\ 0 & C \end{bmatrix} - xI\right|$$

$$= \left|\begin{bmatrix} A - xI & B \\ 0 & C - xI \end{bmatrix}\right| \quad \left(\begin{array}{c}\text{Here } A = \text{matrix of } T_w \text{ on } W \\ C = \text{matrix of } \hat{T} \text{ on } \dfrac{V}{W}\end{array}\right)$$

$$= |A - xI||C - xI|$$

$$= (\text{characteristic polynomial of } T_w) \times (\text{characteristic polynomial of } \hat{T})$$

$$= p_{T_w}(x)\ p_{\hat{T}}(w).$$

A natural question arises 'What is the minimal polynomial for T in terms of minimal polynomial for T_w'? As we saw in above problem that the characteristic polynomial of T_w divides the characteristic polynomial of T, we have a similar result about minimal polynomial of T. We prove

Theorem 2.2: *The minimal polynomial of T_w divides the minimal polynomial for T, where W is an invariant subspace of V and T is a linear operator on V.*

Proof: Let $p(x)$ be the minimal polynomial for T.

Let $p(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_{n-1} x^{n-1} + x^n$

Since $T(w) = T_w(w)$ for all $w \in W$

$T^2(w) = T(T_w(w))$

$= T_w(T_w(w))$ as $T_w(w) \in W$

In this way $T^r(w) = T_w^r(w)$ for all $w \in W$

$\therefore\quad p(T_w)(w) = p(T)(w)$ for all $w \in W$

$= 0$ as $p(T) = 0$ for all $w \in W$

$\therefore\quad p(T_w) = 0$

Let $q(x)$ be the minimal polynomial for T_w. Then $p(x) = q(x)\, r(x) + h(x)$

where $h(x) = 0$ or $\deg h(x) < \deg q(x)$.

$\therefore\quad 0 = p(T_w) = q(T_w)\, r(T_w) + h(T_w)$

$\therefore\quad h(T_w) = 0$

If $h(x) \neq 0$, then $h(x)$ is non zero polynomial satisfied by T_w of degree less than deg $q(x)$, a contradiction as $q(x)$ is minimal.

$\therefore\quad h(x) = 0 \Rightarrow q(x)$ divides $p(x)$.

Definition: A linear operator T on a *F.D.V.S.* $V(F)$ is said to be *triangulable* or *triangularizable* over F if there exists an ordered basis β of V such that $[T]_\beta$ is triangular.

NOTES

Theorem 2.3: *Let T be a linear operator on a F.D.V.S. V(F). Then T is triangulable if and only if the characteristic polynomial for T is a product (not necessarily distinct) of linear factors on F[x]. (Equivalently, T is triangulable if and only if the eigen values of T are all in F).*

Proof: Let the characteristic polynomial of T be product of linear factors in $F[x]$.

Let $c_1, c_2, ..., c_n$ be eigen values of T in F.

We use induction on n.

If $n = 1$, then the result is obvious as 1×1 matrix is always triangular.

Let $n > 1$. Assume that the result is true for all vector spaces over F of dimension less than n.

Let dim $V = n$. Let v_1 be an eigen vector of T with respect to c_1, then $T(v_1) = c_1 v_1$

Let $W = < v_1 >$.

Then W is T-invariant subspace of V. Consider $\frac{V}{W}$. dim $\frac{V}{W} = n - 1$

Then $\hat{T} : \frac{V}{W} \to \frac{V}{w}$ such that,

$$\hat{T}\ (W + v) = W + T(v)$$

is well defined linear operator on $\frac{V}{W}$. Let $f(x)$ be the characteristic polynomial for T and $g(x)$ be the characteristic polynomial for $\hat{T}$. Then $g(x)$ divides $f(x)$ by Example 2.4.

So, $g(x)$ is also product of linear factors in $F[x]$.

By induction hypothesis $\exists$ a basis $\bar{\beta} = \{W + v_2, ..., W + v_n\}$ of $\frac{V}{W}$ such that,

$$\left[\hat{T}\right]_{\bar{\beta}} = \begin{bmatrix} a_{22} & \cdots & \cdots & a_{2n} \\ 0 & \ddots & & \\ \vdots & & \ddots & \\ 0 & \cdots & 0 & a_{nn} \end{bmatrix},\ a_{ij} \in F$$

$$\therefore \quad \hat{T}\,(W + v_j) = a_{2j}\,(W + v_2) + ... + a_{nj}(W + v_n)$$

$$\Rightarrow\ W + T(v_j) = a_{2j}(W + v_2) + ... + a_{nj}(W + v_n)$$

$$= W + a_{2j}\,v_2 + ... + a_{nj}v_n$$

$$\Rightarrow \quad T(v_j) = a_{2j}v_2 + ... + a_{nj}\,v_n + a_{1j}\,v_1 \quad , \quad a_{1j} \in F$$

Now, $\beta = \{v_1, v_2, ..., v_n\}$ is a basis of V

NOTES

$$\therefore \qquad [T]_\beta = \begin{bmatrix} a_{11} & a_{12} & a_{1n} \\ \vdots & a_{22} & a_{2n} \\ \vdots & \vdots & \vdots \\ 0 & 0 & a_{nn} \end{bmatrix}, \text{ where } a_{11} = c_1$$

which is triangular matrix and so T is triangulable. So, result follows by induction.

Conversely, if T is triangulable then $\exists$ a basis β of V such that, $[T]_\beta = A$ is triangular and eigen values of T are diagonal entries in A.

$\therefore$ Characteristic polynomial for A or T is product of linear factors in $F[x]$.

Note: We thus realise that T is triangulable if and only if minimal polynomial for T is product of linear factors in $F[x]$.

Corollary.: If A is $n \times n$ matrix over the field of complex numbers, then A is triangulable.

Proof: By fundamental theorem of algebra (i.e., Every polynomial over the field **C** of complex numbers has all roots in **C**), the minimal polynomial $p(x)$ of A has the form $p(x) = (x - c_1)^{r1} \ldots (x - c_k)^{rk}$, where $c_i \in \mathbf{C}$. By above theorem A is triangulable.

Example 2.5: *Let T be a linear operator on a finite dimensional vector space V(F). Suppose all eigen values of T are in F. Show that every non zero. T-invariant subspace of V contains an eigen vector of T.*

Solution: Let W be a non-zero T-invariant subspace of V. Then the restriction T_w of T on W is a linear operator on W. Since the characteristic polynomial of T_w divides the characteristic polynomial of T, eigen values of T_w also belong to F. Let $c \in F$ be an eigen value of T_w. Then $\exists\ 0 \neq x \in W$ such that $T_w(x) = cx \Rightarrow T(x) = cx \Rightarrow x$ is also an eigen vector of T.

Example 2.6: *Let T be a linear operator on V. If every subspace of V is invariant under T, show that T is a scalar multiple of the identity operator.*

Solution: Let $0 \neq v \in$ V. Let W be a subspace of V spanned by V. Since W is invariant under T, $v \in W \Rightarrow T(v) \in W \Rightarrow T(v) = \alpha v$. $w \in W \Rightarrow w = av \Rightarrow T(w) = aT(v) = a\alpha v = \alpha av = \alpha w$. Let $v' \notin W$, $v' \in V$. Then, v, v' are linearly independent. Let W' be the subspace spanned by v'. Since W' is invariant under T, $T(v') \in W'$.

$\therefore$ $T(v') = \alpha' v'$. Let V' be the subspace spanned by $v - v'$. Then as before $T(v - v') = \beta(v - v')$

$\Rightarrow T(v) - T(v') = \beta v - \beta v' \Rightarrow \alpha v - \alpha' v' = \beta v - \beta v'$

$\Rightarrow (\alpha - \beta)v = (\alpha' - \beta)v' \Rightarrow \alpha = \beta = \alpha'$ as v, v' are linearly independent

$\Rightarrow T(v') = \alpha(v')$.

$\therefore$ for all $v \in V$, $T(v) = \alpha v$

$\Rightarrow T = \alpha I$.

NOTES

Example 2.7: *Let T be a linear operator on a finite dimensional vector space over the field of complex numbers. Prove that T is diagonalisable, if and only if T is annihilated by some polynomial over* **C** *which has distinct roots.*

Solution: Suppose T is a diagonalisable. Let $p(x)$ be the minimal polynomial for T. By theorem 10 $p(x)$ has distinct roots and $p(T) = 0$.

Conversely, let $q(x)$ be a polynomial over **C** such that, $q(T) = 0$ and roots of $q(x)$ are distinct.

$\therefore$ $p(x)$ divides $q(x)$

and thus roots of $p(x)$ are distinct.

Hence T is diagonalizable.

Example 2.8: *If A is nilpotent, show that A is similar to a triangular matrix whose entries on the diagonal are all zero.*

Solution: A is nilpotent $\Rightarrow A^m = 0 \Rightarrow$ the minimal polynomial $p(x)$ of A is x^r, $r \leq m$. So, 0 is only eigen value of A. Since $0 \in F$, by theorem 2.3, A is similar to a triangular matrix B. $\therefore A = P^{-1}BP$

Since eigen value of A is only 0, eigen value of B is only 0 and these are diagonal entries on B.

Projections

We recall, by a projection E of a vector space V, we mean a linear operator on V such that, $E^2 = E$.

Let now E be a projection on V, then $E : V \to V$.

We show $V = R \oplus N$, where R = Range of E and

N = Null space of E = Ker E.

Let $v \in V$ be any element, then

$$E^2 = E$$

$$\Rightarrow \quad E^2(v) = E(v)$$

$$\Rightarrow \quad E(v - E(v)) = 0$$

$$\Rightarrow \quad v - E(v) \in \text{Ker } E = N$$

Thus $v = E(v) + (v - E(v)) \in R + N$

i.e., $V = R + N$

Again, let $x \in R \cap N$ then $x \in R$ and $x \in N$

$x \in R \Rightarrow \exists\, y \in V$ such that, $E(y) = x$

$x \in N \Rightarrow E(x) = 0$

So, $E^2(y) = E\,E(y) = E(x) = 0$

$\Rightarrow E(y) = 0 \Rightarrow x = 0 \Rightarrow R \cap N = \{0\}$

Hence $V = R \oplus N$.

NOTES

Suppose now $V = A \oplus B$, where A, B are subspaces of V.

Define $E : V \to V$, such that,

$E(v) = a$

where $v \in V \Rightarrow v = a + b$ (uniquely) $a \in A$, $b \in B$

Then E is easily seen to be a linear operator

Also $E^2(v) = EE(v) = E(a) = E(a + 0) = a = E(v) \quad \forall\ v \in V$

shows that $E^2 = E$ and thus E is a projection.

We claim A = range of E and B = Ker E

$v \in \text{Ker } E \Rightarrow E(v) = 0 \Rightarrow E(a + b) = 0$ where $v = a + b$

$\Rightarrow a = 0 \Rightarrow v = a + b = b \in B$

Again $b \in B \Rightarrow b = 0 + b \Rightarrow E(b) = E(0 + b) = 0 \Rightarrow b \in \text{Ker } E$

So $B = \text{Ker } E$

It is easy to see that A = range of E.

We thus notice that when there is projection E on V, then V is direct sum of range E and Ker E and *conversely*, if V is direct sum of two subspaces then there exists a projection E on V such that these subspaces are range and Ker of E.

If $V = R \oplus N$ corresponding to a projection E, we say E is projection on R along N (R = range E, N = Ker E).

Suppose again that $V = A \oplus B$ and let's define

$F : V \to V$ such that,

$F(v) = b$ where $v \in V$ is such that, $v = a + b$

then as before we can check that F is a projection on V and A = Ker F, B = Range F.

Hence if E was projection on A along B, then F is projection on B along A. Is there a direct relation between E and F?

Consider $(E + F)(v) = E(v) + F(v) = a + b = v,$

$= I(v) \quad \forall\ v$

and thus $E + F = I$

or that $E = I - F$

We can sum up and say that E is a projection iff $I - E$ is a projection and if E is a projection on R along N then $I - E$ is a projection on N along R.

We give another 'Proof' of this result in Example 2.9.

Let us now consider the general result through

Theorem 2.4: *If $V = W_1 \oplus ... \oplus W_k$, then $\exists\, k$ linear operators $E_1, ..., E_k$ on V such that,*

(*i*) *Each* E_i *is a projection*

(*ii*) $E_iE_j = 0$ *for all* $i \neq j$

(*iii*) $I = E_1 + ... + E_k$

(*iv*) *the range of* E_i *is* W_i

and conversely.

NOTES

Proof: Let $v \in V$ be any element then

$$v = x_1 + x_2 + ... + x_k,\ x_i \in W_i \text{ being uniquely determined}$$

Define $E_i : V \to V$, such that,

$$E_i(x_1 + ... + x_k) = x_i \text{ for all } i$$

Then E_i is linear operator such that,

$$E_i^2 (x_1 + ... + x_k) = E_i(x_i) = x_i = E_i(x_1 + ... + x_k)$$

$$\Rightarrow E_i^2 = E_i \text{ for all } i$$

This proves (*i*).

Let $i \neq j$. Then $E_iE_j(x_1 + ... + x_k) = E_i(x_j) = 0$

$$\therefore \quad E_iE_j = 0 \text{ for all } i \neq j.$$

This proves (*ii*).

Let $v \in V$. Then $v = x_1 + ... + x_k,\ x_i \in W_i$

$$\therefore \quad (E_1 + ... + E_k)v = E_1v + ... + E_kv$$

$$= x_1 + ... + x_k$$

$$= v = I(v)$$

$$\therefore \quad E_1 + ... + E_k = I$$

This proves (*iii*).

By definition of E_i, range of E_i is W_i which proves (*iv*).

Conversely, let $v \in V$. By (*iii*) $I = E_1 + ... + E_k$

$$\Rightarrow \quad v = I(v) = E_1(v) + ... + E_k(v) = x_1 + ... + x_k,\ x_i \in W_i\ (x_i = E_iv)$$

$$\therefore \quad V = W_1 + ... + W_k$$

Let $v = y_1 + ... + y_k,\quad y_i \in W_i$ = Range of E_i

$$\Rightarrow \quad y_i = E_i(z_i)$$

$$\therefore \quad E_j(v) = E_j(y_1) + ... + E_j(y_k)$$

$$= E_j\, E_1(z_1) + ... + E_jE_k(z_k)$$

$$= E_j^2(z_j) = E_j(z_j) = y_j$$

$$\therefore \quad x_j = y_j \text{ for all } j = 1, ..., k$$

$\therefore$ Each $v \in V$ can be uniquely written as sum of elements of $W_1, ..., W_k$.

Hence, $V = W_1 \oplus ... \oplus W_k$.

NOTES

Example 2.9: *Prove that if E is the projection on R along N, then I–E is the projection on N along R.*

Solution: Let $x \in R$ then $x = Ey$, $y \in V$

$$\Rightarrow (I - E)\, x = x - Ex = Ey - Ey = 0$$

$$\Rightarrow x \in \text{null space of } I - E$$

Also $x \in N \Rightarrow Ex = 0$

$$\Rightarrow (I - E)\, x = x \text{ for all } x \in N$$

$\therefore \quad v \in V \Rightarrow v = r + n,\ r \in R,\ n \in N$

$$\Rightarrow (I - E)\, v = (I - E)\, r + (I - E)\, n$$

$$= 0 + n = n$$

$\therefore$ Range space of $I - E$ is N

Also $(I - E)^2 = I + E^2 - 2E = I - E$

$\therefore$ $I - E$ is the projection on N along R.

Example 2.10: *Let V(F) be a vector space. Let E_1 be a projection on R_1 along N_1 and E_2 be a projection on R_2 along N_2. Assuming that $1 + 1 \neq 0$ in F, show that*

(a) *$E_1 + E_2$ is projection iff $E_1E_2 = E_2E_1 = 0$.*

(b) $E_1 + E_2$ is a projection on $R_1 \oplus R_2$ along $N_1 \cap N_2$.

Solution: (a) We have $V = R_1 \oplus N_1$ and $V = R_2 \oplus N_2$

Let $E_1 + E_2$ be a projection. Then $(E_1 + E_2)^2 = E_1 + E_2$

$$\Rightarrow E_1^2 + E_2^2 + E_1E_2 + E_2E_1 = E_1 + E_2$$

$$\Rightarrow E_1E_2 + E_2E_1 = 0 \quad (i)$$

$$\Rightarrow E_1E_1E_2 + E_1E_2E_1 = 0 \Rightarrow E_1E_2 = -E_1E_2E_1$$

and $E_1E_2E_1 + E_2E_1E_1 = 0 \Rightarrow E_2E_1 = -E_1E_2E_1$

Thus $E_1E_2 = E_2E_1$ and so (a) gives

$$(1 + 1)\, E_1E_2 = 0 \Rightarrow E_1E_2 = 0$$

Hence $E_1E_2 = E_2E_1 = 0$

Conversely, $E_1E_2 = E_2E_1 = 0$ gives

$$E_1E_2 + E_2E_1 = 0$$

$$\Rightarrow E_1^2 + E_2^2 + E_1E_2 + E_2E_1 = E_1 + E_2$$

$$\Rightarrow (E_1 + E_2)^2 = E_1 + E_2.$$

(b)We have to show that Range of $E_1 + E_2$ is $R_1 \oplus R_2$ and Ker $(E_1 + E_2) = N_1 \cap N_2$.

Let $x \in \text{Ker}\,(E_1 + E_2) \Rightarrow (E_1 + E_2)x = 0$

$$\Rightarrow E_1x + E_2\, x = 0 \Rightarrow E_1E_1(x) + E_1E_2(x) = 0$$

NOTES

$$\Rightarrow E_1(x) + E_1E_2(x) = 0$$

$$\Rightarrow E_1(x) = 0 \text{ as } E_1E_2(x) = 0$$

Similarly we get $E_2(x) = 0$

Hence $x \in \text{Ker } E_1 = N_1, x \in \text{Ker } E_2 = N_2$

and so $x \in N_1 \cap N_2 \Rightarrow \text{Ker } (E_1 + E_2) \subseteq N_1 \cap N_2$

Again, $y \in N_1 \cap N_2 \Rightarrow y \in N_1 \ \& \ y \in N_2$

$$\Rightarrow E_1(y) = 0, E_2(y) = 0$$

$$\Rightarrow (E_1 + E_2)y = 0 \Rightarrow y \in \text{Ker } (E_1 + E_2)$$

So $N_1 \cap N_2 \subseteq \text{Ker } (E_1 \cap E_2)$

or that $\text{Ker } (E_1 + E_2) = N_1 \cap N_2$

We leave the rest of the proof for the reader as an exercise.

Theorem 2.5: *Any projection E on a vector space V is diagonalizable.*

Proof: Suppose $\{v_1, v_2, ..., v_k\}$ is a basis of range space R of E and $\{v_{k+1}, ..., v_n\}$ is a basis of null space N of E.

Then $\{v_1, v_2, ..., v_k, v_{k+1}, ..., v_n\}$ is a basis of $R \oplus N = V$

Now, $E(v_1) = E(r_1 + n_1) \quad r_1 \in R, n_1 \in N$

$$\Rightarrow E^2(v_1) = E(v_1) = E(r_1 + n_1) = E(r_1) + E(n_1) = E(r_1)$$

$$\Rightarrow E(v_1) = E(r_1)$$

$$\Rightarrow E(v_1 - r_1) = 0 \Rightarrow v_1 - r_1 \in \text{Ker } E = N$$

Also $v_1 \in R, r_1 \in R \Rightarrow v_1 - r_1 \in R$

and thus $v_1 - r_1 \in R \cap N = \{0\}$

$$\Rightarrow v_1 = r_1$$

Again $n_1 = v_1 - r_1 = 0$

Thus $E(v_1) = v_1$. Similarly $E(v_i) = v_i \quad \forall\ i = 1, 2, ..., k$

Also $E(v_j) = 0 \quad \forall\ j = k + 1, ..., n.$

Showing matrix of E with respect to this basis is $\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$

which is clearly a diagonal matrix.

Hence the result follows.

Example 2.11: *If diagonal operator has eigen values 0 and 1 only then show that it is a projection.*

Solution: Since T is diagonal operator, $\exists$ a basis $\beta = \{v_1, ..., v_n\}$ of V such that $[T]_\beta$ = diagonal. Since eigen values of T are 0 and 1, let first m entries in diagonal be 1 and others be 0.

Let $v \in V$. Then $v = \alpha_1 v_1 + ... + \alpha_m v_m + \alpha_{m+1} v_{m+1} + ... + \alpha_n v_n$

NOTES

$\therefore \quad T^2(v) = T(Tv)$

$= T(\alpha_1 v_1 + ... + \alpha_m v_m)$ as $Tv_i = v_i$ for all i, $1 \le i \le m$

$Tv_j = 0$ for all $j > m$

$= T(\alpha_1 v_1 + ... + \alpha_m v_m + \alpha_{m+1} v_{m+1} + ... + \alpha_n v_n)$

$= T(v)$ for all $v \in V$

$\therefore \quad T^2 = T$

Hence T is a projection.

Theorem 2.6: *Let T be a linear operator on the space V and* $V = W_1 \oplus ... \oplus W_k$. *Define* $E_i(v) = E_i(x_1 + ... + x_k) = x_i \in W_i$. *Then each* E_i *is a projection on V such that,* $E_i E_j = 0$ *for all* $i \neq j$ *and* $I = E_1 + ... + E_k$. *Also then each* W_i *is invariant under T iff* $TE_i = E_i T$ *for all* $i = 1, 2, ..., k$.

Proof: Let $TE_i = E_i T$

Let $x_i \in W_i$. Then by definition, $E_i(x_i) = x_i$

$\therefore \quad T(x_i) = T(E_i x_i)$

$= E_i(Tx_i)$

$\Rightarrow T(x_i) \in$ Range of $E_i = W_i$

$\therefore W_i$ is invariant under T for all $i = 1, ..., k$

Conversely, let W_i be invariant under T. Then $v \in V$

$\Rightarrow \quad I(v) = (E_1 + ... + E_k)(v)$

$\Rightarrow \quad v = E_1(v) + ... + E_k(v)$

$\Rightarrow \quad T(v) = TE_1(v) + ... + TE_k(v)$

Since $E_i(v) \in W_i$ and W_i is T- invariant $\Rightarrow T(E_i(v)) \in W_i$.

So, $E_j[T(E_i(v)] = T(E_i(v))$ if $j = i$

$= 0$ if $j \neq i$

$\therefore \quad E_j(T(v)) = T(E_j(v)) \quad \forall\ v \in V$

$\Rightarrow \quad E_j T = TE_j \ \forall\ j.$

Definition: Let V be a vector space and $E_1, E_2, ..., E_k$ be a collection of projections on V, then this collection is called *orthogonal* collection if $E_i E_j = 0\ \forall$ $i \neq j$. Consider the space $\mathbf{R}^2$. Define

$E_1 : \mathbf{R}^2 \to \mathbf{R}^2$, such that, and $E_2 : \mathbf{R}^2 \to \mathbf{R}^2$, such that,

$E_1(a, b) = (a, 0)$ $\qquad$ $E_2(a, b) = (0, b)$

then clearly E_1, E_2 are projections and

$E_1 E_2(a, b) = E_1(0, b) = (0, 0)$

$E_2 E_1(a, b) = E_2(a, 0) = (0, 0)$

Shows $E_1E_2 = E_2E_1$ and thus E_1, E_2 is an orthogonal set of projections.

The above theorem could be restated as

Let T be a linear operator on the space V and let $V = W_1 \oplus W_2 \oplus ... \oplus W_k$ *be determined by orthogonal projections* $E_1, E_2, ..., E_k$ *on V. Then each* W_i *is T-invariant if and only if* $E_iT = TE_i$, $i = 1, 2, ..., k$.

Theorem 2.7: *Let T be a linear operator on a F.D.V.S. V. If T is diagonalizable and* $c_1, ..., c_k$ *are distinct eigen values of T, then* $\exists$ *linear operators* $E_1, ..., E_k$ *on V such that,*

(*i*) $T = c_1E_1 + ... + c_kE_k$

(*ii*) $I = E_1 + ... + E_k$

(*iii*) $E_iE_j = 0$ *for all* $i \neq j$

(*iv*) $E_i^2 = E_i$

(*v*) *Range of* E_i *is the eigen space of T associated with eigen value* c_i *of T.*

Conversely, if $\exists$ *distinct scalars* $c_1, ..., c_k$ *and k non-zero linear operators* $E_1, ..., E_k$ *satisfying* (*i*), (*ii*), (*iii*) *then T is diagonalizable,* $c_1, ..., c_k$ *are eigen values of T and* (*iv*) *and* (*v*) *are also satisfied.*

Proof: Let T be diagonalizable and $c_1, ..., c_k$ be distinct eigen values of T. Let W_i be eigen spaces of T corresponding to eigen values c_i.

Then dim $V = \dim W_1 + ... + \dim W_k$

and $V = W_1 + ... + W_k$

Hence $V = W_1 \oplus ... \oplus W_k$

As in Theorm 2.7, let $E_1, ..., E_k$ be the projections associated with this decomposition. Then (*ii*) to (*v*) are satisfied. Let $v \in V$

Then, $I(v) = v = (E_1 + ... + E_k)v$

$= E_1(v) + ... + E_k(v)$

$\Rightarrow T(v) = TE_1(v) + ... + TE_k(v)$

$= c_1E_1(v) + ... + c_kE_k(v)$ as $E_i(v) \in$ Range of $E_i = W_i$

$= (c_1E_1 + ... + c_kE_k)v$

$\Rightarrow T = c_1E_1 + ... + c_kE_k$

This proves (*i*).

Conversely, suppose T along with distinct scalars c_i and non-zero operators E_i satisfy (*i*), (*ii*) and (*iii*). Also $T = c_1E_1 + ... + c_kE_k$

Then $TE_i = c_iE_i$ for all i

$\Rightarrow (T - c_iI)\, E_i = 0$ for all i

Since $E_i \neq 0 \; \exists \; v_i \in V$ such that, $E_i(v_i) \neq 0$

NOTES

$\therefore \quad (T - c_i I)\,(E_i(v_i)) = 0$ for all i

$\Rightarrow \; T(E_i(v_i)) = c_i(E_i(v_i)),\ E_i(v_i)) \neq 0$ for all i

$\Rightarrow c_i$ is an eigen value of T for all i, ($E_i\, v_i$ is an eigen vector).

If c is any scalar, then

$$(T - cI) = (c_1E_1 + ... + c_kE_k) - c(E_1 + ... + E_k)$$
$$= (c_1 - c)E_1 + ... + (c_k - c)E_k$$

If c is an eigen value of T, then $\exists\ 0 \neq v \in V$ such that,

$$Tv = cv \Rightarrow (T - cI)v = 0$$

$\therefore \quad (c_1 - c)\, E_jE_1(v) + ... + (c_k - c)\, E_jE_k(v) = 0$

$$\Rightarrow (c_j - c)\, E_j(v) = 0 \text{ for all } j = 1, ..., k$$

If $E_j(v) = 0$ for all j, then $I = E_1 + ... + E_k$

$$\Rightarrow v = I(v) = E_1(v) + ... + E_k(v) = 0$$

$\therefore \quad E_j(v) \neq 0$ for some j

$\therefore \quad c_j = c$ for some j

$\therefore c_1, ..., c_k$ are only eigen values of T.

Let W_i = range of E_i, $i = 1, ..., k$.

By (ii) $\quad I = E_1 + ... + E_k$

$\Rightarrow \; v = Iv = E_1v + ... + E_kv \in W_1 + ... + W_k$ for all $v \in V$

$\Rightarrow \; V = W_1 + ... + W_k$

As in Theorem 2.4, $V = W_1 \oplus ... \oplus W_k$

$\therefore \quad \dim V = \dim W_1 + ... + \dim W_k$

$\Rightarrow T$ is diagonalisable if W_i = eigen space of T corresponding to c_i.

Let $x \in$ eigen space of T. Then $T(x) = c_ix$, $1 \leq i \leq k$

$\Rightarrow \; (c_1E_1 + ... + c_kE_k)x = c_iI(x) = c_i(E_1 + ... + E_k)x$

$\Rightarrow \; c_1E_1(x) + ... + c_kE_k(x) = c_iE_1(x) + ... + c_iE_k(x)$

$\Rightarrow \; (c_1 - c_i)\, E_1(x) + ... + (c_k - c_i)\, E_k(x) = 0$

$\Rightarrow \; (c_j - c_i)\, E_j(x) = 0$ for all $j = 1, ..., k$

as $\quad E_j(x) \in$ Range of $E_j = W_j$

and $\quad W_1, ..., W_k$ are independent.

we get $\quad E_j(x) = 0,\ j \neq i$ as $c_j - c_i \neq 0$ for all $j \neq i$

Since $\quad I = E_1 + ... + E_k$,

$$x = E_1(x) + ... + E_k(x) = E_i(x)$$

$\Rightarrow \; x \in$ Range of $E_i = W_i$

$\therefore$ eigen space corresponding to c_i is contained in W_i.

NOTES

Also $0 \neq x \in W_i \Rightarrow x = E_i(y_i) \neq 0$

But $(T - c_i I)E_i = 0$

$\Rightarrow TE_i(y_i) = c_i E_i(y_i)$

$\Rightarrow T(x) = c_i x \Rightarrow x \in$ eigen space corresponding to c_i

$\therefore$ W_i = eigen space corresponding to c_i.

Suppose T is a linear operator with minimal polynomial $p(x) = (x - c_1) ... (x - c_k)$ such that, $c_1, ..., c_k \in F$ are distinct. To show T is diagonalizable.

Proof: Let $p_j(x) = \prod_{i \neq j} \frac{(x - c_i)}{(c_j - c_i)}$, $j = 1, ..., k$

Then $p_j(c_i) = \delta_{ij}$

Let V = space of all polynomials over F of degree less than k.

Then $p_1, ..., p_k \in V$ and are linearly independent as $\alpha_1 p_1 + ... + \alpha_k p_k = 0$

$\Rightarrow \alpha_1 p_1(c_i) + ... + \alpha_k p_k(c_i) = 0$

$\Rightarrow \alpha_i = 0$ for all i

Since dim $V = k$, $\{p_1, ..., p_k\}$ is a basis of V.

Now $1 \in V \Rightarrow 1 = \alpha_1 p_1 + ... + \alpha_k p_k$

Put $x = c_i$ on both sides to get

$1 = \alpha_i$ for all i

$\Rightarrow 1 = p_1 + ... + p_k$ ...(2.1)

$x \in V \Rightarrow x = \beta_1 p_1 + ... + \beta_k p_k$

Put $x = c_i$

Then $c_i = \beta_i$ for all i

$\Rightarrow x = c_1 p_1 + ... + c_k p_k$ ...(2.2)

Let $p_j(T) = E_j$

Put $x = T$ in Equation (2.1) and Equation (2.2) above to get

$I = p_1(T) + ... + p_k(T) = E_1 + ... + E_k$

$T = c_1 E_1 + ... c_k E_k$

Since $p(x)$ divides $p_i(x)p_j(x)$ for all $i \neq j$

$p_i(T)p_j(T) = p(T)q(T)$ for all $i \neq j$

$\Rightarrow E_i E_j = 0$ for all $i \neq j$

If $E_j = 0$ for some j, then $p_j(T) = 0$ and

degree of $p_j(x) <$ deg $p(x)$, a contradiction

$\therefore E_j \neq 0$ for all $j = 1, ..., k$

$\therefore$ T is diagonalizable.

NOTES

Example 2.12: *Let E be a projection of V and let T be a linear operator on V. Prove that the range of E is invariant under T if and only if ETE = TE. Prove that both the range and null space of E are invariant under T if and only if ET = TE.*

Solution: Let R = range of E

N = null space of E

Then $V = R \oplus N$

We have shown before that $I - E$ is also a projection. $x \in N \Rightarrow Ex = 0 \Rightarrow (I - E)x = x \Rightarrow x \in$ range of $I - E$. $\therefore$ Range of $E = R$, Range of $(I - E) = N$.

Also $E(I - E) = E - E^2 = E - E = 0$. Suppose R is invariant under T then $\Rightarrow T(EV) \subseteq EV \Rightarrow T(I - E)V = T(V - EV) \subseteq V - EV = (I - E)V \Rightarrow N = (I - E)V$ is invariant under T.

$\therefore$ By Theorem 2.6, $TE = ET$

$\Rightarrow \quad ETE = E^2T = ET = TE.$

Conversely, suppose $ETE = TE$

Let $E(v) \in R$ = range of E

Then $E\,(TE(v)) \in R$ as $T : V \to V$, $E : V \to V$

$\Rightarrow \quad TE(v) \in R$ since $ETE = TE$

$\Rightarrow \quad R$ is invariant under T.

Further, if both R and N are invariant under T, then by Theorem 2.7, $TE = ET$.

Conversely, suppose $TE = ET \Rightarrow ETE = TE$

From above then, R is invariant under T.

Also $n \in N \Rightarrow E(n) = 0 \Rightarrow (ET)(n) = (TE)(n)$

$= T(E(n))$

$= T(0) = 0$

$\therefore \quad E(T(n)) = 0$ for all $n \in N$

$\Rightarrow T(n) \in$ null space of E for all $n \in N$

$\Rightarrow N$ is invariant under T.

Example 2.13: *Let $V = \mathbf{R}^2$ and T be the linear operator on V whose matrix relative to standard ordered basis is* $\begin{bmatrix} a & b \\ 0 & a \end{bmatrix}$ *for same non-zero a, b* $\in \mathbf{R}$. *Show that*

(a) W_1 *the subspace generated by* (1, 0) is *T-invariant*

(b) W_2 *the subspace generated by* (0, 1) *is not T-invariant*

(c) $\exists$ *no T-invariant subspace W of* $\mathbf{R}^2$ *such that,* $\mathbf{R}^2 = W_1 \oplus W$.

NOTES

Solution: We have $W_1 = \{(x, 0) \mid x \in \mathbf{R}\}$

$$T(W_1) = \{a(x, 0) \mid x \in \mathbf{R}\} \subseteq W_1$$

and thus W_1 is invariant under T.

Suppose now W is T-invariant subspace of $\mathbf{R}^2$ such that, $\mathbf{R}^2 = W_1 \oplus W$. Since dim $W_1 = 1$, dim W must also be 1.

Define $E : \mathbf{R}^2 \to \mathbf{R}^2$, such that,

$$E(x, y) = (x, 0)$$

then E is a projection of $\mathbf{R}^2$ onto W_1.

By Example 2.12, we should have $TE = ET$.

But $TE(1, 1) = T(1, 0) = (a, 0)$

$$ET(1, 1) = E(a + b, a) = (a + b, 0)$$

Showing that $ET \neq TE$ and thus there does not exist any T-invariant subspace W such that, $\mathbf{R}^2 = W_1 \oplus W$. We leave part (b) and (c) for students to complete.

Theorem 2.8: *Let T be a linear operator on the F.D.V.S. V(F). Suppose that the minimal polynomial for T decomposes over F into a product of linear polynomials. Then* $\exists$ *a diagonalizable operator D on V and a nilpotent operator N on V such* that (*i*) $T = D + N$ (*ii*) $DN = ND$.

Proof: Let $p(x) = (x - c_1)^{r1} \ldots (x - c_k)^{rk}$ be the minimal polynomial for T where $c_1, \ldots, c_k$ are distinct scalars in F.

By Primary decomposition theorem, $V = W_1 \oplus \ldots \oplus W_k$, where W_i = null space of $(T - c_i I)^{ri}$. Let $E_1, \ldots, E_k$ be the corresponding projections. Then W_i = range of E_i.

Let $D = c_1E_1 + \ldots + c_kE_k$

By Theorem 2.7, D is diagonalizable.

Since $I = E_1 + \ldots + E_k$

$$T = TE_1 + \ldots + TE_k,\ D = c_1E_1 + \ldots + c_kE_k$$

Let $N = T - D = (T - c_1I)E_1 + \ldots + (T - c_kI)E_k$

Then $N^2 = (T - c_1I)^2E_1 + \ldots + (T - c_kI)^2 E_k$ as $TE_i = E_iT\ \forall\ i$

and in general that, $N^r = (T - c_1I)^rE_1 + \ldots + (T - c_kI)^r E_k$

Since $(x - c_i)^{ri}$ is the minimal polynomial of T on W_i, $(T - c_iI)^{ri} = 0$ on W_i for all i.

$\Rightarrow$ $(T - c_iI)^r = 0$ on W_i for all $r \geq r_i$

$\therefore$ $N^r = 0$ for all $r \geq r_i$ for each i

$\therefore$ N is nilpotent operator.

$\therefore$ $T = D + N$, D is diagonalizable and N, nilpotent operator.

Now $DT = (c_1E_1 + \ldots + c_kE_k)\ (TE_1 + \ldots + TE_k)$

$$= c_1TE_1 + \ldots + c_kTE_k$$

NOTES

as $W_i s$ are invariant under $T \Rightarrow TE_i = E_i T$ for all i

$$= T(c_1E_1) + ... + (c_k E_k)$$
$$= (TE_1)\ (c_1E_1) + ... + (TE_k)\ (c_k E_k)$$
$$= (TE_1 + ... + TE_k)\ (c_1E_1 + ... + c_kE_k)$$
$$= TD$$

$\therefore \quad D(D + N) = (D + N)D$

$\Rightarrow \quad DN = ND.$

2.4 NILPOTENT TRANSFORMATIONS

A linear transformation $N : U \to U$ is called nilpotent if there exists a $k \in N$ such that $N^k = 0$ for some positive integer k. The smallest such k is sometimes called the degree of N.

A nilpotent transformation is a linear transformation L of a vector space such that $L^k = 0$ for some positive integer k. A nilpotent transformation naturally determines a flag of subspaces

$\{0\} \subset \ker N^1 \subset \ker N^2 \subset .... \ker N^{k-1} \subset \ker N^k = U$ and a signature,

$0 = n_0 < n_1 < n_2 < n_{k-1} < n_k = \dim U$, $n_i = \dim \ker N^i$.

The signature is governed by the following constraint, and characterizes N up to linear isomorphism.

Theorem 2.9: A sequence of increasing natural numbers is the signature of a nilpotent transformation if and only if

$$n_{j+1} - n_j \leq n_j - n_{j-1}$$

for all $j = 1, ..., k-1$. Equivalently, there exists a basis of U such that the matrix of N relative to this basis is block diagonal

$$\begin{pmatrix} N_1 & 0 & 0 & ... & 0 \\ 0 & N_2 & 0 & ... & 0 \\ 0 & 0 & N_3 & ... & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & ... & N_k \end{pmatrix}$$

with each of the blocks having the form

$$N_i = \begin{pmatrix} 0 & 1 & 0 & ... & 0 & 0 \\ 0 & 0 & 1 & ... & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & 0 & ... & 1 & 0 \\ 0 & 0 & 0 & ... & 0 & 1 \\ 0 & 0 & 0 & ... & 0 & 0 \end{pmatrix}$$

Letting d_i denote the number of blocks of size i, the signature of N is given by

$$n_i = n_{i-1} + d_i + d_{i+1} + \ + d_k,\ i = 1, \ldots, k$$

Theorem 2.10: Theorem NJB

Nilpotent Jordan Blocks

The Jordan block $J_n(0)$ is nilpotent of index n.

Proof: While not phrased as an if-then statement, the statement in the theorem is understood to mean that if we have a specific matrix $J_n(0)$) then we need to establish it is nilpotent of a specified index. The first column of J_n is the zero vector, and the remaining $n-1$ columns are the standard unit vectors e_i, $1 \le i \le n-1$.

which are also the first $n-1$ columns of the size n identity matrix I_n. As shorthand, write $J = J_n$

$$J = \left[0 \,|e_1\, |e_2\, |e_3\, |\ldots\, |e_{n-1}\right]$$

We will use the definition of matrix multiplication together with a proof by induction to study the powers of J. Our claim is that

$$J^k = \left[0 \,|0\, |\ldots\, |0\, |e_1\, |e_2\, |\ldots\, |e_{n-k}\right]$$

for $1 \le k \le n$. For the base case, $k = 1$ and the definition of $J^1 = J_n(0)$ establishes the claim. For the induction step, first note that $Je_1 = 0$ and $Je_i = e_{i-1}$ for $2 \le i \le n$. Then, assuming the claim is true for k, we examine the k + 1 case,

$$\begin{aligned} J^{k+1} &= JJ^k \\ &= J\left[0 \,|0\, |\ldots\, |0\, |e_1\, |e_2\, |\ldots\, |e_{n-k}\right] \\ &= \left[J0 \,|J0\, |\ldots\, |J0\, |Je_1\, |Je_2\, |\ldots\, |Je_{n-k}\right] \\ &= \left[0 \,|0\, |\ldots\, |0\, |0\, |e_1\, |e_2\, |\ldots\, |e_{n-k-1}\right] \\ &= \left[0 \,|0\, |\ldots\, |0\, |e_1\, |e_2\, |\ldots\, |e_{n-(k+1)}\right] \end{aligned}$$

This concludes the induction. So J_k has a nonzero entry (a one) in row $n-k$ and column n, for $1 \le k \le n-1$, and is therefore a nonzero matrix. However, $J^n = \left[0 \,|0\, |\ldots\, |0\right] = O$. J is nilpotent of index n.

Theorem 2.11: ENLT

Eigenvalues of Nilpotent Linear Transformations

Suppose that $T : V \to V$ is a nilpotent linear transfromation and λ is an eigenvalue of T. Then $\lambda = 0$

Proof: Let x be an eigenvector of T for the eigenvalue λ, and suppose that T is nilpotent with index p.

NOTES

Then

$$0 = T^p(x)$$
$$= \lambda^p x$$

Because x is an eigenvector, it is nonzero, and therefore Theorem SMEZV tells us that $\lambda^p = 0$ and so $\lambda = 0$.

Paraphrasing, all of the eigenvalues of a nilpotent linear transformation are zero. So in particular, the characteristic polynomial of a nilpotent linear transformation, T, on a vector space of dimension n, is simply $p_T(x) = x^n$.

Theorem 2.12: DNLT

Diagonalizable Nilpotent Linear Transformations

Suppose the linear transformation $T: V \to V$ is nilpotent. Then T is diagonalizable if and only T is the zero linear transformation.

Proof: We start with the easy direction. Let $n = dim\ (V)$.

($\Leftarrow$) The linear transformation $Z: V \to V$ defined by $Z(v) = 0$ for all $v \in V$ is nilpotent of index $p = 1$ and a matrix representation relative to any basis of V is the $n \times n$ zero matrix, O. Quite obviously, the zero matrix is a diagonal matrix and hence Z is diagonalizable.

($\Rightarrow$) Assume now that T is diagonalizable, so $\gamma_T(\lambda) = \alpha_T(\lambda)$ for every eigenvalue λ. By Theorem ENLT, T has only one eigenvalue (zero), which therefore must have algebraic multiplicity n (Theorem NEM). So the geometric multiplicity of zero will be n as well $\gamma_T(0) = n$.

Let B be a basis for the eigenspace $e_T(0)$. Then B is a linearly independent subset of V of size n, and by will be a basis for V. For any $x \in B$ we have

$$T(x) = 0x$$
$$= 0$$

So T is identically zero on a basis for B, and since the action of a linear transformation on a basis determines all of the values of the linear transformation, it is easy to see that $T(v) = 0$ for every $v \in V$.

So, other than one trivial case (the zero matrix), every nilpotent linear transformation is not diagonalizable.

Theorem 2.13: KPLT

Kernels of Powers of Linear Transformations

Suppose $T: V \to V$ is a linear transformation, where $dim\ (V) = n$. Then there is an integer m, $0 \le m \le n$, such that

$$\{0\} = K(T^0) \subsetneq K(T^1) \subsetneq K(T^2) \subsetneq \cdots \subsetneq K(T^m) = K(T^{m+1}) = K(T^{m+2}) = \cdots$$

NOTES

Proof:

There are several items to verify in the conclusion as stated. First, we show that $K(T^k) \subseteq K(T^{k+1})$ for any k. Choose $z \in K(T^k)$. Then

$$\begin{aligned} T^{k+1}(z) &= T(T^k(z)) \\ &= T(0) \\ &= 0 \end{aligned}$$

So by Definition KLT, $z \in K(T^{k+1})$ and by Definition SSET we have $K(T^k) \subseteq K(T^{k+1})$.

Second, we demonstrate the existence of a power m where consecutive powers result in equal kernels. A byproduct will be the condition that m can be chosen so that $m \leq n$. To the contrary, suppose that

$$\{0\} = K(T^0) \subsetneq K(T^1) \subsetneq K(T^2) \subsetneq \cdots \subsetneq K(T^{n-1}) \subsetneq K(T^n) \subsetneq K(T^{n+1}) \subsetneq \cdots$$

Since

$$K(T^k) \subsetneq K(T^{k+1}),\ dim\left(K(T^{k+1})\right) \geq dim\left(K(T^k)\right) + 1.$$

Repeated application of this observation yields

$$\begin{aligned} dim\left(K(T^{n+1})\right) &\geq dim\left(K(T^n)\right) + 1 \\ &\geq dim\left(K(T^{n-1})\right) + 2 \\ &\vdots \\ &\geq dim\left(K(T^0)\right) + (n+1) \\ &= dim\left(\{0\}\right) + n + 1 \\ &= n + 1 \end{aligned}$$

Thus, $K(T^{n+1})$ has a basis of size at least $n + 1$, which is a linearly independent set of size greater than n in the vector space of dimension n.

This contradiction yields the existence of an integer k such that $K(T^k) = K(T^{k+1})$, so we can define m to be smallest such integer with this property. From the argument above about dimensions resulting from a strictly increasing chain of subspaces, it should be clear that $m \leq n$.

NOTES

It remains to show that once two consecutive kernels are equal, then all of the remaining kernels are equal. More formally, if $K\left(T^m\right) = K\left(T^{m+1}\right)$, then $K\left(T^m\right) = K\left(T^{m+j}\right)$ for all $J \geq 1$.

Theorem 2.14: KPNLT

Kernels of Powers of Nilpotent Linear Transformations

Suppose $T : V \to V$ is a nilpotent linear transformaton with index p and $dim\ (V) = n$. Then $0 \leq p \leq n$, and

$$\{0\} \subsetneq K\left(T^0\right) \subsetneq K\left(T^1\right) \subsetneq K\left(T^2\right) \subsetneq \cdots \subsetneq K\left(T^p\right) = K\left(T^{p+1}\right) = \cdots = V$$

Proof: Since $T^p = 0$ it follows that $T^{p+j} = 0$ all $j \geq 0$ and thus $K(T^{p+j}) = V$ for $j \geq 0$. So the value of m guaranteed by Theorem KPLT is at most p. The only remaining aspect of our conclusion that does not follows form Theorem KPLT is that $m = p$. To see this we must show that $K(T^k) \subseteq K(T^{k+1})$ for $0 \leq k \leq p-1$. If $K(T^k) = 1$. For some $k < p$, then $K(T^k) = K(T^p) = V$. This implies that $T^k = 0$, violating the fact that T has index p. So the smallest value of m is indeed p, and we learn that $p < n$.

Theorem 2.15: CFNLT

Canonical Form for Nilpotent Linear Transformations

Suppose that $T : V \to V$ is a nilpotent linear transformation of index p. Then there is a basis for V so that the matrix representation $M^T_{B,B}$, is block diagonal with eachblock being a Jardan block, $J_n(0)$. The size of the largest block is the ndex p, and the total number of block si the nullity of T, $n(T)$.

Proof: We will explicitly construct the desired basis, so the proof is constructive and can be used in practice. As we begin, the basis vectors will not be in the proper order, but we will rearrange them at the end of the proof. For convenience, define $n_i = n(T^i)$, so for example, $n_0 = 0$, $n_1 = n(T)$, and $n_p = n(T_p)$. Define $s_i = n_i - n_{i-1}$, for $1 \leq i \leq p$, so we can think of as "how much bigger" $K(T_i)$ is than $K(T^{i-1})$. In particular, Theorem KPNLT implies that si > 0 for $1 \leq i \leq p$.

We are going to build a set of vectors $z_{i,j}$, $1 \leq i \leq p$, $1 \leq j \leq s_i$. Each $z_{i,j}$, will be an element of $K(T^i)$ and not an element of $K(T^{i-1})$. In total, we will obtain a linearly independent set of $\sum_{i=1}^{p} s_i = \sum_{i=1}^{p} n_i - n_{i-1} = n_p - n_0 = dim\left(V\right)$ vectors that form a basis of V. We construct this set in pieces, starting at the "wrong" end. Our procedure will build a series of subspaces, Z_i, each lying in between $K(T^{i-1})$ and $K(T^i)$, having bases $z_{i,j}$, $1 \leq j \leq s_i$, and which together equal V as a direct sum.

We build the subspace Z_p first (this is what we meant by "starting at the wrong end"). $K(T^{p-1})$ is a proper subspace of $K(T^p) = V$ (Theorem KPNLT). There is a subspace of V that will pair with the subspace $K(T^{p-1})$ to form a direct sum of V. Call this subspace Z_p, and choose vectors $z_{p,j}$, $1 \leq j \leq s_p$, as a basis of

NOTES

Z_p, which we will denote as B_p. Note that we have a fair amount of freedom in how to choose these first basis vectors. Several observations will be useful in the next step. First $V = K(T^{p-1)\oplus Z_p}$. The basis $B_p = \left\{z_{p,1}, z_{p,2}, z_{p,3}, \ldots z_{p,s_p}\right\}$ is linearly independent. For $1 \le j \le s_p$, $z_{p,j} \in K(T^p)$. Since the two subspaces of a direct sum have no nonzero vectors in common, for $1 \le j \le s_p$, $z_{p,j} \in K(T^{p-1})$. That was comparably easy.

If obtaining Z_p was easy, getting Z_{p-1} will be harder. We will repeat the next step $p-1$ times, and so will do it carefully the first time. Eventually, Z_{p-1} will have dimension S_{p-1}. However, the first s_p vectors of a basis are straightforward. Define $z_{p-1,j} = T(z_{p,j})$, $1 \le j \le s_p$. Notice that we have no choice in creating these vectors, they are a consequence of our choices for $z_{p,j}$. In retrospect (i.e. on a second reading of this proof), you will recognize this as the key step in realizing a matrix representation of a nilpotent linear transformation with Jordan blocks. We need to know that this set of vectors in linearly independent, so start with a relation of linear dependence and massage it,

$$\begin{aligned} 0 &= a_1 z_{p-1,1} + a_2 z_{p-1,2} + a_3 z_{p-1,3} + \cdots + a_{s_p} z_{p-1,s_p} \\ &= a_1 T\left(z_{p,1}\right) + a_2 T\left(z_{p,2}\right) + a_3 T\left(z_{p,3}\right) + \cdots + a_{s_p} T\left(z_{p,s_p}\right) \\ &= T\left(a_1 z_{p,1} + a_2 z_{p,2} + a_3 z_{p,3} + \cdots + a_{s_p} z_{p,s_p}\right) \end{aligned}$$

Define $x = a_1 z_{p,1} + a_2 z_{p,2} + a_3 z_{p,3} + \cdots + a_{s_p} z_{p,s_p}$. The statement just above means that $x \in K(T) \subseteq K(T^{p-1})$. As defined, x is a linear combination of the basis vectors B_p, and therefore $x \in Z_p$. Thus $x \in K(T^{p-1}) \cap Z_p$. Because $V = K(T^{p-1}) \oplus Z_p$, that $x = 0$. Now we recognize the definition of x as a relation of linear dependence on the linearly independent set B_p, and therefore $a_1 = a_2 = \cdots = a_{s_p} = 0$. This establishes the linear independence of $z_{p-1,j}$, $1 \le j \le s_p$.

2.4.1 Index of Nilpotency

Nilpotent element is an element a of a ring or semi-group with zero A such that, $a^n = 0$ for some natural number n. The smallest such n is called the nilpotency index of a. For example, in the residue ring modulo p^n (under multiplication), where p is a prime number, the residue class of p is nilpotent of index n.

In the ring of (2×2) matrices with coefficients in a field K the matrix,

$$\begin{Vmatrix} 0 & 1 \\ 0 & 0 \end{Vmatrix}$$

is nilpotent of index 2.

In the group algebra $Fp[G]$, where F_p is the field with p elements and G the cyclic group of order p generated by σ, the element $1-\sigma$ is nilpotent of index p.

NOTES

2.4.2 Invariants of Nilpotent Transformations

Lemma 1: A linear transformation whose only eigenvalue is zero is nilpotent.

Proof: If $t : V^n \to V^n$ has all $\lambda = 0$, then $c(x) = x^n$ and from Cayley-Hamilton Theorem t^n = zero map.

Note: If T is a square matrix with characteristic polynomial $c(x)$ then $c(T) = \mathbf{O}$.

Canonical form for nilpotent matrices is one that is all zeroes except for blocks of subdiagonal ones. This can be made unique by setting some rules for the arrangement of blocks.

Lemma 2: If the matrices $T - \lambda I$ and N are similar, then T and $N + \lambda I$ are also similar, via the same change of basis matrices.

Proof: $N = P\,(T - \lambda I)\,P^{-1} = PTP^{-1} - \lambda I$

$\Rightarrow PTP^{-1} = N + \lambda I$

Definition: Invariant Subspace: Let $t : V \to V$ be a transformation. Then a subspace M is t invariant if $m \in M \Rightarrow t(m) \in M$

Example 2.14: $N_\infty(t)$ and $R_\infty(t)$ are both t invariant.

Solution: If $v \in N\infty(t)$, then $\exists\, k$ such that, $t^n(v) = 0 \;\forall\; n \geq k$.

$t^{n+1}(v) = t^n(t(v)) = 0 \;\rightarrow\; t(v) \in \mathrm{N}_\infty(t)$.

If $v \in \mathrm{R}_\infty(t)$, then $\exists\, w$ such that, $v = t^n(w)$.

Then $t(v) = t^{n+1}(w) = t^n(t(w)) \in R_\infty(t)$.

Hence, $N_\infty(t - \lambda_i)$ and $R_\infty(t - \lambda_i)$ are both $t - \lambda_i$ invariant. By definition, $t - \lambda_i$ is nilpotent on $N_\infty(t - \lambda_i)$.

Lemma 3: A subspace M is t invariant iff it is $t - \lambda$ invariant for any scalar λ. In particular, where λ_i is an eigenvalue of a linear transformation t, then for any other eigenvalue λ_j, the spaces $N_\infty(t - \lambda_i)$ and $R_\infty(t - \lambda_i)$ are both $t - \lambda_j$ invariant.

Proof: If M is $t - \lambda$ invariant for any scalar λ, then setting $\lambda = 0$ means W is t invariant.

If M is t invariant, then $m \in M \;\Rightarrow\; t(m) \in M$.

Since M is a subspace, it is closed under all linear combinations of its members. Hence, $t(m)\,\lambda m \in M$, i.e., $m \in M \;\Rightarrow\; (t - \lambda)(m) \in M$.

Now since $N_\infty(t - \lambda_i)$ and $R_\infty(t - \lambda_i)$ are $t - \lambda_i$ invariant, they are t invariant and hence also t-1_j invariant.

Lemma 4: Given $t : V \to V$ and let N and R be t invariant complementary subspaces of V. Then t can be represented by a matrix with blocks of square submatrices T_1 and T_2 :

$$\left(\begin{array}{c|c} T_1 & O \\ \hline O & T_2 \end{array}\right) \begin{array}{l} (\dim N) \text{ rows} \\ (\dim R) \text{ rows} \end{array}$$

NOTES

Proof: Let the bases of N and R be $B_N = \langle v_1,....,v_p \rangle$ and $B_R = \langle \rho_1,....,\rho_q \rangle$, respectively. N and R are complementary

$\Rightarrow B = \langle v_1,....,v_p,\rho_1,....,\rho_q \rangle$ is a basis for V.

Then $t_{B\to B} = \left(t(v_1)_B \mid .... \mid t(\rho_q)_B\right)$ has the desired form.

Lemma 5: If T is a matrix with submatrices T_1 and T_2 such that, $T = \left(\begin{array}{c|c} T_1 & O \\ \hline O & T_2 \end{array}\right)$

Then $|T| = |T_1||T_2|$

Proof: Let the dimensions of T, T_1 and T_2 be $n\times n$, $r \times r$ and $(n–r) \times (n–r)$ respectively.

Then

$$\sum_P (-)^P t_{1P(1)}....t_{rP(r)}t_{r+1P(r+1)}....t_{nP(n)}$$
$$= \sum_{P_1}(-)^{P_1} t_{1P_1(1)}....t_{rP_1(r)} \sum_{P_2}(-)^{P_2} t_{r+1P_2(r+1)}....t_{nP_2(n)}$$
$$= | T_1 \| T_2 |$$

Example 2.15: *Solve* $\begin{vmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{vmatrix}$

Solution: $\begin{vmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{vmatrix} = \begin{vmatrix} 2 & 0 \\ 1 & 2 \end{vmatrix}\begin{vmatrix} 3 & 0 \\ 0 & 3 \end{vmatrix} = 36$

Lemma 6: If a linear transformation $t : V \to V$ has the characteristic polynomial

$c(x) = (x-\lambda_1)^{p_1}....(x-\lambda_k)^{p_k}$

Then,

$V = N_\infty(t-\lambda_1) \oplus .... \oplus N_\infty(t-\lambda_k)$ and dim $N_\infty(t-\lambda_i) = p_i$

Proof: Since $\dim(V) - p_1 + ... + p_k$

Therefore $V = N_\infty(t-\lambda_1) \oplus .... \oplus N_\infty(t-\lambda_k)$ if dim $N_\infty(t-\lambda_i) = p_i$ and

$N_\infty(t-\lambda_i) \cap N_\infty(t-\lambda_j) = \{0\} \quad \forall i \neq j$

According to Lemma 3, $N_\infty(t-\lambda_i)j$ and $N_\infty(t-\lambda_j)j$ are t invariant.

Since the intersect of t invariant subspaces is t invariant, the restriction of t to $M = N_\infty(t-\lambda_i) \cap N_\infty(t-\lambda_j)$ is a linear transformation.

Now both $t–\lambda_i$ and $t–\lambda_j$ are nilpotent on M.

NOTES

Therefore the only eigenvalue λ of t on M must satisfy $\lambda = \lambda_i$ and $\lambda = \lambda_j$.

To prove dim $N_\infty(t-\lambda_i) = p_i$ fix the index I and write

$V = N_\infty(t-\lambda_i) \oplus P_\infty(t-\lambda_i)$

From Lemma 4 $T = \left(\begin{array}{c|c} T_1 & O \\ \hline O & T_2 \end{array}\right)$ and from Lemma 5 $|T-xI| = |T_1-xI||T_2-xI|$

From the uniqueness clause of the fundamental theorem of arithmetic,

$$|T_1 - xI| = (x-\lambda_1)^{q_1} \ldots\ldots (x-\lambda_k)^{q_k}$$

And

$$|T_2 - xI| = (x-\lambda_1)^{r1} \ldots\ldots (x-\lambda_k)^{r_k}$$

Then, $q_j + r_j = p_j j = 1, \ldots, k$

i.e., the restriction of $t-\lambda i$ M to is $N_\infty(t-\lambda_i)$ nilpotent on M.

The only eigen value of t on M is hence λi.

Hence $c(x) = (x-\lambda_i)$ on M, i.e., $q_j = 0 \; \forall \; j \neq i$

Consider next the restriction of $t-\lambda_i$ to $R = R_\infty(t-\lambda_i)$.

Since $t-\lambda_i$ is nonsingular on R, λ_i is not an eigenvalue of R.

Hence, $q_i = p_i$.

Check Your Progress

1. Define linear transformation.
2. Write the hypercompanion matrix.
3. What do you mean by rational cononical form?
4. What is nilpotent element?

2.5 PRIMARY DECOMPOSITION THEOREM

Theorem 2.16: *Let T be a liner operator on a finite dimensional space V over F. Let p(x) be the minimal polynomial for T such that,*

$$p(x) = p_1(x)^{r1} \ldots\ldots p_k(x)^{rk}$$

where the $p_i(x)$ are distinct irreducible monic polynomials over F and r_i are +ve integers. Let W_i be the null spaces of $p_i(T)^{ri}$, i = 1, . . ., k. Then

(*i*) $V = W_1 \oplus \ldots . \oplus W_k$

(*ii*) *Each W_i is invariant under T (i.e., $T(W_i) \subseteq W_i \; \forall i$)*

(*iii*) *If T_i is operator induced on W_i by T, then the minimal polynomial $q_i(x)$ for T_i is $p_i(x)^{ri}$.*

Proof: Let $f_i(x) = \dfrac{p(x)}{p_i(x)^{r_i}}$, $i = 1, 2, ..., k$

Then g.c.d. $(f_1(x), ..., f_k(x)) = 1$

$\therefore \exists\ g_1(x), ..., g_k(x) \in F[x]$ such that,

$$g_1(x) f_1(x) + ... + g_k(x) f_k(x) = 1$$

$$\Rightarrow \quad g_1(T) f_1(T) + ... + g_k(T) f_k(T) = I$$

Let, $v \in V$, then

$$v = g_1(T) f_1(T)(v) + ... + g_k(T) f_k(T)(v)$$

Now $p_i(T)^{r_i} f_i(T)\, g_i(T) = p(T)\, g_i(T) = 0$

$\therefore$ $g_i(T) f_i(T)(v) = f_i(T) g_i(T)(v)$

$\Rightarrow$ $p_i(T)^{r_i} g_i(T) f_i(T)(v) = 0$

$\Rightarrow$ $g_i(T) f_i(T)(v) \in \text{Ker}\ p_i(T)^{r_i} = W_i$

$\Rightarrow$ $v \in W_1 + ... + W_k$

$\Rightarrow$ $V = W_1 + ... + W_k$

or that $V = W_1 \oplus ... \oplus W_k$

For let $x_1 + ... + x_k = 0, \quad x_i \in W_i$

then $x_1 = -(x_2 + ... + x_k)$

$\Rightarrow f_1(T)\, x_1 = 0$ as $\forall\ i \neq 1, f_1(T)\, x_i = 0$

Now g.c.d. $(f_1(x), p_1(x)^{r_1}) = 1$

So $\exists\ q_1(x), r_1(x) \in F[x]$ such that,

$$f_1(x)\, q_1(x) + p_1(x)^{r_1}\, r_1(x) = 1$$

$$\Rightarrow I = q_1(T) f_1(T) + r_1(T)\, p_1(T)^{r_1}$$

$$\Rightarrow x_1 = 0$$

Similarly $x_i = 0\ \forall\ i$

This proves (*i*).

Let $x_i \in W_i = \text{Ker}\ p_i(T)^{r_i}$

Then $p_i(T)^{r_i}(x_i) = 0$

$\Rightarrow\ Tp_i(T)^{r_i}(x_i) = 0$

$\Rightarrow\ p_i(T)^{r_i}(T(x_i)) = 0$

$\Rightarrow\ T(x_i) \in W_i\ \forall\ i$

$\Rightarrow\ W_i$ is T-invariant $\forall\ i$

which proves (*ii*).

Again, since $p_i(T)^{r_i}(x_i) = 0 \quad \forall\ x_i \in W_i$

$$p_i(T)^{r_i} = 0 \text{ on } W_i$$

NOTES

NOTES

$\Rightarrow \quad p_i(T_i)^{r_i} = 0$ as T restricted to W_i is T_i.

$\Rightarrow q_i(x) \mid p_i(x)^{r_i} \Rightarrow q_i(x) = p_i(x)^{s_i}$, $s_i \leq r_i$

Let $f(x) = f_i(x)p_i(x)^{si}$

and let $v \in V$ then $v = w_1 + ... + w_k$, $w_i \in W_i$

$\therefore \quad f(T)(v) = f_i(T)\, p_i(T)^{si} w_i$

$\because \quad f_i(T)\, w_j = 0 \quad \forall\, j \neq i$

$\therefore \quad f(T)(v) = f_i(T) q_i(T) w_i$

$= f_i(T)\, q_i(T_i)\, w_i$

$= 0$ as $q_i(T_i) = 0$

$\Rightarrow \quad f(T) = 0$

$\therefore \quad p(x) \mid f(x)$

$\Rightarrow p_i(x)^{r_i} \mid p_i(x)^{s_i}$

$\Rightarrow \quad r_i \leq s_i$

$\therefore \quad r_i = s_i$

So, $q_i(x) = p_i(x)^{r_i}$

which proves (*iii*).

Corollary: If T is a linear operator on a finite dimensional space V over F and minimal polynomial $p(x)$ of T is a product of distinct linear factors, then T is diagonalisable.

Proof: Let $p(x) = (x - c_1) ... (x - c_r)$, where c_i are distinct roots of $p(x)$ in F. By primary decomposition theorem

$V = W_1 \oplus ... \oplus W_r$, where each W_i = Null space of $(T - c_i I)$

$\therefore \quad v \in W_i \Rightarrow (T - c_i I)v = 0$

$\Rightarrow T(v) = c_i v$

$\therefore$ every non-zero vector in W_i is an eigen vector of T corresponding to eigen value c_i of T. If β_i is a basis of W_i, then $\{\beta_1, ..., \beta_r\}$ is a basis of V. β_i consists of eigen vectors of $T \Rightarrow \{\beta_1, ..., \beta_r\} = \beta$ consists of eigen vectors of T and is a basis of $V \Rightarrow T$ is a diagonalisable.

Example 2.16: *Let T and S be linear operators on V(F), each having all its eigen values in F such that, TS = ST.*

Show that they have a common eigen vector.

Solution: Let c be an eigen value of T. Let $W_c = \{v \in V \mid T(v) = cv\}$ be the eigen space with respect to eigen value c.

Let $v \in W_c$.

Then $T(S(v)) = (TS)(v)$

$= (ST)(v)$

NOTES

$= S(T(v))$

$= S(cv)$

$= cS(v)$

$\therefore S(v) \in W_c \ \forall\ v \in W_c \Rightarrow S : W_c \to W_c$.

$\therefore S$ is a linear operator on W_c.

Let $\alpha \in F$ be an eigen value of S as linear operator on W_c.

$\therefore \exists\ w \in W_c$ such that,

$S(w) = \alpha w,\ w \neq 0$

$w \in W_c \Rightarrow T(w) = cw$

$\therefore w$ is a common eigen vector of T and S.

Example 2.17: *Let N be 2×2 complex matrix such that $N^2 = 0$. Prove that either $N = 0$ or N is similar over* **C** *to* $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$.

Solution: Let $T : V \to V$ be a linear operator such that,

$[T]_\beta = N,\ \beta = \{v_1, v_2\}$ is a basis of V.

Now, $0 = N^2 = N \,.\, N = [T]_\beta\ [T]_\beta = [T^2]_\beta$

$\Rightarrow \quad T^2 = 0.$

Suppose $N \neq 0$, *i.e.*, $T \neq 0$.

Let λ be an eigen value of T.

Then there exists $0 \neq v \in V$ such that,

$T(v) = \lambda v$

$\Rightarrow \quad T^2(v) = \lambda(T(v)) = \lambda^2 v$

$\Rightarrow \quad 0 = \lambda^2 v$

$\Rightarrow \quad \lambda^2 = 0$ as $v \neq 0$

$\Rightarrow \quad \lambda = 0$

$\Rightarrow$ 0 is the only eigen value of T.

Let W_o be the eigen space of T with respect to eigen value 0.

Then $W_o = \{x \in V \mid T(x) = 0\} = \text{Ker } T$

Since $0 \neq v \in W_o,\ W_o \neq \{0\}$

So, dim $W_o = 1$ or 2.

If dim $W_o = 2$, then dim $W_o = \dim V$

$\Rightarrow W_o = V \Rightarrow \text{Ker } T = V \Rightarrow T = 0$, which is not true.

Therefore, dim $W_o = 1$.

Let $W_o = < w_2 >$

NOTES

There exists a subspace W' of V such that,

$$V = W' \oplus W_o$$

Since dim $V = 2$, dim $W_o = 1$, dim $W' = 1$.

Let $W' = < w_1 >$

Then $\{w_1, w_2\}$ is a basis of V.

Let $T(w_1) = \alpha_1 w_1 + \alpha_2 w_2$

$T(w_2) = 0w_1 + 0w_2$ as $w_2 \in \text{Ker } T$.

But $T^2 = 0 \Rightarrow 0 = T^2(w_1)$

$= \alpha_1 T(w_1) + \alpha_2 T(w_2)$

$= \alpha_1 T(w_1)$

$= \alpha_1(\alpha_1 w_1 + \alpha_2 w_2)$

$= \alpha_1^2 w_1 + \alpha_1 \alpha_2 w_2$

$\Rightarrow \alpha_1 = 0$ ($\alpha_2 \neq 0$ as $\alpha_2 = 0$ $\Rightarrow w_1 \in \text{Ker } T$

$\Rightarrow w_1 \in W_o \cap W' = \{0\}$

$\Rightarrow w_1 = 0$ which is not true).

So, $T(w_1) = \alpha_2 w_2$.

Now $\{\alpha_2^{-1} w_1, w_2\} = \beta'$ is also a basis of V

as $a\alpha_2^{-1} w_1 + bw_2 = 0$

$\Rightarrow a\alpha_2^{-1} = 0, b = 0$

$\Rightarrow a = 0 = b$.

$\Rightarrow \{\alpha_2^{-1} w_1, w_2\} = \beta'$ is a *L.I.* set

$\Rightarrow \beta'$ is a basis of V as dim $V = 2$

Therefore, $T(\alpha_2^{-1} w_1) = \alpha_2^{-1} T(w_1)$

$= \alpha_2^{-1} \alpha_2 w_2 = w_2$

$= 0\alpha_2^{-1} w_1 + 1w_2$

$\Rightarrow [T]_{\beta'} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$

Also $[T]_\beta = N$

$\Rightarrow N$ is similar to $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ over **C**.

Example 2.18: *Show that if A is a* 2×2 *matrix over* **C** *then A is similar to a matrix of the type* $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ *or* $\begin{bmatrix} a & 0 \\ 1 & a \end{bmatrix}$ *over* **C**.

NOTES

Solution: Let $f(x)$ be the characteristic polynomial of A. If the roots of $f(x)$ are distinct, then A is diagonalisable.

So, $A = P^{-1}BP$, $B = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ $a, b \in \mathbf{C}$.

$\Rightarrow A$ is similar over $\mathbf{C}$ to $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$

If the roots of $f(x)$ are same, let $f(x) = (x - \alpha)^2$

Then $0 = f(A) = (A - \alpha I)^2$

Let $N = A - \alpha I$

By above problem either $N = 0$ or N is similar over $\mathbf{C}$ to $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$.

If $N = 0$, then $A = \alpha I = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix}$

$\Rightarrow A$ is similar over $\mathbf{C}$ to $\begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix}$

If $N = Q^{-1}\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}Q$

Then $A - \alpha I = Q^{-1}\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}Q$

$\Rightarrow A = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix} + Q^{-1}\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}Q$

$= Q^{-1}\left\{\begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right\}Q$

$= Q^{-1}\begin{bmatrix} \alpha & 0 \\ 1 & \alpha \end{bmatrix}Q$

$\Rightarrow A$ is similar over $\mathbf{C}$ to the matrix of the type $\begin{bmatrix} \alpha & 0 \\ 1 & \alpha \end{bmatrix}$.

Example 2.19: *Give an example to show that AB is diagonalisable and BA is not diagonalisable, where A and B are n × n matrices over F.*

Solution: Let $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

Then $AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

So, AB is a diagonal matrix. AB is a diagonalisable matrix.

Now $BA = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $(BA)^2 = 0$

$\Rightarrow$ minimal polynomial of BA is x^2.

So, the minimal polynomial of BA is not product of distinct linear factors.

$\therefore$ BA is not diagonalisable.

NOTES

Example 2.20: *If T is an idempotent linear operator (i.e., $T^2 = T$) then show that 0 or 1 are only eigen values of T and T is diagonalisable.*

Solution: Let $f(x) = x\ (x - 1) = x^2 - x$

then $f(T) = T^2 - T = 0$

If $p(x)$ is the minimal polynomial of T, then $p(x) \mid f(x)$.

$p(x) = x$ or $x - 1$ or $x(x - 1)$

The eigen values of T are the roots of the minimal polynomial of T.

$\therefore$ 0 or 1 are only eigen values of T.

In each case $p(x) = x$ or $x - 1$ or $x(x - 1)$,

$p(x)$ is product of distinct linear factors. So, T is diagonalisable.

Example 2.21: *Give an example of a linear operator T having eigen values 0 and 1 but T is not idempotent.*

Solution: Let T be a linear operator on V where dim $V = 3$ such that matrix of T with respect to a basis of V is

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

Then eigen values of A (or T) are entries on the diagonal as A is a triangular matrix.

$\therefore$ eigen values of T are 0, 1, 1.

But $$A^2 = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

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

$\therefore$ A is not idempotent.

So, T is not idempotent.

2.6 JORDAN BLOCKS AND JORDAN FORMS

A Jordan block is a matrix with zeros everywhere except along the diagonal and superdiagonal, with each element of the diagonal consisting of a single number λ, and each element of the superdiagonal consisting of a 1. For example,

NOTES

$$\begin{bmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \ddots & 0 & 0 \\ 0 & 0 & \lambda & \ddots & 0 & 0 \\ 0 & 0 & 0 & \ddots & 1 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}$$

The degenerate case of a 1 ×1 matrix is considered a Jordan block even though it lacks a superdiagonal to be filled with 1s. Any Jordan block is thus specified by its dimension n and its eigenvalue λ and is indicated as $J_{\lambda,n}$.

For an arbitrary square matrix A over an algebraically closed field k there always exists a square non-singular matrix C over k such that $C^{-1}AC$ is a Jordan matrix or A is similar over k to a Jordan matrix. This assertion is valid under weaker restrictions on k. For a matrix A to be similar to a Jordan matrix it is necessary and sufficient that k contains all roots of the minimum polynomial of A. The matrix $C^{-1}AC$ mentioned above is called a Jordan form or Jordan normal form of the matrix A.

The Jordan form of a matrix is determined only up to the order of the Jordan blocks. More exactly, two Jordan matrices are similar over k if and only if they consist of the same Jordan blocks and differ only in the distribution of the blocks along the main diagonal. The number $C_m(\lambda)$ of Jordan blocks of order m with eigen value λ in a Jordan form of a matrix A is given by the formula

$$C_m(\lambda) = rk\,(A - \lambda E)^{m-1} - 2\,rk\,(A - \lambda E)^m + rk\,(A - \lambda E)^{m+1},$$

where E is the unit matrix of the same order n as A, rk B is the rank of the matrix B, and $rk\,(A - \lambda E)^0$ is n, by definition.

Theorem 2.17: There exists a basis of V such that the matrix of T is in block-diagonal form with Jordan blocks. If a is an eigenvalue of T, t the sequence $(t_0, t_1,\ldots,t_n,\ldots)$ with $t_i = \dim ker(T - a)^i$ and $(s_0, s_1,\ldots, s_n,\ldots) = -R(L - 1)^2(t)$, where R and L are the left and right-shift operators on R^∞, then s_i is the number of Jordan blocks of size i with eigenvalue a.

As an illustration, let T be the linear operator on F_2^6 whose matrix with respect the standard basis of F_2^6 is

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

NOTES

We have,

$$A-1=\begin{bmatrix}0&1&1&1&0&0\\0&0&1&1&0&1\\0&0&0&0&1&1\\0&0&0&0&1&1\\0&0&0&0&0&0\\0&0&0&0&0&0\end{bmatrix},\quad (A-1)^2=\begin{bmatrix}0&0&1&1&0&1\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\end{bmatrix},\ (A-1)^3=0.$$

It follows that $(T-1)^3=0$ and 1 is the only eigenvalue of T. We also have rank$(T-1)=3$, rank$(T-1)^2=1$, rank$(T-1)^3=0$ so that $t_1=\dim\ker(T-1)=3$, $t_2=\dim\ker(T-1)^2=5$, $t_3=\dim\ker(T-1)^3=6$. Hence t is the sequence $(0, 3, 5, 6, 6, ..., 6, ...)$.

Now $(L-1)(t)=(3, 2, 1, 0, 0, ..., 0, ...)$, $(L-1)^2(t)=(-1,-1,-1, 0, 0, ..., 0, ...)$ and so $-R(L-1)^2(t)=(0, 1, 1, 1, 0, 0, ..., 0, ...)$ which, according to the above Theorem , implies that there is one Jordan block of size 1, one of size 2 and one of size 3. Hence there is a basis of F_2^6 such that the matrix of T with respect to this basis is

$$\begin{bmatrix}1&0&0&0&0&0\\0&1&1&0&0&0\\0&0&1&0&0&0\\0&0&0&1&1&0\\0&0&0&0&1&1\\0&0&0&0&0&1\end{bmatrix}$$

If W is a T-invariant subspace of V and $f=(f_1, f_2, ..., f_n)$ is a basis of W, the matrix (with respect to this basis) of the restriction of T to W is the Jordan matrix $J_n(a)$ iff $T(f_1)=af_1$, $T(f_2)=af_2+f_1$, ..., $T(f_i)=af_i+f_{i-1}$, ..., $T(f_n)=af_n+f_{n-1}$ or, equivalently, $(T-a)(f_1)=0$, $(T-a)(f_2)=f_1$, ..., $(T-a)(f_i)=f_{i-1}$, ..., $(T-a)(f_n)=f_{n-1}$.

For such a basis we have $f_i=(T-a)^{n-i}(f_n)$ with $f_n\in Ker((T-a)^n)-Ker((T-a)^{n-1})$. Conversely, if $g\in Ker((T-a)^n)-Ker((T-a)^{n-1})$ the sequence $g, (T-a)(g), (T-a)^2(g), ..., (T-a)^{n-1}(g)$ is a basis for a T-invariant subspace of V such that the matrix of this mapping with respect to the basis $f_1=(T-a)^{n-1}(g)$, $f_2=(T-a)^{n-2}(g)$, ..., $(T-a)(g)$, g is the Jordan matrix $Jn(a)$. The vector g is called a cyclic vector of cycle length n for the eigenvalue a. Each Jordan block corresponds to a cyclic vector. The subspace generated by a cyclic vector g and its images under the powers of T is called the cyclic subspace generated by g.

We now illustrate how to find cyclic vectors that give decomposition into a direct sum of cyclic subspaces in the case of the above illustration. We first find bases for $ker(T-1)$, $ker(T-1)^2$, $ker(T-1)^3$.

$Ker(T-1) = \text{Span}(e_1, e_3 + e_4, e_2 + e_3 + e_5 + e_6)$,

$Ker((T-1)^2) = \text{Span}(e_1, e_2, e_5, e_3 + e_4, e_3 + e_6)$,

$Ker((T-1)^3) = \text{Span}(e_1, e_2, e_3, e_4, e_5, e_6)$.

The next step is to complete the basis of $Ker(T-1)^2$ to a basis of $(T-1)^3$. We find that $g_1 = e_6$ completes the given basis of $Ker((T-1)^2)$ to a basis of $Ker((T-1)^3)$. Now $(T-1)(e_6) = e_2+e_3+e_4$ is in the kernel of $(T-1)^2$ but not in the kernel of $T-1$. Thus $e_1, e_3 + e_4, e_2 + e_3 + e_5 + e_6, e_2 + e_3 + e_4$ is linearly independent and we can complete this sequence to a basis of $Ker((T-1)^2)$ with the vector $g_2 = e_5$. Now $(T-1)^2(g_1) = e_1, (T-1)(g_2) = e_3 + e_4$ are in the kernel of $T-1$ and are linearly independent. We complete these two vectors to a basis of $\ker(T-1)$ by means of the vector

$g_3 = e_2 + e_3 + e_5 + e_6$. Now, the sequence of vectors

$g_1 = e_6, (T-1)(g_1) = e_2 + e_3 + e_4, (T-1)^2(g_1) = e_1,$

$g_2 = e_5, (T-1)(g_2) = e_3 + e_4, g_3 = e_2 + e_3 + e_5 + e_6$

is linearly independent and the basis

$f_1 = g_3, f_2 = (T-1)(g_2), f_3 = g_2, f_4 = (T-1)^2(g_1), f_5 = (T-1)(g_1), f_6 = g_1$

yields the above Jordan canonical form for T. If $v_1, v_2, ..., v_n \in V$ and W is a subspace of V, we say that the sequence $v_1, v_2, ..., v_n$ is linearly independent mod W if

$a_1v_1 + a_2v_2 + ...+ a_nv_n \in W \Rightarrow a_1 = a_2 = ... = a_n = 0.$

This is equivalent to saying that the images of the vectors v_i in the quotient space $V=W$ form a linearly independent sequence. Similarly, we say that $v_1, v_2, ..., v_n$ generate V mod W if every $v \in V$ can be written in the form $v = a_1v_1 + a_2v_2 + ... + a_nv_n$ with $w \in W$. This is equivalent to saying that the images of the vectors v_i in $V=W$ span $V=W$.

Lemma 7: If $Ker((T-a)^i) = Ker((T-a)^{i+1})$ then $Ker((T-a)^{i+1}) = Ker((T-a)^{i+2})$.

Proof: Let $v \in Ker((T-a)^{i+2})$. Then $(T-a)(v) \in Ker((T-a)^{i+1}) = Ker((T-a)^i)$ which implies that $(T-a)^{i+1}(v) = (T-a)^i(T-a)(v) = 0$ and hence that $v \in Ker((T-a)^{i+1})$.

This lemma shows that, for an eigenvalue a of T, there is an integer $p > 0$ such that

$0 = t_0 < t_1 < ..... < t_p = t_{p+1} = t_{p+2} = ……,$

where $t_i = \dim(T-a)^i$.

Lemma 8: If $i \geq 2$ and $v \in Ker((T-a)^i) - Ker((T-a)^{i-1})$ then

$(T-a)(v) \in Ker((T-a)^{i-1}) - ker((T-a)^{i-2})$.

Proof: If $v \in Ker((T-a)^i)$ and $(T-a)(v) \in Ker((\text{T} - \text{a})^{i-2})$ then

$(T-a)^{i-1}(v) = (T-a)^{i-2}(T-a)(v) = 0$

which implies that $v \in Ker((T-a)^{i-1})$.

NOTES

This Lemma is simply the assertion that the linear mapping

$S_{i-1} : Ker((T-a)^i)=Ker((T-a)^{i-1}) \rightarrow Ker((T-a)^{i-1})=Ker((T-a)^{i-2})$

defined by $S_{i-1}(v+Ker((T-a)^{i-1}) = (T-a)(v)+Ker((T-a)^{i-1})$ is injective. This yields the following result.

Lemma 9: If $i \geq 2$ and $v_1, v_2, ..., v_n \in Ker(T-a)i$ is linearly independent mod $Ker((T-a)^{i-1})$ then

$(T-a)(v_1), (T-a)(v_2), ..., (T-a)(v_n) \in Ker((T-a)^{i-1})$

is a linearly independent sequence mod $Ker((T-a)^{i-2})$.

If $r = (r_0, r_1, ..., r_i, ...) = (L-1)(t)$ then

$r_i = \dim(Ker(T-a)_{i+1}) - \dim(Ker(T-a)^i) = \dim(Ker((T-a)^{i+1})=Ker((T-a)^i)$.

Lemma 7 shows that r is a decreasing sequence of natural numbers which are zero for $i > p$, i.e.,

$r_0 \geq r_1 \geq r_2 \geq ..... \geq r_p = r_{p+1} = r_{p+2} = ... = 0.$

The above Theorem states that the number of Jordan blocks of size $i \geq 1$ is

$-(r_i - r_i - 1) = r_i - 1 - r_i = \dim(Ker((T-a)^i)=Ker((T-a)^{i-1})-\dim(Ker((T-a)^{i+1})=Ker((T-a)^i))$

Following is the proof of the above stated Theorem:

Proof: Without loss of generality, we can assume that the minimal polynomial of T is

$$(\lambda - a_1)^{k_1}\left(\lambda - a_2\right)^{k_2} ...(\lambda - a_m)^{k_m} = 0$$

By the primary decomposition theorem, V is a direct sum of the subspaces $V(a_i) = \text{Ker}((T-a_i)^{k_i})$ with $\{a_1, \ldots, a_m\}$ being the set of eigenvalues of T. The integer k_i is the smallest integer > 0 such that, $\text{Ker}((T-a_i)^{k_i}) = \text{Ker}((T-a_i)^{k_i+1})$ and so $V(a_i) = \bigcup_{j \geq 0} \ker(T-a_i)^j$,

This subspace is called the generalized eigenspace for the eigenvalue a_i.

Let a be any eigenvalue of T. If $t_i = \dim ker(T-a)^i$, then we have

$0 = t_0 < t_1 < \ldots < t_p = t_{p+1}$

for a unique $p \geq 1$.

Given below is an algorithm for decomposing $V(a)$ into a direct sum of cyclic Subspaces:

Step 1: Find a basis for $Ker((T-a)^p)$ mod $Ker((T-a)^{p-1})$, i.e., find a sequence of vectors in $Ker((T-a)^p)$ which complete some basis of $ker(T-a)^{p-1}$ to a basis of $Ker((T-a)^p)$.

Step 2: If $p = 1$ stop, if $p > 1$ take the image, under $T-a$, of the basis of $Ker((T-a)^p)$ mod $Ker((T-a)^{p-1})$ obtained in the previous step and complete it to a basis of $Ker((T-a)^{p-1})$ mod $Ker((T-a)^{p-2})$.

Step 3: Repeat Step 2 with p replaced by $p - 1$.

The vectors obtained in this way are a basis of $V(a)$ and the vectors which, for each $i \geq 1$ complete to a basis of $Ker((T-a)^i)$ mod $Ker((T-a)^{i-1})$ the image of the basis of $Ker((T-a)^{i+1})$ mod $Ker((T-a)^i)$ obtained in the previous step, are cyclic vectors of cycle length i. The number of these cyclic vectors is dim($ker((T-a)^i)$=$Ker((T-a)^{i-1})$) – dim($Ker((T-a)^{i+1})$=$Ker((T-a)^i)$).

Moreover, V is the direct sum of the cyclic subspaces generated by the cyclic vectors so obtained.

Corollary 1: Let V be a finite-dimensional vector space over a field K and let T be a linear operator on V whose minimal polynomial is a product of linear factors. If dim(V) = n, there are T-invariant subspaces

$$\{0\} = V_0 \subset V_1 \subset V_2 \subset \ldots\ldots \subset V_n = V \text{ with } \dim(V_i) = i.$$

Corollary 2: If A is an $n \times n$ matrix over a field K whose minimal polynomial is a product of linear factors then there is an invertible matrix $P \in K^{n \times n}$ such that $P^{-1}AP$ is upper triangular.

Corollary 3: (Cayley-Hamilton) If $\Delta_A(\lambda)$ is the characteristic polynomial of the matrix $A \in C^{n \times n}$ then $\Delta_A(A) = 0$.

Corollary 3 is true for a matrix A over any field K since it is possible to find a field F, containing K as a subfield, such that the minimal polynomial of A is a product of linear factors $\lambda - c$ with $c \in F$.

Let $(\lambda - a_1)^{n_1}(\lambda - a_2)^{n_2} \ldots (\lambda - x_\ell)^{n_\ell}$ be the characteristic polynomial of a linear operator T on a finite-dimensional vector space V with $a_1, a_2, \ldots, a_l$ distinct. The integer n_i is called the algebraic multiplicity of the eigenvalue a_i. n_i is the dimension of the generalized eigenspace $V(a_i)$ for the eigenvalue a_i. The dimension of the eigenspace $Ker((T-a_i))$ is called the geometric multiplicity of the eigenvalue a_i. Thus T is diagonalizable if and only if the geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity.

Example 2.22: *If $A \in C^{5 \times 5}$ with characteristic polynomial $\Delta(\lambda) = (\lambda - 1)^2 (\lambda - 2)^3$ and minimal polynomial $m(\lambda) = (\lambda - 1)(\lambda - 2)^2$, what is the Jordan form for A?*

Solution: The generalized eigenspace for the eigenvalue 2 has dimension 3 and there is a cyclic vector of cycle length 2. It follows that there is one Jordan block of size 1 and one of size 2. On the other hand the cyclic vectors for the eigenvalue 1 have cycle length 1 and so there must be 2 Jordan blocks of size 1 for the eigenvalue 1 since the generalized eigenspace for this eigenvalue has dimension 2. The Jordan form (up to order of the blocks) is therefore

$$\begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

NOTES

Example 2.23: *Find the possible Jordan normal forms for a complex 6 × 6 matrix with minimal polynomial λ^3. Show that two such matrices having the same nullity are similar.*

Solution: The only eigenvalue is 0 and there must be one Jordan block of size 3. It follows that there must be either (i) 2 Jordan blocks of size 3 or (ii) 1 of size 3, one of size 2 and one of size 1 or (iii) one of size 3 and 3 of size 1. The corresponding possible Jordan forms for A are:

$$(i)\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (ii)\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (iii)\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Since the nullity of A is respectively 2, 3, 4 in cases (i), (ii), (iii), we get that two such matrices with the same nullity are similar.

Example 2.24: *If N is an $n \times n$ matrix with $n \geq 2$, $N^n = 0$, $N^{n-1} \neq 0$, show that there is no complex $n \times n$ matrix A with $A^2 = N$.*

Solution: Suppose that $A^2 = N$ for some A. Then $A^{2n} = N^n = 0$ and so the characteristic polynomial of A must be λ^n. Hence $A^n = 0$ which implies $N^{n-1} = A^{2n-2} = 0$ since $2n - 2 \geq n$.

This contradicts the assumption that $N^{n-1} \neq 0$.

Check Your Progress

5. When is a linear operator on a finite dimensional space V over F diagonalizable?
6. What is Jordan block?

2.7 CYCLIC MODULES

A group we noticed is a system with a non-empty set and a binary composition. One can of course talk about non-empty sets with two binary compositions also, the set of integers under usual addition and multiplication being an example. Though this set forms a group under addition and not under multiplication, it does have certain specific properties satisfied with respect to multiplication as well. We single out some of these and generalize the concept in the form of a ring. We start with the formal definition.

Definition 1: A non-empty set R, together with two binary compositions + and . is said to form a *Ring* if the following axioms are satisfied:

(*i*) $a + (b + c) = (a + b) + c$ for all $a, b, c \in R$

(*ii*) $a + b = b + a$ for $a, b \in R$

(*iii*) $\exists$ some element 0 (called zero) in R, such that, $a + 0 = 0 + a = a$ for all $a \in R$

(*iv*) For each $a \in R$, $\exists$ an element $(-a) \in R$, such that, $a + (-a) = (-a) + a = 0$

(*v*) $a \,.\, (b \,.\, c) = (a \,.\, b) \,.\, c$ for all $a, b, c \in R$

(*vi*) $a \,.\, (b + c) = a \,.\, b + a \,.\, c$

$(b + c) \,.\, a = b \,.\, a + c \,.\, a$ for all $a, b, c \in R$

Notes: 1. Since we say that + and . are binary compositions on R, it is understood that the closure properties with respect to these hold in R. In other words, for all $a, b \in R$, $a + b$ and $a \,.\, b$ are unique in R.

2. One can use any other symbol instead of + and ., but for obvious reasons, we use these two symbols (the properties look so natural with these). In fact, in future, the statement that R is a ring would mean that R has two binary compositions + and . defined on it and satisfies the above axioms.
3. Axiom (*v*) is named associativity with respect to . and axiom (*vi*) is referred to as distributivity (left and right) with respect to . and +.
4. Axioms (*i*) to (*iv*) could be restated by simply saying that $< R, + >$ forms an abelian group.
5. Since 0 in axiom (*iii*) is identity with respect to +, it is clear that this element is unique (see groups).

Definitions 2: A ring R is called a *commutative* ring if $ab = ba$ for all $a, b \in R$. Again if $\exists$ an element $e \in R$ such that,

$$ae = ea = a \quad \text{for all } a \in R$$

we say, R is a ring with *unity*. Unity is generally denoted by 1. (It is also called unit element or multiplicative identity).

It would be easy to see that if unity exists in a ring then it must be unique.

Note: We recall that in a group by a^2 we meant $a \,.\, a$ where '.' was the binary composition of the group. We continue with the same notation in rings as well. In fact, we also introduce similar notation for addition, and shall write na to mean $a + a + ...+ a$ (n times), n being an integer.

Case 8: Sets of real numbers, rational numbers, integers form rings with respect to usual addition and multiplication. These are all commutative rings with unity.

Case 9: Set **E** of all even integers forms a commutative ring, without unity (under usual addition and multiplication).

Case 10: (*a*) Let M be the set of all 2×2 matrices over integers under matrix addition and matrix multiplication. It is easy to see that M forms a ring with unity $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, but is not commutative.

NOTES

(*b*) Let *M* be set of all matrices of the type $\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}$ over integers under matrix addition and multiplication. Then *M* forms a non commutative ring without unity.

Case 11: The set $\mathbf{Z}_7 = \{0, 1, 2, 3, 4, 5, 6\}$ forms a ring under addition and multiplication modulo 7. (In fact, we could take *n* in place of 7).

Case 12: Let *F* be the set of all continuous functions $f: \mathbf{R} \to \mathbf{R}$, where **R** = set of real numbers. Then *F* forms a ring under addition and multiplication defined by:

for any $f, g \in F$

$$(f + g)x = f(x) \qquad \text{for all } x \in \mathbf{R}$$

$$(f\ g)x = f(x)g(x) \qquad \text{for all } x \in \mathbf{R}$$

zero of this ring is the mapping $O: \mathbf{R} \to \mathbf{R}$, such that,

$$O(x) = 0 \text{ for all } x \in \mathbf{R}$$

Also additive inverse of any $f \in F$ is the function $(-f): \mathbf{R} \to \mathbf{R}$ such that, $(-f)x = -f(x)$

In fact, *F* would have unity also, namely the function $i: \mathbf{R} \to \mathbf{R}$ defined by $i(x) = 1$ for all $x \in \mathbf{R}$.

Note: Although the same notation *fg* has been used for product here it should not be mixed up with *fog* defined earlier.

Case 13: Let **Z** be the set of integers, then $\mathbf{Z}[i] = \{a + ib \mid a, b \in \mathbf{Z}\}$ forms a ring under usual addition and multiplication of complex numbers. $a + ib$ where a, $b \in \mathbf{Z}$ is called a Gaussian integer and $\mathbf{Z}[i]$ is called the ring of Guassian integers.

We can similarly get $\mathbf{Z}_n[i]$ the ring of Gaussian integers modulo *n*. For instance,

$$\mathbf{Z}_3[i] = \{a + ib \mid a, b \in \mathbf{Z}_3 = \{0, 1, 2\} \bmod 3\}$$

$$= \{0, 1, 2, i, 1 + i, 2 + i, 2i, 1 + 2i, 2 + 2i\}$$

Case 14: Let *X* be a non-empty set. Then $\mathscr{P}(X)$ the power set of *X* (*i.e.*, set of all subsets of *X*) forms a ring under + and · defined by

$$A + B = (A \cup B) - (A \cap B)$$

$$A \,.\, B = A \cap B$$

In fact, this is a commutative ring with unity and also satisfies the property $A^2 = A$ for all $A \in \mathscr{P}(X)$.

Case 15: Let *M* = set of all 2 × 2 matrices over members from the ring of integers modulo 2. *It would be a finite non-commutative ring*. *M* would have $2^4 = 16$ members as each element *a*, *b*, *c*, *d* in matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ can be chosen in 2 ways. Compositions in *M* are given by

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} x & y \\ z & u \end{bmatrix} = \begin{bmatrix} a \oplus x & b \oplus y \\ c \oplus z & d \oplus u \end{bmatrix}$$

where $\oplus$ denotes addition modulo 2 and

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x & y \\ z & u \end{bmatrix} = \begin{bmatrix} a\otimes x \oplus b\otimes z & a\otimes y \oplus b\otimes u \\ c\otimes x \oplus d\otimes z & c\otimes y \oplus d\otimes u \end{bmatrix}$$

$\otimes$ being multiplication modulo 2.

NOTES

That M is non commutative follows as $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

But $\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

Case 16: Let $R = \{0, a, b, c\}$. Define + and . on R by

+	0	a	b	c
0	0	a	b	c
a	a	0	c	b
b	b	c	0	a
c	c	b	a	0

.	0	a	b	c
0	0	0	0	0
a	0	a	b	c
b	0	a	b	c
c	0	0	0	0

Then one can check that R forms a non commutative ring without unity. In fact it is an example of the smallest non commutative ring.

Theorem 2.18: *In a ring R, the following results hold*

(*i*) $a \,.\, 0 = 0 \,.\, a = 0$ *for all* $a \in R$

(*ii*) $a(-b) = (-a)b = -ab$ *for all* $a, b \in R$

(*iii*) $(-a)\,(-b) = ab$. $\forall\ a, b \square R$

(*iv*) $a(b - c) = ab - ac$. $\forall\ a, b, c \in R$

Proof: (*i*) $a \,.\, 0 = a \,.\, (0 + 0)$

$\Rightarrow \quad a \,.\, 0 = a \,.\, 0 + a \,.\, 0$

$\Rightarrow a \,.\, 0 + 0 = a \,.\, 0 + a \,.\, 0$

$\Rightarrow \quad 0 = a \,.\, 0$

using cancellation w.r.t + in the group $< R, + >$.

(*ii*) $a \,.\, 0 = 0$

$\Rightarrow \quad a\,(-b + b) = 0$

$\Rightarrow \quad a\,(-b) + ab = 0$

$\Rightarrow \quad a\,(-b) = -\,(ab)$

similarly, $(-a)\,b = -ab$.

(*iii*) $(-a)\,(-b) = -\,[a\,(-b)] = -[-ab] = ab$

(*iv*) $a\,(b - c) = a\,(b + (-c))$

$= ab + a\,(-c)$

$= ab - ac$.

NOTES

Notes: 1. If R is a ring with unity and $1 = 0$, then since for any $a \in R$, $a = a.1 = a.0 = 0$, we find $R = \{0\}$ which is called the *trivial* ring. We generally exclude this case and thus whenever, we say R is a ring with unity, it will be understood that $1 \neq 0$ in R.

2. If n, m are integers and a, b elements of a ring, then it is easy to see that

$$n(a + b) = na + nb$$
$$(n + m)a = na + ma$$
$$(nm)a = n(ma)$$
$$a^m \; a^n = a^{m+n}$$
$$(a^m)^n = a^{mn}$$

We are so much used to the property that whenever $ab = 0$ then either $a = 0$ or $b = 0$ that it may need more than a bit of convincing that the result may not always be true. Indeed in the ring of integers (or reals or rationals) this property holds. But if we consider the ring of 2×2 matrices over integers, we notice, we can have two non zero elements A, B s.t, $AB = 0$, but $A \neq 0$ $B \neq 0$. In fact, take $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ then $A \neq 0$, $B \neq 0$. But $AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. We formalise this notion through

Definition 1: Let R be a ring. An element $0 \neq a \in R$ is called a *zero-divisor*, if $\exists$ an element $0 \neq b \in R$ such that, $ab = 0$ or $ba = 0$.

Definition 2: A commutative ring R is called an *Integral domain* if $ab = 0$ in R $\Rightarrow$ either $a = 0$ or $b = 0$. In other words, a commutative ring R is called an integral domain if R has no zero divisors.

An obvious example of an integral domain is $< \mathbf{Z}, +, \cdot >$ the ring of integers whereas the ring of matrices, talked about above is an example of a ring which is not an integral domain.

Note: Some authors do not insist upon the condition of commutativity as a part of the definition of an integral domain. One can have non commutative rings without zero divisors.

The following theorem gives us a necessary and sufficient condition for a commutative ring to be an integral domain.

Theorem 2.19: *A commutative ring R is an integral domain iff for all a, b, c* $\in R$ $(a \neq 0)$

$$ab = ac \Rightarrow b = c.$$

Proof: Let R be an integral domain

Let $\quad ab = ac \quad (a \neq 0)$

Then $\quad ab - ac = 0$

$\Rightarrow \quad a(b - c) = 0$

$\Rightarrow$ $a = 0$ or $b - c = 0$

Since $a \neq 0$, we get $b = c$.

Conversely, let the given condition hold.

Let $a, b \in R$ be any elements with $a \neq 0$.

Suppose $ab = 0$

then $ab = a.0$

$\Rightarrow$ $b = 0$ using given condition

Hence $ab = 0 \Rightarrow b = 0$ whenever $a \neq 0$ or that R is an integral domain.

Note: A ring R is said to satisfy *left cancellation law* if for all $a, b, c \in R, a \neq 0$

$$ab = ac \Rightarrow b = c.$$

Similarly we can talk of *right cancellation law*. It might, of course, be noted that cancellation is of only non zero elements.

Definition 1: An element a in a ring R with unity, is called invertible (or a *unit*) with respect to multiplication if $\exists$ some $b \in R$ such that $ab = 1 = ba$.

Notice, unit and unit element (unity) are different concepts and should not be confused with each other.

Definition 2: A ring R with unity is called a *Division ring* or a *skew field* if non zero elements of R form a group with respect to multiplication.

In other words, a ring R with unity is a Division ring if non zero elements of R have multiplicative inverse.

Definition 3: A commutative division ring is called a *field*.

Real numbers form a field, whereas integers do not, under usual addition and multiplication. Since a division ring (field) forms groups with respect to two binary compositions, it must contain two identity elements 0 and 1 (with respect to addition and multiplication) and thus a division ring (field) has at least two elements.

Case 17: A division ring which is not a field. Let M be the set of all 2×2 matrices of the type $\begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix}$ where a, b are complex numbers and $\bar{a}, \bar{b}$ are their conjugates, *i.e.*, if $a = x + iy$ then $\bar{a} = x - iy$. Then M is a ring with unity $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ under matrix addition and matrix multiplication.

Any non zero element of M will be $\begin{bmatrix} x+iy & u+iv \\ -(u-iv) & x-iy \end{bmatrix}$

where x, y, u, v are not all zero.

NOTES

NOTES

One can check that the matrix $\begin{bmatrix} \frac{x-iy}{k} & -\frac{u+iv}{k} \\ \frac{u-iv}{k} & \frac{x+iy}{k} \end{bmatrix}$

where $k = x^2 + y^2 + u^2 + v^2$, will be multiplicative inverse of the above non zero matrix, showing that M is a division ring. But M will not be a field as it is not commutative as

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

But $$\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix}.$$

Case 18: Consider

$D = \{a + bi + cj + dk \mid a, b, c, d \in \mathbf{R}\}$ with $i^2 = j^2 = k^2 = -1$, then D forms a ring under multiplication.

Since $i = 0 + 1i + 0j + 0k$, $j = 0 + 0i + 1j + 0k$ gives $ij = k$, $ji = -k$, we find D is not commutative and hence is not a field. D has unity $1 = 1 + 0i + 0j + 0k$.

If $a + bi + cj + dk$ be any non zero element of D (*i.e.*, at least one of a, b, c, d is non zero) then $(a + bi + cj + dk)\frac{(a - bi - cj - dk)}{a^2 + b^2\ c^2 + d^2} = 1$.

Hence D is a division ring but not a field.

Theorem 2.20: *A field is an integral domain.*

Proof: Let $< R, +, . >$ be a field, then R is a commutative ring.

Let $ab = 0$ in R. We want to show either $a = 0$ or $b = 0$. Suppose $a \neq 0$, then a^{-1} exists (definition of field)

thus $ab = 0$

$\Rightarrow$ $a^{-1}(ab) = a^{-1}0$

$\Rightarrow$ $b = 0.$

which shows that R is an integral domain.

A 'Partial Converse' of the above result also holds.

Theorem 2.21: *A non-zero finite integral domain is a field.*

Proof: Let R be a non zero finite integral domain.

Let R' be the subset of R containing non zero elements of R.

Since associativity holds in R, it will hold in R'. Thus R' is a finite semi group.

Again cancellation laws hold in R (for non zero elements) and therefore, these hold in R'.

Hence R' is a finite semi group with respect to multiplication in which cancellation laws hold.

$\therefore < R', \cdot >$ forms a group.

In other words $< R, +, \cdot >$ is a field (it being commutative as it is an integral domain).

NOTES

Aliter: Let $R = \{a_1, a_2, ..., a_n\}$ be a finite non zero integral domain. Let $0 \neq a \in R$ be any element then $aa_1, aa_2, ..., aa_n$ are all in R and if $aa_i = aa_j$ for some $i \neq j$, then by cancellation we get $a_i = a_j$ which is not true. Hence $aa_1, aa_2, ..., aa_n$ are distinct members of R.

Since $a \in R$, $a = aa_i$ for some i

Let $x \in R$ be *any* element, then $x = aa_j$ for some j

Thus $ax = (aa_i)x = a(a_ix)$

i.e., $x = a_ix$

Hence using commutativity we find

$$x = a_ix = xa_i$$

or that a_i is unity of R. Let $a_i = 1$

Thus for $1 \in R$, since $1 = aa_k$ for some k

We find a_k is multiplicative inverse of a. Hence any non zero element of R has multiplicative inverse or that R is a field.

Case 19: An infinite integral domain which is not a field is the ring of integers.

Definition: A ring R is called a *Boolean ring* if $x^2 = x$ for all $x \in R$.

Case 20: The ring $\{0, 1\}$ under addition and multiplication mod 2 forms a Boolean ring.

Example 2.25: *Show that a Boolean ring is commutative.*

Solution: Let $a, b \in R$ be any elements

Then $a + b \in R$ (closure)

By given condition

$$(a + b)^2 = a + b$$
$$\Rightarrow a^2 + b^2 + ab + ba = a + b$$
$$\Rightarrow a + b + ab + ba = a + b$$
$$\Rightarrow ab + ba = 0$$
$$\Rightarrow ab = -ba \quad ...(1)$$
$$\Rightarrow a(ab) = a(-ba)$$
$$\Rightarrow a^2b = -aba$$
$$\Rightarrow ab = -aba \quad ...(2)$$

NOTES

Again Equation (1) gives,

$$(ab)a = (-ba)a$$

$$\Rightarrow \quad aba = -ba^2 = -ba \qquad ...(3)$$

Equations (2) and (3) give,

$$ab = ba \ (= -aba)$$

or that R is commutative.

Example 2.26: (*a*) *Show that a non zero element a in* $\mathbf{Z}_n$ *is a unit iff a and n are relatively prime.*

(*b*) *If a is not a unit then it is a zero divisor.*

Solution: (*a*) $\mathbf{Z}_n = \{0, 1, 2, ..., n-1\} \bmod n$

Let $a \in \mathbf{Z}_n$ be a unit, then $\exists\ b \in \mathbf{Z}_n$ such that,

$$a \otimes b = 1$$

i.e., when ab is divided by n, remainder is 1, in other words,

$$ab = nq + 1$$

or $\quad ab - nq = 1$

$\Rightarrow$ a and n are relatively prime.

Conversely, let $(a, n) = 1$, then $\exists$ integers u, v such that,

$$au + nv = 1$$

$$\Rightarrow \quad au = n(-v) + 1$$

Suppose, $\quad u = nq + r, \quad 0 \le r < n, \quad r \in \mathbf{Z}_n,$

Then $\quad au = anq + ar = n(-v) + 1$

$$\Rightarrow \quad ar = n(-v - aq) + 1, \quad r \in \mathbf{Z}_n$$

i.e., $\quad a \otimes r = 1, \quad r \in \mathbf{Z}_n$

i.e., $\quad a$ is a unit.

(*b*) Let a be not a unit and suppose g.c.d$(a, n) = d > 1$

Since $d \mid a$, $a = dk$ for some k. Also $d \mid n \Rightarrow n = dt$

$$\Rightarrow a.t = dk\ \frac{n}{d} = kn = 0 \bmod n$$

i.e., a is a zero divisor.

Example 2.27: *Show that* $\mathbf{Z}_p = \{0, 1, 2, ..., p-1\}$ *modulo p is a field iff p is a prime.*

Solution: Let $\mathbf{Z}_p$ be a field. Suppose p is not a prime, then $\exists\ a, b$, such that $p = ab$, $1 < a, b < p$

$\Rightarrow a \otimes b = 0$ where a, b are non zero $\Rightarrow \mathbf{Z}_p$ has zero divisors.

i.e. $\mathbf{Z}_p$ is not an integral domain, a contradiction as $\mathbf{Z}_p$ being a field is an integral domain.

Hence p is prime.

Conversely, let p be a prime. We need show that $\mathbf{Z}_p$ is an integral domain (it being finite will then be a field).

Let $a \otimes b = 0 \quad a, b \in \mathbf{Z}_p$

Then ab is a multiple of p

$\Rightarrow p \mid ab$

$\Rightarrow p \mid a$ or $p \mid b$ (p being prime)

$\Rightarrow a = 0$ or $b = 0$ (Notice $a, b \in \mathbf{Z}_p \Rightarrow a , b < p$)

$\Rightarrow \mathbf{Z}_p$ is an integral domain and hence a field.

NOTES

Example 2.28: *If in a ring R, with unity,* $(xy)^2 = x^2y^2$ *for all* $x, y \in R$ *then show that R is commutative.*

Solution: Let $x, y \in R$ be any elements

then $y + 1 \in R$ as $1 \in R$

By given condition

$$(x(y + 1))^2 = x^2 (y +1)^2$$

$$\Rightarrow (xy + x)^2 = x^2 (y + 1)^2$$

$$\Rightarrow (xy)^2 + x^2 + xyx + xxy = x^2(y^2 + 1 + 2y)$$

$$\Rightarrow x^2y^2 + x^2 + xyx + xxy = x^2y^2 + x^2 + 2x^2y$$

$$\Rightarrow xyx = x^2y \qquad ...(1)$$

Since Equation (1) holds for all x, y in R, it holds for $x + 1, y$ also. Thus replacing x by $x + 1$, we get

$$(x + 1)\, y(x +1) = (x +1)^2y$$

$$\Rightarrow (xy + y)\, (x +1) = (x^2 +1 +2x)y$$

$$\Rightarrow xyx + xy + yx + y = x^2y + y + 2xy$$

$\Rightarrow yx = xy$ using Equation (1)

Hence R is commutative.

Example 2.29: *Show that the ring R of real valued continuous functions on* [0, 1] *has zero divisors.*

Solution: Consider the functions f and g defined on [0, 1] by

$$f(x) = \frac{1}{2} - x, \quad 0 \le x \le \frac{1}{2}$$

$$= 0, \quad \frac{1}{2} \le x \le 1$$

and $g(x) = 0, \quad 0 \le x \le \frac{1}{2}$

$$= x - \frac{1}{2}, \quad \frac{1}{2} \le x \le 1$$

NOTES

then f and g are continuous functions and $f \neq 0, g \neq 0$

whereas $g f(x) = g(x)f(x) \quad = 0 \cdot \left(\frac{1}{2} - x\right)$ if $0 \leq x \leq \frac{1}{2}$

$$= \left(x - \frac{1}{2}\right) . 0 = 0 \text{ if } \frac{1}{2} \leq x \leq 1$$

i.e., $gf(x) = 0$ for all x

i.e., $gf = 0$ but $f \neq 0, g \neq 0$.

Definition: A non-empty subset S of a ring R is said to be a *subring* of R if S forms a ring under the binary compositions of R.

The ring $< \mathbf{Z}, +, \cdot >$ of integers is a subring of the ring $< \mathbf{R}, +, \cdot >$ of real numbers.

If R is a ring then $\{0\}$ and R are always subrings of R, called *trivial* subrings of R.

It is obvious that a subring of an integral domain will be an integral domain.

In practice it would be difficult and lengthy to check all axioms in the definition of a ring to find out whether a subset is a subring or not. The following theorem would make the job rather easy.

Theorem 2.22: *A non-empty subset S of a ring R is a subring of R iff $a, b \in S \Rightarrow ab, a - b \in S$.*

Proof: Let S be a subring of R

then $a, b \in S \Rightarrow ab \in S$ (closure)

$a, b \in S \Rightarrow a - b \in S$

as $< S, + >$ is a subgroup of $< R, + >$.

Conversely, since $a, b \in S \Rightarrow a - b \in S$, we find $< S, + >$ forms a subgroup of $< R, + >$. Again for any $a, b \in S$, since $S \subseteq R$

$a, b \in R$

$\Rightarrow a + b = b + a$

and so we find S is abelian.

By a similar argument, we find that multiplicative associativity and distributivity hold in S.

In other words, S satisfies all the axioms in the definition of a ring.

Hence S is a subring of R.

Definition: A non-empty subset S of a field F is called a *subfield*, if S forms a field under the operations in F. Similarly, we can define a *subdivision ring* of a division ring.

The simple modules over a ring R are the (left or right) modules over R, which have no non zero proper submodules.

NOTES

Module

A **left *R*-module** M over the ring R consists of an abelian group $(M, +)$ and an operation $R \times M \to M$ called scalar multiplication, such that for all $r, s \in R$ and $x, y \in M$, we have:

1. $r(x + y) = rx + ry$
2. $(r + s)x = rx + sx$
3. $(rs)x = r(sx)$
4. $1_R x = x$, if R has multiplicative identity 1_R.

A **right *R*-module** is defined in the similar way but the ring acts on the right, i.e., we have a scalar multiplication of the form $M \times R \to M$, and the axioms are written with scalars r and s on the right of x and y. If R is commutative, then left R-modules are the same as right R-modules and are called R-modules.

Submodule

Suppose M is a left R-module and N is a subgroup of M. Then N is a **submodule** or ***R*-submodule** if, for any $n \in N$ and any $r \in R$, the product $rn \in N$ or $nr \in N$ in the case of right R-module.

Quotient Module

Given a module A over a ring R, and a submodule B of A, the quotient space A/B is defined by the equivalence relation

$a \sim b$ if and only if $b - a \in B$,

for any a and $b \in A$. The elements of A/B are the equivalence classes $[a] = \{ a + b : b \text{ in } B \}$.

The addition operation on A/B is defined for two equivalence classes as the equivalence class of the sum of two representatives from these classes as,

$[a] + [b] = [a + b]$ for $a, b \in A$ and $r \in R$

and the multiplication by elements of R as,

$r \cdot [a] = [r \cdot a]$, for all $a, b \in A$ and $r \in R$

In this way, A/B becomes itself a module over R, called the **quotient module**.

2.7.1 Simple Modules

Definition 1: A module is an algebraic object in which things can be added together commutatively by multiplying coefficients and in which most of the rules of manipulating vectors hold. If a module takes its coefficients in a ring R then it is called a module over R or an R-module. If a and b are two integers then the smallest module containing a and b is the module for their greatest common divisor.

Definition 2: The left R-module M is said to be finitely generated if there exist $m_1, m_2, \ldots, m_n \in M$ such that, $M = \sum_{i=1}^{n} Rm_i$.

NOTES

In this case, we say that $\{ m_1, m_2, \ldots, m_n \}$ is a set of generators for M. The module M is called cyclic if there exists $m \in M$ such that $M = Rm$. The module M is called a free module if there exists a subset $X \leq M$ such that each element $m \in M$ can be expressed uniquely as a finite sum $m = \sum_i^n = 1\ a_i x_i$, with $a_1, \ldots, a_n \ R$ and $x_1, \ldots, x_n \in X$.

Definition 3: Let R be a ring and let M be a left R-module. For any element $m \in M$, the left ideal

$$\text{Ann}(m) = \{ r \in R \mid r\, m = 0 \}$$

is called the annihilator of m. The ideal

$$\text{Ann}\ (M) = \{ r \in R \mid r\, m = 0 \text{ for all } m \in M \}.$$

is called the **annihilator** of M.

The module M is called **faithful** if $\text{Ann}(M)=(0)$.

A module is *simple* if it is non-zero and does not admit a proper non-zero submodule. If a module M is simple then the following are equivalent:

- $Am = M$ for every m non-zero in M. simple module
- $M \cong A/m$ for some maximal left ideal of A.

In particular, simple modules are cyclic and the annihilator of any non-zero element of a simple module is a maximal left ideal.

The annihilator of a simple module is called a primitive ideal. The ring A is primitive if the zero ideal is primitive or equivalently, if A admits a faithful simple module.

- A module may have no simple submodules. Simple submodules of ${}_AA$ are minimal left ideals.
- The module ${}_AA$ is simple if and only if A is a division ring. In this case, any simple module is isomorphic to ${}_AA$.
- The Z-module Z/p^nZ where p is a prime is indecomposable. It is simple if and only if $n = 1$.
- Let $A = \text{End}_k V$ for a field k and a k-vector space V. The set a of finite rank endomorphisms is a two-sided ideal of A. Let B be the subring A generated by the identity endomorphism and a. Then V is a simple B-module, in particular a simple A-module and $B \ A$ if $\dim_k V$ is infinite. Let W be a codimension 1 subspace of V. The endomorphisms killing W form a minimal left ideal in A and in B. Thus A and B when $\dim_k V$ is infinite give examples of primitive rings that admit non-trivial proper two-sided ideals.

Definition 4: A *uniform* module is a non-zero module M such that the intersection of any two non zero submodules of M is non-zero or equivalently such that every non zero submodule of M is essential in M.

Note: An essential submodule of a module B is any submodule A which has non-zero intersection with every non-zero submodule of B.

NOTES

2.7.2 Semi-Simple Module

In mathematics, particularly in the area of abstract algebra known as **module theory**, a semi-simple module or completely reducible module is a type of module that can be understood easily from its parts. A ring that is a semi simple module over itself is known as an Artinian semi simple ring. Some important rings, such as group rings of finite groups over fields of characteristic zero, are semi-simple rings. An Artinian ring is initially understood via its largest semi-simple quotient. The structure of Artinian semi-simple rings is well understood by the **Artin–Wedderburn theorem**, which exhibits these rings as finite direct products of matrix rings.

Definition: A module over a (not necessarily commutative) ring is said to be semi simple (or completely reducible) if it is the direct sum of simple (irreducible) submodules.

For a module M, the following are equivalent:

1. M is semi-simple; i.e., a direct sum of irreducible modules.

2. M is the sum of its irreducible submodules.

3. Every submodule of M is a direct summand: for every submodule N of M, there is a complement P such that $M = N \oplus P$.

The most basic example of a semi simple module is a module over a field, i.e., a vector space. On the other hand, the ring $\boldsymbol{Z}$ of integers is not a semi simple module over itself, since the submodule $2\boldsymbol{Z}$ is not a direct summand.

Semi-simple is stronger than completely decomposable, which is a direct sum of indecomposable submodules.

Let A be an algebra over a field K. Then a left module M over A is said to be absolutely semi simple if, for any field extension F of K, $F \otimes_K M$ is a semi-simple module over $F \otimes_K A$.

Properties of Semi-Simple Module

- If M is semi simple and N is a submodule, then N and M/N are also semi simple.
- An arbitrary direct sum of semi-simple modules is semi-simple.
- A module M is finitely generated and semi-simple if and only if it is Artinian and its radical is zero.

2.7.3 Schur's Lemma

Schur's lemma is a fundamental result in representation theory, an elementary observation about irreducible modules, which is nonetheless noteworthy because of its profound applications.

Lemma 10: Let G be a finite group and let V and W be irreducible G-modules. Then, every G-module homomorphism $f: V \to W$ is either invertible or the trivial zero map.

NOTES

Proof: Both the kernel, ker f and the image, im f are G-submodules of V and W, respectively. Since V is irreducible, ker f is either trivial or all of V. In the former case, im f is all of W also because W is irreducible and hence f is invertible. In the latter case, f is the zero map.

Given below is one of the most important consequences of Schur's lemma:

Corollary: Let V be a finite-dimensional, irreducible G-module taken over an algebraically closed field. Then, every G-module homomorphism $f: V \to V$ is equal to a scalar multiplication.

Proof: Since the ground field is algebraically closed, the linear transformation $f: V \to V$ has an eigenvalue λ, say. By definition, $f - \lambda$ is not invertible, and hence equal to zero by Schur's lemma. In other words, $f = \lambda$, i.e., a scalar.

2.7.4 Free Modules Fundamental Structure Theorem

In a principal ideal domain, the generators of an ideal is unique up to associates. If $a \in R$, then the generator of ann(a) (= $\{r \in R | ra = 0\}$) is called the order of a, denoted by $o(a)$. Now we attach a weight $P(a)$ to $a \in$ R. Since R is a unique factorization domain, we denote the number of prime factors (counting multiplicity) of a by $P(a)$. By convention, $P(0) = 1$. Thus, $a|b$ in R implies that $P(a) \leq P(b)$, where the equality holds if and only if a, b are associates.

Lemma 11: Let M be a finitely generated module over a principal ideal domain R, say $M = \{m_1, \ldots, m_n\}$. Suppose that there is a relation $a_1 m_1 + \ldots + a_n m_n = 0$, where not all the a_i are zero. Then there are elements $m'_1, \ldots, m'_n \in M$, such that, $M = \{m'_1, \ldots, m'_n\}$, and the order of m'_1 divides every a_i.

Proof: If one of the a_i is a unit then the proof follows.

If a_1 is a unit, then m_1 is a linear combination of the other m_i. So take $m'_1 = 0$, $m'_i = m_i$, $i > 1$.

Let $s = \sum P(a_i)$ where $a_i \neq 0$. We will prove this by induction on s. If $s = 0$, every a_i is zero or a unit and at least one a_i is a unit.

If only one a_i is non-zero, the result is easy to establish, so let us assume a_1, a_2 are nonzero and non-unit. Let b = g.c.d.(a_1, a_2), $a_1 = bc_1$, $a_2 = bc_2$, and $b_1 c_1 + b_2 c_2 = 1$.

Now,

$M = \{m_1, m_2, \ldots, m_n\}$

$$= \left\{ (m_1, m_2) \begin{pmatrix} c_2 & b_1 \\ -c_1 & b_2 \end{pmatrix}, m_3, \ldots, m_n \right\}$$

$0 = b(b_1 m_1 + b_2 m_2) + a_3 m_3 + \ldots + a_n m_n$

Now $P(b) \leq P(a_1 < P(a_1) + P(a_2)$. By induction, $M = \{m'_1, \ldots, m'_n\}$, with $o(m'_1) \mid b$, and $o(m'_1) | a_i$, for $i \geq 3$. But $b|a_1$, $b|a_2$, hence $o(m'_1)|a_i$, for all i.

Theorem 2.23: *Every n-generated module M over a principal ideal domain R is a direct sum of n cyclic modules* $M \simeq \oplus_{i=1}^{n} Rm_i$. *Equivalently,* $M = \{m_1, \ldots, m_n\}$, *and* $\sum a_i m_i = 0$ *implies* $a_i m_i = 0$, *for all i.*

Proof: If $n = 1$, this is true, as R is a principal ideal domain. Now let $n > 1$. We induct on n.

Amongst all possible set of generators of M having n elements choose one which has an element m with least $P(m)$. Let $M = \{m = m_1, m'_2, \ldots, m'_n\}$. If $M = R_m \oplus \sum_{i\geq 2} Rm'_i$, then by induction the submodule $\sum_{i\geq 2} Rm'_i$ has a basis $\{m_2, \ldots, m_n\}$. But then $\{m_1, \ldots, m_n\}$ is a basis of M.

We show that Rm is indeed a direct summand of M: If not, one has a relation $a_1 m_1 + \ldots + a_n m_n = 0$, with $a_1 m_1 \neq 0$. Let b =g.c.d.$(a_1, o(m_1)) = c_1 a_1 + c_2 o(m_1)$. Since $a_1 m_1 \neq 0$, a_1 and $o(m_1)$ are not associates. Hence, $P(b) < P(o(m_1))$.

Note that $bm_1 + c_1 a_2 m_2 + \ldots + c_1 a_n m_n = 0$. By above Lemma $M = \{m'_1, \ldots, m'_n\}$, with $o(m'_1) | b$, $o(m'_1) | c_1 a_i$, for $i \geq 2$. Since $P(o(m'_1)) \leq P(b) < P(o(m_1))$, this contradicts the minimality of $\{m_1, \ldots, m_n\}$. Thus, R, m is a summand of M and the result follows.

Check Your Progress

7. When a oring R is called a Boolean ring?
8. State the Schur's lemma.

2.8 ANSWERS TO 'CHECK YOUR PROGRESS'

1. Let V and U be two vector spaces over the same field F, then a mapping $T: V \to U$ is called a linear transformation if

   $T(x+y) = T(x) + T(y)$ for all $x, y \in V$

   $T(\alpha x) = \alpha T(x)$, $\alpha \in F$

2. Let $\{p(\lambda)\}^q$ be one of the elementary divisors of the characteristic matrix of some λ-matrix and $C(p)$ be the companion matrix of $p(\lambda)$. The hypercompanion matrix H associated with the elementary divisor $\{p(\lambda)\}^q$ is given by

$$H = c(p) \quad if\ q = 1 \quad H = \begin{bmatrix} C(p) & M & 0 & \cdots & 0 & 0 \\ 0 & C(p) & M & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & C(p) & M \\ 0 & 0 & 0 & \cdots & 0 & C(p) \end{bmatrix} \quad if\ q > 1$$

NOTES

where M is a matrix of the same order as $C(p)$ having the element 1 in the lower left-hand corner and zeros elsewhere. The diagonal of the hypercompanion matrix H consists of q identical $C(p)$ matrices.

3. Let A be an $n \times n$ matrix A and let C_i, C_i, C_{i+1}, ..., C_n be the companion matrices of the non-trivial invariant factors of $\lambda I - A$.
4. Nilpotent element is an element a of a ring or semi-group with zero A such that $a^n = 0$ for some natural number n.
5. If T is a linear operator on a finite dimensional space V over F and minimal polynomial $p(x)$ of T is a product of distinct linear factors then T is diagonalizable.
6. A Jordan block is a matrix with zeros everywhere except along the diagonal and superdiagonal, with each element of the diagonal consisting of a single number λ, and each element of the superdiagonal consisting of a 1.
7. A ring R is called a *Boolean ring* if $x^2 = x$ for all $x \in R$.
8. Let G be a finite group and let V and W be irreducible G-modules. Then, every G-module homomorphism f: $V \to W$ is either invertible or the trivial zero map.

2.9 SUMMARY

- A similarity transformation is a conformal mapping whose transformation matrix A^2 can be written in the form

  $$A' = BAB^{-1},$$

  where A and A^2 are called similar matrices.
- Every square matrix A over F is similar to the direct sum of the hypercompanion matrices of the elementary divisors over F of $\lambda I - A$.
- The Jacobson canonical form of a square matrix A consists of the direct sum of the hypercompanion matrices of the elementary divisors over F of $\lambda I - A$.
- An n-square matrix A is similar to a diagonal matrix if and only if the elementary divisors of $\lambda I - A$ are linear polynomials, i.e., if and only if the minimum polynomial of A is the product of distinct linear polynomials.
- Let T be a linear operator on a vector space V. If W is a subspace of V such that, $T(W) \subseteq W$, we say W is *invariant under T* or is *T-invariant*.
- Every square matrix A is similar to the direct sum of the companion matrices of the non-trivial invariant factors of $\lambda I - A$.
- A linear transformation $N: U \to U$ is called nilpotent if there exists a $k \in N$ such that $N^k = 0$ for some positive integer k.

NOTES

- Nilpotent element is an element a of a ring or semi-group with zero A such that $a^n = 0$ for some natural number n. The smallest such n is called the nilpotency index of a.
- Canonical form for nilpotent matrices is one that is all zeroes except for blocks of subdiagonal ones.
- A Jordan block is a matrix with zeros everywhere except along the diagonal and superdiagonal, with each element of the diagonal consisting of a single number λ, and each element of the superdiagonal consisting of a 1.
- The Jordan form of a matrix is determined only up to the order of the Jordan blocks. More exactly, two Jordan matrices are similar over k if and only if they consist of the same Jordan blocks and differ only in the distribution of the blocks along the main diagonal.
- A non empty set R, together with two binary compositions + and . is said to form a *Ring* if the following axioms are satisfied:

  (*i*) $a + (b + c) = (a + b) + c$ for all $a, b, c \in R$

  (*ii*) $a + b = b + a$ for $a, b \in R$

  (*iii*) $\exists$ some element 0 (called zero) in R, such that, $a + 0 = 0 + a = a$ for all $a \in R$

  (*iv*) For each $a \in R$, $\exists$ an element $(-a) \in R$, such that, $a + (-a) = (-a) + a = 0$

  (*v*) $a \,.\, (b \,.\, c) = (a \,.\, b) \,.\, c$ for all $a, b, c \in R$

  (*vi*) $a \,.\, (b + c) = a \,.\, b + a \,.\, c$

  $(b + c) \,.\, a = b \,.\, a + c \,.\, a$ for all $a, b, c \in R$

2.10 KEY TERMS

- **Canonical form:** The canonical form of a positive integer in decimal representation is a finite sequence of digits that does not begin with zero.
- **Jacobson canonical form:** The Jacobson canonical form of a square matrix A consists of the direct sum of the hypercompanion matrices of the elementary divisors over F of $\lambda I - A$
- **Nilpotent froms formations:** A linear transformation $N : U \to U$ is called nilpotent if there exists a $k \in N$ such that $N^k = 0$ for some positive integer k. The smallest such k is sometimes called the degree of N.
- **Jordan blocks:** A Jordan block is a matrix with zeros everywhere except along the diagonal and superdiagonal, with each element of the diagonal consisting of a single number λ, and each element of the superdiagonal consisting of a 1.

- **Simple modules:** A module is an algebraic object in which things can be added together commutatively by multiplying coefficients and in which most of the rules of manipulating vectors hold.

NOTES

2.11 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. What is the significance of linear transformations?
2. What does Jordan canonical form
3. How will you define a minimal polynomial?
4. Define nilpotent transformations.
5. State the primary decomposition theorem.
6. What is Jordan blocks used for?
7. What are simple modules?
8. Specify the term submodule.
9. What is the significance of Schur's lemma?
10. State the fundamental structure theorem for modules.

Long-Answer Questions

1. Let V be the vector space of all polynomials of degree ≤ 6 over F. Let W be the subspace of V spanned by $\{1, x^2, x^4, x^6\}$. Let D be the differential operator on V. (i.e., $D(f(x)) = \frac{d}{dx} f(x)$. Show that W is not invariant under D.

2. In (1) show that W is invariant under D^2 where $D^2(f(x)) = \frac{d^2}{dx^2} f(x)$. Let $T = D^2$. Find
   (*i*) The matrix of T_w in a suitable basis of W.
   (*ii*) The matrix of $\hat{T}$ in a suitable basis of $\frac{V}{W}$
   (*iii*) The matrix of T in a suitable basis of V.

   (*i*) $A = \begin{bmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 12 & 0 \\ 0 & 0 & 0 & 30 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ (*ii*) $C = \begin{bmatrix} 0 & 6 & 0 \\ 0 & 0 & 20 \\ 0 & 0 & 0 \end{bmatrix}$ (*iii*) $\begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix}$

3. Let V be the vector space of all polynomials over the field of real numbers R. Let W be the subspace of V spanned by $\{1, x, x^2\}$. Let T be the linear

operator on V defined by $T(f(x)) = xf(x)$. Show that W is not invariant under T.

4. Let T be a linear operator on a vector space V over F. If $W_1, ..., W_k$ are T-invariant subspaces of V, prove that $\sum_{i=1}^{k} W_i$ and $\bigcap_{i=1}^{k} W_i$ are T-invariant subspaces of V.

NOTES

5. Let c be a characteristic value of T and W be the space of characteristic vectors associated with the characteristic value c. What is the restriction operator T_w? ($T_w = cI$)
6. Let T be a linear operator on a finite dimensional vector space V. Prove that T is diagonalisable if and only if V is a direct sum of one dimensional T-invariant subspaces.
7. Let T be a linear operator on a finite dimensional vector space V and let W be a T-invariant subspace of V.
   (*i*) Show that if λ is an eigen value of T_w, then λ is an eigen value of T.
   (*ii*) Show that the eigen space of T_w corresponding to eigen value λ of T_w is $W_\lambda \cap W$, where W_λ denotes the eigen space of T corresponding to λ.
   (*iii*) Prove that if T is diagonalizable, then so is T_w.
   (Hint: T is diagonalizable $\Leftrightarrow V = W_1 + ... + W_k$ where W_i denotes eigen space corresponding to eigen value λ_i of T. Use (*ii*)).
8. Let W be a proper T-invariant subspace of V, where T is a linear operator on a finite dimensional vector space V.

   Let $\eta : V \to \frac{V}{W}$ such that,

   $\eta(v) = W + v$ be a linear transformation. Show that $\eta T = \hat{T}\eta$ where $\hat{T}$ is a linear operator on $\frac{V}{W}$ defined by $\hat{T}(W + v) = W + T(v)$.

   Further, if T is diagonalizable, show that $\hat{T}$ is also diagonalizable.
   (Hint: T is diagonalizable $\Rightarrow \exists$ a basis $\{x_1, ..., x_n\}$ of V consisting of eigen vectors of T. Also $\eta T = \hat{T}\eta \Rightarrow \{\eta x_1, ..., \eta x_n\}$ are eigen vectors of $\hat{T} \Rightarrow \{W + x_1, ..., W + x_n\}$ are eigen vectors of $\hat{T}$. If $\{W + v_1, ..., W + v_r\}$ is a basis of $\frac{V}{W}$, then it can be replaced by $\{W + x_1, ..., W + x_r\}$ such that it forms a basis of $\frac{V}{W}$ consisting of eigen vectors of $\frac{V}{W}$).
9. Let T be a linear operator on a finite dimensional vector space and suppose that $V = W_1 \oplus ... \oplus W_k$, where W_i is a T-invariant subspace of V for each

NOTES

$i = 1, ..., k$. If $f(t)$ denotes the characteristic polynomial of T and $f_i(t)$ denotes the characteristic polynomial of T_{w_i} $(1 \le i \le k)$, then show that $f(t) = f_1(t) \cdot f_2(t) \ldots f_k(t)$ (Hint: Use induction on k).

10. If E_1, E_2 are projections onto independent subspaces, show that $E_1 + E_2$ is also a projection.
11. Let T be a linear operator on a finite dimensional vector space V. Let R be the range of T and let N be the null space of T. Prove that R and N are independent if and only if $V = R \oplus N$.
12. Let T be a linear operator on a *F.D.V.S.V*. Suppose T is diagonalizable. Show that $T = \text{Ker } T \oplus ImT$
13. Show that the eigen values of $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$ are the fourth roots of unity.
14. Let T be a linear operator on V such that T is diagonalizable. Show that $(T - \lambda I)^n(v) = 0$, $v \in V$, $\lambda \in F \Rightarrow (T - \lambda I)(v) = 0$.
15. Let T be a linear operator on V such that, $T^m = I$. Let char $F = 0$. Suppose T has all eigenalues in F. Show that T is diagonalizable.

    [Hint: If g.c.d. $(f, f') = 1$, then roots of f are simple.]
16. Show that a ring R is commutative iff

    $(a + b)^2 = a^2 + b^2 + 2ab$ for all $a, b \in R$.
17. If in a ring R, $x^2 = x$ for all x then show that $2x = 0$ and $x + y = 0 \Rightarrow x = y$.
18. If R is a ring with unity and $(ab)^2 = (ba)^2$ for all $a, b \in R$ and $2x = 0$

    $\Rightarrow x = 0$ then show that R is commutative.

2.12 FURTHER READING

Herstein, I.N. 1975. *Topics in Algebra*, 3rd Edition. New Delhi: Wiley Eastern Ltd.

Khanna, V.K. and S.K. Bhambri. 2008. *A Course in Abstract Algebra*, 3rd Edition. New Delhi: Vikas Publishing Hous Pvt. Ltd.

Bhattacharya, P.B., S.K. Jain and S.R. Nagpaul. 1997. *Basic Abstract Algebra*, 2nd Edition. New Delhi: Cambridge University Press.

Artin, M.1991. *Algebra*. New Delhi: Prentice-Hall of India.

Lang, S. 1993. *Algebra*, 3rd Edition. New York: Addison-Wesley.

Datta, K.B. 2000. *Matrix and Linear Algebra*. New Delhi: Prentice-Hall of India.

UNIT 3 FIELD THEORY

NOTES

Structure

3.0 INTRODUCTION

In mathematics, a field theory studies the properties of fields. A field is a mathematical entity for which addition, subtraction, multiplication and division are well defined. Fields are important in algebra since they provide the proper generalization of number domains, such as, the sets of rational numbers, real numbers and complex numbers. Field extensions are an object of study in field theory in which we start with a base field and construct a larger field containing the base field and satisfying additional properties. A field extension L/K is called algebraic if every element of L is algebraic over K, i.e., if every element of L is a root of some non-zero polynomial with coefficients in K. Field extensions that are not algebraic, i.e., which contain transcendental elements, are called transcendental.

In this unit, you will study about the field theory, algebraic and transcendental extensions, separable and inseparable extensions, normal extensions, finite fields, algebraically closed fields, automorphism of extensions, Galois extension and solution of polynomial equations by radicals.

NOTES

3.1 OBJECTIVES

After going through this unit, you will be able to:

- Know about the field theory
- Define algebraic, transcendental, separable and inseparable extensions
- Describe perfect fields, normal extension, finite fields and algebraically closed fields
- Understand automorphism of extensions, primitive elements, Galois extensions and fundamental theorem of Galois theory
- Solve polynomial equations by radicals
- Justify the insolvability of the general equation of degree 5

3.2 FIELD THEORY

In mathematics, a field is a set on which addition, subtraction, multiplication, and division are defined and behave as the corresponding operations on rational and real numbers do. A field is thus a fundamental algebraic structure, which is widely used in algebra, number theory, and many other areas of mathematics.

The best known fields are the field of rational numbers, the field of real numbers and the field of complex numbers. Many other fields, such as fields of rational functions, algebraic function fields, algebraic number fields, and *p*-adic fields are commonly used and studied in mathematics, particularly in number theory and algebraic geometry. Most cryptographic protocols rely on finite fields, i.e., fields with finitely many elements.

The relation of two fields is expressed by the notion of a field extension. **Galois Theory**, initiated by **Évariste Galois** in the 1830s, is devoted to understanding the symmetries of field extensions. Among other results, this theory shows that angle trisection and squaring the circle cannot be done with a compass and straightedge. Moreover, it shows that **quintic equations** are, in general, algebraically unsolvable.

Fields serve as foundational notions in several mathematical domains. This includes different branches of mathematical analysis, which are based on fields with additional structure. Basic theorems in analysis hinge on the structural properties of the field of real numbers. Most importantly for algebraic purposes, any field may be used as the scalars for a vector space, which is the standard general context for linear algebra. Number fields, the siblings of the field of rational numbers, are studied in depth in number theory. Function fields can help describe properties of geometric objects.

NOTES

Definition: Informally, a field is a set, along with two operations defined on that set: an addition operation written as $a + b$, and a multiplication operation written as $a + b$, both of which behave similarly as they behave for rational numbers and real numbers, including the existence of an additive inverse $-a$ for all elements a, and of a multiplicative inverse b^{-1} for every non-zero element b. This allows one to also consider the so-called inverse operations of subtraction, $a - b$, and division, a / b, by defining:

$$a - b = a + (-b),$$

$$a / b = a \,.\, b - 1.$$

3.2.1 Extension Field

Definition: Let K be a field and suppose F is a subfield of K, then K is called an *extension* of F.

Suppose S is a non-empty subset of K. Let $F(S)$ denote the smallest subfield of K which contains both F and S. (In fact $F(S)$ would be the intersection of all subfields of K that contain F and S). The following theorem is then an easy consequence.

Theorem 3.1: *If S, T are non-empty subsets of a field K and K is an extension of a field F then* $F(S \cup T) = F(S)\,(T)$ (*where, of course, if* $F(S) = E$, *then by* $F(S)(T)$ *we mean* $E(T)$).

Proof: $F(S \cup T)$ is the smallest subfield of K containing $S \cup T, F$

i.e., $S, T, F \subseteq F(S \cup T)$

$\Rightarrow$ $F(S) \subseteq F(S \cup T),\ T \subseteq F(S \cup T)$

$\Rightarrow$ $F(S)(T) \subseteq F(S \cup T)$

Again, $F, S, T \subseteq F(S)(T)$

$\Rightarrow$ $F, S \cup T \subseteq F(S)\,(T)$

$\Rightarrow$ $F(S \cup T) \subseteq F(S)(T)$

or that $F(S \cup T) = F(S)(T)$

Corollary: $F(S \cup T) = F(T \cup S) = F(S)(T)$ follows clearly as $S \cup T = T \cup S$.

Note: If S is a finite subset $\{a_1, a_2, ..., a_n\}$ of K we write $F(S) = F(a_1, a_2, ..., a_n)$. The order in which a_i appear is immaterial in view of the above Corollary as

$$\begin{aligned} F(a_1, a_2, ..., a_n) &= F(\{a_1\}\{a_2, a_3, ..., a_n\}) \\ &= F(\{a_2, a_3, ..., a_n\} \cup \{a_1\}) \\ &= F(a_2, a_3, ..., a_n, a_1) \end{aligned}$$

Also then, $F(a)(b) = F(a, b) = F(b, a) = F(b)(a)$

Again, if $K = F(a)$, K is called *simple extension* of F and we say K is obtained by adjoining the element a to F.

Example 3.1: *Let* $\mathbf{Q}$ *be the field of rationals then show that*

$$\mathbf{Q}(\sqrt{2}, \sqrt{3}) = \mathbf{Q}(\sqrt{2} + \sqrt{3}).$$

NOTES

Solution: By definition

$$\sqrt{2}, \sqrt{3} \in \mathbf{Q}(\sqrt{2}, \sqrt{3})$$

$$\Rightarrow \quad \sqrt{2}+\sqrt{3} \in \mathbf{Q}(\sqrt{2}, \sqrt{3}) \quad \text{(Closure)}$$

$$\Rightarrow \quad \mathbf{Q}(\sqrt{2}+\sqrt{3}) \subseteq \mathbf{Q}(\sqrt{2}, \sqrt{3})$$

Now, $\quad \sqrt{2}+\sqrt{3} \in \mathbf{Q}(\sqrt{2}+\sqrt{3})$

$$\Rightarrow \quad (\sqrt{2}+\sqrt{3})^2 \in \mathbf{Q}(\sqrt{2}+\sqrt{3})$$

$$\Rightarrow \quad 2+3+2\sqrt{2}\sqrt{3} \in \mathbf{Q}(\sqrt{2}+\sqrt{3})$$

Also $\quad 5 \in \mathbf{Q}(\sqrt{2}+\sqrt{3})$

$$\Rightarrow \quad 5+2\sqrt{2}\sqrt{3}-5 = 2\sqrt{2}\sqrt{3} \in \mathbf{Q}(\sqrt{2}+\sqrt{3})$$

Again, $\quad 2 \in \mathbf{Q}(\sqrt{2}+\sqrt{3})$

$$\therefore \quad 2 \times \frac{1}{2}\sqrt{2}\sqrt{3} = \sqrt{2}\sqrt{3} \in \mathbf{Q}(\sqrt{2}+\sqrt{3})$$

$$\Rightarrow \quad (\sqrt{2}+\sqrt{3})\sqrt{2}\sqrt{3} = 2\sqrt{3}+3\sqrt{2} \in \mathbf{Q}(\sqrt{2}+\sqrt{3}) \qquad ...(1)$$

Also $\quad \sqrt{2}+\sqrt{3} \in \mathbf{Q}(\sqrt{2}+\sqrt{3})$

$$\Rightarrow \quad 2(\sqrt{2}+\sqrt{3}) \in \mathbf{Q}(\sqrt{2}+\sqrt{3})$$

$$\Rightarrow \quad 2\sqrt{2}+2\sqrt{3} \in \mathbf{Q}(\sqrt{2}+\sqrt{3})$$

$\Rightarrow \quad (2\sqrt{3}+3\sqrt{2}) - (2\sqrt{2}+2\sqrt{3}) \in \mathbf{Q}(\sqrt{2}+\sqrt{3})$ by using Equation (1)

$$\Rightarrow \quad \sqrt{2} \in \mathbf{Q}(\sqrt{2}+\sqrt{3})$$

Again, $\quad \sqrt{2}+\sqrt{3} \in \mathbf{Q}(\sqrt{2}+\sqrt{3}) \Rightarrow 3(\sqrt{2}+\sqrt{3}) \in \mathbf{Q}(\sqrt{2}+\sqrt{3})$

and using Equation (1) we get

$$(3\sqrt{2}+3\sqrt{3})-(2\sqrt{3}+3\sqrt{2}) \in \mathbf{Q}(\sqrt{2}+\sqrt{3})$$

i.e., $\quad \sqrt{3} \in \mathbf{Q}(\sqrt{2}+\sqrt{3})$

Hence $\quad \sqrt{2}, \sqrt{3} \in \mathbf{Q}(\sqrt{2}+\sqrt{3}) \Rightarrow \mathbf{Q}(\sqrt{2}, \sqrt{3}) \subseteq \mathbf{Q}(\sqrt{2}+\sqrt{3})$

or that $\quad \mathbf{Q}(\sqrt{2}, \sqrt{3}) = \mathbf{Q}(\sqrt{2}+\sqrt{3})$.

If K is an extension of F, then we know that K can be regarded as a vector space over F. In that case dimension of K over F is called *degree of K over F* and we denote it by $[K : F]$. Our next theorem is about the degree of extension fields. If $[K : F]$ is finite, we say K is finite extension of F.

Theorem 3.2: *Let K be a finite extension of F and L, a finite extension of K. Then L is a finite extension of F and* $[L : F] = [L : K]\,[K : F]$.

Proof: Let $[L : K] = m$, $[K : F] = n$

Let $\{a_1, ..., a_m\}$ be a basis of L over K and $\{b_1, ..., b_n\}$ be a basis of K over F.

NOTES

We show that $\{a_i b_j \mid 1 \le i \le m, 1 \le j \le n\}$ is a basis of L over F.

$$a_i \in L, b_j \in K \Rightarrow b_j \in L. \quad \therefore \quad a_i b_j \in L \text{ for all } i, j$$

Let, $$\sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} a_i b_j = 0, \quad \alpha_{ij} \in F$$

Then $$\sum_{i=1}^{m} \sum_{j=1}^{n} (\alpha_{ij} b_j) a_i = 0, \quad \sum_{j=1}^{n} \alpha_{ij} b_j \in K$$

Since $\{a_1, ..., a_m\}$ are linearly independent over K,

$$\sum_{j=1}^{n} \alpha_{ij} b_j = 0 \quad \text{for all } i = 1, ..., m$$

Also $b_1, ..., b_n$ are linearly independent over F.

$$\alpha_{ij} = 0 \quad \text{for all } i = 1, ..., m \quad j = 1, ..., n$$

$\therefore$ $\{a_i b_j \mid 1 \le i \le m, 1 \le j \le n\}$ is a linearly independent subset of L over F. Let $a \in L$. Since $\{a_1, ..., a_m\}$ is a basis of L over K, $a = \alpha_1 a_1 + ... + \alpha_m a_m$, $\alpha_i \in K$ and $\{b_1, ..., b_n\}$ is a basis of K over F

$$\Rightarrow \quad \alpha_i = \beta_{i1} b_1 + ... + \beta_{in} b_n, \quad \beta_{ij} \in F$$

$$\therefore \quad a = \sum_{i=1}^{m} \alpha_i a_i = \sum_{i=1}^{m} (\beta_{i1} b_1 ... + \beta_{in} b_n) a_i$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} \beta_{ij} a_i b_j, \quad \beta_{ij} \in F$$

$\therefore$ $\{a_i b_j \mid 1 \le i \le m, 1 \le j \le n\}$ spans L over F and so forms a basis of L over F.

$$\therefore \quad [L : F] = mn = [L : K] \, [K : F]$$

Note: If $[L : K]$ is infinite, then $[L : F]$ is also infinite because $[L : F] = r \Rightarrow$ every subset of L having $r + 1$ elements is linearly dependent over F. Since $[L : K]$ is infinite, $\exists\, a_1, ..., a_{r+1} \in L$ which are linearly independent over K. Now $1 \in K$ and 1 is linearly independent over F as $1 \neq 0$. As in Theorem 3.2, $a_1.1, a_2.1, ...$ $a_{r+1}.1$ are linearly independent over F. We find $a_1, ..., a_{r+1} \in L$ are linearly independent over F, a contradiction.

$\therefore$ $[L : F]$ is infinite. Similarly, $[K : F]$ is infinite.

Corollary 1: If L is a finite extension of F and K is a subfield of L which contains F, then $[K : F]$ divides $[L : F]$.

Proof: By remark above $[K : F]$ is finite as $[L : F]$ = finite. Also $[L : K]$ is finite.

By Theorem 3.2, $[L : F] = [L : K] \, [K : F]$

$\therefore$ $[K : F]$ divides $[L : F]$

Corollary 2: If K is an extension of F, then $K = F$ if and only if $[K : F] = 1$.

Proof: If $K = F$, then $[K : F] = [K : K] = 1$

NOTES

If $[K : F] = 1$, let $\{a\}$ be a basis of K over F.

$\therefore \quad 1 \in K \Rightarrow 1 = \alpha a,\ \alpha \in F,\ \alpha \neq 0$ as $1 \neq 0$

$\Rightarrow a = \alpha^{-1} \in F$

Let, $b \in K \Rightarrow b = \beta a, \quad \beta \in F, a \in F$

$\Rightarrow b \in F \Rightarrow K \subseteq F \Rightarrow K = F.$

Corollary 3: If L is an extension of F and $[L : F]$ is a prime number p, then there is no field K such that, $F \subset K \subset L$.

Proof: Suppose $\exists$ a field K such that, $F \subset K \subset L$.

Then $p = [L : F] = [L : K]\ [K : F]$ by Theorem 3.2

$\Rightarrow [L : K] = 1$ or $[K : F] = 1$

$\Rightarrow K = L$ or $K = F$ by Corollary 2 a contradiction.

Hence the result.

Trivially then, if K is an extension of F of prime degree then for every $a \in K$, $F(a) = F$ or $F(a) = K$.

Example 3.2: *Let D be an integral domain. Let F be a field such that, $F \subseteq D$. Suppose unity 1 of F is also unity of D. Then D can be regarded as a vector space over F. Show that D is a field if $[D : F]$ = finite.*

Solution: Let $[D : F] = r$. Let $\{a_1, ..., a_r\}$ be a basis of D over F.

Let $0 \neq a \in D$. We show that a is invertible. Consider $\{aa_1, ..., aa_r\}$.

Let $\alpha_1(aa_1) + ... + \alpha_r(aa_r) = 0,\ \alpha_i \in F.$

Then $a(\alpha_1 a_1 + ... + \alpha_r a_r) = 0$

$\Rightarrow \alpha_1 a_1 + ... + \alpha_r a_r = 0$, as $a \neq 0$ and D is an integral domain.

$\Rightarrow \alpha_i = 0$ for all $i = 1, ..., r$ as $\{a_1, ..., a_r\}$ is linearly independent over F.

$\Rightarrow \{aa_1, ..., aa_r\}$ is linearly independent over F.

But $[D : F] = r \Rightarrow \{aa_1, ..., aa_r\}$ is a basis of D over F.

$\therefore \quad 1 \in D \Rightarrow 1 = \beta_1 aa_1 + ... + \beta_r aa_r, \quad \beta_i \in F$

$= a(\beta_1 a_1 + ... \beta_r a_r)$

$= ab, \quad b = \beta_1 a_1 + ... + \beta_r a_r \in D$

$\Rightarrow$ a is invertible.

$\Rightarrow$ D is a field.

3.3 ALGEBRAIC AND TRANSCENDENTAL EXTENSIONS

Suppose K is an extension of F and $a \in K$.

Let $F[a] = \{f(a) \mid f(x) = a_0 + a_1 x + ... + a_n x^n \in F[x]\}, \quad a_i \in F$

then as $f(a) = a_0 + a_1 a + ... + a_n a^n \in K$, we find $F[a] \subseteq K$

One can show that $F[a]$ is an integral domain.

Let E be its field of quotients. Then E is the smallest field containing $F[a]$. We show

$$F[a] \subseteq F(a) \subseteq E.$$

Now, $x = 0 + 1.x + 0.x^2 + \ldots \in F[x]$ and so

$$a = 0 + 1.a + 0.a^2 + \ldots \in F[a]$$

i.e., $a \in F[a] \subseteq E$

Again if $\alpha \in F$ be any element then

$$\alpha = \alpha + 0x + 0x^2 + \ldots \in F[x]$$

gives $\alpha \in F[a]$ or that $F \subseteq F[a] \subseteq E$

Hence $F(a) \subseteq E$, as $F(a)$ is the smallest field containing F and a.

If $f(a) \in F[a]$ be any member where

$$f(a) = \alpha_0 + \alpha_1 a + \ldots + \alpha_n a^n,\ \alpha_i \in F$$

then as $a \in F(a)$, $\alpha_i \in F \subseteq F(a)$, we find $f(a) \in F(a)$

Hence $F[a] \subseteq F(a)$ and so

$$F[a] \subseteq F(a) \subseteq E$$

But E is the smallest field containing $F[a]$.

$\therefore$ $E \subseteq F(a)$. Hence $F(a) = E$.

So, we have explicitly determined the field $F(a)$. It is the field of quotients of $F[a]$.

We write, $F(a) = \left\{ \dfrac{f(a)}{g(a)} \middle| g(a) \neq 0,\ f(x), g(x) \in F[x] \right\}$

In general, one can show that

$$F(a_1, \ldots, a_n) = \left\{ \frac{f(a_1, \ldots, a_n)}{g(a_1, \ldots, a_n)} \middle| g(a_1, \ldots, a_n) \neq 0, \begin{matrix} f(x_1, \ldots, x_n) \in F[x] \\ g(x_1, \ldots, x_n) \in F[x] \end{matrix} \right\}$$

A natural question arises. When is $F[a] = F(a)$? To answer this, we first define what is an algebraic element. Let K be an extension of F. $a \in K$ is said to be *algebraic* over F if $\exists$ non-zero polynomial $f(x) \in F[x]$ such that, $f(a) = 0$. Otherwise, it is called transcendental element. For example, $\sqrt{2} \in \mathbf{R}$ = real field, is algebraic over $\mathbf{Q}$ = rational field as $\sqrt{2}$ satisfies non-zero polynomial $f(x) = x^2 - 2 \in \mathbf{Q}[x]$. However, $\pi, e \in \mathbf{R}$ are not algebraic over $\mathbf{Q}$. An extension K of F is called an *algebraic extension* if every $a \in K$ is algebraic over F.

If for some $a \in K$, a is not algebraic over F, then K is called transcendental extension of F. For example, $\mathbf{R}$ is transcendental extension of $\mathbf{Q}$. We shall see in the following theorem that finite extensions are algebraic. So, $\mathbf{C}$ = the field of complex numbers is algebraic over $\mathbf{R}$ as $[\mathbf{C} : \mathbf{R}] = 2$, $\{1, i\}$ being a basis of $\mathbf{C}$ over $\mathbf{R}$.

We sometimes use the notation K/F to express the fact that K is an extension of F. Similarly, K/F is algebraic would mean K is an algebraic extension of F.

NOTES

NOTES

Theorem 3.3: *A finite extension is algebraic.*

Proof: Let K be a finite extension of F. Let $[K : F] = n$. Let $a \in K$. Then $1, a, ..., a^n$ are linearly dependent over F. Thus $\exists\ \alpha_0, \alpha_1, ..., \alpha_n \in F$ such that, $\alpha_0.1 + \alpha_1 a + ... + \alpha_n a^n = 0$ for some $\alpha_i \neq 0$.

Let $f(x) = \alpha_0 + \alpha_1 x + ... + \alpha_n x^n$. Then $f(x)$ is non-zero polynomial in $F[x]$ as some $\alpha_i \neq 0$. Also $f(a) = \alpha_0 + \alpha_1 a + ... + \alpha_n a^n = 0$

$\therefore$ a is algebraic over F.

$\therefore$ K is algebraic over F.

Note: Converse of Theorem 3.3 is not true.

Corollary: $a \in K$ is algebraic over F if $[F(a) : F]$ = Finite.

Proof: By Theorem 3.3, $F(a)$ is algebraic over F.

$\therefore$ $a \in F(a)$ is algebraic over F.

Converse of the above corollary is also true. But we will prove it after the next theorem.

Theorem 3.4: *Let $a \in K$ be algebraic over F. Then*

(*i*) $\exists$ *a unique monic irreducible polynomial* $p(x) \in F[x]$ *such that,* $p(a) = 0$

(*ii*) $\exists$ *non-zero polynomial* $q(x) \in F[x]$ *such that,* $q(a) = 0$, *then* $p(x)$ *divides* $q(x)$,

(*iii*) $F(a) = F[a]$.

Proof: (*i*) Since a is algebraic over F, $\exists$ a non-zero polynomial $f(x) \in F[x]$, such that,

$$f(a) = 0.$$

Let $t(x)$ be the non-zero polynomial of smallest degree over F such that, $t(a) = 0$ and suppose

$$t(x) = a_0 + a_1 x + ... + a_n x^n, \quad a_i \in F$$

If $t(x)$ is not monic [By monic polynomial, we mean a polynomial in which coefficient of highest degree term is 1], then let

$$p(x) = a_n^{-1} a_0 + a_n^{-1} a_1 x + ... + x^n = a_n^{-1} t(x)$$

Now deg $p(x) = n$ = deg $t(x)$ and $p(a) = 0$ and $p(x)$ is a monic polynomial. Thus $\exists$ a monic polynomial $p(x)$ of least degree such that, $p(a) = 0$.

Suppose $p(x) = p_1(x)p_2(x)$, where p_1 and p_2 are polynomials with lesser degree than deg p.

Then $$0 = p(a) = p_1(a)p_2(a)$$

$\Rightarrow$ $$p_1(a) = 0 \quad \text{or} \quad p_2(a) = 0 \quad [\text{as } F[a] \text{ is an I.D.}]$$

But that would lead to a contradiction as $p(x)$ is such polynomial with least degree.

Hence $p(x)$ is irreducible polynomial.

NOTES

To show uniqueness of $p(x)$, suppose $q(x)$ is any irreducible monic polynomial over F such that, $q(a) = 0$. Since $F[x]$ is a Euclidean domain, $\exists\ h(x)$ and $r(x)$ such that, $q(x) = p(x)h(x) + r(x)$

where either $r(x) = 0$ or $\deg r < \deg p$

Now, $0 = q(a) = p(a)h(a) + r(a)$

$\Rightarrow \quad r(a) = 0$ as $p(a) = 0$

Since $p(x)$ is of least degree such that, $p(a) = 0$, we find $\deg r < \deg p$ is not possible. Hence $r(x) = 0$

i.e., $q(x) = p(x)h(x)$ ...(3.1)

Since $q(x)$ is irreducible, $h(x)$ must be a constant polynomial, say $h(x) = c$

Then $q(x) = cp(x)$

Since $q(x)$ is monic, coefficient of highest degree term in L.H.S. is 1 and therefore it is 1 on R.H.S. also

R.H.S. being $cp(x) = ca_n^{-1}a_0 + ca_n^{-1}a_1x + ... + cx^n$ gives $c = 1$

Hence $q(x) = p(x)$, proving the uniqueness of $p(x)$

(*ii*) Follows by Equation (3.1)

(*iii*) Define a mapping $\theta : F[x] \to F[a]$, such that,

$$\theta(f(x)) = f(a)$$

then θ is onto homomorphism (verify!)

By fundamental theorem then

$$F[a] \cong \frac{F[x]}{\text{Ker } \theta}$$

Since $F[a]$ is an integral domain, so would be $\frac{F[x]}{\text{Ker } \theta}$ which implies Ker θ is a prime ideal. Since a is algebraic over K, $\exists$ a *non-zero polynomial* $f(x) \in F[x]$ such that, $f(a) = 0$.

$\Rightarrow \quad \theta(f(x)) = f(a) = 0$

$\Rightarrow \quad f(x) \in \text{Ker } \theta \Rightarrow \text{Ker } \theta \neq (0)$

i.e., Ker θ is a non-zero prime ideal of $F[x]$ which being a Euclidean domain is a PID.

Thus Ker θ is a maximal ideal.

$\Rightarrow \quad \frac{F[x]}{\text{Ker } \theta}$ is a field.

$\Rightarrow \quad F[a]$ is a field.

But $F(a)$ is the smallest field containing F and a and thus $F(a) \subseteq F[a]$

Also $F[a] \subseteq F(a)$

Hence $F(a) = F[a]$.

Note $F(a)$ is field of quotients of $F[a]$ and when $F[a]$ is itself a field, $F[a] = F(a)$.

NOTES

Note: 1. $p(x)$ determined in Theorem 3.4 is denoted by $p(x) = Irr(F, a)$. It is the unique monic irreducible polynomial over F satisfied by a. Since $p(x)$ is of least degree such that, $p(a) = 0$, $p(x)$ is called the minimal polynomial for a.

where $F(x) = \left\{ \dfrac{f(x)}{g(x)} \middle| f, g \in K[x], \underset{g \neq 0}{} \right\}$.

2. If $a \in K$ is transcendental over F then $F(x) \cong F(a)$.

Proof: Define $\varphi : F(x) \to F(a)$ such that,

$$\varphi\left(\frac{f(x)}{g(x)}\right) = \frac{f(a)}{g(a)},$$

Then φ is well defined onto homomorphism.

Also $\varphi\left(\dfrac{f(x)}{g(x)}\right) = 0$

$\Rightarrow$ $\dfrac{f(a)}{g(a)} = 0$

$\Rightarrow$ $f(a) = 0$

$\Rightarrow$ $f(x) = 0$, for otherwise a would be algebraic over F.

$\Rightarrow$ $\dfrac{f(x)}{g(x)} = 0$

$\Rightarrow$ φ is 1–1.

Hence $F(x) \cong F(a)$.

Corollary 1: Let $a \in K$ be algebraic over F. Then $[F(a) : F]$ = finite = deg Irr (F, a) and so $F(a)$ is an algebraic extension of F.

Proof: Let $p(x) = Irr\ (F, a)$. Let $n = \deg p(x)$.

We show that $1, a, a^2, ..., a^{n-1}$ form a basis of $F(a)$ over F.

Let $0 \neq f(a) \in F[a] = F(a)$. Then $f(x) \in F[x]$.

Now for $f(x)$, $p(x) \in F[x]$, $\exists\ q(x), r(x) \in F[x]$ such that,

$f(x) = p(x)q(x) + r(x)$ where either $r(x) = 0$

or $\deg r < \deg p$.

But $r(x) = 0 \Rightarrow f(x) = p(x)q(x)$

$\Rightarrow f(a) = p(a)q(a) = 0$ as $p(a) = 0$

which is not possible as $f(a) \neq 0$

Thus $r(x) \neq 0$. Hence $\deg r < \deg p$.

Suppose $r(x) = \beta_0 + \beta_1 x + ... + \beta_{n-1} x^{n-1}$, $\beta_i \in F$, where some β_i could be zero.

Again as $f(a) = p(a)\ q(a) + r(a)$ and $p(a) = 0$

we find $f(a) = r(a)$

Thus $f(a) = \beta_0 + \beta_1 a + \beta_2 a^2 + ... + \beta_{n-1} a^{n-1}$

i.e., $\{1, a, a^2, ..., a^{n-1}\}$ spans $F[a] = F(a)$ over F.

We show these are *L.I.*

Suppose these are *L.D.*, then $\exists\ \gamma_i$, not all zero, such that,

$$\gamma_0 + \gamma_1 a + \gamma_2 a^2 + ... + \gamma_{n-1} a^{n-1} = 0$$

$$\Rightarrow \quad t(x) = \gamma_0 + \gamma_1 x + ... + \gamma_{n-1} x^{n-1}$$

is non zero polynomial (some $\gamma_i \neq 0$) with $t(a) = 0$.

A contradiction to the fact that $p(x)$ is such polynomial with least degree. Hence $1, a, ..., a^{n-1}$ are Linearly Independent (L.I.) and thus form a basis of $F(a)$.

Hence $[F(a) : F] = n$.

3. Using Corollary to theorem 3.3 we conclude $a \in K$ is algebraic over F iff $[F(a) : F]$ = finite.

Definition: An element $a \in K$ is said to be *algebraic of degree n* over F if it satisfies a polynomial of degree n over F and does not satisfy any polynomial of lesser degree (than n).

Thus a is algebraic of degree n over F if deg *Irr* $(F, a) = n$. Also in that case, $[F(a) : F] = n$ and $\{1, a, a^2, ..., a^{n-1}\}$ is a basis of $F(a)$ over F.

Corollary 2: If $a_1, ..., a_n \in K$ are algebraic over F then $F(a_1, ..., a_n)$ is finite extension of F and so is algebraic over F.

Proof: We prove the result by induction on n. If $n = 1$, result follows from Corollary 1. Assume it to be true for naturals less than n. Let $a_1, ..., a_n \in K$ be algebrai cover F. Now a_n is algebraic over $F \Rightarrow a_n$ is algebraic over F $(a_1, ..., a_{n-1})$ as $F \subseteq F(a_1, a_2, ..., a_{n-1})$.

$\therefore$ By Corollary 1, $[F(a_1, ..., a_{n-1})(a_n) : F(a_1, ..., a_{n-1})]$ is finite. By induction hypothesis, $[F(a_1, ..., a_{n-1}) : F]$ is finite.

$\therefore$ $[F(a_1, ..., a_n) : F] = [F(a_1, ..., a_n) : F(a_1, ..., a_{n-1})]\ [F(a_1, ..., a_{n-1}) : F]$ = finite

$\therefore$ Result is true for n also.

By induction, result is true for all $n \geq 1$.

Corollary 3: If $a, b \in K$ are algebraic over F, then $a \pm b$, ab, ab^{-1} (if $b \neq 0$) are algebraic over F. In other words, the elements of K which are algebraic over F form a subfield of K (and this subfield is called the *algebraic closure* of F over K).

Proof: By Corollary 2, $F(a, b)$ is algebraic over F.

$\therefore$ $a \pm b, ab, ab^{-1} \in F(a, b)$ are algebraic over F.

NOTES

NOTES

Notes 1.: If K is an extension field of a field F and $S \subseteq K$, then

$$F(S) = \left\{ \frac{f(u_1, ..., u_n)}{g(u_1, ..., u_n)} \middle| \begin{array}{l} f, g \in F[x_1, ..., x_n] \\ g(u_1, ..., u_n) \neq 0, n \in \mathbf{N} \\ u_1, ..., u_n \in S \end{array} \right\}$$

Proof: Let L denote the R.H.S. We first show that L is a subfield of K.

Let $\frac{f(u_1, ..., u_m)}{g(u_1, ..., u_m)} \in L, \frac{f_1(v_1, ..., v_n)}{g_1(v_1, ..., v_n)} \in L$

Let $Y = \frac{f(u_1, ..., u_m)}{g(u_1, ..., u_m)} - \frac{f_1(v_1, ..., v_n)}{g_1(v_1, ..., v_n)}$

$$= \frac{f(u_1, ..., u_m)\, g_1(v_1, ..., v_n) - f_1(v_1, ..., v_n)\, g(u_1, ..., u_n)}{g(u_1, ..., u_m)\, g_1(v_1, ..., v_n)}$$

Define $h(x_1, ..., x_{m+n}) = f(x_1, ..., x_m)\, g_1(x_{m+1}, ..., x_{m+n})$
$- g(x_1, ..., x_m)\, f_1(x_{m+1}, ..., x_{m+n})$

$r(x_1, ..., x_{m+n}) = g(x_1, ..., x_m)\, g_1(x_{m+1}, ..., x_{m+n})$

Then $Y = \frac{h(u_1, ..., u_m, v_1, ..., v_n)}{r(u_1, ..., u_m, v_1, ..., v_n)} \in L$

Suppose $\frac{f_1(v_1, ..., v_n)}{g_1(v_1, ..., v_n)} \neq 0$

Let $Z = \frac{f(u_1, ..., u_m)}{g(u_1, ..., u_m)} \cdot \frac{g_1(v_1, ..., v_n)}{f_1(v_1, ..., v_n)}$

Define $h_1(x_1, ..., x_{m+n}) = f(x_1, ..., x_m)\, g_1(x_{m+1}, ..., x_{m+n})$;
$r_1(x_1, ..., x_{m+n}) = g(x_1, ..., x_m)\, f_1(x_{m+1}, ..., x_{m+n})$.

Then $Z = \frac{h_1(u_1, ..., u_m, v_1, ..., v_n)}{r_1(u_1, ..., u_m, v_1, ..., v_n)} \in L$

So, L is subfield of K.

Let $u_1 \in S$. Define $f(x) = x$, $g(x) = 1$.

Then $f(u_1) = u_1$, $g(u_1) = 1$

$$\Rightarrow \frac{f(u_1)}{g(u_1)} \in L \Rightarrow \frac{u_1}{1} \in L \Rightarrow u_1 \in L$$

So, $S \subseteq L$.

Let $\alpha \in F$. Define $f(x) = \alpha$, $g(x) = 1$.

Let $u \in S$. Then $f(u) = \alpha$, $g(u) = 1$.

Now, $\frac{f(u)}{g(u)} \in L \Rightarrow \frac{\alpha}{1} = \alpha \in L$.

So, $F \subseteq L$.

But $F(S)$ is the smallest field containing F and S, $F(S) \subseteq L$.

Let $Y \in L$. Then $Y = \frac{f(u_1,...,u_n)}{g(u_1,...,u_n)}$, $u_i \in S$.

Since $u_i \in S$ and coefficients in f, g belong to F, $f(u_1, ..., u_n) \in F(S)$ and $g(u_1, ..., u_n) \in F(S)$.

So, $Y \in F(S)$, then $L \subseteq F(S)$.

Hence $F(S) = L$.

2. If K is an extension field of F, and K is generated by algebraic elements (i.e., $K = F(S)$, where $S \subseteq K$ is a set of algebraic elements over K), then K is an algebraic extension of F.

Proof: Let $C \in K$, then $C = \frac{f(u_1, ..., u_n)}{g(u_1, ..., u_n)}$, $u_i \in S$.

where $f(x_1, ..., x_n)$, $g(x_1, ..., x_n) \in F[x_1, ..., x_n]$.

Clearly $C \in F(u_1, ..., u_n)$. But $u_1, ..., u_n$ are algebraic over $F \Rightarrow F(u_1, ..., u_n)$ is an algebraic extension of $F \Rightarrow C$ is algebraic over F.

Hence K/F is algebraic.

Theorem 3.5: *If L is an algebraic extension of K and K, an algebraic extension of F, then L is an algebraic extension of F.*

Proof: Let $a \in L$. Since L is algebraic over K, a is algebraic over K.

$\therefore \exists\ 0 \neq f(x) \in K[x]$ such that, $f(a) = 0$. Let $f(x) = \alpha_0 + \alpha_1 x + ... + \alpha_n x^n$, $\alpha_i \in K$.

Since K is algebraic over F, each $\alpha_i \in K$ is algebraic over F. By Corollary 2 Theorem 3.4, $[F(\alpha_0, \alpha_1, ..., \alpha_n) : F]$ = finite.

Let $M = F(\alpha_0, \alpha_1, ..., \alpha_n)$

Then $[M : F]$ is finite and so M is algebraic over F. Clearly, each $\alpha_i \in M$. Thus, $f(x) \in M[x]$.

i.e., a is algebraic over M.

By Corollary 1, $M(a)$ is finite extension of M.

$\Rightarrow$ $[M(a) : F] = [M(a) : M]\ [M : F]$ = Finite.

$\Rightarrow$ $M(a)$ is algebraic over F.

$\Rightarrow$ $a \in M(a)$ is algebraic over F.

Since a is an arbitrary element of L, L is an algebraic extension of F.

Definition: A complex number is said to be an *algebraic number* if it is algebraic over the field of rational numbers.

An algebraic number is said to be an *algebraic integer* if it satisfies an equation of the form $x^n + \alpha_1 x^{n-1} + ... + \alpha_n$ where $\alpha_1, ..., \alpha_n$ are integers (i.e., a monic polynomial over integers).

Example 3.3: *If a is any algebraic number, prove that* $\exists$ *a* positive *integer n such that na is an algebraic integer.*

NOTES

NOTES

Solution: Since a is an algebraic number, a is algebraic over the field of rationals. Thus $\exists$ a non-zero monic polynomial $f(x) \in \mathbf{Q}[x]$ such that, $f(a) = 0$, where $\mathbf{Q}$ = Field of rationals.

Let $f(x) = x^m + \alpha_1 x^{m-1} + \ldots + \alpha_{m-1} x + \alpha_m$, $\alpha_i \in \mathbf{Q}$

Let $\alpha_i = \frac{p_i}{q_i}$ where p_i, q_i are integers, $q_i > 0$

$\therefore \quad a^m + \frac{p_1}{q_1} a^{m-1} + \ldots + \frac{p_{m-1}}{q_{m-1}} a + \frac{p_m}{q_m} = 0$

Let $n = q_1 \ldots q_m$. Then n is a positive integer

and $na^m + p_1 q_2 \ldots q_m a^{m-1} + \ldots + p_m q_1 \ldots q_{m-1} = 0$

$\Rightarrow \quad n^m a^m + p_1 q_2 \ldots q_m a^{m-1} n^{m-1} + \ldots + p_m q_1 \ldots q_{m-1} n^{m-1} = 0$

$\Rightarrow$ na satisfies the polynomial

$$x^m + p_1 q_2 \ldots q_m x^{m-1} + \ldots + p_m q_1 \ldots q_{m-1} n^{m-1} = 0$$

where coefficients are integers.

$\therefore$ na is an algebraic integer.

Example 3.4: *If the rational number r is also an algebraic integer, prove that r must be an ordinary integer.*

Solution: Let $r = \frac{p}{q}$, where $q > 0$, $(p, q) = 1$

Since r is an algebraic integer

$$r^m + \alpha_1 r^{m-1} + \ldots + \alpha_{m-1} r + \alpha_m = 0$$

α_is are integers.

$\therefore$ — ——

$\therefore 2\cos\frac{m\pi}{180}$ is algebraic number for all integers m.

$\therefore \cos\frac{m\pi}{180}$ is algebraic number for all integers m.

$\therefore \cos m^{\circ}$ is algebraic number for all integers m.

Also $\cos\frac{m\pi}{180}$ and $\cos\frac{m\pi}{180} + i\sin\frac{m\pi}{180}$ is algebraic number $\Rightarrow i\sin\frac{m\pi}{180}$ is algebraic number $\Rightarrow \sin\frac{m\pi}{180}$ is algebraic number as i is also algebraic number $\Rightarrow$ $\sin m^{\circ}$ is algebraic number.

NOTES

Example 3.6: *Find a basis of* $\mathbf{Q}(\sqrt{3}, \sqrt{5})$ over $\mathbf{Q}$.

Solution: We have,

$$\begin{aligned}[\mathbf{Q}(\sqrt{3}, \sqrt{5}) : \mathbf{Q}] &= [\mathbf{Q}(\sqrt{3})(\sqrt{5}) : \mathbf{Q}] \\ &= [\mathbf{Q}(\sqrt{3})(\sqrt{5}) : \mathbf{Q}(\sqrt{3})]\,[\mathbf{Q}\sqrt{3} : \mathbf{Q}] \\ &= [L(\sqrt{5}) : L]\,[\mathbf{Q}(\sqrt{3}) : \mathbf{Q}] \text{ where } L = \mathbf{Q}(\sqrt{3}) \\ &= \deg Irr\,(L, \sqrt{5}) \times \deg Irr\,(\mathbf{Q}, \sqrt{3}) \\ &= \deg (x^2 - 5) \times \deg (x^2 - 3) \\ &= 2 \times 2 = 4.\end{aligned}$$

Thus basis has 4 elements.

Also if $[(F(a) : F)] = n$ then $1, a, a^2, ..., a^{n-1}$ is basis of $F(a)$ over F, and thus

Basis of $L(\sqrt{5})$ over L is $\{1, \sqrt{5}\}$

Basis of $\mathbf{Q}(\sqrt{3})$ over $\mathbf{Q}$ is $\{1, \sqrt{3}\}$

Thus basis of $[L(\sqrt{5}) : L]\,[\mathbf{Q}(\sqrt{3}) : \mathbf{Q}] = [(\mathbf{Q}(\sqrt{3}, \sqrt{5}) : \mathbf{Q}]$

is $1.1,\ 1.\sqrt{3},\ 1.\sqrt{5},\ \sqrt{3}\sqrt{5}$ [Refer Theorem 3.2]

i.e., $1,\ \sqrt{3}, \sqrt{5}, \sqrt{15}$.

Example 3.7: *Find the minimal polynomial for* $\sqrt{2} + \sqrt{3}$ *and use it to show that* $\mathbf{Q}(\sqrt{2}, \sqrt{3}) = \mathbf{Q}(\sqrt{2} + \sqrt{3})$. *Find a basis for* $\mathbf{Q}(\sqrt{2}, \sqrt{3})$.

Solution: Now, $(\sqrt{2} + \sqrt{3})^2 = 5 + 2\sqrt{6}$;

$(\sqrt{2} + \sqrt{3})^4 = 49 + 20\sqrt{6}$.

So, $(\sqrt{2} + \sqrt{3})^4 - 10(\sqrt{2} + \sqrt{3})^2 + 1 = 0$

Therefore, $a = \sqrt{2} + \sqrt{3}$ satisfies

$f(x) = x^4 - 10x^2 + 1$ over $\mathbf{Q}$.

Let $p(x) = Irr(\mathbf{Q}, a)$

Then $\sqrt{2}-\sqrt{3},\ -\sqrt{2}+\sqrt{3},\ -\sqrt{2}-\sqrt{3}$ are also roots of $p(x)$. So, degree of $p(x)$ is at least 4. But $f(a) = 0$ and $f(x) \in \mathbf{Q}[x]$

$\Rightarrow$ $p(x)$ divides $f(x)$

NOTES

$\Rightarrow$ $p(x) = f(x)$.

So, $f(x)$ is the minimal polynomial for $\sqrt{2}+\sqrt{3}$.

Therefore, $[\mathbf{Q}(\sqrt{2}+\sqrt{3}) : \mathbf{Q}] = 4$.

Also, $[\mathbf{Q}\sqrt{2} : \mathbf{Q}] = \deg Irr\ (\mathbf{Q}, \sqrt{2})$
$= \deg\ (x^2 - 2) = 2.$

Now, $\mathbf{Q} \subseteq \mathbf{Q}(\sqrt{2}) \subseteq \mathbf{Q}(\sqrt{2}, \sqrt{3})$.

Consider $g(x) = x^2 - 3 \in \mathbf{Q}(\sqrt{2})\ [x]$.

Then $g(\sqrt{3}) = 0$

$\therefore$ $\deg Irr\ (\mathbf{Q}(\sqrt{2}), \sqrt{3}) \leq \deg g(x) = 2$

$\Rightarrow$ $[(\mathbf{Q}(\sqrt{2}, \sqrt{3}) : \mathbf{Q}] \leq 2$.

So, $[\mathbf{Q}(\sqrt{2}, \sqrt{3}) : \mathbf{Q}] \leq 4$.

Clearly, $\mathbf{Q} \subseteq \mathbf{Q}(\sqrt{2} + \sqrt{3}) \subseteq \mathbf{Q}(\sqrt{2}, \sqrt{3})$.

$\therefore$ $[\mathbf{Q}(\sqrt{2}, \sqrt{3})] : \mathbf{Q}] = [\mathbf{Q}(\sqrt{2}, \sqrt{3}) : \mathbf{Q}(\sqrt{2} + \sqrt{3})]$
$\times [\mathbf{Q}(\sqrt{2} + \sqrt{3}) : \mathbf{Q}]$

$\Rightarrow$ $[\mathbf{Q}(\sqrt{2}, \sqrt{3})] : \mathbf{Q}(\sqrt{2} + \sqrt{3})] = 1$

$\Rightarrow$ $\mathbf{Q}(\sqrt{2}, \sqrt{3}) = \mathbf{Q}(\sqrt{2} + \sqrt{3})$

Since $[\mathbf{Q}(\sqrt{2} + \sqrt{3}) : \mathbf{Q}] = 4$

$\{1, \sqrt{2} + \sqrt{3}, (\sqrt{2} + \sqrt{3})^2, (\sqrt{2} + \sqrt{3})^3\}$ is a basis for $\mathbf{Q}(\sqrt{2} + \sqrt{3}) = \mathbf{Q}(\sqrt{2}, \sqrt{3})$ over $\mathbf{Q}$.

Example 3.8 : *Let $F(x)$ be the field of rational functions in an indeterminate x. Show that every element of $F(x)$ which is not in F is transcendental over F.*

Solution: Let $0 \neq \frac{f}{g} \in F(x), \frac{f}{g} \notin F,\ \ (f, g) = 1.$

Suppose $\frac{f}{g}$ is not transcendental over F.

Then $\frac{f}{g}$ is algebraic over F.

So $F\left(\frac{f}{g}\right) = F\left[\frac{f}{g}\right]$.

Consider $\frac{g}{f} \in F\left[\frac{f}{g}\right] = F\left(\frac{f}{g}\right)$.

NOTES

(Note $0 \neq \frac{f}{g} \in F\left[\frac{f}{g}\right]$ and $F\left[\frac{f}{g}\right]$ is a field, $\frac{g}{f} \in F\left(\frac{f}{g}\right) = F\left[\frac{f}{g}\right]$)

Therefore, $\frac{g}{f} = \alpha_0 + \alpha_1 \left(\frac{f}{g}\right) + ... + \alpha_n \left(\frac{f}{g}\right)^n$, $\alpha_i \in F$.

So, $g^{n+1} = (\alpha_0 g^n + \alpha_1 f g^{n-1} + ... a_n f^n)\ f$.

Since $(f, g) = 1$, $f \mid g^{n+1} \Rightarrow f \mid g \Rightarrow f =$ unit

$\Rightarrow$ $g =$ unit $\Rightarrow \frac{f}{g} =$ unit $\in F$, a contradiction.

So, $\frac{f}{g}$ is transcendental over F.

Example 3.9: *Let K be an extension of F and let a ∈K. Then F[a] can be regarded as a vector space over F. If the dimension of F[a] over F is finite, show that F[a] = F(a).*

Solution: Let $0 \neq c \in F[a]$. Define

$T : F[a] \to F[a]$ such that,

$T(b) = bc$

Then T is a linear transformation.

Let $b \in$ Ker T. Then $T(b) = 0 \Rightarrow bc = 0 \Rightarrow b = 0$ as $c \neq 0$ and $F[a]$ is an integeral domain.

Thus Ker $T = \{0\}$ forcing T to be 1–1.

Since $F[a]$ is a FDVS over F, T is also onto.

Now $1 \in F[a] \Rightarrow \exists\ b \in F[a]$ st., $T(b) = 1$

$\Rightarrow bc = 1$ or that c is invertible.

So $F[a]$ is a field containing F and a. But $F(a)$ is the smallest field containing F & a and so $F(a) \subseteq F[a]$, However $F[a] \subseteq F(a)$ giving $F[a] = F(a)$.

Example 3.10 : *Let K be an extension of F. Show that K/F is algebraic if and only if every ring R, such that, $F \subseteq R \subseteq K$ is a field.*

Solution: Let K/F be algebraic and let R be a ring such that, $F \subseteq R \subseteq K$.

Since $R \subseteq K$, R will be commutative and also unity of K will be unity of R as $F \subseteq R \subseteq K$.

Let $0 \neq a \in R$, then $a \in K \Rightarrow a^{-1} \in K$

K/F algebraic $\Rightarrow$ a is algebraic over F

$\Rightarrow$ $\exists\ 0 \neq f(x) \in F(x)$ such that, $f(a) = 0$

Let $f(x) = \alpha_0 + \alpha_1 x + ... + \alpha_n x^n$, $\alpha_i \in F$

Then $\alpha_0 + \alpha_1 a + \alpha_2 a^2 + ... + \alpha_n a^n = 0$ with some $\alpha_i \neq 0$. Suppose $\alpha_o \neq 0$

Then $\alpha_0 a^{-1} = -(\alpha_1 + \alpha_2 a + ... + \alpha_n a^{n+1}) \in R$

NOTES

$\Rightarrow \quad a^{-1} \in R$ as $\alpha_0^{-1} \in F \subseteq R$

So, every non zero element is invertible in R.

Conversely, let $a \in K$. Let $R = F[a]$, then R is a ring such that, $F \subseteq R \subseteq K$. By hypothesis R is a field.

Suppose $a \neq 0$, then $a^{-1} \in R = F[a]$

Thus $a^{-1} = \alpha_0 + \alpha_1 a + ... + \alpha_n a^n, \quad \alpha_i \in F$

Let $f(x) = \alpha_0 + \alpha_1 x + ... + \alpha_n x^n \in F[x]$

Now $1 = \alpha_0 a + \alpha_1 a^2 + ... + \alpha_n a^{n+1}$

gives $\alpha_0 a + \alpha_1 a^2 + ... + \alpha_n a^{n+1} - 1 = 0$

showing that a satisfies $x f(x) - 1 \in F[x]$.

Clearly $x f(x) - 1$ is a non zero polynomial.

Hence, a is algebraic over F and so K/F is algebraic.

3.3.1 Separable and Inseparable Extensions

This section deals with those polynomials which have simple roots and the fields generated by these simple roots.

A root α of $f(x) \in K[x]$ is called simple if $x - \alpha$ divides $f(x)$ and $(x - \alpha)^2$ does not divide $f(x)$. Similarly, a root α of $f(x) \in K[x]$ is said to be a root with multiplicity m, if $(x - \alpha)^m$ divides $f(x)$ but $(x - \alpha)^{m+1}$ does not divide $f(x)$.

Let $f(x) = a_0 + a_1 x + ... + a_n x^n \in K[x]$.

Define $f'(x) = a_1 + 2a_2 + ... + n a_n x^{n-1} \in K[x]$.

Then $f'(x)$ or f' is called the *derivative* of f.

If $f, g \in K[x]$, then it can be easily proved that

(*i*) $(f \pm g)' = f' \pm g'$

(*ii*) $(fg)' = f'g + fg'$

(*iii*) $(af)' = a f', \ a \in K$

(*iv*) $x' = 1$.

It can be easily checked that α is a simple root of $f(x) \in K[x]$ iff $f'(\alpha) \neq 0$. In other words, α is not a simple root of $f \in K[x]$ iff $f'(\alpha) = 0$.

Theorem 3.6: *Suppose all roots of $f(x) \in K[x]$ in a minimal spitting field of f over K are simple. Then the roots of f in any minimal splitting field of f over K are simple.*

Proof: Let $f(x) = \alpha_0(x - \alpha_1) ... (x - \alpha_n), \quad \alpha_i \in E$.

where $E = K(\alpha_1, ..., \alpha_n)$ is a minimal splitting field of f over K.

Suppose each α_i is a simple root of f.

Let E' be another minimal splitting field of f over K.

Then $E' = K(\beta_1, ..., \beta_n)$ where $\beta_i s$ are roots of f.

Then there exists a K-isomorphism $\sigma : E \to E'$.

Since α_i is a root of f, $\sigma(\alpha_i)$ is also a root of f in E'.

Therefore $\{\sigma(\alpha_1), ..., \sigma(\alpha_n)\} = \{\beta_1, ..., \beta_n\}$.

Since σ is 1-1 and $\alpha_i s$ are distinct roots of f, $\sigma(\alpha_i)s$ are all distinct. So $\beta_i s$ are all distinct.

Thus, the roots of f in E' are also simple roots.

NOTES

Note: By the above arguments, we can also prove that if there is a root of multiplicity m in a minimal splitting field of f over K, then every minimal splitting field of f over K will have a root of f of multiplicity m.

Theorem 3.7: *Let F be an extension of K. Let $f, g \in K[x]$. Then the g.c.d. of f and g regarded as polynomials in $K[x]$ is same as that of f and g regarded as polynomials in $F[x]$, upto associates.*

Proof: Let d be the g.c.d. of $f, g \in K[x]$ and d_1 be the g.c.d. of $f, g \in F[x]$.

Now $\quad d \mid f, d \mid g$ in $K[x] \Rightarrow d \mid f, d \mid g$ in $F[x]$

$\Rightarrow \quad d \mid d_1$ in $F[x] \quad \Rightarrow d_1 = du, \quad u \in F[x]$.

Also, $\quad d = ff_1 + gg_1, \quad f, f_1 \in K[x]$.

Since $\quad d_1 \mid f, d_1 \mid g, d_1 \mid ff_1, d_1 \mid gg_1$.

Therefore, $\quad d_1 \mid ff_1 + gg_1 = d$ in $F[x]$.

$\Rightarrow \quad d = d_1 v \quad v \in F[x]$.

So, $d = duv \Rightarrow uv = 1 \Rightarrow u, v$ are units in $F \Rightarrow d, d_1$ are associates. Thus d and d_1 are same upto associates.

Theorem 3.8: *Let F be an extension of K.Then f and g are relatively prime regarded as elements of $K[x]$ iff f and g are relatively prime regarded as elements of $F[x]$.*

Proof: Suppose f and g are relatively prime regarded as elements of $F[x]$.

Then (f, g) = g.c.d. of $f, g \in F[x]$ is a unit $d \in F$.

Let (f, g) = g.c.d. of $f, g \in K[x]$ be d_1

Then d and d_1 are associates

$\Rightarrow \quad d = ud_1$, u = Unit in F

$\Rightarrow \quad d_1 = u^{-1} d$ = Unit in F

Since $d_1 \in K$, d_1 is a unit in K.

The converse follows similarly.

Theorem 3.9: *Let F be an extension of K. Let $f(x) \in K[x]$, $\alpha \in F$. Then f can be written as $f = (x - \alpha)^2 g + (x - \alpha) f'(\alpha) + f(\alpha)$ for some $g \in F[x]$.*

Proof: Now $(x - \alpha) \in F[x]$.

Let $\quad f = (x - \alpha)^2 g + h, \quad g, h \in F[x]$

and $\quad h = (x - \alpha) g_1 + h_1, \quad g, h_1 \in F$

So, $\quad f(\alpha) = h(\alpha) = h_1$ (deg $h < 2$)

NOTES

Also, $f' = 2(x-\alpha)g + (x-\alpha)^2 g' + h'$

and $h' = g_1$

$\Rightarrow \quad f'(\alpha) = h'(\alpha) = g_1.$

Theorem, $f = (x-\alpha)^2 g + (x-\alpha) f'(\alpha) + f(\alpha).$

Theorem 3.10: *Let $f \in K[x]$. Then the roots of f are simple iff f and f' are relatively prime.*

Proof: Suppose the roots of f are simple. Let $(f, f') = d$.

If d is a non-constant polynomial in $K[x]$, then d has a root α in some extension F of K.

Now $f = df_1, f' = dg_1, \quad f_1, g_1 \in K[x]$

$\Rightarrow \quad f(\alpha) = d(\alpha) f_1(\alpha), \quad f'(\alpha) = d(\alpha) g_1(\alpha)$

$\Rightarrow \quad f(\alpha) = 0 = f'(\alpha).$

Using above result, we get

$$f = (x-\alpha)^2 g + (x-\alpha) f'(\alpha) + f(\alpha)$$
$$= (x-\alpha)^2 g$$

$\Rightarrow \alpha$ is not a simple root of f, a contradiction.

So, d = constant $\in K$.

Since $f \neq 0$, d is a non zero element in K.

Therefore, d is a unit $\Rightarrow f, f'$ are relatively prime.

Conversely, let f and f' be relatively prime. Then $(f, f') = d$ = unit in K.

Let α be a root of f' such that α is not a simple root of f. Let $\alpha \in F \supseteq K$.

Then $f(\alpha) = 0 = f'(\alpha)$

$\Rightarrow \quad x - \alpha$ divides f and f' in $F[x] \supseteq K[x]$

$\Rightarrow \quad x - \alpha$ divides d

But $d \in K \Rightarrow \deg d = 0 \ (d \neq 0)$.

and $x - \alpha$ divides d

$\Rightarrow \deg (x-\alpha) \leq \deg d = 0$, a contradiction.

So all roots of f are simple.

Definition: A polynomial is said to be *separable* if all its roots are simple. In view of the above theorem, the following result follows.

Theorem 3.11: *A polynomial $f(x) \in F[x]$ is separable iff f and f' are relatively prime.*

Corollary 1: If $f(x) \in F[x]$ is irreducible over F such that, $f' \neq 0$, then f is separable.

Proof: Let g.c.d. $(f, f') = d$ then $\deg d \leq \deg f' < \deg f$.

Since f is irreducible over F and d is a factor of f such that $\deg d < \deg f$, we find d is (non zero) constant and thus a unit. So, f and f' are relatively prime.

By above theorem, f is separable.

Corollary 2: Let $f(x) \in F[x]$ be irreducible over F. If characteristic of F is zero, then f is separable. (In other words, an irreducible polynomial over a field of characteristic zero is separable).

NOTES

Proof: Let $f = a_0 + a_1x + ... + a_nx^n \in F[x]$.

Then $f' = a_1 + 2a_2 + ... + na_nx^{n-1}$.

If $f' = 0$, then $ra_r = 0$ for all $r = 1, 2, ..., n$. Since char $F = 0$, $a_r = 0$ for all $r = 1, 2, ..., n \Rightarrow f = a_0$, a contradiction as F is irreducible (deg $f \geq 1$).

Thus, $f' \neq 0$. By Corollary 1, f is separable.

Theorem 3.12: *Let F be a field of characteristic p. Then for any polynomial* $f(x) \in F[x]$, $f' = 0$ *iff* $f(x) = g(x^p)$ *for some polynomial* $g(x) \in F[x]$.

Proof: Let $f(x) = a_0 + a_1x + ... + a_nx^n$ and $f' = 0$.

Then $ra_r = 0 \quad \forall\ r = 1, 2, ..., n$.

$\Rightarrow$ $a_r = 0$ or p divides r as char $F = p$.

Thus, $f = a_0 + a_px^p + ... + a_{sp}x^{sp}$

$= g(x^p)$, where $g(x) = a_0 + a_px + ... + a_{sp}x^s \in F[x]$.

Conversely, let $f = g(x^p)$, where

$g(x) = b_0 + b_1x + ... + b_nx^n \in F[x]$

Then, $f = b_0 + b_1x^p + ... + b_n\, x^{np}$

$\Rightarrow$ $f' = pb_1x^{p-1} + ... + npb_nx^{np-1} = 0$ as $pa = 0\ \forall\ a \in F$.

Theorem 3.13: *Let* $f(x) \in F[x]$ *be irreducible over F. Then all its roots have the same multiplicity.*

Proof: (*i*) Let char $F = 0$. Then by corollary 2 to Theorem 3.11 f is separable. So, all roots of f are simple.

(*ii*) Let char $F = p$. If $f' \neq 0$, then by corollary 1 to Theorem 3.11 f is separable. So, all roots of f are simple.

If $f' = 0$, then $f(x) = g(x^p)$, for some $g \in F[x]$.

Since f is irreducible over F, so is g over F.

If $g' \neq 0$, then g is separable over F. Let α be a root of f.

Then $0 = f(\alpha) = g(\alpha^p) \Rightarrow g(x) = Irr\ (F, \alpha^p)$.

Now, $g(x) = (x - \alpha^p)\ h(x)$, $h(\alpha^p) \neq 0$ as α^p is a simple root of $g(x)$.

So, $f(x) = g(x^p) = (x^p - \alpha^p)\ h(x^p)$

$= (x - \alpha)^p\ h_1(x)$

$[h_1(x) = h(x^p) \Rightarrow h_1(\alpha) = h(\alpha^p) \neq 0]$

$\Rightarrow$ $x - \alpha$ appears exactly p times in $f(x)$.

This is true for all roots of $f(x)$.

NOTES

If $g' = 0$, then $g(x) = q(x^p)$, $q \in F[x]$

$\Rightarrow \quad f(x) = q(x^{p^2})$.

Proceeding in this way, since, deg f is finite, after finite number of steps we get $f(x) = r(x^{p^e})$, $r' \neq 0$. Then r is separable over F and every root of f appears exactly p^e times.

Hence all roots of f have same multiplicity p^e $(e \geq 0)$.

Aliter: Let α be a root of f of multiplicity m.

Then $f(x) = (x - \alpha)^m \; g(x)$, $g(\alpha) \neq 0 \quad g(x) \in K[x]$, $K = k(\alpha)$

Let β the another root of f. Then $\exists$ an F-isomorphism

$\sigma : F(\alpha) \to F(\beta)$ such that,

$$\sigma(\alpha) = \beta$$

Now $\quad f = \sigma(f) = (x - \beta)^m \; \sigma(g(x))$

Let $\quad g(x) = a_0 + a_1 x + ... + a_n x^n$, $a_i \in K$

Then $\quad \sigma(g(x)) = \sigma(a_0) + \sigma(a_1)\, x + ... + \sigma(a_n)\, x^n$

$\Rightarrow \quad \sigma(g(\beta)) = \sigma(a_0) + \sigma(a_1)\beta + ... + \sigma(a_n)\, \beta^n$

$= \alpha(a_0) + \alpha(a_1)\, \sigma(\alpha) + ... + \sigma(a_n)\, \sigma(\alpha^n)$

$= \sigma(a_0 + a_1\alpha_1 + ... + a_n\alpha^n)$

$= \sigma(g(\alpha)) \neq 0$ as $g(\alpha) \neq 0$

$\Rightarrow$ β is a root of f of multiplicity m, showing that all roots of f have same multiplicity.

Corollary: If $f \in F_p[x]$ is irreducible over F_p and f is not separable, then p divides n, where $n = \deg f$. (F_p deotes the field $\{0, 1, 2, ..., p - 1\}$ mod p).

Proof: By above theorem, all roots of f have same multiplicity p^e, $e > 0$ as f is not separable.

So, $\quad \deg f = rp^e$

$\Rightarrow \quad p$ divides $n = \deg f$. (Note, char $F_p = p$).

Theorem 3.14: *Let $x^p - a \in F[x]$, where p = char F. Then either $x^p - a$ is irreducible over F or $x^p - a$ is a p-th power of a linear polynomial in F.*

Proof: Let $\quad f(x) = x^p - a$.

If b is a root of $f(x)$, then $f(b) = 0 \Rightarrow a = b^p$.

$\Rightarrow \quad f(x) = x^p - b^p = (x - b)^p$.

If $b \in F$, then $f(x)$ is p-th power of linear polynomial $x - b \in F[x]$.

Suppose $b \notin F$. Let $p(x)$ be a monic irreducible factor of $f(x)$ in $F[x]$.

Since $p(x)$ divides $f(x)$, $p(x) = (x - b)^m$ for some m, $1 \leq m \leq p$.

So, $\quad p(b) = 0$. Thus, $p(x) = Irr\ (F, b)$.

If $q(x)$ is another monic irreducible factor of $f(x)$ in $F[x]$, then

$$q(x) = Irr\ (F, b) = p(x).$$

So, $f(x) = (p(x))^r$.

Since $\deg f = p$, $p = rm$.

If $r > 1$, then $m = 1 \Rightarrow p(x) = x - b \in F[x] \Rightarrow b \in F$, a contradiction.

So, $r = 1 \Rightarrow f(x) = p(x)$ is irreducible over F.

NOTES

Case 1: We give an example of an irreducible polynomial which does not have distinct roots.

Let $K = F_2(t)$, $F_2 = \{0, 1\}$ mod 2 and t is an indeterminate over F_2. Let $f(x) = x^2 - t \in K[x]$.

If f is reducible over K, then there would be an element $a \in K$ such that, $f(a) = 0$.

$$\Rightarrow \quad t = a^2. \text{ But } a \in K \Rightarrow a = \frac{g(t)}{h(t)}.$$

$$\text{So,} \quad t = \frac{(g(t))^2}{(h(t))^2} \Rightarrow \deg (g(t))^2 = \deg t(h(t))^2.$$

$\Rightarrow \quad 2 \deg g(t) = \deg t + 2 \deg h(t) = 1 + 2 \deg h(t)$, which is not true.

So, f is irreducible over K.

If α is a root of f, then $f'(\alpha) = 0$ (as char K = 2 = char F_2) $\Rightarrow \alpha$ is not a simple root of f.

So, $f = (x - \alpha)^2$.

Thus, f is an irreducible polynomial having no simple roots.

Definition: Let F be an algebraic extension of K. Then $a \in F$ is called separable over K if $Irr\ (K, a)$ is separable.

Thus, $a \in F$ is separable over K iff a is a simple root of $Irr\ (K, a)$. Further, if each $a \in F$ is separable over K, then F is called a separable extension of K. (We write F/K is separable).

In the case above, $x^2 - t = Irr\ (K, \alpha)$ and α is not a simple root of $x^2 - t$.

If F is a minimal splitting field of $f = x^2 - t$ over K, containing α then F/K is algebraic and $\alpha \in F$ is not separable over K.

So, $F = K(\alpha)$ is not separable over F.

However, if char $K = 0$ then every algebraic extension of K is separable by Corollary 2. to Theorem 3.11.

Theorem 3.15: *Let char $K = p$. Then every algebraic extension of K is separable iff $K = K^p$.*

Proof: Suppose every algebraic extension of K is separable. Let $a \in K$. Let $f(x) = x^p - a$ and b be a zero of $f(x)$. Then $0 = f(b) = b^p - a \Rightarrow a = b^p \Rightarrow f(x) = x^p - b^p = (x - b)^p$. If $b \notin K$ then $f(x)$ is irreducible over K.

So, $x^p - a = Irr(K, b)$.

But $f(x) = x^p - a$

$\Rightarrow \quad f'(x) = px^{p-1}$

NOTES

$\Rightarrow \quad f'(b) = 0$ as char $K = p$

$\Rightarrow b$ is not a simple root of $f(x)$

$\Rightarrow K(b)/K$ is not separable, contradicting the given fact that every algebraic extension of K is separable.

So, $b \in K$ and $a = b^p \in K^p \Rightarrow K \subseteq K^p$.

However, $K^p \subseteq K$. So $K = K^p$ (Note, $K^p = \{a^p \mid a \in K\}$).

Conversely, let $K = K^p$. Let F/K be algebraic.

Let $\alpha \in F$, $f(x) = Irr\ (K, \alpha)$. If f is not separable, then $f' = 0$. So, $f = g(x^p)$ for some $g \in K[x]$.

Let $g = a_0 + a_1x + ... + a_nx^n,\ a_i \in K.$

Then $f = g(x^p) = a_0 + a_1x^p + ... + a_nx^{np}$

Since $K = K^p,\ a_i = b_i^p,\ b_i \in K.$

So, $f = b_0^p + b_1^p x^p + ... + b_n^p x^{np}$

$= (b_0 + b_1x + ... + b_nx^n)^p,\ b_i \in K$

contradicting that f is irreducible over K.

Thus f is separable $\Rightarrow \alpha$ is separable.

Since α is an arbitrary element of F, F/K is separable.

3.4 PERFECT FIELDS

Definition: *A* field K is called *perfect field* if every algebraic extension of K is separable.

A field of characteristic zero is perfect by Corollary 2 to Theorem 3.11. So, **Q**, **R**, **C**, are perfect fields.

Theorem 3.16: *Let char $K = p$. Then the following are equivalent:*

(*i*) *K is perfect.*

(*ii*) $K = K^p$

(*iii*) *Every element in K is a p-th power of some element in K.*

(*iv*) $\theta : K \to K$ *such that* $\theta(a) = a^p$ *is an automorphism.*

Proof: (*i*) $\Rightarrow$ (*ii*) follows by Theorem 3.15

(*ii*) $\Rightarrow$ (*iii*) obvious

(*iii*) $\Rightarrow$ (*iv*): Since char $K = p$, θ is clearly a homomorphism and is 1-1.

Also, $b \in K \Rightarrow b = a^p,\ a \in K$ by (*iii*).

$\Rightarrow b = \theta(a) \Rightarrow \theta$ is onto. So, θ is an automorphism.

(*iv*) $\Rightarrow$ (*i*): Now $\theta(K) = \{\theta(a) \mid a \in K\}$

$= \{a^p \mid a \in K\}$

$= K^p.$

Since θ is onto, $K = K^p$.

By Theorem 3.15 then K is perfect.

Theorem 3.17: *Let $F \subseteq K \subseteq L$ be a tower of fields. Suppose L/F is separable. Then L/K is separable.*

NOTES

Proof: Let $a \in L, p(x) = Irr(K, a)$

$q(x) = Irr(F, a)$

Then $q(x) \in K[x]$ and $q(a) = 0$.

So, $p(x)$ divides $q(x)$ in $K[x]$

$\Rightarrow \quad q(x) = p(x)\, r(x), \quad r(x) \in K[x]$

$\Rightarrow \quad q'(x) = p'(x)\, r(x) + p(x)\, r'(x)$

$\Rightarrow \quad q'(a) = p'(a)\, r(a).$

Since L/F is separable, a is separable over F.

So a is a simple root of $q(x) \Rightarrow q'(a) \neq 0$

$\Rightarrow \quad p'(a) \neq 0 \Rightarrow a$ is a simple root of $p(x)$

$\Rightarrow \quad a$ is separable over K

$\Rightarrow \quad L$/K is separable.

Corollary: Every finite extension of a perfect field is perfect.

Proof: Let F be a perfect field. Let K/F be finite extension. Then K/F is algebraic. Let L/K be algebraic. Then L/F is algebraic. Since F is perfect, L/F is separable. From above, L/K is separable. So, K is perfect.

Example 3.11: *Let F be a perfect field. Show that the set of elements fixed under all automorphisms of F is a perfect subfield.*

Solution: Let char $F = p$, $K = \{a \in F \mid \sigma(a) = a \; \forall \; \sigma \in G\}$, where G is the group of all automorphisms of F. Then K is subfield of F.

Define $\theta : F \to F$ such that,

$\theta(\alpha) = \alpha^p$

Then θ is a homomorphism. Since F is perfect, θ is onto. So, $\theta \in G$.

Let $\alpha \in K$. Then $\sigma(\alpha) = \alpha \quad \forall \; \sigma \in G$

$\Rightarrow \quad \theta(\alpha) = \alpha \Rightarrow \alpha^p = \alpha \Rightarrow \alpha \in K^p \Rightarrow K \subseteq K^p.$

$\Rightarrow \quad K = K^p \Rightarrow K$ is perfect.

Example 3.12: Let *K/F be a finite extension and suppose K is perfect then show that F is perfect.*

Solution: Let char $F = p$, then char $K = p$.

Let $[K : F] = n$ and $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ be a basis of K over F.

Since K is perfect, $K = K^p$. We show $F = F^p$.

Now, $F^p \subseteq F \subseteq K$. So we show that

$[K : F^p] = [K : F]$ which would give $F = F^p$.

NOTES

Let $S = \{\alpha_1^p, \alpha_2^p, ..., \alpha_n^p\} \subseteq K^p = K$

If $a_1^p\alpha_1^p + a_2^p\alpha_2^p + ... + a_n^p\alpha_n^p = 0$, $a_i \in F$

then $(a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n)^p = 0$

$\Rightarrow \quad a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n = 0$

$\Rightarrow \quad a_i = 0 \quad \forall\ i$

$\Rightarrow S$ is L.I. set in K (over F^p)

Let $b \in K$, then $b = a^p$, $a \in K$ as $K = K^p$

Now, $a \in K \Rightarrow a = b_1\alpha_1 + b_2\alpha_2 + ... + b_n\alpha_n$, $b_i \in F$

$\Rightarrow \quad b = a^p = b_1^p\ \alpha_1^p + b_2^p\ \alpha_2^p + ... + b_n^p\ \alpha_n^p$

$\Rightarrow$ S spans K over F^p

Hence S is a basis of K over F^p

$\Rightarrow \quad [K : F^p] = o(S) = n = [K : F]$

$\Rightarrow \quad F = F^p$ or that F is perfect.

3.4.1 Normal Extensions

If $f(x) \in K[x]$ is irreducible over K, then $\exists$ an extension E of K containing a root of $f(x)$. In this section we consider those extensions of K which contain all roots of $f(x)$ and study properties of such extensions.

Definition: Let E be an extension of K. E is called *normal extension* of K if (*i*) E/K is algebraic (*ii*) $\alpha \in E \Rightarrow p(x) = Irr\ (K, \alpha)$ splits in $E[x]$ or E.

Case 2: A quadratic extension is a normal extension.

Let E be a quadratic extension of K. Then $[E : K] = 2$.

Since E/K is finite, E/K is algebraic.

Let $\alpha \in E$, $\quad p(x) = Irr\ (K, \alpha)$.

Now $\quad K \subseteq K(\alpha) \subseteq E$. Since $2 = [E : K] = [E : K(\alpha)]\ [K(\alpha) : K]$.

Either $\quad [E : K(\alpha)] = 1$ or $[K(\alpha) : K] = 1$.

If $\quad [K(\alpha) : K] = 1$, then $K(\alpha) = K \Rightarrow \alpha \in K$

$\Rightarrow \quad p(x) = x - \alpha$ splits in $K[x] \subseteq E[x]$.

If $\quad [E : K(\alpha)] = 1$, then $E = K(\alpha)$.

So, $\quad 2 = [E : K] = [K(\alpha) : K] = \deg Irr\ (K, \alpha) = \deg p(x)$.

Now α is a root of $p(x) \Rightarrow x - \alpha$ divides $p(x)$ in $E[x]$.

$\Rightarrow \quad p(x) = (x - \alpha)\ q(x), \quad q(x) \in E[x]$.

Since $\quad \deg p(x) = 2$, $\deg q(x) = 1$. So $q(x) = (x - \beta)$, $\beta \in E$.

Therefore, $\quad p(x) = (x - \alpha)\ (x - \beta)$ splits in $E[x]$.

Thus, E/K is normal.

Case 3: Let $f(x) = x^3 - 2 \in \mathbf{Q}[x]$. Let α be the real root of $f(x)$. Consider $\mathbf{Q}(\alpha)/\mathbf{Q}$. We show that $\mathbf{Q}(\alpha)/\mathbf{Q}$ is not normal.

NOTES

Now $f(x)$ is irreducible over **Q** by Eisenstein's criterion (take $p = 2$). So, $f(x) = Irr\ (\mathbf{Q}, \alpha)$.

Since α is algebraic over **Q**(being root of $f(x) \in \mathbf{Q}[x]$), $\mathbf{Q}(\alpha)/\mathbf{Q}$ is algebraic.

If $f(x)$ splits in $\mathbf{Q}(\alpha)$, then $\mathbf{Q}(\alpha)$ contains a minimal splitting field E of $f(x)$ over **Q**.

So, $\mathbf{Q} \subseteq E \subseteq \mathbf{Q}(\alpha)$.

But $[E : \mathbf{Q}] = 6$ and $[\mathbf{Q}(\alpha) : \mathbf{Q}] = \deg Irr\ (\mathbf{Q}, \alpha) = \deg f(x) = 3$.

Since $3 = [\mathbf{Q}(\alpha) : \mathbf{Q}] \geq [E : \mathbf{Q}] = 6$, we get a contradiction.

So, $\mathbf{Q}(\alpha)/\mathbf{Q}$ is not normal.

Similarly, $\mathbf{Q}(\alpha w)/\mathbf{Q}$ and $\mathbf{Q}(\alpha w^2)/\mathbf{Q}$ are not normal extensions.

Note: We have seen in above case that an extension of degree 3 need not be normal. We can, however, have a normal extension of degree 3. Consider $f(x) = x^3 + x^2 + 1 \in F_2[x]$, where $F_2 = \{0, 1\}$ mod 2. Let α be a root of $f(x)$. Then α^2, $1 + \alpha + \alpha^2$ are also roots of $f(x)$. So $F_2(\alpha)$ is a minimal splitting field of $f(x)$ over F_2. Thus $F_2(\alpha)/F_2$ is normal and $[F_2(\alpha) : F_2] = \deg \text{Irr}\ (F_2, \alpha) = \deg f(x) = 3$.

Theorem 3.18: *Let $F \subseteq K \subseteq E$ be a tower of fields. If E/F is normal, then so is E/K.*

Proof: Since E/F is normal, E/F is algebraic. So, E/K is algebraic.

Let $\alpha \in E$, $p(x) = Irr\ (K, \alpha)$, $q(x) = Irr(F, \alpha)$.

Then $q(x) \in F[x] \subseteq K[x] \Rightarrow q(x) \in K[x]$ and $q(\alpha) = 0$.

So, $p(x)$ divides $q(x)$ in $K[x]$.

Since E/F is normal and $\alpha \in E$, $q(x)$ splits in $E[x]$.

So, $p(x)$ splits in $E[x]$. Thus, E/K is normal.

Note: In above theorem K/F need not be normal. Consider $f(x) = x^3 - 2 \in \mathbf{Q}[x]$. Let $\alpha \in \mathbf{R}$ be a root of $f(x)$. Then $\mathbf{Q}(\alpha)/\mathbf{Q}$ is not normal by Case 3. However, $\mathbf{Q}(\alpha, w)/\mathbf{Q}$ is normal by Theorem 3.19 and $\mathbf{Q} \subseteq \mathbf{Q}(\alpha) \subseteq \mathbf{Q}(\alpha, w)$. Notice $\mathbf{Q}(\alpha, w)$ is a minimal splitting field of $f(x)$ over **Q**.

Theorem 3.19: *A minimal splitting field of a non-constant polynomial $f(x) \in K[x]$ over K is normal extension of K.*

Proof: Let E be a minimal splitting field of $f(x)$ over K. Then E/K is algebraic and finite. Let $f(x) = \alpha_0(x - \alpha_1) \ldots (x - \alpha_n)$, $\alpha_i \in E$.

Then $E = K(\alpha_1, \alpha_2, \ldots, \alpha_n)$

Let $\alpha \in E$, $p(x) = Irr\ (K, \alpha) \in K[x] \subseteq E[x]$.

Then $p(x)$ splits in some extension of E.

Let β be a root of $p(x)$ in some extension of E. We show that $\beta \in E$.

Now α, β are roots of $p(x) \Rightarrow \exists$ a K-isomorphism $\sigma : K(\alpha) \to K(\beta)$ such that, $\sigma(\alpha) = \beta$.

NOTES

Then, a minimal splitting field of f over $K(\alpha)$ is $K(\alpha)$ $(\alpha_1, \alpha_2, ..., \alpha_n)$

$= K(\alpha_1, \alpha_2, ..., \alpha_n)$ (α)

$= E(\alpha)$

$= E$ as $\alpha \in E$

Also, a minimal splitting field of $\sigma(f) = f$ over $K(\beta)$ is

$K(\beta)$ $(\alpha_1, \alpha_2, ..., \alpha_n)$

$= K(\alpha_1, \alpha_2, ..., \alpha_n)$ (β)

$= E(\beta)$.

So, $\exists$ an isomorphism $\theta : E \to E(\beta)$ such that, $\theta(a) = \sigma(a)$ $\forall$ $a \in K(\alpha)$

$\Rightarrow$ $\theta(\alpha) = \sigma(\alpha) = \beta$.

Now, $K \subseteq K(\alpha) \subseteq E \subseteq E(\beta)$

$\Rightarrow$ $[E : K(\alpha)] = [\theta(E) : \theta(K(\alpha)]$

$= [E(\beta) : \sigma(K(\alpha)]$

$= [E(\beta) : K(\beta)]$

So, $[E(\beta) : K] = [E(\beta) : K(\beta)]\ [K(\beta) : K]$

$= [E : K(\alpha)]$ deg $p(x)$

$= [E : K(\alpha)]\ [K(\alpha) : K]$

$= [E : K]$.

Since $E \subseteq E(\beta)$ and E, $E(\beta)$ as vector spaces over K have same dimension, $E = E(\beta)$. So, $\beta \in E$. Thus, $p(x)$ splits in E. This proves E/K is normal.

Theorem 3.20: *A finite normal extension is a minimal splitting field of some polynomial.*

Proof: Let E/K be a finite normal extension.

E/K is finite $\Rightarrow E = K(\alpha_1, \alpha_2, ..., \alpha_n)$.

Let $p_i(x) = Irr\ (K, \alpha_i)$. Since $\alpha_i \in E$ and E/K is normal, each $p_i(x)$ splits in E.

Let $f = p_1 p_2 \ ... \ p_n \in K[x]$.

Then, a minimal splitting field of f over K is

$K(\alpha_1, \alpha_2, ..., \alpha_n,$ roots of $p_i s$ in $E) = E$.

So, E is a minimal splitting field of f over K.

Corollary: Let $K \subseteq E_1 \subseteq E$, $K \subseteq E_2 \subseteq E$ be towers of fields such that, E_1/K, E_2/K are finite normal extensions. Then E_1E_2, the smallest subfield of E containing $E_1 \cup E_2$ is finite normal extension of K.

Proof: Since E_1/K is finite, $E_1 = K(\alpha_1, ..., \alpha_n)$.

So, $E_1E_2 = K(\alpha_1, ..., \alpha_n)E_2$

$= E_2(\alpha_1, ..., \alpha_n)$, as $K \subseteq E_2$

$\Rightarrow$ $KE_2 = E_2$

NOTES

Thus $[E_1E_2: E_2] = [E_2(\alpha_1, ..., \alpha_n) : E_2]$

$= [E_2(\alpha_1, ..., \alpha_n) : E_2(\alpha_1, ..., \alpha_{n-1})] .... [E_2(\alpha_1) : E_2]$

$\leq [K(\alpha_1, ..., \alpha_n) : K(\alpha_1, ..., \alpha_{n-1})] ... [K(\alpha_1) : K]$

$= [K(\alpha_1, ..., \alpha_n) : K]$

$= [E_1 : K]$.

Therefore,

$[E_1 E_2 : K] = [E_1E_2 : E_2] [E_2 : K]$

$\leq [E_1 K] [E_2 : K]$ = Finite

$\Rightarrow \quad [E_1E_2 : K]$ = Finite.

Now E_1/K is finite normal $\Rightarrow E_1$ is a minimal splitting field of f_1 over K

Also, E_2/K is finite normal $\Rightarrow E_2$ is a minimal splitting field of f_2 over K

Let $f = f_1 f_2$, $E_1 = K(a_1, ..., a_r)$, $E_2 = K(b_1, ..., b_s)$.

Then, a minimal splitting field of f over K is $K(a_1, ..., a_r, b_1, ..., b_s)$

$= E_1(b_1, ..., b_s)$

$= E_1K(b_1, ..., b_s)$ as $E_1K = E_1$

$= E_1E_2$.

Thus, E_1E_2/K is finite normal extension.

(Note, we have also shown above that E_1/K, E_2/K are finite $\Rightarrow E_1E_2/K$ is finite).

Case 4: We now give an example to show that a normal extension of a normal extension need not be a normal extension.

Consider the tower of fields $\mathbf{Q} \subseteq \mathbf{Q}(\sqrt{2}) \subseteq \mathbf{Q}(2^{1/4})$.

Now $[\mathbf{Q}(\sqrt{2}) : \mathbf{Q}] = \deg Irr\ (\mathbf{Q}, \sqrt{2}) = \deg(x^2 - 2) = 2$

and $[\mathbf{Q}(2^{1/4}) : \mathbf{Q}(2^{1/2})] = \deg Irr(\mathbf{Q}(\sqrt{2}), 2^{1/4}) = \deg (x^2 - \sqrt{2}) = 2$

So, $\mathbf{Q}(\sqrt{2})/\mathbf{Q}$, $\mathbf{Q}(2^{1/4})/\mathbf{Q}(\sqrt{2})$ are normal.

If $\mathbf{Q}(2^{1/4})/\mathbf{Q}$ is normal, then

$f(x) = Irr(\mathbf{Q}, 2^{1/4}) = x^4 - 2$ must split in $\mathbf{Q}(2^{1/4})$.

So, $\mathbf{Q}(2^{1/4})$ contains a minimal splitting field E of $f(x)$.

But $[E : \mathbf{Q}] = 8$ and $\mathbf{Q} \subseteq E \subseteq \mathbf{Q}(2^{1/4})$

$\Rightarrow \quad [\mathbf{Q}(2^{1/4}) : \mathbf{Q}] = 4 \geq [E : \mathbf{Q}] = 8$, a contradiction.

Therefore, $\mathbf{Q}(2^{1/4})/\mathbf{Q}$ is not normal, proving our assertion.

Theorem 3.21: *Let $K \subseteq F \subseteq E$ be a tower of fields such that, E/K is finite normal. Then any K-homomorphism of F into E can be extended to K-automorphism of E.*

Proof: Since E/K is finite, $E = K(\alpha_1, \alpha_2, ..., \alpha_n)$. Also E/K is finite normal $\Rightarrow E$ is a minimal splitting field of some $f(x) \in K[x]$ over K. Let σ be a K-

NOTES

homomorphism of F into E. Then σ is a K-isomorphism from F onto $\sigma(F) = F'$.
$f = p_1 p_2 \ldots p_n$, where $p_i = Irr\ (K, \alpha_i)$ splits in E.

So, a minimal splitting field of f over F is

$F(\alpha_1, \alpha_2, \ldots, \alpha_n$, roots of f in $E)$

$= E$(roots of f in E) $= E$

$(E = K(\alpha_1, \alpha_2, \ldots, \alpha_n) \subseteq F(\alpha_1, \alpha_2, \ldots, \alpha_n) \subseteq E$

$\Rightarrow \quad E = F(\alpha_1, \alpha_2, \ldots, \alpha_n))$

Also, a minimal splitting field of $\sigma(f(x)) = f$ over F' is

$F'(\alpha_1, \alpha_2, \ldots, \alpha_n$, roots of f in $E)$

$= E$(roots of $f(x)$ in E)

$= E$

$[E = K(\alpha_1, \alpha_2, \ldots, \alpha_n) \subseteq F(\alpha_1, \alpha_2, \ldots, \alpha_n)$

$\cong F'(\alpha_1, \alpha_2, \ldots, \alpha_n)$

$\subseteq F(\alpha_1, \alpha_2, \ldots, \alpha_n) = E$

$\Rightarrow \quad E = F'(\alpha_1, \alpha_2, \ldots, \alpha_n)]$

Therefore, $\exists$ an isomorphism $\theta : E \to E$ such that,

$\theta(a) = \sigma(a)\ \forall\ a \in F$

$\Rightarrow \quad \theta(\alpha) = \sigma(\alpha) = \alpha\ \forall\ \alpha \in K \Rightarrow \theta$ is a K-automorphism of E extending σ. This proves the result.

Normal Closure: Let E/K be a finite extension. Then $E = K(\alpha_1, \alpha_2, \ldots, \alpha_n)$.

Let $p_i = Irr(K, \alpha_i)$ and $f = p_1 p_2 \ldots p_n \in K[x]$.

Then E' the minimal splitting field of f over K is

$K(\alpha_1, \ldots, \alpha_n$, root of f in some extension of $E)$

$= E$(roots of f in some extension of E)

$\Rightarrow E \subseteq E'$ and E'/K is finite normal

(as a minimal splitting field of f over K is finite normal extension of K)

Suppose $K \subseteq E \subseteq F$ such that, F/K is finite normal.

We show that E' can be embedded in F.

$\alpha_i \in E \subseteq F \Rightarrow \alpha_i \in F\ \forall\ i$. Also F/K is normal.

So, $p_i(x)$ splits in $F[x]\ \forall\ i \Rightarrow f$ splits in $F[x]$

$\Rightarrow F$ contains a minimal splitting field E_1 of f over K.

$\Rightarrow E_1 \subseteq F$. But E' is also a minimal splitting field of f over K.

Therefore, $E' \cong E_1 \subseteq F \Rightarrow E'$ can be embedded in F.

Thus, E' is the least finite normal extension of K such that, $K \subseteq E \subseteq E'$.

E' is called the normal closure of E/K.

Case 5: Let $f(x) = x^3 - 2$

$= (x - \alpha)\ (x - \alpha w)\ (x - \alpha w^2)$

We find the normal closure of $\mathbf{Q}(\alpha)/\mathbf{Q}$.

Now $[\mathbf{Q}(\alpha) : \mathbf{Q}] = \deg f(x) = 3.$

where $f = Irr\ (\mathbf{Q}, \alpha).$

Then, a minimal splitting field of f over $\mathbf{Q}$ is $\mathbf{Q}(\alpha, \alpha w, \alpha w^2) = \mathbf{Q}(\alpha, w)$.

So, $\mathbf{Q}(\alpha w)/\mathbf{Q}$ is the normal closure of $\mathbf{Q}$.

NOTES

3.4.2 Finite Fields

A field having finite number of elements is called a finite field or a Galois field.

Theorem 3.22: *If F is a finite field, then $o(F) = p^n$ for some prime p and an integer $n \geq 1$.*

Proof: Let P be the prime subfield of F.

Since F is finite, so is P. Therefore, $P \cong \dfrac{\mathbf{Z}}{< p >}$ for some prime p.

But $\dfrac{\mathbf{Z}}{< p >} \cong \{0, 1, 2, ..., p-1\} \bmod p = F_p \Rightarrow P \cong F_p$.

Since $P \subseteq F$, we can regard $F_p \subseteq F$. Now F is a vector space over F_p. Since F is finite, $[F : F_p] = n$ = finite.

Let $\{u_1, ..., u_n\}$ be a basis of F/F_p.

Then $F = \{\alpha_1 u_1 + ... + \alpha_n u_n \mid \alpha_i \in F_p\}$.

Now each α_i can be chosen in p ways and $\Sigma\alpha_i u_i = \Sigma\beta_i u_i \Rightarrow \alpha_i = \beta_i$, therefore $o(F) = p^n$.

Theorem 3.23: *Let p be a prime and $n \geq 1$ be an integer. Then there exists a field with p^n elements.*

Proof: Let $f(x) = x^q - x \in F_p[x]$, $q = p^n$. Let F be a minimal splitting field of $f(x)$ over F_p.

Then $F = F_p$ (zeros of f in F).

Let $S = \{\text{zeros of } f \text{ in } F\}$.

Now $f' = qx^{q-1} - 1 = -1$ as char $F = p$

$\Rightarrow$ $q - 1 = p^n - 1 = -1.$

Therefore, $(f, f') = 1$

$\Rightarrow$ all zeros of f in F are simple and so distinct.

So, $o(S) = q.$

Now $0 \in S \Rightarrow S \neq \varphi.$

Also $a, b \in F_q \Rightarrow a^q = a,\ b^q = b$

$\Rightarrow$ $(a \pm b)^q = a^q \neq b^q = a \pm b,$

$(ab)^q = a^q b^q = ab,\ (ab^{-1})^q = a^q b^{-q} = ab^{-1}$

$\Rightarrow a \pm b, ab, ab^{-1}$ (if $b \neq 0$) $\in S.$

Thus, S is a subfield of F.

NOTES

Let $a \in F_p$. Then $a^{p-1} = 1$

$\Rightarrow \quad a^p = a \Rightarrow a^{p^n} = a \Rightarrow a^q = a.$

$\Rightarrow$ a is a zero of f in $F \Rightarrow a \in S \Rightarrow F_p \subseteq S$.

So S is a field containing F_p and S.

But F is the smallest field containing F_p and S.

$\Rightarrow F \subseteq S$. Also $S \subseteq F$. So, $S = F \Rightarrow o(F) = o(S) = q$.

We now prove the following results from group theory.

Lemma 1: *Let G be an abelian group under multiplication. Let a, b ∈ G be such that o(a) = m, o(b) = n and (m, n) = 1. Then o(ab) = mn*

Proof: Now $(ab)^{mn} = a^{mn}\, b^{mn} = (a^m)^n (b^n)^m = 1$ = identity of G

If $(ab)^t = 1$, then $a^t b^t = 1$.

$\Rightarrow \quad a^t = b^{-t} \Rightarrow a^{mt} = b^{-mt} \Rightarrow b^{-mt} = 1$

$\Rightarrow \quad b^{mt} = 1 \Rightarrow o(b) \mid mt \Rightarrow n \mid mt \Rightarrow n \mid t$ as $(n, m) = 1$.

Similarly, $m \mid t$. So, $mn \mid t \Rightarrow t \geq mn \Rightarrow o(ab) = mn$.

Lemma 2: *Let G be an abelian group under multiplication. Let a, b ∈ G be such that, o(a) = m, o(b) = n. Then there exists c ∈ G such o(c) = l.c.m. of m and n.*

Proof: Let $(m, n) > 1$.

Let $\quad m = p_1^{\alpha_1} \ldots p_r^{\alpha_r}$

$\quad n = p_1^{\beta_1} \ldots p_r^{\beta_r}$

where $p_1, \ldots, p_r$ are distinct primes and α_i, β_i are non negative integers.

Let $\quad l = p_1^{\alpha_1} \ldots p_s^{\alpha_s}\; p_{s+1}^{\beta_{s+1}} \ldots p_r^{\beta_r}$

where $\alpha_i \geq \beta_i$ for $i = 1, \ldots, s$ and $\beta_j \geq \alpha_j$ for $j = s + 1, \ldots, r$.

Then l is the l.c.m of m and n.

Let $\quad x = a^{p_{s+1}^{\alpha_{s+1}}} \ldots p_r^{\alpha_r}$, $y = b^{p_1^{\beta_1}} \ldots p_s^{\beta_s}$

Then $\quad o(x) = p_1^{\alpha_1} \ldots p_s^{\alpha_s}$

$\quad o(y) = p_{s+1}^{\beta_{s+1}} \ldots p_r^{\beta_r}$

and $(o(x)), o(y)) = 1$.

By Lemma 1,

$\quad o(xy) =$ l.c.m. of m and n

$\quad = p_1^{\alpha_1} \ldots p_s^{\alpha_s}\; p_{s+1}^{\beta_{s+1}} \ldots p_r^{\beta_r}$.

Lemma 3: *With the hypothesis of lemma 2, if n ∤ m, then the l.c.m. l of m and n is greater than m.*

Proof: Now $m \mid l \Rightarrow m \leq l$. If $m = l$, then $n \mid l \Rightarrow n \mid m$, a contradiction. So $l > m$.

NOTES

Lemma 4: *Let G be a finite abelian group under multiplication. Let* $\alpha \in G$ *be of maximum order. Then* $o(\beta) \mid o(\alpha)$ *for all* $\beta \in G$.

Proof: Let $o(\alpha) = m$, $o(\beta) = n$.

Suppose $n \nmid m$. By lemma 3, l = l.c.m. of $m, n > m$. By Lemma 2, there is $\gamma \in G$ such that $o(\gamma) = l > m$ contradicting $\alpha \in G$ is of maximum order. So, $n \mid m \Rightarrow o(\beta) \mid o(\alpha)$ for all $\beta \in G$.

Theorem 3.24: *Let F be a finite field. Then F*, the set of non zero elements of F forms a cyclic group under multiplication in F.*

Proof: Now F^* is an abelian group under multiplication.

Let $\alpha \in F^*$ be an element of maximum order m.

Then by Lemma 4, $o(\beta) \mid m$ for all $\beta \in F^*$.

So, $m = o(\beta)r$

$\Rightarrow \beta^m = \beta^{o(\beta)r} = 1$ for all $\beta \in F^*$.

$\Rightarrow \beta$ satisfies $x^m - 1$ over F.

Since F can't have more than m zeros of $x^m - 1$, $o(F^*) \leq m$.

But $\alpha \in F^*$ and $o(\alpha) = m$

$\Rightarrow 1, \alpha, \alpha^2, ..., \alpha^{m-1}$ are distinct elements of F^*

$\Rightarrow \quad o(F^*) \geq m \Rightarrow o(F^*) = m = o(\alpha) \Rightarrow F^* = <\alpha>$.

The generators of F^* are called *primitive elements* of F.

Theorem 3.25: *Let F be a finite field of order* p^n. *Then F is a minimal splitting field of* $x^{p^n} - x$ *over* F_p.

Proof: We can regard F as an extension of F_p. Let $q = p^n$.

Now $\quad F^* = <\alpha>$, $o(\alpha) = o(F^*) = q - 1$. Also $\alpha^{q-1} = 1$.

$\Rightarrow \quad \alpha^q = \alpha$.

$\Rightarrow$ Elements of F are zeros of $f(x) = x^q - x$ over F_p.

So, $f(x)$ splits in F.

Therefore, $\quad f(x) = x(x - \alpha) ... (x - \alpha^{q-1})$

$\Rightarrow$ Minimal splitting field of f over F_p is $F_p(\alpha, \alpha^2, ..., \alpha^{m-1}, 1, 0) = F_p(F) = F$.

Theorem 3.26: *Any two finite fields with the same number of elements* p^n *are* F_p*-isomorphic.*

Proof: Let F_1, F_2 be finite fields such that $o(F_1) = p^n = o(F_2)$. Then, by above theorem F_1, F_2 are minimal splitting fields of $f(x) = x^{p^n} - x$ over $F_p \Rightarrow F_1, F_2$ are F_p-isomorphic.

The above theorem shows that there is unique field of order $q = p^n$ upto an isomorphism. It is denoted by $GF(p^n)$ or $GF(q)$ or Fq.

NOTES

Example 3.13: *Show that $x^m - 1$ divides $x^n - 1$ over a field F if and only if m divides n.*

Solution: Let $n = km + r$, $0 \le r < m$.

The $x^n - 1 = x^r\left(\sum_{i=0}^{k-1} x^{im}\right)(x^m - 1) + (x^r - 1)$.

Therefore, $x^m - 1$ divides $x^n - 1$ if and only if $x^r - 1 = 0$.

Also $x^r - 1 = 0$ if and only if $r = 0$.

So $x^m - 1$ divides $x^n - 1$ if and only if m divides n.

Example 3.14: *Show that $x^{p^m} - x$ divides $x^{p^n} - x$ if m divides n.*

Solution: Let $n = mu$.

Then $p^n - 1 = p^{mu} - 1$

$= (p^m)^u - 1$

$= (p^m - 1)$ (integer)

$\Rightarrow$ $p^m - 1$ divides $p^n - 1$

By above problem

$x^{p^m - 1} - 1$ divides $x^{p^n - 1} - 1$

$\Rightarrow$ $x^{p^m} - x$ divides $x^{p^m} - x$.

Theorem 3.27: *Let F be a field with p^n elements. Then F has a subfield k with p^m elements if and only if m divides n.*

Proof: Suppose k is a subfield of F. Then k can be regarded as an extension of F_p such that $[k : F_p] = m$. Similarly, F can be regarded as an extension of F_p such that $[F : F_p] = n$. Now $[F : F_p] = [F : k]\,[k : F_p] \Rightarrow m$ divides n.

Conversely, let F be a field such that, $o(F) = p^n$. Suppose m divides n. Now F is a minimal splitting field of $x^{p^n} - x$ over F_p.

Let $f(x) = x^{p^n} - x$ and $g(x) = x^{p^m} - x$.

Since m divides n, by above problem $g(x)$ divides $f(x)$.

Consider $F' = \{$zeros of $g(x)$ in $F\}$.

Then F' is a subfield of F.

Since $g(x)$ has p^m distinct zeros, F' is a subfield of F with p^m elements.

If k is another subfield of F such that $o(k) = p^m$, then $o(k) = o(F') = p^m$.

$\Rightarrow k, F'$ are F_p-isomorphic.

Thus, there is exactly one subfield of F (up to isomorphism) with p^m elements.

Example 3.15: *Determine the algebraic closure of F_p.*

Solution: We know $m!$ divides $n!$ for all positive integers $m < n$. By above theorem $F_p m!$ is a subfield of $F_p n!$. Thus, there is an ascending chain of subfields

$$F_p \subseteq F_{p^{2!}} \subseteq F_{p^{3!}} \subseteq \ldots$$

and $F_{p^\infty} = \cup_n F_{p^{n!}}$ is a field such that $F_{p^n} \subseteq F_{p^{n!}} \subseteq F_{p^\infty}$ for any positive integer n.

Let S be the set of all polynomials over F_p. Let $f \in S$.

Then the minimal splitting field of f over F_p is a finite field F_{p^n}.

So, each $f \in S$ splits in F_{p^∞}.

Thus, the minimal splitting field of S over F_p is

F_p (zeros of $f \in S$ in F_{p^∞}) $\subseteq F_{p^\infty}$.

Also, $a \in F_{p^\infty} \Rightarrow a \in F_{p^n}$ for some $n \Rightarrow a$ is zero of $x^{p^n} - x$ over F_p.

Now $f = x^{p^n} - x \in S \Rightarrow a$ is zero of $f \in S$ in F_{p^∞}

$\Rightarrow$ $F_{p^\infty} \subseteq F_p$ (zeros of $f \in S$ in F_{p^∞})

$\Rightarrow$ Minimal splitting field of S over F_p is F_{p^∞}

$\Rightarrow$ F_{p^∞} is the algebraic closure of F_p.

Theorem 3.28: *Every finite extension of a finite field is Galois.*

Proof: Let K be a finite extension of a finite field k. Then K is also a finite field. So, char k = char $K = p$, for some prime p. Let $o(k) = p^m$, $o(K) = p^n$.

Now K is a minimal splitting field of $x^{p^n} - x$ over $F_p \Rightarrow K/F_p$ is finite normal.

Also F_p is finite $\Rightarrow F_p$ is perfect $\Rightarrow$ every algebraic extension of F_p is separable $\Rightarrow K/F_p$ is separable $\Rightarrow K/F_p$ is Galois. Now, $F_p \subseteq k \subseteq K$ and K/F_p is Galois $\Rightarrow K/k$ is Galois.

Corollary: F_q/F_p is Galois, $q = p^n$.

Theorem 3.29: *Let F be a finite field. Then there exists an irreducible polynomial of any given degree n over k.*

Proof: Let $o(F) = p^m$, p being a prime.

Let $q = p^{nm}$ and let $f(x) = x^q - x$

Then F_q is the minimal splitting field of $f(x)$ over F_p.

Since m/nm, $F_{p^m} = F$ can be imbedded in F_q.

Now $F_p \subseteq F = F_{p^m} \subseteq F_{p^{mn}} = E$.

Then $[E : F] = n$.

Let E^* be the multiplicative group of non zero elements of E and let $E^* = < \alpha >$

Then $E = F(\alpha)$ as $F \subseteq E$, $\alpha \in E$

So, $n = [E : F] = [F(\alpha) : F] = \deg Irr\ (F, \alpha)$

$\Rightarrow Irr\ (F, \alpha)$ is an irreducible polynomial of degree n over F.

Theorem 3.30: *Let G be the group of F_p-automorphisms of F_q. Then G is a cyclic group generated by Frobenius map of order n, where $q = p^n$.*

NOTES

NOTES

Proof: Let $\theta : F_q \to F_q$ such that,

$\theta(b) = b^p$.

Then θ is called Frobenius map.

Since char F_p = char $F_q = p$, θ is a homomorphism.

Also θ is 1-1.

Since F_q is finite, θ is onto.

If $b \in F_p$, then $b^p = b$

$\Rightarrow \quad \theta(b) = b$ for all $b \in F_p$.

So, θ is an F_p-automorphism of $F_q \Rightarrow \theta \in G$.

By Artin's theorem, $o(G) = [F_q : F_p]$ as F_p is the fixed field of G.

$\Rightarrow \quad o(G) = n$. We show that $o(\theta) = n$.

Let $\theta^r = I$, let $F_q^* = < a >$.

Then $a^{q-1} = 1 \Rightarrow a^q = a \Rightarrow a^{p^n} = a$.

Now, $\theta^r = I \Rightarrow \theta^r(a) = a \Rightarrow a^{p^r} = a \Rightarrow a^{p^r - 1} = 1$.

$\Rightarrow \quad o(a) \mid p^r - 1 \Rightarrow q - 1 \mid p - 1$

$\Rightarrow \quad p^n - 1 \mid p^r - 1 \Rightarrow p^n - 1 \le p^r - 1 \Rightarrow n \le r$.

Also $\theta^r(b) = b^{p^n} = b$ for all $b \in F_p \Rightarrow \theta^n = I$.

So, $o(\theta) = n \Rightarrow G = < \theta >$.

Example 3.16: *Show that for any integer a and prime p, $a^p \equiv a \pmod p$.*

Solution: Let $a = pq + r, \quad 0 \le r \le p$.

Then $a \equiv r \pmod p$

Now, $0 \le r < p \Rightarrow r \in F_p$

$\Rightarrow r_o r_{o \ldots o} r = r$

p *times*

$\Rightarrow \quad r^p - pu = r$

$\Rightarrow \quad r^p \equiv r \pmod p$

$\Rightarrow \quad r^p \equiv a \pmod p$

So, $a \equiv r \pmod p$

$\Rightarrow \quad a^p \equiv r^p \pmod p$

$\Rightarrow \quad a^p \equiv a \pmod p$

(The above result is known as Fermat's theorem)

Example 3.17: *Show that every irreducible polynomial $f(x) \in F_p[x]$ is a divisor of $x^{p^n} - x$ for some n.*

Solution: Let deg $f(x) = d$ and α be a zero of $f(x)$ in an extension of F_p.

Then, $[F_p(\alpha) : F_p] = \deg Irr(F_p, \alpha) = \deg f(x) = d$.

So, $o(F_p(\alpha)) = p^d$. Then $\alpha \in F_p(\alpha)$

$\Rightarrow \quad \alpha^{p^d} = \alpha \Rightarrow \alpha$ is zero of $x^{p^d} - x \in F_p[x]$

$\Rightarrow f(x)$ divides $x^{p^d} - x$.

Example 3.18: *Show that $x^{p^n} - x$ is the product of monic irreducible polynomials in $F_p[x]$ of degree d, d dividing n.*

Solution: Let $f(x) = x^q - x$, $q = p^n$. Let $p(x)$ be a monic irreducible factor of $f(x)$ over F_p. Let α be a zero of $p(x)$ in F, where F is a minimal splitting field of $f(x)$ over F_p. Then $F = F_q$ and $p(x) = Irr\ (F_p, \alpha)$

Now $\quad F_p \subseteq F_p(\alpha) \subseteq F_q$

and $\quad n = [F_q : F_p] = [F_q : F_p(\alpha)]\ [F_p(\alpha) : F_p]$

$= [F_q : F_p(\alpha)]$ deg $Irr\ (F_q, \alpha)$

$= [F_q : F_p(\alpha)]$ deg $p(x)$

$\Rightarrow$ deg $p(x)$ divides n.

$\Rightarrow$ Any monic irreducible polynomial dividing $x^{p^n} - x$ is of degree dividing n.

Example 3.19: *Contruct a field of order* 9.

Solution: Let F_9 be the field of order 9. Let $F_3 = \{0, 1, 2\}$ mod 3. Then $[F_9 : F_3] = 2$. Let $f(x) = x^9 - x$. Then F_9 is a minimal splitting field of $f(x)$ over F_3. Let $p(x)$ be an irreducible factor of $f(x)$ over F_3. Let α be a zero of $p(x)$ in F_9. Then α is a zero of $f(x)$. If $\alpha \in F_3$, then $p(x) = x - \alpha \Rightarrow$ deg $p(x) = 1$. If $\alpha \notin F_3$, then $F_3 \subseteq F_3(\alpha) \subseteq F_9 \Rightarrow [F_9 : F_3] = 2 = [F_9 : F_3(\alpha)]\ [F_3(\alpha) : F_3]$.

Since $\alpha \notin F_3$, $[F_3(\alpha) : F_3] \neq 1$

$\Rightarrow \quad [F_3(\alpha) : F_3] = 2$

But $\quad [F_3(\alpha) : F_3] = \deg Irr\ (F_3, \alpha)$

$= \deg p(x)$

Thus $\quad \deg p(x) = 2$.

Hence any irreducible factor of $f(x)$ over F_3 has degree 1 or 2.

Now, $\quad x^9 - x = x(x^8 - 1)$

$= x(x^4 - 1)\ (x^4 + 1)$

$= x(x - 1)\ (x + 1)\ (x^2 + 1)\ (x^2 - x - 1)\ (x^2 + x - 1)$

Note, $x^2 + 1$, $x^2 - x - 1$, $x^2 + x - 1$ are irreducible over F_3 as none of 0, 1, 2 are zeros of these factors.

Let $\quad p(x) = x^2 + 1$. Let α be a zero of $p(x)$.

Then $\{1, \alpha\}$ is a basis of $F_9 = F_3(\alpha)$ over F_3.

So, $\quad F_9 = \{a + b\alpha \mid a, b \in F_3\}$

$= \{0, 1, 2, \alpha, \alpha + 1, \alpha + 2, 2\alpha, 2\alpha + 1, 2\alpha + 2\}$.

Let $u = \alpha + 1$. Then $u^2 = 2\alpha$, $u^4 = -1$, $u^8 = 1$. So, $o(u) = 8$

$\Rightarrow F_9^* = < u >$.

NOTES

NOTES

Therefore, $F_9 = \{0, 1 = u^8, 2 = u^4, \alpha = u^6, \alpha + 1 = u,$
$\alpha + 2 = u^7, 2\alpha = u^2, 2\alpha + 1 = u^3, 2\alpha + 2 = u^5\}$

Now multiplication is defined by element u^i in F_9. We wish to define addition in F_9 with the help of u^i.

If $u^n + 1 \neq 0$, let $u^n + 1 = u^{z(n)}$.

Define $u^a + u^b = u^{z(a-b)+b}$ if $u^{a-b} + 1 \neq 0$ where $a \geq b$

$= 0$ if $u^{a-b} + 1 = 0$

Let's find $u^7 + u^1$

Now $u^6 + 1 = \alpha + 1 = u^1 \neq 0$.

So, $\mathbf{z}(6) = 1$. Therefore, $u^7 + u^1 = u^{\mathbf{Z}(6)+1} = u^2$

Also, $u^6 + u^2 = 0$ as $u^4 + 1 = -1 + 1 = 0$. In this way addition is defined in terms of u^i.

Let $a = u^i$. Then write $\log a = i$. If $b = u^j$, then $ab = u^{i \oplus j}$, where $\oplus$ denotes the addition modulo 9.

So, $\log ab = i \oplus j = \log a \oplus \log b$.

Such a logarithm is known as *Zech logarithm.*

3.4.3 Algebraically Closed Fields

In this section, we give a characterization of normal extensions. Also, we show that given a tower of fields $k \subseteq F \subseteq K$ such that K/k is normal, any k-homomorphism of F into K can be extended to a k-automorphism of K. We have already seen this result when K/k is finite normal. We also show that given a field k, there is an algebraic extension $\bar{k}$ of k such that $\bar{k}$ has no algebraic extension other that $\bar{k}$ itself. $\bar{k}$ is called an *algebraic closure* of k. We define the product of two subfields of a field and show that the product and the intersection of two normal extensions of k is again a normal extension of k.

Let S be a set of polynomials over k. Suppose each $f \in S$ splits in a field E containing k. Then E is called a *splitting field of S over k* and k(zeros of $f \in S$ in E) is called a *minimal splitting field of S over k.* For a finite set S, it is very easy to show the existence of a minimal splitting field of S over k. For, let

$$S = \{f_1, f_2, ..., f_n \mid f_i \in k[x]\}.$$

Let E_1 be a minimal splitting field of f_1 over k, E_2 be a minimal splitting field of f_2 over E_1 and so on, E_n be a minimal splitting field of f_n over E_{n-1}. Then $E_1 \subseteq E_2 \subseteq ... \subseteq E_n$ and each f_i splits in $E_i \subseteq E_n \Rightarrow S$ splits in E_n. So, k(zeros of f_i in E_n) is a minimal splitting field of S over k. It is also a minimal splitting field of $f = f_1 f_2 ... f_n$ over k.

Definition: A field k is called *algebraically closed* if every polynomial f over k splits in k.

By fundamental theorem of algebra, every polynomial over **C**, the field of complex numbers splits in **C**. So, **C** is an algebraically closed field. However, **R** the field of reals is not algebraiclly closed as $x^2 + 1 \in \mathbf{R}[x]$ does not split in **R**. We have the following characteriszations of algebraically closed fields.

NOTES

Theorem 3.31: *A field k is algebrically closed iff every irreducible polynomial over k has degree one.*

Proof: Suppose k is algebraically closed.

Let f be an irreducible polynomial over k. Since k is algebrically closed, f splits in k.

So, $f = f_1 f_2 \ldots f_n$ where each f_i is linear over k.

Since f is irreducible over k, $f = f_1 \Rightarrow f$ is linear over $k \Rightarrow \deg f = 1$.

Conversely, let $g \in k[x]$.

Then $g = g_1 g_2 \ldots g_m$, where each g_i is irreducible over k.

By hypothesis, $\deg g_i = 1 \Rightarrow g_i$ is linear over k for each i

$\Rightarrow g$ is a product of linear factors over $k \Rightarrow g$ splits in k.

So, k is algebraically closed.

Theorem 3.32: *A field k is algebraically closed iff every algebraic extension of k is k itself.*

Proof: Let k be algebraically closed. Let K/k be algebraic.

Let $\alpha \in K$, $p(x) = Irr\ (k, \alpha)$.

By above theorem $\deg p(x) = 1 \Rightarrow p(x) = x - \alpha \in k[x] \Rightarrow \alpha \in k \Rightarrow K = k$.

Conversely, let $f \in k[x]$. Let K be a minimal splitting field of f over k.

Then K/k is algebraic. By hypothesis, $K = k$.

So, $f(x)$ splits in $k[x] \Rightarrow k$ is algebraically closed.

Summarizing the last two results, we have the following

Theorem 3.33: *Let k be a field. Then following are equivalent*

(*i*) *k is algebraically closed.*

(*ii*) *Every irreducible polynomial over k has degree one.*

(*iii*) *Every algebraic extension over k is k itself.*

Theorem 3.34: *A finite field is not algebraically closed.*

Proof: Let k be the finite field $\{a_1, a_2, \ldots, a_n\}$

Let $f = 1 + (x - a_1)(x - a_2) \ldots (x - a_n) \in k[x]$.

Since $f(a_i) \neq 0$ for all i, we find f does not split in k.

Hence k is not algebraically closed.

Definition: Let k be a field. An extension E of k is called *algebraic closure* of k if

(*i*) E/k is algebraic.

(*ii*) E is algebraically closed.

The following result is now an immediate consequence of Theorem 3.32.

NOTES

Theorem 3.35: *Let E be an algebraic extension of k. Then E is algebraically closed iff E has no algeraic extension other than E itself.*

Example 3.20: Since [**C** : **R**] = 2, **C**/**R** is algebraic. Also, **C** is algebraically closed, So, **C** is an algebraic closure of **R**. However, **C** is not an algebraic closure of **Q** as **C**/**Q** is not algebraic ($\pi \in \mathbf{C}$ is not algebraic over **Q**).

Theorem 3.36: *Let K/k be algebraic. Let $\bar{k}$ denote an algebraic closure of K. Then $\bar{k}$ is an algebraic closure of k such that*

$$k \subseteq K \subseteq \bar{k}.$$

Proof: Since $\bar{k}$ is an algebraic closure of K, $\bar{k}/K$ is algebraic. Also, K/k is algebraic. So, $\bar{k}/k$ is algebraic. But $\bar{k}$ is algebraically closed. Thus $\bar{k}$ is also an algebraic closure of k.

Theorem 3.37: *Let K be an algebraically closed field such that K is an extension of k. Let $F = \{a \in K \mid a \text{ is algebraic over } k\}$.*

Then F is an algebraic closure of k.

Proof: We know that

$k \subseteq F \subseteq K$ is a tower of fields.

Also, by definition of F, F/k is algebraic.

Let $f \in F[x]$. Then $f \in K[x]$. Since K is algebraically closed, f splits in K.

Let $f = \alpha(x - \alpha_1) \ldots (x - \alpha_n)$, $\alpha_i \in K$.

Since α_i is algebraic over F, $F(\alpha_i)/F$ is algebraic for all i.

Also F/k is algebraic. So, $F(\alpha_i)/k$ is algebraic for all i.

$\Rightarrow$ $\alpha_i \in K$ is algebraic over k

$\Rightarrow$ $\alpha_i \in F$

$\Rightarrow$ f splits in F

$\Rightarrow$ F is algebraically closed

$\Rightarrow$ F is an algebraic closure of k.

From above theorem it follows that $F = \{a \in \mathbf{C} \mid a \text{ is algebraic over } \mathbf{Q}\}$ is an algebraic closure of **Q**.

We now show the existence of a minimal splitting field of a set of polynomials over k.

Theorem 3.38: *Let S be a set of polynomials over k. Then there is a minimal splitting field of S over k.*

Proof: Suppose $S = \{f_i \mid f_i \in k[x], i \in I\}$.

Let $A = \{i_1, i_2, \ldots, i_n\}$ be a finite subset of I.

Put $f_A = f_{i1} f_{i2} \ldots f_{i_n} \in k[x]$.

Let E_A be a minimal splitting field of f_A over k.

Suppose $B \subseteq A$. Then f_B divides f_A. So, f_B splits in E_A.

Let $F_B = k$(zeros of f_B in E_A).

Then F_B is a minimal splitting field of f_B over k. So, $F_B \cong E_B$. But $F_B \subseteq E_A$. Therefore, we can regard $E_B \subseteq E_A$. So, we have $B \subseteq A \Rightarrow E_B \subseteq E_A$.

Let $E = \bigcup_A E_A$. Let $a, b \in E$. Then $a \in E_A$, $b \in E_B$ for some finite sets $A, B \subseteq I$.

Let $C = A \cup B$. Then $A, B \subseteq C$.

So, $E_A, E_B \subseteq E_C \Rightarrow a, b \in E_C$

$\Rightarrow$ $a \pm b$, ab, ab^{-1}, (if $b \neq 0$) are in $E_C \subseteq E$

$\Rightarrow$ E is a field.

Therefore, for each $f_i \in S$, f_i splits in E_A, where $A = \{i\}$.

$\Rightarrow$ each $f_i \in S$ splits in E.

$\Rightarrow$ E is a splitting field of S over k.

$\Rightarrow$ k(zero of f_i in E) is a minimal splitting field of S over k.

Using Zorn's lemma or otherwise one can prove the following result.

Theorem 3.39: *Any two minimal splitting fields of a set of polynomials over k are isomorphic.*

We can now show the existence of an algebraic closure of a field k.

Theorem 3.40: *Let S be the set of all polynomials over k. Then a minimal splitting field of S over k is an algebraic closure of k.*

Proof: Let F be a minimal splitting field of S. Since F is generated by zeros of $f \in S$, F is generated by algebraic elements over k. So, F/k is algebraic.

Let $f = a_0 + a_1x + ... + a_nx^n \in F[x]$.

Let $E = k(a_0, a_1, ..., a_n) \subseteq F$.

Then $f \in E[x]$. Let E' be a minimal splitting field of f over E.

Let $f = \alpha(x - \alpha_1) ... (x - \alpha_n)$, $\alpha_i \in E'$.

Then $E' = E(\alpha_1, ..., \alpha_n)$.

Since each α_i is algebraic over E, E'/E is algebraic. Also, each $a_i \in F$ is algebraic over $k \Rightarrow E/k$ is algebraic. So, E'/k is algebraic.

Let $g_i = Irr(k, \alpha_i)$

Let $g = g_1g_2 ... g_n \in k[x]$

Now $g_i = (x - \alpha_i)f_i$, $f_i \in E'[x]$.

Therefore, $g = (x - \alpha_1) .... (x - \alpha_n)f_1 ... f_n$

$= \alpha(x - \alpha_1) ... (x - \alpha_n)\alpha^{-1} f_1 ... f_n$

$= ff'$, $f' = \alpha^{-1} f_1 ... f_n \in E'[x]$

Let $g = \sum c_i x^i$, $f' = \sum b_i x^i$, $f = \sum a_i x^i$

where $c_i \in k$, $b_i \in E'$, $a_i \in F$.

Now $c_r = \sum_i a_i b_{r-i}$

NOTES

NOTES

Let a_j be the first non zero coefficient in $f(x)$.

Therefore, $c_j = a_j b_0 \Rightarrow b_0 = a_j^{-1} c_j \in F$.

Suppose $b_0, b_1, ..., b_r \in F$.

Then $c_{j+r+1} = a_{j+r+1} b_0 + a_{j+r} b_1 + a_{j+1} b_r + a_j b_{r+1}$

$\Rightarrow \quad b_{r+1} = a_j^{-1} (c_{j+r+1} - a_{j+r+1} b_0 ... - a_{j+i} b_r) \in F$

By induction, each $b_i \in F \Rightarrow f' \in F[x]$.

By hypothesis, $g \in k[x] \Rightarrow g$ splits in F.

Let $g = (x - \beta_1) ... (x - \beta_m) \ \beta_i \in F$

Suppose $f \in F[x]$ splits in some extension F' of F.

Let $f = d(x - d_1) ... (x - d_n), \quad d_i \in F' \subseteq F$.

Now $f' \in F[x] \subseteq F'[x] \Rightarrow f'$ splits in some extension F'' of F'.

Let $f' = e(x - e_1) ... (x - e_r), \quad e_i \in F'' \supseteq F' \supseteq F$

So, $g = ff' \quad \Rightarrow g(d_i) = 0$ for all i

$\Rightarrow \quad d_i - \beta_j = 0$ for some j depending on i

$\Rightarrow \quad d_i = b_j \in F$

$\Rightarrow \quad d_i \in F$ for all i

$\Rightarrow \quad f$ splits in F.

Thus, F is algebraically closed.

Hence F is an algebraic closure of k.

Converse of above theorem is also true.

Theorem 3.41: *Let F be an algebraic closure of k. Then F is a minimal splitting field of the set S of all polynomials over k.*

Proof: Now F is an algebraic closure of k

$\Rightarrow$ F is algebraically closed

$\Rightarrow$ Each $f \in S$ splits in F.

Let $F' = k$(zeros of $f \in S$ in F) $\subseteq F$.

Let $\alpha \in F'$. Then α is algebraic over k as F/k is algebraic.

Let $p(x) = Irr(k, \alpha)$

Then α is a zero of $p(x) \in S$ in F.

So, $\alpha \in F' \Rightarrow F \subseteq F$.

Therefore, $F' = F \Rightarrow F$ is a minimal splitting field of F' of the set of all polynomials over k. The following is then immediate.

Theorem 3.42: *Any two algebraic closures of a field are isomorphic.*

Proof: Let k be a field and F_1, F_2 be algebraic closures of k. Then F_1, F_2 are minimal splitting fields of the set of all polynomials over k. So, F_1, F_2 are isomorphic by Theorem 3.39.

NOTES

Theorem 3.43: *Algebraic closure of a countable field is countable.*

Proof: Let k be a countable field. For each integer $n \geq 1$, there is a countable set of polynomials of degree n over k. Thus, the set S of all polynomials over k is countable. Let $S = \{f_1, f_2, ..., f_n, ...\}$. Let $E_0 = k$, and E_1 be a minimal splitting field of f_1 over $E_0 = k$. In this way, let E_i be a minimal splitting field of f_i over E_{i-1}. Then $E_{n-1} \subseteq E_n$ for all n.

So, $E = \cup_n E_n$ is a field $\Rightarrow$ each f_i splits in E

$\Rightarrow$ E is a splitting field of S over k.

Let $F = k$(zeros of f_i in E) $\subseteq E$.

Then $k \subseteq F \subseteq E$ is a tower of fields and F is a minimal splitting field of S over k. So, F is an algebraic closure of $k \Rightarrow F$ is algebraically closed $\Rightarrow F$ is not finite. Since E is countable, F is also countable. Thus, any algebraic closure F' of k being isomorphic to F is also countable.

Lemma: *Let E be an algebraic extension of k and let $\sigma : E \to E$ be a k-homomorphism. Then σ is a k-automorphism.*

Proof: Let $\alpha \in E$, $p(x) = Irr\ (k, \alpha)$.

Let $\alpha = \alpha_1, \alpha_2, ..., \alpha_r$ be zeros of $p(x)$ lying in E.

Let $E' = k(\alpha_1, \alpha_2, ..., \alpha_r) \subseteq E$.

Then E'/k is finite.

Let $p(x) = (x - \alpha_i)\ q_i(x), \quad q_i(x) \in k(\alpha_i)[x]$.

Since $\sigma(a) = a$ for all $a \in k$, $\sigma(p(x)) = p(x)$.

Therefore, $p(x) = \sigma(p(x)) = (x - \sigma(\alpha_i))\ \sigma(q_i(x))$

$\Rightarrow$ $\sigma(\alpha_i)$ is a zero of $p(x)$ for all i.

But $\sigma : E \to E \Rightarrow \sigma(\alpha_i) \in E$ for all i.

So, $\sigma(\alpha_i)$ is a zero of $p(x)$ in E for all i.

$\Rightarrow$ $\sigma(\alpha_i) \in E'$ for all i.

$\Rightarrow$ $\sigma : E' \to E'$ is k-homomorphism.

Also E'/k is finite. Since σ is also $1 - 1$, $\sigma : E' \to E'$ must be onto (See below).

Therefore, $E' = \sigma(E') \Rightarrow \alpha = \sigma(\beta)$, for some $\beta \in E' \subseteq E$

$\Rightarrow$ $\sigma : E \to E$ is onto $\Rightarrow \sigma$ is a k-automorphism of E.

That $\sigma : E \to E'$ is onto follows from the result

'If V is a finite dimensional vector space over F and $T : V \to V$ is a linear transformaion, then T is 1-1 iff T is onto'. Here $\sigma : E' \to E'$ is a k-homomorphism $\Rightarrow \sigma$ is a linear transformation as $\sigma(a\beta) = \sigma(a)\ \sigma(\beta) = a\sigma(\beta)$ for all $a \in k$, $\beta \in E'$. Also E' as a vector space over k is finite dimensional.

We now give two characterizations of normal extensions. These are very useful in finding whether the given extension is normal or not.

NOTES

Theorem 3.44: *Let K be an algebraic extension of k. Let $\bar{k}$ denote an algebraic closure of k such that $k \subseteq K \subseteq \bar{k}$. Then K/k is normal iff every k-homomorphism of K into $\bar{k}$ is a k-automorphism of K.*

Proof: Let K/k be normal. Let $\sigma : K \to \bar{k}$ be a k-homomorphism. Let $a \in K$. Since K/k is algebraic, a is algebraic over k. Let $p(x) = Irr\ (k, \alpha)$. Let $\sigma(a) = b$. Since $\sigma(p(x)) = p(x)$, b is a zero of $p(x)$ in $\bar{k} \supseteq K$.

Since K/k is normal, $p(x)$ splits in $K[x]$. So, $b \in K$.

Therefore, $\sigma : K \to K$ is k-homomorphism.

By above lemma, σ is a k-automorphism of K.

Conversely, let $\alpha \in K$ and $p(x) = Irr(k, \alpha)$.

Since $\bar{k}$ is an algebraic closure of k, $p(x)$ splits in $\bar{k}[x]$.

Let β be a zero of $p(x)$ in $\bar{k}$.

Then there exists a k-isomorphism $\sigma : k(\alpha) \to k(\beta)$ such that $\sigma(\alpha) = \beta$.

Since $\beta \in k$, $k(\beta) \subseteq k$. So, σ is a k-homomorphism from $k(\alpha)$ into $\bar{k}$.

Thus σ can be extended to k-homomorphism $\bar{\sigma} : K \to K$.

By hypothesis, $\bar{\sigma}$ is a k-automorphism of K.

So, $\bar{\sigma}(K) = K$. Also $\bar{\sigma}(a) = \sigma(a)$ for all $a \in k(\alpha)$. In particular $\bar{\sigma}(\alpha) = \sigma(\alpha) = \beta$.

Since $\alpha \in K$, $\bar{\sigma}(\alpha) \in \bar{\sigma}(K) = K \Rightarrow \beta \in$ K.

Therefore, $p(x)$ splits in $K[x]$.

Hence K/k is normal.

Theorem 3.45: *Let K be an algebraic extension of k. Then K/k is normal iff K is a minimal splitting field over k of a set of polynomials in k[x].*

Proof: Let K/k be normal. Let $\alpha \in K$. Let $f_\alpha(x) = Irr\ (k, \alpha)$. Then $f_\alpha(x)$ splits in $K[x]$ for all $\alpha \in K$. Let $S = \{f_\alpha \mid \alpha \in K\}$. Let $F = k$(zeros of f_α in K, $\alpha \in K$).

Then F is a minimal splitting field of S over k.

Clearly, $F \subseteq K$. Also $\alpha \in K \Rightarrow \alpha$ is a zero of $f_\alpha \Rightarrow \alpha \in F$. So, $F = K$. Thus K is a minimal splitting field of S over k.

*Converse*ly, let K be a minimal splitting field of a set S of polynomials over k. Let $\bar{k}$ be an algebraic closure of k such that, $k \subseteq K \subseteq \bar{k}$.

Let $\sigma : K \to \bar{k}$ be a k-homomorphism.

Let $a \in K$ be a zero of some $f \in k[x]$ in S.

Then $\sigma(a)$ is also a zero of f as σ is a k-homomorphism.

As f splits in $K[x]$, we can write $f = \alpha(x - \alpha_1) \ldots (x - \alpha_n)$, $\alpha_i \in K$, $\alpha \in k$.

Since $\sigma(\alpha_i)$ is a zero of f for all i, $\sigma(\alpha_i) \in \bar{k}$, $\{\alpha_1, \ldots, \alpha_n\} = \{\sigma(\alpha_1), \ldots, \sigma(\alpha_n)\}$ as $\bar{k}$ can't have more than n zeros of f. So, $\sigma(\alpha_i) \in K$ for all i.

Let $T =$ {zeros of f in K, $f \in S$}. Then $\sigma : T \to T$. Also $\sigma : T \to T$ is 1 –1 as $\sigma : K \to \bar{k}$ is 1 – 1.

NOTES

Let $b \in \sigma(K)$. Then $b = \sigma(c)$, $c \in K$.

Now, $c \in K \Rightarrow c = \dfrac{f(\beta_1, ..., \beta_n)}{g(\beta_1, ..., \beta_n)}$, $\beta_i \in T$.

Then $b = \dfrac{f(\sigma(\beta_1), ..., \sigma(\beta_n))}{g(\sigma(\beta_1), ..., \sigma(\beta_n))} = \dfrac{f(\gamma_1, ..., \gamma_n)}{g(\gamma_1, ..., \gamma_n)}$, $\gamma_i \in T$

So, $b \in K \Rightarrow \sigma(K) \subseteq K$.

Also $d \in K \Rightarrow d = \dfrac{f_1(\delta_1, ..., \delta_m)}{g_1(\delta_1, ..., \delta m)}$,

$\delta_i \in T \Rightarrow d = \dfrac{f_1(\sigma(u_1), ..., \sigma(u_m))}{g_1(\sigma(u_1), ..., \sigma(u_m))}$, $u_i \in T$

$\Rightarrow \quad d = \dfrac{\sigma(f_1(u_1, ..., u_m))}{\sigma(g_1(u_1, ..., u_m))} = \sigma\left(\dfrac{f_1(u_1, ..., u_m)}{g_1(u_1, ..., u_m)}\right)$, $u_i \in T$

$\Rightarrow \quad d \in \sigma(K) \Rightarrow K \subseteq \sigma(K) \Rightarrow \sigma(K) = K$.

So, $\sigma : K \to K$ is onto Thus, σ is a k-automorphism of K. By previous result, K/k is normal.

Summarizing, the last two theorems we get

Theorem 3.46: *Let K be an algebraic extension of k. Then following are equivalent:*

(*i*) *K/k is normal.*

(*ii*) *Every k-homomorphism of K into $\bar{k}$ is a k-automorphism of K where $\bar{k}$ is an algebraic closure of k.*

(*iii*) *K is a minimal splitting field of a set of polynomials over k.*

Theorem 3.47: *Let F/k be algebraic. If every finite extension of k admis a k-homomorphism into F, then F is an algebraic closure of k.*

Proof: Let $f = a_0 + a_1x + ... + a_nx^n \in k[x]$. Let E be a minimal splitting field of f over k. Then E/k is finite.

By hypothesis, there is a k-homomorphism $\sigma : E \to F$.

Let $f = \alpha(x - \alpha_1) ... (x - \alpha_n)$, $\alpha_i \in E$.

Then $f = \sigma f = \alpha(x - \sigma(\alpha_1)) ... (x - \sigma(\alpha_n))$

$\Rightarrow$ f splits in F

$\Rightarrow$ every polynomial over k splits in F.

Let F' be a minimal splitting field of the set of all polynomials over k.

Then $F' = k$(zero of $f \in k[x]$ in F) $\subseteq F$

Also, $\alpha \in F \Rightarrow \alpha$ is algerbraic over k.

Let $p(x) = Irr(k, \alpha)$. Then $\alpha \in F$ is a zero of $p(x) \in k[x]$

$\Rightarrow \quad \alpha \in F' \Rightarrow F \subseteq F' \Rightarrow F = F'$.

So, F is an algebraic closure of k.

NOTES

Theorem 3.48: *Let K/k be an algebraic extension. Let $\bar{k}$ be an algebraic closure of k such that $k \subseteq K \subseteq \bar{k}$. Let F be an algebraically closed field such that $k \subseteq F$. Then any k-homomorphism from K into F can be extended to a k-homomorphism from $\bar{k}$ into F.*

Proof: Let $\alpha : K \to F$ be a k-homomorphism.

Let $S = \left\{ (E, g) \;\middle|\; \begin{array}{l} E \text{ is a subfield of } \bar{k} \text{ and } K \subseteq E \\ g : E \to F \text{ is a homomorphism extending } \sigma \end{array} \right\}$.

Define a relation $\leq$ on S as follows:

$(E_1, g_1) \leq (E_2, g_2)$ if $E_1 \subseteq E_2$ and g_2 is an extension of g_1 to E_2.

Then $\leq$ is a partial order on S.

Let $\{(E_i, g_i)\}_i$ be a chain in S. Let $E = \underset{i}{\cup} E_i$ and define $g : E \to F$ such that, $g(\alpha) = g_i(\alpha)$ if $\alpha \in E_i$.

Then $(E, g) \in S$ and is an upper bound of the chain $\{(E_i, g_i)\}$.

By Zorn's lemma S has a maximal element, say (E_0, g_0).

We show that $E_0 = \bar{k}$. Suppose $E_0 \neq \bar{k}$.

Then we can find $a \in \bar{k}$ such that $a \notin E_0$. Since $\bar{k}/k$ algebraic, a is algebraic over k.

Let $f = Irr(k, a)$. Now $k \subseteq E_0 \Rightarrow f \in E_0[x]$. Since F is algebraically closed, $g_0(f) \in F[x]$ splits in $F[x]$.

Let b be a zero of $g_0(f)$ in F. Then there exists an isomorphism $\theta : E_0(a) \to E_0'(b)$ extending g_0, where $E_0' = g_0(E_0)$.

But $b \in F, E_0' \subseteq F \Rightarrow E_0'(b) \subseteq F$. So, $\theta : E_0(a) \to F$ is a homomorphism extending g_0.

Therefore, $(E_0, g_0) \leq (E_0(a), \theta)$ and $E_0 \neq E_0(a) \Rightarrow (E_0, g_0) \neq (E_0(a), \theta)$. This contradicts the maximality of (E_0, g_0).

So, $E_0 = \bar{k}$. Therefore, $g_0 : \bar{k} \to F$ is a homomorphism extending σ.

Corollary: Let K/k be algebraic such that $k \subseteq K \subseteq \bar{k}$. Then any k-homomorphism of K into $\bar{k}$ can be extended to a k-homomorphism of $\bar{k}$ into $\bar{k}$.

Proof: Take $F = \bar{k}$ in above theorem.

Corollary: Any two algebraic closures of a field k are k-isomorphic.

Proof: Let K_1, K_2 be algebraic closures of k.

Now $k \subseteq K_1, K_2$. Let $\sigma : k \to K_1$ be the inclusion map i.e., $\sigma(\alpha) = \alpha$ for all $\alpha \in k$.

By taking $K = \bar{k}$, $k = K_2$, $F = k_1$, in above theorem, σ can be extended to a k-homomorphism $\eta : K_2 \to K_1$.

As $K_2 \cong \eta(K_2)$ and K_2 is algebraically closed we find $\eta(K_2)$ is algebraically closed.

Also $k \subseteq K_2 \Rightarrow \eta(K_2)$ can be regarded as an extension of k.

So, we have $k \subseteq \eta(K_2) \subseteq K_1$.

Since K_1/k is algebraic, $K_1/\eta(K_2)$ is also algebraic.

But $\eta(K_2)$ is algebraically closed $\Rightarrow \eta(K_2)$ has no algebraic extension other than itself $\Rightarrow K_1 = \eta(K_2) \Rightarrow \eta$ is onto $\Rightarrow \eta$ is a k-isomorphism.

Hence, K_1, K_2 are k-isomorphic.

Theorem 3.49: *Let k, E, K be fields such that, $k \subseteq E \subseteq K$ and K/k is normal. Then any k-homomorphism $\sigma : E \to K$ can be extended to a k-automorphism of K.*

Proof: Since K/k is normal, K is minimal splitting field a set of polynomials over k. Let $\bar{k}$ denote an algebraic closure of k.

Then $\bar{k}$ is a minimal splitting field of the set of all polynomials over k.

So K can be regarded as a subfield of $\bar{k}$.

Now $\sigma : E \to K$ is a k-homomorphism.

Thus $\sigma : E \to \bar{k}$ is a k-homomorphism.

Since K/k is algebraic, so is E/k. Now $k \subseteq E \subseteq \bar{k}$, E/k is algebraic.

By previous theorem, σ can be extended to a k-homomorphism $\tau : \bar{k} \to \bar{k}$. Therefore, $\tau : K \to \bar{k}$ is also a k-homomorphism.

Again, K/k is normal $\Rightarrow \tau$ is a k-automorphism of K.

This proves the result.

Product of Fields: Let M, N be extensions of a field k such that M, N are contained in a field L. Then MN is defined as the smallest subfield of L containing M and N.

Let $M[N] = \{a_1b_1 + ... + a_nb_n \mid a_i \in M, b_i \in N\}$.

n = Finite

Then $M[N]$ is an integral domain. Let K be field of quotients of $M[N]$. Clearly, $M \subseteq M[N]$, $N \subseteq M[N]$.

So, $M, N \subseteq M[N] \subseteq K$.

But MN is the smallest field containing M, N, $MN \subseteq K$.

Also, $\sum_{i=1}^{n} a_ib_i \in M[N]$, for all $a_i \in M$, $b_i \in N$

$\Rightarrow$ $\sum_{i=1}^{n} a_ib_i \in MN$, as $a_i \in M \Rightarrow a_i \in MN$, $b_i \in N \Rightarrow b_i \in MN$

$\Rightarrow$ $M[N] \subseteq MN$.

But K is the smallest field containing $M[N] \Rightarrow K \subseteq MN$

$\Rightarrow$ $K = MN$

NOTES

$\Rightarrow MN$ is a quotient field of $M[N]$.

Lemma: *Let K_1, K_2 be extensions of a field k contained in a field K and let σ be a k-homomorphism of K in some field L. Then*

$$\sigma(K_1K_2) = \sigma(K_1)\,\sigma(K_2).$$

Proof: Let $\alpha = \dfrac{a_1b_1 + \ldots + a_nb_n}{a_1'\,b_1' + \ldots + a_n'\,b_n'} \in K_1K_2$

where $a_i, a_i' \in K_1,\ b_i, b_i' \in K_2$

Then $\sigma(\alpha) = \dfrac{\sigma(a_1)\sigma(b_1) + \ldots + \sigma(a_n)\,\sigma(b_n)}{\sigma(a_1')\sigma(b_1') + \ldots + \sigma(a_n')\,\sigma(b_n')} \in \sigma(K_1)\sigma(K_2)$

$\Rightarrow \sigma(K_1K_2) \subseteq \sigma(K_1)\,\sigma(K_2)$

Let $\beta \in \sigma(K_1)\,\sigma(K_2)$.

Then $\beta = \dfrac{\sigma(c_1)\sigma(d_1) + \ldots + \sigma(c_r)\sigma(d_r)}{\sigma(c_1')\sigma(d_1') + \ldots + \sigma(c_r')\sigma(d_r')}$

$= \sigma(\alpha)$, where $\alpha = \dfrac{c_1d_1 + \ldots + c_rd_r}{c_1'd_1' + \ldots + c_r'd_r'} \in K_1K_2$

$\Rightarrow \beta \in \sigma(K_1K_2)$

$\Rightarrow \sigma(K_1)\sigma(K_2) \subseteq \sigma(K_1K_2)$

$\Rightarrow \sigma(K_1K_2) = \sigma(K_1)\sigma(K_2)$.

Theorem 3.50: *If E, F are normal extensions of k, then EF and $E \cap F$ are normal over k.*

Proof: (*i*) Let $\bar{k}$ denote an algebraic closure of *k*. Let σ be a *k*-homomorphism from *EF* into $\bar{k}$ such that, $k \subseteq EF \subseteq \bar{k}$.

Now $\sigma(EF) = \sigma(E)\sigma(F)$ by above lemma.

Since $E, F \subseteq EF$, σ is also *k*-homomorphism from *E* into $\bar{k}$ and *F* into $\bar{k}$. Also *E*, *F* are normal over $k \Rightarrow \sigma : E \to E$ and $\sigma : F \to F$ are *k*-automorphisms

$\Rightarrow \sigma(E) = E,\ \sigma(F) = F$

$\Rightarrow \sigma(EF) = EF$

Now $\sigma : EF \to \bar{k}$ is also a *k*-homomorphism from *EF* into *EF*. But $\sigma(EF) = EF$

$\Rightarrow \sigma : EF \to EF$ is onto.

So, $\sigma : EF \to EF$ is a *k*-automorphism.

$\Rightarrow EF/k$ is normal.

(*ii*) Let σ be a *k*-homorphism from $E \cap F$ into $\bar{k}$ such that $k \subseteq E \cap F \subseteq \bar{k}$. Then σ can be extended to $\bar{k}$-homomorphism $\eta : \bar{k} \to \bar{k}$.

Since *E*/*k* is normal, *E* is a minimal splitting field of a set of polynomials over *k*. However, $\bar{k}$ is a minimal splitting field of the set of all polynomials over *k*. So, *E* can be regarded as a subfield of $\bar{k}$. Therefore, $k \subseteq E \subseteq \bar{k}$. Similarly $k \subseteq F \subseteq \bar{k}$.

Let $\eta \mid E = \eta_1$, $\eta \mid F = \eta_2$.

Now $\eta_1 : E \to \bar{k}$, $\eta_2 : F \to \bar{k}$ are k-homomorphisms. Since E/k, F/k are normal, η_1 and η_2 are k-automorphism of E and F respectively. So, $\eta_1(E) = E$, $\eta_2(F) = F$. Now $E \cap F \subseteq E$, $F \subseteq \bar{k}$.

NOTES

Thus, $\eta(E \cap F) = \eta(E) \cap \eta(F)$

$= \eta_1(E) \cap \eta_2(F)$

$= E \cap F$.

But $\eta \mid E \cap F = \sigma$

$\Rightarrow$ $\sigma(E \cap F) = E \cap F$

$\Rightarrow \sigma$ is a k-automorphism of $E \cap F$.

$\Rightarrow E \cap F/k$ is normal.

Check Your Progress

1. What do you mean by extension of a field F?
2. When is a complex number said to be an algebraic number?
3. What is a prime subfield?
4. Define normal extension.
5. What is a finite field?
6. What is a splitting field?

3.5 AUTOMORPHISM OF EXTENSIONS

The purpose of this section is to find conditions under which a finite extension F/K is separable in terms of k-automorphisms of F. We first show that the number of k-automorphisms of F is at most $n = [F : K]$. We then show that the upper bound n is achieved iff F/K is both normal and separable.

Definition: Let $\sigma_1, \sigma_2, ..., \sigma_n$ be homomorphisms from a field E into a field E'. Then, $\sigma_i s$ are called *linearly independent* over E' if $\alpha_1\sigma_1 + ... + \alpha_n\sigma_n = 0, \Rightarrow \alpha_i = 0 \ \forall \ i$ where $\alpha_i \in E'$.

Note, $\alpha_i\sigma_i : E \to E'$ such that, $(\alpha_i\sigma_i)(a) = \alpha_i(\sigma_i(a)) \quad \forall \ a \in E$.

In the following result, we show that any family of homomorphisms from a field into another field is linearly independent.

Theorem (Dedekind). 3.51: *Let $(\sigma_i)_i$ be a family of distinct homomorphism from a field E into a field E′. Then $\{\sigma_i\}_i$ is linearly independent over E′.*

Proof: Suppose $\{\sigma_i\}_i$ is not linearly independent over E'. Then $\exists$ finite subset of $\{\sigma_i\}_i$ which is not linearly independent over E'. (*i.e.*, it is linearly dependent over E'.). Let $\{\sigma_1, \sigma_2, ..., \sigma_r\}$ be a minimal linearly dependent subset of $\{\sigma_i\}_i$ over E'.

NOTES

So, $\exists\ \alpha_1, \alpha_2, ..., \alpha_r \in E'$ such that,

$\alpha_1\sigma_1 + ... + \alpha_r\sigma_r = 0$ and some $\alpha_i \neq 0$.

$\Rightarrow \quad (\alpha_1\sigma_1 + ... + \alpha_r\sigma_r)\ (a) = 0 \quad \forall\ a \in E$

$\Rightarrow \quad \alpha_1\sigma_1(a) + ... + \alpha_r\sigma_r(a) = 0 \quad \forall\ a \in E$

Suppose $\alpha_1 \neq 0$.

Now $\sigma_1(a) = (-\alpha_1^{-1}\alpha_2)\sigma_2(a) + ... + (-\alpha_1^{-1}\alpha_r)\sigma_r(a) \quad \forall\ a \in E$

$\sigma_1(a) = \beta_2\sigma_2(a) + ... + \beta_r\sigma_r(a)$,

$\beta_i = -\alpha_1^{-1}\ \alpha_i \in E', \quad \forall\ a \in E$ ...(3.2)

So, $\sigma_1(ab) = \beta_2\sigma_2(ab) + ... + \beta_r\sigma_r(ab) \quad \forall\ a, b \in E$

$\Rightarrow \quad \sigma_1(a)\ \sigma_1(b) = \beta_2\sigma_2(a)\sigma_2(b) + ... + \beta_r\sigma_r(a)\sigma_r(b)\ \forall\ a, b \in E$ ...(3.3)

Consider Equation (3.3) – $\sigma_1(b)$ Equation (3.2).

Then $0 = \beta_2\sigma_2(a)\ (\sigma_2(b) - \sigma_1(b)) + ... + \beta_r\sigma_r(a)\ (\sigma_r(b) - \sigma_1(b))$

$$= \sum_2^r \beta_i\ (\sigma_i(b) - \sigma_1(b))\ \sigma_i(a) \quad \forall\ a \in E$$

$$\Rightarrow \quad 0 = \sum_2^r \beta_i\ (\sigma_i(b) - \sigma_1(b))\ \sigma_i$$

$\Rightarrow \quad \beta_i(\sigma_i(b) - \sigma_1(b)) = 0 \quad \forall\ i = 2, 3, ..., r, \forall\ b \in E$

as $\{\sigma_1, \sigma_2, ..., \sigma_r\}$ is a minimal linearly dependent subset of $\{\sigma_i\}_i$.

Since $\sigma_i \neq \sigma_1\ \forall\ i > 1, \exists\ c_i \in E$ such that, $\sigma_i(c_i) \neq \sigma_1(c_i)$.

Now $\beta_i(\sigma_i(c_i) - \sigma_1(c_i) = 0 \quad \forall\ i = 2, 3, ...,$

$\Rightarrow \quad \beta_i = 0 \quad \forall\ i = 2, 3, ..., r.$

So, $\sigma_1(a) = 0 \quad \forall\ a \in E$, by (3.2)

$\Rightarrow \quad \sigma_1(1) = 0$

$\Rightarrow 1 = 0$, which is not true.

Thus $\{\sigma_i\}_i$ is a linearly independent set over E'.

Theorem 3.52: *Let E, E′ be extensions of K. Let* $[E : K] = n$. *Then, there are at most n K-homomorphisms from E into E′.*

Proof: Let $\{u_1, u_2, ..., u_n\}$ be a basis of E/K. Let $\sigma_0, \sigma_1, ..., \sigma_n$ be $n + 1$ distinct K-homomorphisms from E into E'.

Consider the system of equations $\sum_{i=0}^{n} \sigma_i(u_j)x_i = 0, \quad j = 1, 2, ..., n.$

Then, we have n equation in $n + 1$ unknowns $x_i s \in E'$. Since the number of equations is less than number of unknowns, the above system of equations has a non zero solution, say $c_0, c_1, ..., c_n \in E'$ where some $c_i \neq 0$.

Let $a \in E$. Since $\{u_1, u_2, ..., u_n\}$ spans E/K, $a = \alpha_1u_1 + ... + \alpha_nu_n$, $\alpha_i \in K$

NOTES

$$\text{Thus,} \quad \sum_{i=0}^{n} \sigma_i(a)\, c_i = \sum_i \sigma_i \left(\sum_j \alpha_j u_j \right) c_i$$

$$= \sum_i \sum_j (\sigma_i(\alpha_j)\, \sigma_i(u_j)) c_i$$

$$= \sum_i \sum_j (\alpha_j \sigma_i(u_j)) c_i$$

$$= \sum_j \alpha_j \left(\sum_i \sigma_i(u_j)) c_i \right)$$

$$= 0 \quad \text{as } \sum_i \sigma_i(u_j)\, c_i = 0$$

$$\Rightarrow \quad \sum_{i=0}^{n} c_i \sigma_i(a) = 0 \quad \forall\ a \in E$$

$$\Rightarrow \quad \sum_{i=0}^{n} c_i \sigma_i = 0 \Rightarrow c_i = 0 \quad \forall\ i \text{ by above theorem.}$$

But some $c_i \neq 0$. So, we get a contradiction. Thus, there are at most n K-homomorphisms from E into E'.

Corollary: There are at most n K-automorphisms of E, where $n = [E : K]$.

Proof: Take $E' = E$ in above theorem. By automorphism of E, we mean isomorphism of E into E. Now any K-homomorphism from E into E is a linear transformation from E into E as vector space over K. Also, any homomorphism from E into E is 1–1 and so onto as $[E : K]$ = finite. By above theorem, there are at most n K-automorphisms of E where $n = [E : K]$.

Case 6: Define $\theta : \mathbf{C} \to \mathbf{C}$ such that,

$$\theta(z) = \bar{z}\text{, where } \bar{z} = \text{Conjugate of } z$$

Then θ is **R**-homomorphism and $\theta \neq I$. So, θ, I are two distinct **R**-homomorphisms of **C** into **C**. But $[\mathbf{C} : \mathbf{R}] = 2 \Rightarrow$ there are at most two **R**-automorphismss of **C**. Also, any **R**-homomorphism of **C** into **C** is an **R**-automorphism of **C**. So, θ, I are only **R**-automorphisms of **C**. Note, **C**/**R** is normal as $[\mathbf{C} : \mathbf{R}] = 2$ and **C**/**R** is separable as char $\mathbf{R} = 0 \Rightarrow$ **R** is perfect $\Rightarrow$ Every algebraic extension of **R** is separable.

Case 7: Let α be the real cube root of $f(x) = x^3 - 2$. Let $F = \mathbf{Q}(\alpha) \subseteq \mathbf{R}$. Let θ be a **Q**-automorphism of F.

Since α is a root of $f(x)$ in **R**, $\theta(\alpha)$ is a root of $\theta(f(x)) = f(x)$ in **R**.

So, $\theta(\alpha) = \alpha$. But $[\mathbf{Q}(\alpha) : \mathbf{Q}] = \deg Irr\ (\mathbf{Q}, \alpha) = \deg f(x) = 3$ and $\{1, \alpha, \alpha^2\}$ is a basis of $\mathbf{Q}(\alpha)/\mathbf{Q}$.

$\Rightarrow \quad \mathbf{Q}(\alpha) = \{a_0 + a_1\alpha + a_2\alpha^2 \mid a_i \in \mathbf{Q}\}$.

Since $\theta(a_i) = a_i$ and $\theta(\alpha) = \alpha$, θ fixes every element of $\mathbf{Q}(\alpha)$.

So, $\theta = I \Rightarrow$ Identity map is the only **Q**-automorphism of $F = \mathbf{Q}(\alpha)$.

Note $\mathbf{Q}(\alpha)/\mathbf{Q}$ is separable as char $\mathbf{Q} = 0 \Rightarrow$ **Q** is perfect $\Rightarrow$ Every algebraic extension of **Q** is separable.

As seen before $\mathbf{Q}(\alpha)/\mathbf{Q}$ is not normal.

Thus, we notice that if E/K is separable but not normal, then one may not get the *full quota* (i.e., $[E : K]$) of K-automorphisms of E.

NOTES

Case 8: Let char $K = p$ and $F = K(t)$. Then $x^p - t$ is irreducible over F. Let α be a root of $f(x)$ in some extension of F.

Now $f(x) = x^p - t$ is irreducible over $F \Rightarrow [F(\alpha) : F] = p$.

$\Rightarrow \{1, \alpha, ..., \alpha^{p-1}\}$ is a basis of $F(\alpha)/F$.

So, $$F(\alpha) = \{\sum_{i=0}^{p-1} a_i \alpha^i \mid a_i \in F\}.$$

If θ is F-automorphism of $F(\alpha)$, then $\theta(\alpha)$ is a root of $f(x) = \theta(f(x))$ in $F(\alpha)$.

But α is the only root of $f(x)$ in any extension of F.

$\Rightarrow \theta(\alpha) = \alpha \Rightarrow \theta$ fixes every element of $F(\alpha)$.

$\Rightarrow \theta$ is the identity map.

Thus, identity map is the only F-automorphism of $F(\alpha)$.

Since α is not a simple root of $f(x)$, α is not separable over F.

Therefore, if E/K is not sparable then one may not get $[E : K]$, K-automorphisms of E.

The above two examples clearly demonstrate that in order that an extension E/K has $[E : K]$, K-automomorphisms of E, E/K should be both normal and separable. In the first example, we saw that we do get $[E : K]$, K-automorphisms of E when E/K is both normal and separable. We would like to prove this in general.

Theorem 3.53: *Let $K \subseteq L \subseteq F \subseteq E$ be a tower of fields. Suppose E/K is finite normal. If r is the number of K-homorphisms from L into E and s the number of L-homomorphisms from F into E, then the number of K-homomorphisms from F into E is rs.*

Proof: Let $\sigma_1, ..., \sigma_r$ be the K-homomorphisms of L into E and $\tau_1, \tau_2, ..., \tau_s$ be the L-homomorphisms from F into E. Since E/K is finite normal, each σ_i can be extended to K-automorphisms $\bar{\sigma}_i$ of E.

We show that $\{\bar{\sigma}_i \tau_j \mid 1 \le i \le r,\ 1 \le j \le s\}$ is the set of distinct K-homomorphisms from F into E.

Suppose $\bar{\sigma}_i \tau_j = \bar{\sigma}_p \tau_q$. Then $\bar{\sigma}_i \tau_j(a) = \bar{\sigma}_p \tau_q(a), \quad \forall\ a \in F$

$\Rightarrow \quad \bar{\sigma}_i \tau_i(l) = \bar{\sigma}_p \tau_q(l) \quad \forall\ l \in L$

$\Rightarrow \quad \bar{\sigma}_i(l) = \bar{\sigma}_p(l) \quad \forall\ l \in L$

$\Rightarrow \quad \sigma_i = \sigma_p \Rightarrow i = p \Rightarrow \tau_j = \tau_q \Rightarrow j = q.$

Let σ be any K-homomorphisms from F into E. Then $\sigma \mid L$ is a K-homomorphisms from L into E.

$\Rightarrow \quad \sigma \mid L = \sigma_i$ for some i.

Then $\bar{\sigma}_i^{-1}\sigma$ is K-homomorphisms from F into E.

So, $\bar{\sigma}_i^{-1}\sigma(l) = \bar{\sigma}_i^{-1}\sigma_i(l) = \bar{\sigma}_i^{-1}\ \bar{\sigma}_i(l) = l \quad \forall\ l \in L$

$\Rightarrow\ \bar{\sigma}_i^{-1}\sigma$ is L-homomorphism from F into E

$\Rightarrow\ \sigma_i^{-1}\sigma = \tau_j$ for some $j \Rightarrow \sigma = \bar{\sigma}_i\tau_j$

Thus, $\bar{\sigma}_i\,\tau_j$ are the only K-homomorphisms from F into E and so, there are exactly rs K-homomorphisms from F into E.

Theorem 3.54: *Let $K \subseteq E \subseteq E'$ be a tower of fields. Suppose E'/K is finite normal. Then E/K is separable if and only if the number of K-homomorphisms from E into E' is* $[E : K]$.

Proof: Suppose E/K is separable. We prove the result by induction on $n = [E : K]$.

If $n = 1$, then $E = K$ and $I : E \to E'$ such that, $I(a) = a$ is K-homomorphisms from E into E'.

So, the result is true for $n = 1$.

Let $n > 1$. Assume that the result is true for all integers $< n$.

Let $a \in E$, $a \notin K$.

Now $K \subseteq K(a) \subseteq E \subseteq E'$ and E'/K is finite normal $\Rightarrow E'/K(a)$ is finite normal.

Also, $\quad [E : K] = [E : K(a)]\ [K(a) : K]$ and $[K(a) : K] > 1$

$\Rightarrow \quad [E : K(a)] < [E : K] = n$.

Since E/K is separable $E/K(a)$ is also separable.

By induction hypothesis (applied to tower of fields $K(a) \subseteq E \subseteq E'$), the number of $K(a)$-homomorphisms from E into E' is $[E : K(a)]$.

Let $p(x) = Irr(K, a)$. Since $a \in E$, a is separable over K. So, all roots of $p(x)$ are simple.

Let deg $p(x) = r$. Since E'/K is normal, $p(x)$ splits in E' as $a \in E \subseteq E'$.

Let $a = a_1, a_2, ..., a_r$ be distinct roots of $p(x)$ in E'. Then $\exists$ K-isomorphisms

$\sigma_i : K(a) \to K(a_i)$ such that, $\sigma_i(a) = a_i \quad \forall\ i = 1, 2, ..., r$. σ_i's, a_i's being distinct.

Since $a_i \in E'$, σ_i's are r K-homomorphisms from $K(a)$ into E'.

Also as $[K(a) : K] = \deg Irr\ (K, a) = \deg p(x) = r$, these $\sigma_i's$ are only K-homomorphisms from $E(a)$ into E'.

By previous theorem these are excatly $[E : K(a)]\ [K(a) : K] = [E : K]$, K-homomorphisms from E into E'.

So, the result is true in this case. By induction the result is true for all $n \geq 1$.

Conversely, let there be $n = [E : K]$ K-homomorphisms from E into E'. Let $a \in E$.

NOTES

NOTES

Now, the number m of K-homomorphisms from $K(a)$ into E' is at most $r = [K(a) : K]$.

Let $m < r$. Let s be the number of $K(a)$-homomorphisms from E into E'. Then

$$s \leq [E : K(a)] = \frac{[E : K]}{[K(a) : K]} = \frac{n}{r}.$$

By above theorem, the number of K-automorphisms from E into E' is $ms < r\frac{n}{r} = n$, a contradiction. So, $m = r$. That is, the number of K-homomorphisms from $K(a)$ into E' is $[K(a) : K] = \deg Irr\ (K, a)$.

Let $p(x) = Irr\ (K, a)$, $\deg p(x) = r$.

Since E'/K is normal, $p(x)$ splits in E' as $a \in E \subseteq E'$.

Let $a = a_1, a_2, ..., a_t$ be distinct roots or $p(x)$ in E'.

Then, for each $i \exists$ K-isomorphisms $\theta_i : K(a) \to K(a_i)$ such that, $\theta_i(a) = a_i$.

Since $a_i \in E'$, $K(a_i) \subseteq E'$. So, $\theta_i : K(a) \to E'$ is K-homomorphism.

Again as $a_i s$ are distinct, θ is are also distinct K-homomorphisms from $K(a)$ into E'.

If θ is a K-homomorphisms from $K(a)$ into E', then a is a root of $p(x)$ in E

$\Rightarrow$ $\theta(a)$ is a root of $\theta(p(x)) = p(x)$ in E'

$\Rightarrow$ $\theta(a) = a_i$ for some i

$\Rightarrow$ $\theta(a) = \theta_i(a)$ for some $i \Rightarrow \theta = \theta_i$ for some i.

So, $\theta_1, \theta_2, ..., \theta_t$ are the only K-homomorphisms from $K(a)$ into E

$\Rightarrow$ $t = [K(a) : K] = \deg p(x) = r$.

$\Rightarrow$ all roots of $p(x)$ are distinct and so simple.

$\Rightarrow$ a is separable over K. Thus, E/K is separable.

Corollary 1: Let E/K be finite normal. Then E/K is separable if and only if the number of K-automorphisms of E is $[E : K] = n$.

Proof: Since E/K is finite, a K-homomorphism of E is K-automorphism of E and conversely. The result then follows by above theorem.

Corollary 2: Let $K \subseteq E \subseteq E'$ be a tower of fields such that, E/K and E'/E are finite separable. Then E'/K is also finite separable.

Proof: Let $[E : K] = r$, $[E' : E] = s$. Since E/K, E'/E are finite so is E'/K.

Thus $\exists$ an extension F of K such that, F/K is finite normal and $K \subseteq E \subseteq E' \subseteq F$.

By above theorem since E/K is separable, there are r K-homomorphisms from E into F.

Now F/K is normal $\Rightarrow$ F/E is also normal.

As E'/E is separable, there are s E-homomorphisms from E' into F.

Therefore, there are rs K-homomorphisms from E' into F. But $rs = [E':K]$.

By above theorem then E'/K is finite seperable.

NOTES

Corollary 3: Let E be an extension of K. Let $a_1, a_2, ..., a_n \in E$ be separable over K. Then $K(a_1, a_2, ..., a_n)/K$ is separable.

Proof: We prove the result by induction on n. Since $a_1, a_2, ..., a_n$ are separable over K, $a_1, a_2, ..., a_n$ are algebraic over K. So, $K(a_1, a_2, ..., a_n)/K$ is finite. Let E'/K be finite normal extension such that, $K \subseteq K(a_1, ..., a_n) \subseteq E'$. Let $n = 1$. Let $p(x) = Irr\ (K, a_1)$, deg $p(x) = r$. Then $\exists$ r K-homomorphisms from $K(a_1)$ into E' as seen in above theorem. But $r = [K(a_1) : K]$. By above theorem, $K(a_1)/K$ is separable. So, the result is true for $n = 1$. Let $n > 1$. Assume that the result is true for all integers $< n$. By induction hypothesis, $K(a_1, ..., a_n)/K$ is finite separable. Also, a_n is separable over K and $K \subseteq K(a_1, ..., a_{n-1}) \subseteq K(a_1, ..., a_n)$

$\Rightarrow a_n$ is separable over $K(a_1, ..., a_{n-1})$

$\Rightarrow K(a_1, ..., a_n) \mid K(a_1, ..., a_n)$ is finite separable.

By above corollary, $K(a_1, ..., a_n)/K$ is separable. By induction the result is true $\forall\ n \geq 1$.

Corollary 4: Let $F \subseteq K \subseteq E$ be a tower of fields such that, E/K and K/F are separable. Then E/F is also separable.

Proof: Let $a \in E$.

Let $p(x) = Irr\ (K, a)$

$= b_0 + b_1x + ... + b_rx^r, \quad b_i \in K$

Let $K' = F(b_0, b_1, ..., b_r) \subseteq K$

$b_i \in K \Rightarrow b_i$ is separable over F

$\Rightarrow K'/F$ is separable by above Corollary

Since $p(x)$ is irreducible over K, it is also irreducible over K'.

So, $p(x) = Irr\ (K', a)$

Now $K' \subseteq K \subseteq E$ and $a \in E$ is separable over $K \Rightarrow p'(a) \neq 0 \Rightarrow a$ is separable over $K' \Rightarrow K'(a)/K'$ is separable and finite. Also, K'/F is finite separable.

So, $K'(a)/F$ is finite separable.

$\Rightarrow$ a is separable over F.

Thus, E/F is separable.

Theorem 3.55: *Let $K \subseteq E \subseteq E'$ be a tower of fields such that, E'/K is finite normal. Then following are equivalent:*

(*i*) *There are exactly $n = [E : K]$ K-homomorphisms from E into E'.*

(*ii*) *E/K is separable.*

(*iii*) *E/K is generated by separable elements.*

Proof: $(i) \Leftrightarrow (b)$ follows from previous theorem

$(ii) \Rightarrow (c)$ $[E : K] = n \Rightarrow E = K(a_1, ..., a_n)$. Since $a_i \in E$, a_i is separable over K. So, E is generated by separable elements over K.

NOTES

$(iii) \Rightarrow (b)$. Let $E = K(S)$, where $S \subseteq E$ is a set of separable elements over K. Let $a \in E$, then $a = \frac{f(u_1, ..., u_n)}{g(u_1, ..., u_n)}$, $f, g \in K[x_1, ..., x_n]$, $u_i \in S$. So, $a \in K(u_1, u_2, ..., u_n)$. Since $u_1, u_2, ..., u_n$ are separable over K, $K(u_1, u_2, ..., u_n)/K$ is separable. Therefore, a is separable over K. Thus, E/K is separable. This proves (b).

Theorem (Artin's) 3.56: *Let E be a field, G the group of automorphisms of E and suppose K is the set of elements of E fixed by G. Then K is a subfield of E, called the fixed field of G. E/K is finite if and only if G is finite. In that case,* $[E : K] = o(G)$.

Proof: $K = \{a \in E \mid \sigma(a) = a \quad \forall\ \sigma \in G\}$

$0, 1 \in K \Rightarrow K \neq \varphi$.

Let $a, b \in K$. Then $\sigma(a \pm b) = \sigma(a) \pm \sigma(b) = a \pm b \Rightarrow a \pm b \in K$. Also $\sigma(ab) = \sigma(a)\ \sigma(b) = ab \Rightarrow ab \in K$. If $b \neq 0$, then $\sigma(ab^{-1}) = \sigma(a)\ \sigma(b)^{-1} = ab^{-1} \Rightarrow ab^{-1} \in K$. So, K is a subfield of E.

Clearly, G is a group of K-automorphism of E. If E/K is finite, then the number of K-automorphisms of E is at most $[E : K]$. So, G is finite. Suppose $o(G) = r$. Let $u_0, u_1, ..., u_r \in E$ be linearly independent over K. Consider the r equations (in $r + 1$ unknowns x_js in E)

$$\sum_{j=0}^{r} \sigma(u_j)x_j = 0 \quad \text{for all } \sigma \in G$$

Since the number of equations is less than the number of unknowns, the system of equations has a non-zero solution.

Let $(a_0, a_1, ..., a_s, 0, 0, ..., 0)$ be a non zero solution of least length $s + 1$ $(a_i \neq 0\ \forall\ i = 0, 1, ..., s)$

Then $\quad \sigma(u_0)a_0 = -\sigma(u_1)a_1 + ... + -\sigma(u_s)a_s$

$\Rightarrow \quad \sigma(u_0) = \sigma(u_1)b_1 + ... + \sigma(u_s)b_s$ for all $\sigma \in G$ ...(3.4)

Take $\quad \sigma = I$. Then $u_0 = u_1b_1 + ... + u_sb_s$

If $b_i \in K$ for all i, then $(-1)u_0 + b_1u_1 + ... + b_su_s = 0$, contradicting that $u_0, u_1, ..., u_s$ linearly independent over K.

So, some $b_i \notin K$. Let $b_1 \notin K$.

Then $\exists\ \tau \in G$ such that, $\tau(b_1) \neq b_1$.

Replace σ by $\tau^{-1}\sigma$ in (i) to get

$$\tau^{-1}\sigma(u_0) = \sum_{j=1}^{r} \tau^{-1}\sigma(u_j)b_j \quad \text{for all } \sigma \in G$$

$$\Rightarrow \quad \tau(\tau^{-1}\sigma(u_0)) = \sigma(u_0) = \sum_{j=1}^{r} \sigma(u_j)\,\tau(b_j) \qquad \text{for all } \sigma \in G \quad ...(3.5)$$

Then Equation (3.5) – Equation (3.4) gives

$$\sum_{j=1}^{r} \sigma(u_j)\,(\tau(b_j - b_j) = 0, \quad \text{for all } \sigma \in G$$

$$\Rightarrow \quad \sum_{j=1}^{r} \sigma(u_j)\,c_j = 0, \quad \text{for all } \sigma \in G, \text{ where } c_j = \tau(b_j) - b_j$$

Since $c_1 = \tau(b_1) - b_1 \neq 0.$

We have a non zero solution $(0, c_1, ..., c_s, 0, ..., 0)$ of length less than $s + 1$, a contradiction.

Therefore, $r + 1$ elements in E are not linearly independenent over K

$\Rightarrow$ $[E : K] \leq r \Rightarrow E/K$ is finite.

So, $[E : K] \leq o(G)$. But $o(G) \leq [E : K]$

$\Rightarrow$ $o(G) = [E : K].$

Example 3.21: *Let E be a field with n distinct automorphisms and suppose K is the fixed field of the set of automorphisms. Show that* $[E : K] \geq n$.

Solution: Let $\sigma_1, \sigma_2, ..., \sigma_n$ be distinct automorphisms of E. Let G be the group generated by $\sigma_1, \sigma_2, ..., \sigma_n$. Then $o(G) \geq n$. If F is the fixed field of G, then $K \subseteq F \subseteq E$. By Artin's theorem, $[E : F] = o(G) \geq n$.

So, $[E : K] \geq [E : F] \geq n$.

Example 3.22: *Find the fixed field F of* $K(x)$ *under the automorphisms* $x \to 1 - x$, $x \to \frac{1}{x}$. *Show that the degree is* 6. *Verify that* $\frac{(x^2 - x + 1)^3}{(x^2 - x)^2}$ *lies in F and use this to find an equation for x over F.*

Solution: Let $\sigma(x) = 1 - x$, $\eta(x) = \frac{1}{x}$. Then $\sigma, \eta, \sigma\eta, \eta\sigma, \sigma\eta\sigma, I$ are six distinct automorphisms of $E = K(x)$. Let F' be the fixed field of these 6 automorphisms of E. So, $F \subseteq F' \subseteq E$. By previous example, $[E : F'] \geq 6 \Rightarrow [E : F] \geq 6$.

Let $g(x) = \frac{(x^2 - x + 1)^3}{(x^2 - x)^2}$

Then $\eta(g(x)) = g(x)$, $\sigma(g(x)) = g(x)$

$\Rightarrow$ $g(x) \in F$

Let $L = K(g(x)) \subseteq F \subseteq E$

Then $[E : L] = [E : F]\,[F : L] \geq 6.$

Now $L(x) = K(x) = E.$

Also, $(x^2 - x + 1)^3 - g(x)\,x^2(x - 1)^2 = 0$

$\Rightarrow x$ is a root of a polynomial of degree 6 with coefficients in L

NOTES

NOTES

$\Rightarrow \quad [L(x) : L] \leq 6$

$\Rightarrow \quad [E : L] \leq 6 \Rightarrow [E : L] = 6$

So, $\quad [E : F]\ [F : L] = 6 \leq [E : F]$

$\Rightarrow \quad [F : L] \leq 1$

$\Rightarrow \quad F = L = K(g(x))$

$\Rightarrow (x^2 - x + 1)^3 - g(x)x^2(x - 1)^2 = 0$ is an equation for x over F.

3.5.1 Primitive Elements

Theorem 3.57: *Let K/F be a finite separable extension. Then K = F (a) for some a ∈ K.*

Proof: Since K/F is finite, $K = F(a_1, ..., a_n)$ for some $a_1, ..., a_n \in K$. It is enough to prove the theorem for $n = 2$.

Let $K = F(\alpha, \beta)$. Then α, β are separable over F.

Case (i): Let F be an infinite field.

Let $\quad p(x) = Irr\ (F, \alpha)$

$\quad q(x) = Irr\ (F, \beta)$

Let $\alpha = \alpha_1, ..., \alpha_n$, $\beta = \beta_1, ..., \beta_m$ be the roots of $p(x)$, $q(x)$ respectively in a splitting fields of $p(x)$ and $q(x)$. Since K is finite, there exists $a \in K$ such that $a \neq 0$ and $a \neq \dfrac{\alpha_i - \alpha}{\beta - \beta_j}$ for $1 \leq i \leq n$, $2 \leq j \leq m$.

Since α, β are separable over F, $\alpha_i s$ and $\beta_j s$ are distinct roots of $p(x)$, $q(x)$ respectively.

Let $\quad \theta = a\beta + \alpha.$

We show that $\quad F(\theta) = F(\alpha, \beta).$

Clearly $\quad F(\theta) \subseteq F(\alpha, \beta).$

Define $\quad g(x) = p(\theta - ax).$

Then $\quad g(\beta) = p(\theta - a\beta) = p(\alpha) = 0.$

Also, $\quad g(\beta_j) = p(\theta - a\beta_j) \neq 0$ for all $j = 2, ..., m$.

(For, $p(\theta - a\beta_j) = 0 \Rightarrow \theta - a\beta_j - \alpha_i = 0$ for some i

$\Rightarrow a\beta + \alpha - a\beta_j - \alpha_i = 0$

$\Rightarrow \quad a = \dfrac{\alpha_i - \alpha}{\beta - \beta_j}$, a contradiction

Now β is a root of $g(x)$ and $q(x)$ and no β_j $(j \neq 1)$ is a root of $g(x)$ $\Rightarrow \beta$ is the only common root of $g(x)$ and $q(x)$. Let $f(x) = Irr\ (F(\theta), \beta)$.

Since $g(x) \in F(\theta)\ [x]$ and $g(\beta) = 0$, $f(x)$ divides $g(x)$. Similarly $f(x)$ divides $q(x)$

So, $f(x)$ divides g.c.d. of $g(x)$ and $q(x)$.

$\Rightarrow f(x)$ divides $x - \beta$

NOTES

$\Rightarrow \quad f(x) = x - \beta$

Since $\quad f(x) \in F(\theta)[x], \ \beta \in F(\theta)$

Also, $\quad \alpha = \theta - a\beta \in F(\theta)$

$\Rightarrow \quad F(\alpha, \beta) \subseteq F(\theta).$

Thus, $\quad F(\theta) = F(\alpha, \beta).$

Case (ii): K is finite. We shall prove later that $K^* = K - \{0\}$ is a cyclic group. If $K^* = <a>$, then $K = F(a)$.

Note: An extension K/F a called a *simple extension* if $K = F(a)$ for some $a \in K$. In the above theorem, we have shown that a finite separable extension is a simple extension. a is called a *primitive element* of K over F if $K = F(a)$.

Example 3.23: *Find a primitive element for* $\mathbf{Q}(i, 2^{1/2})$ *over* $\mathbf{Q}$.

Solution: Since char $\mathbf{Q} = 0$, $\mathbf{Q}$ is perferct. So, $\mathbf{Q}(i, 2^{1/2})/\mathbf{Q}$ is separable. Therefore, primitive element of $\mathbf{Q}(i, 2^{1/2})$ over $\mathbf{Q}$ exists.

Let $\quad p(x) = Irr\,(\mathbf{Q}, 2^{1/2}) = x^2 - 2 = (x - 2^{1/2})\,(x + 2^{1/2})$

$\quad q(x) = Irr\,(\mathbf{Q}, i) = x^2 + 1 = (x - i)\,(x + i).$

Consider $\dfrac{-2^{\frac{1}{2}} - 2^{\frac{1}{2}}}{i - (-i)} = \dfrac{-2^{\frac{1}{2}}}{i} = -2^{1/2}i.$

Take $\quad a = 1.$

Then $\quad \theta = a\beta + \alpha = i + 2^{1/2}.$

By above theorem $\mathbf{Q}(i, 2^{1/2}) = \mathbf{Q}(\theta) = \mathbf{Q}(i + 2^{1/2})$.

3.6 GALOIS EXTENSIONS

Definition: An extension E of F is called a *Galois extension* if

(i) E/F is finite

(ii) F is the fixed field of a group of automorphisms of E.

We first find a necessary and sufficient condition for a finite extension to be Galois.

Theorem 3.58: *Let E/F be a finite extension. Then E/F is a Galois extension if and only if it is both normal and seperable.*

Proof: Let E/F be a Galois extension.Then F is the fixed field of a group G of automorphisms of E. By Artin's theorem, since E/F is finite, G is also finite.

Let $\quad G = \{\sigma_1 = I, \sigma_2, ..., \sigma_n\}.$

Let $\quad a \in E.$

Let $\quad \sigma_i(a) = a_i, \ i = 1, 2, ..., n.$

Suppose $a_1 = a, a_2, ..., a_r$ are distinct elements of $\{a_1, a_2, ..., a_n\}$.

Let $\quad S = \{a_1, a_2, ..., a_r\}.$ Then $S \subseteq E$.

NOTES

Now $\sigma_j(a_i) = \sigma_j\,\sigma_i(a) = \sigma_k(a) = a_k \in S$.

So, $\sigma_j : S \to S$ for all $j = 1, 2, ..., n$. Since $\sigma_j : E \to E$ is 1-1, so is $\sigma_j : S \to S$. Also, S is finite $\Rightarrow \sigma_j : S \to S$ is onto. Therefore, σ_j is a permutation of S for all j.

Let $f(x) = (x - a_1) ... (x - a_r)$

$= x^r + \alpha_1 x^{r-1} + ... + \alpha_r x^r$

Now $\sigma_t(f(x)) = (x - \sigma_t(a_1)) ... (x - \sigma_t(a_r))$

$= (x - a_1) ... (x - a_r) = f(x)$ for all t.

So, $x^r + \sigma_t(\alpha_1)x^{r-1} + ... + \sigma_t(\alpha_r) = x^r + \alpha_1 x^{r-1} + ... + \alpha_r$

$\Rightarrow \sigma_t(\alpha_i) = \alpha_i$ for all t and i

$\Rightarrow \alpha_i$ belongs to the fixed field of G

$\Rightarrow \alpha_i \in F$, for all i

$\Rightarrow f(x) \in F[x]$.

Let $g(x)$ be a monic irreducible factor of $f(x)$ in $F[x]$.

Let a_i be a zero of $g(x)$ in E.

Now $a_j = \sigma_j(a) = \sigma_j\,\sigma_i^{-1}(a_i) = \sigma_t(a_i)$. So, a_i is a zero of $g(x)$ in E.

$\Rightarrow \sigma_t(a_i)$ is a zero of $\sigma_t(g(x)) = g(x)$ in E

$\Rightarrow a_j$ is a zero of $g(x)$ in E for all j

$\Rightarrow g(x) = f(x)$

$\Rightarrow f(x) = Irr\,(F, a)$.

Since a is a simple zero of $f(x)$, a is separable over F. So, E/F is separable. Also, $f(x)$ splits in $E[x]$.

$\Rightarrow E/F$ is normal.

Conversely, let G be the group of all F-automorphisms of E. Let F' be the fixed field of G.

Then $F \subseteq F' \subseteq E$ and $o(G) = [E : F]$.

Since E/F is finite, So is E/F'.

Also, E/F is seperable normal $\Rightarrow E/F'$ is separable, normal.

Therefore, there are exactly $n = [E : F]$ F-automorphisms of E.

$\Rightarrow o(G) = n \Rightarrow [E : F'] = n$

$\Rightarrow [F' : F] = 1 \Rightarrow F' = F$.

$\Rightarrow F$ is the fixed field of $G \Rightarrow E/F$ is Galois.

Corollary 1: Let E/F be finite extension. Then E/F is Galois if and only if F is the fixed field of the group of all F-automorphisms of E.

Proof: Let E/F be Galois. Then from above E/F is finite, normal, separable. Again by converse part of the above result, F is the fixed field of the group of all F-automorphisms of E. Converse, follows by definition.

NOTES

Corollary 2: Let char $k = 0$. Then k is contained in some Galois extension of k.

Proof: Let $f(x)$ be a non constant polynomial in $k[x]$. Let E be a minimal splitting field of $f(x)$ over k. Then E/K is finite normal. Since char $k = 0$, k is perfect $\Rightarrow$ E/K is separable. So, E/K is Galois.

Note: When E/F is Galois, the group of all F-automorphisms of E is denoted by $Gal(E/F)$ or $G(E/F)$ called the *Galois group* of E/F.

Theorem 3.59: *Let E/F be a finite extension. Then E/F is contained in a Galois extension if and only if it is separable.*

Proof: Let E/F be a contained in a Galois extension E'/F. Then $F \subseteq E \subseteq E'$.

Now E'/F is Galois $\Rightarrow$ E'/F is separable $\Rightarrow$ E/F' is separable.

Conversely, let E/F be separable. Since E/F is finite,

$$E = F(\alpha_1, \alpha_2, ..., \alpha_n).$$

Let $p_i = Irr(F, \alpha_i), \quad \alpha_i \in E$

$\alpha_i \in E \Rightarrow \alpha_i$ is separable over F

$\Rightarrow \alpha_i$ is a simple zero of p_i, for all i

$\Rightarrow$ Each zero of p_i in a splitting field is simple

Let $f = \prod_{i=1}^{n} p_i$. Then $f \in k[x] \subseteq E[x]$, and f splits in some extension of E.

Let L be a minimal splitting field of $f(x)$ over F.

Then $L = F$ (zeros of f in an extension of E)

$= F$ ($\alpha_1, \alpha_2, ..., \alpha_n$, zeros of f other than $\alpha_i s$ in an extension of E)

$= E$ (zeros of f other than $\alpha_i s$ in an extension of E)

$\Rightarrow \quad F \subseteq E \subseteq L$

Also, L is generated by separable elements over F (as each zero of f in an extension of E is simple and is a zero of an irreducible polynomial of $p_i \in F[x]$) $\Rightarrow$ L/F is separable $\Rightarrow$ E/F is contained in a separable extension L/F.

Theorem 3.60: *Let E/k be Galois and F be any extension of k. Then EF/F is Galois and G(EF/F) is isomorphic to a subgroup of G(E/k).*

Proof: Since E/k is Galois, E/k is finite normal. So, E is a minimal splitting field of some polynomial $f(x) \in k[x]$.

Let $f(x) = \alpha(x - \alpha_1)(x - \alpha_2) ... (x - \alpha_n), \ \alpha_i \in E, \ \alpha \in k.$

Then $E = k(\alpha_1, \alpha_2, ..., \alpha_n).$

Also, E/k is separable

$\Rightarrow$ Each α_i is separable over k. Now $k \subseteq F \subseteq EF$ and α_i is separable over $k \Rightarrow \alpha_i$ is separable over F.

Again, $E = k(\alpha_1, \alpha_2, ..., \alpha_n).$

$\Rightarrow \quad EF = FE = Fk(\alpha_1, \alpha_2, ..., \alpha_n)$

$= F(\alpha_1, \alpha_2, ..., \alpha_n)$ as $k \subseteq F$

$\Rightarrow$ EF is a minimal splitting field of $f(x)$ over F

$\Rightarrow$ EF/F is finite normal

NOTES

Also, EF is generated by separable elements over F

$\Rightarrow EF/F$ is separable.

So, EF/F is Galois.

Let $\sigma \in G(EF/F)$.

Let $f = \alpha f_1 f_2 ..f_r$ where each f_i is monic irreducible polynomial in $k[x]$.

So, each α_i is a zero of some $f_j \in k[x]$.

Since α_i is separable over k, α_i is a simple zero.

Let $S = \{\alpha_1, \alpha_2, ..., \alpha_n\}$. Then α_i is a zero of f in $E \subseteq EF$

$\Rightarrow$ $\sigma(\alpha_i)$ is a zero of $\sigma(f) = f$ in $EF \Rightarrow \sigma(\alpha_i) \in S$.

So, $\{\sigma(\alpha_1), \sigma(\alpha_2), ..., \sigma(\alpha_n)\} = \{\alpha_1, \alpha_2, ..., \alpha_n\}$

$\Rightarrow$ $\sigma(E) = k(\sigma(\alpha_1), \sigma(\alpha_2), ..., \sigma(\alpha_n))$

$= k(\alpha_1, \alpha_2, ..., \alpha_n) = E$

$\Rightarrow$ σ restricted to E belongs to $G(E/k)$

Define $\theta : G(EF \mid F) \to G(E/k)$ such that,

$\theta(\sigma) = \sigma \mid E$

Then θ is a homomorphism.

Also θ is 1 – 1 as $\sigma \mid E = I$

$\Rightarrow$ $\sigma(\alpha_i) = \alpha_i$ for all i

$\Rightarrow$ $\sigma(a) = a$ for all $a \in EF$ as $EF = F(\alpha_1, \alpha_2, ..., \alpha_n)$ and σ fixes each element of F

$\Rightarrow \sigma = I$ on EF.

So, $G(EF/F) \cong \theta(G(EF/F)) \leq G(E/F)$.

Corollary: If E/k is Galois and F, an extension of k, then $[EF : F]$ divides $[E : k]$.

Proof: By above theorem, EF/F is Galois

$\Rightarrow$ $[EF : F] = o(G(EF/F))$

Also, $[E : k] = o(G(E/k))$

But $\theta(G(EF/F)) \leq G(E/F)$

$\Rightarrow$ $o(\theta(G(EF/F))$ divides $o(G(E/F))$

$\Rightarrow$ $o(G(EF/F))$ divides $o(G(E/F))$

$\Rightarrow$ $[EF : F]$ divides $[E : k]$.

Note: The above corollary need not be true if E/k is not Galois. For example, let $k = \mathbf{Q}$, let α be the real cube root of 2. Then α, αw, αw^2 are roots of $f(x) = x^3 - 2$ in $\mathbf{C}$.

Let $E = \mathbf{Q}(\alpha w)$, $F = \mathbf{Q}(\alpha)$.

Then $EF = \mathbf{Q}(\alpha w)\ \mathbf{Q}(\alpha) = \mathbf{Q}(\alpha, \alpha w) = \mathbf{Q}(\alpha, \sqrt{3}i)$

$= F(\sqrt{3}i)$

So, $[EF : F] = [F(\sqrt{3}i) : F] = 2$

while $[E : k] = [\mathbf{Q}(\alpha w) : \mathbf{Q}] = \deg Irr\ (\mathbf{Q}, \alpha w)$

$= \deg f(x) = 3.$

NOTES

3.6.1 Fundamental Theorem of Galois Theory

Theorem (The fundamental theorem of Galois Theory) 3.61: *Let E/k be Galois. Let G = G(E/k) be the group of all k-automorphisms of E. Then*

(*i*) *There is one-one correspondence between the sets* $\mathtt{A} = \{F \mid F = \text{field}, k \subseteq F \subseteq E\}$ *and* $\mathtt{B} = \{H \mid H \leq G\}$ *which is an order inverting bijection.*

(*ii*) $F \in \mathtt{A}$ *is the fixed field of the subgroup* $H \in \mathtt{B}$ *corresponding to F and* $H \in \mathtt{B}$ *is the group of H*-automorphisms of E, where H* is the fixed field of H.*

(*iii*) *If H is the subgroup of* B *corresponding to the field F in* A*, then* $o(H) = [E : F]$ *and* $[G : H] = [F : k]$.

(*iv*) *If* $H_1, H_2 \in \mathtt{B}$ *corresponding to* $F_1, F_2 \in \mathtt{A}$ *respectively, then* F_1, F_2 *are conjugate under an automorphism* $\sigma \in G$ *if and only if* $\sigma^{-1} H_1 \sigma = H_2$.

(*v*) *If* $H \in \mathtt{B}$ *corresponds to* $F \in \mathtt{A}$, *then F/k is normal if and only if H is normal subgroup of G and in that case,* $G(F/k) \cong \frac{G}{H}$.

Proof: Define $\theta : \mathtt{A} \to \mathtt{B}$ such that,

$$\theta(F) = F^*$$

where $F^* = \{\sigma \in G \mid \sigma(x) = x \text{ for all } x \in F\}$. Then $F^* \in \mathtt{B}$.

Similarly, define $\theta : \mathtt{B} \to \mathtt{A}$ such that,

$$\theta(H) = H^*$$

where $H^* = \{x \in E \mid \sigma(x) = x \text{ for all } \sigma \in H\}$

Then $H^* \in \mathtt{A}$ is the fixed field of H.

Let $F_1, F_2 \in \mathtt{A}$ such that $F_1 \subseteq F_2$.

Let $\sigma \in F_2^*$. Then $\sigma(x) = x$ for all $x \in F_2$

$\Rightarrow$ $\sigma(x) = x$ for all $x \in F_1$ as $F_1 \subseteq F_2$

$\Rightarrow$ $\sigma \in F_1^* \Rightarrow F_2^* \subseteq F_1^*$

$\Rightarrow$ $\theta(F_2) \subseteq \theta(F_1) \Rightarrow \theta$ is an order inverting map.

Similarly, φ is an order inverting map.

Let $H \in \mathtt{B}$. Then $\sigma \in H \Rightarrow \sigma(x) = x$ for all $x \in H^* \Rightarrow \sigma \in H^{**} \Rightarrow H \subseteq H^{**}$.

NOTES

Also $x \in F (F \in \mathrm{A}) \Rightarrow \sigma(x) = x$ for all $\sigma \in F^*$

$\Rightarrow x$ belongs to the fixed field of F^*

$\Rightarrow \quad x \in F^{**} \Rightarrow F \subseteq F^{**}$ for all $F \in \mathrm{A}$.

Let $F \in \mathrm{A}$ and $F^* = H$. Then $H^{**} = F^{***}$.

Now $H \subseteq H^{**} \Rightarrow F^* \subseteq F^{***}$ for all $F \in \mathrm{A}$.

Also, $F \subseteq F^{**} \Rightarrow \theta(F^{**}) \subseteq \theta(F)$

$\Rightarrow F^{***} \subseteq F^*$ for all $F \in \mathrm{A}$. So, $F^* = F^{***}$. Similarly, $H^* = H^{***}$ for all $H \in \mathrm{B}$.

Now θ is 1-1 onto if and only if $\theta\varphi$ = Identity and $\varphi\theta$ = Identity if and only if $H = H^{**}$ for all $H \in \mathrm{B}$ and $F = F^{**}$ for all $F \in \mathrm{A}$.

Let $H \in \mathrm{B}$. Then $H^* = F$ is the fixed field of H.

By Artin's theorem $o(H) = [E : F]$.

Also, $o(H^{**}) = [E : H^{***}] = [E : H^*] = [E : F]$.

So, $o(H) = o(H^{**})$. But $H \subseteq H^{**}$. Therefore, $H = H^{**}$.

Let $F \in \mathrm{A}$. Then $k \subseteq F \subseteq E$.

Now E/k is Galois $\Rightarrow E/F$ is Galois $\Rightarrow F$ is the fixed field of the group H of all F-automorphisms of E.

$\Rightarrow \quad H \leq G \Rightarrow H \in \mathrm{B}$.

Now H^* = fixed field of $H = F$

$\Rightarrow \quad H^{***} = F^{**} \Rightarrow H^* = F^{**} \Rightarrow F = F^{**}$ for all $F \in \mathrm{A}$.

Thus, θ is 1-1 onto.

This proves (*i*).

(*ii*) Let $F \in \mathrm{A}$. Let $\theta(F) = H$. Then $F^* = H \Rightarrow F^{**} = H^* \Rightarrow F = H^* \Rightarrow F$ is the fixed field of H.

Let $H \in \mathrm{B}$. Then there exists $F \in \mathrm{A}$ such that $\theta(F) = H \Rightarrow H = F^*$.

Let $\sigma \in H$. Then $\sigma \in F^* \Rightarrow \sigma(x) = x$ for all $x \in F \Rightarrow \sigma$ is an F-automorphism of E.

Conversely, let σ be an F-automorphism of E.

Then $\sigma(x) = x$ for all $x \in F \Rightarrow \sigma \in F^* = H$.

So, H is the group of all $F = H^*$-automorphisms of E.

(*iii*) By Artin's theorem

$$o(H) = [E : H^*] = [E : F]$$

$$[G : H] = \frac{o(G)}{o(H)} = \frac{[E : k]}{[E : F]} = [F : k].$$

(*iv*) Suppose $F_1, F_2 \in \mathrm{A}$ are conjugate under $\sigma \in G$. Then $\sigma(F_1) = F_2$.

Let $y \in F_2$. Then $y = \sigma(z)$, $z \in F_1$. Therefore, $\sigma^{-1}(y) = z$.

$\Rightarrow \quad \tau\sigma^{-1}(y) = \tau(z)$, for all $\tau \in H_1$

NOTES

$\Rightarrow \quad \sigma\tau\sigma^{-1}(y) = \sigma\tau(z) = \sigma(z), \quad$ for all $\tau \in H_1$

$\Rightarrow \quad \sigma\tau\sigma^{-1}(y) = y, \quad$ for all $\tau \in H_1, y \in F_2$

$\Rightarrow \quad \sigma\tau\sigma^{-1} \in H_2, \quad$ for all $\tau \in H_1$

$\Rightarrow \quad \sigma H_1\sigma^{-1} \subseteq H_2$

Let $a \in F_1$. Then $\sigma(a) = b \in F_2$

$\Rightarrow \quad \eta\sigma(a) = \eta(b), \quad$ for all $\eta \in H_2$

$\Rightarrow \quad \eta\sigma(a) = b, \quad$ for all $\eta \in H_2$

$\Rightarrow \quad \sigma^{-1}\eta\sigma(a) = \sigma^{-1}(b) = a,$ for all $\eta \in H_2, a \in F_1$

$\Rightarrow \quad \sigma^{-1}\eta\sigma \in H_1, \quad$ for all $\eta \in H_2$

$\Rightarrow \quad \sigma^{-1}H_2\,\sigma \subseteq H_1$

$\Rightarrow \quad H_2 \subseteq \sigma H_1\sigma^{-1}$

So, $H_2 = \sigma H_1\sigma^{-1}$.

Conversely, let $H_2 = \sigma H_1\sigma^{-1}$ for $\sigma \in G$.

Let $y \in F_2$. Now $\sigma\tau\sigma^{-1} \in H_2, \quad$ for all $\tau \in H_2$

$\Rightarrow \quad \sigma\tau\sigma^{-1}(y) = y$

$\Rightarrow \quad \tau\sigma^{-1}(y) = \sigma^{-1}(y) = z$

$\Rightarrow \quad \tau(z) = z, \quad$ for all $\tau \in H_1$

$\Rightarrow \quad z \in F_1$

$\Rightarrow \quad y = \sigma(z) \in \sigma(F_1)$

$\Rightarrow \quad F_2 \subseteq \sigma(F_1)$

Let $x \in F_1$. Now $\sigma^{-1}\eta\sigma \in H_1, \quad$ for all $\eta \in H_2$

$\Rightarrow \quad \sigma^{-1}\eta\sigma(x) = x$

$\Rightarrow \quad \eta\sigma(x) = \sigma(x) = x'$

$\Rightarrow \quad \eta(x') = x', \quad$ for all $\eta \in H_2$

$\Rightarrow \quad x' \in F_2$

$\Rightarrow \quad \sigma(x) \in F_2$

$\Rightarrow \quad \sigma(F_1) \subseteq F_2$.

So, $\sigma(F_1) = F_2 \Rightarrow F_2$ are conjugate under σ.

(*v*) Suppose F/k is normal. Since E/k is finite, so is F/k. Therefore, F/k is finite normal $\Rightarrow F$ is a minimal splitting field of some $f \in k[x]$.

Let $f = \alpha(x - \alpha_1) \ldots (x - \alpha_n), \quad \alpha_i \in E, \alpha \in k.$

Then $F = k(\alpha_1, \alpha_2, \ldots, \alpha_n)$.

Let $\sigma \in G$. Then σ is a k-autmorphism of $E \Rightarrow \sigma(f) = f$.

$\Rightarrow \quad f = \alpha(x - \sigma(\alpha_1)) \ldots (x - \sigma(\alpha_n))$

$\Rightarrow \quad \sigma(\alpha_1), \ldots, \sigma(\alpha_n)$ are zeros of f in E

$\Rightarrow \quad \{\alpha_1, \alpha_2, \ldots, \alpha_n\} = \{\sigma(\alpha_1), \ldots, \sigma(\alpha_n)\}$.

NOTES

So, $\sigma(F) = k(\sigma(\alpha_1), ..., \sigma(\alpha_n))$

$= k(\alpha_1, ..., \alpha_n) = F$ for all $\sigma \in G$.

By (*iv*), $\sigma^{-1}H\sigma = H$ for all $\sigma \in G$

$\Rightarrow H$ is a normal subgroup of G.

Conversely, let $H = F^*$ be normal subgroup of G. Then $\sigma^{-1} H\sigma = \sigma$ for all $\sigma \in G$.

$\Rightarrow \sigma(F) = F$ by (*iv*) for all $\sigma \in G$

Let $\alpha \in F$, $p(x) = Irr\ (k, \alpha)$.

Since E/k is normal and $\alpha \in E$, we find $p(x)$ splits in E.

Let β be a zero of $p(x)$ in E.

Then α, β are zeros of $p(x)$ in E.

$\Rightarrow$ There is an isomorphism $\theta : k(\alpha) \to k(\beta)$ such that,

$\theta(\alpha) = \beta$, $\theta(a) = a$ for all $a \in k$.

Since $\beta \in E$, $k(\beta) \subseteq E$. So θ is a k-homomorphisms from $k(\alpha)$ to E.

Since E/K is finite normal, θ can be extended to k-automorphism σ of E. So, $\sigma \in G$.

Now $\sigma(\alpha) = \theta(\alpha) = \beta$ and $\sigma(\alpha) \in \sigma(F) = F \Rightarrow \beta \in F$.

Thus, $p(x)$ splits in $F \Rightarrow F/K$ is normal.

Let H be a normal subgroup of G. Then the corresponding field F is normal over k from above. Since E/k is Galois, so is F/k. Let $N = \text{Gal}(F/k)$

Define $\psi : G \to N$ such that,

$\psi(\sigma) = \bar{\sigma}$, where $\bar{\sigma}$ is the restriction of σ on F.

(Since $H \leq G$, $\sigma^{-1}H\sigma = H \Rightarrow \sigma(F) = F$)

Let $\sigma, \eta \in G$.

Then $\bar{\sigma}\ \bar{\eta}(\alpha) = (\sigma\eta)(\alpha)$, $\alpha \in F$

$= \sigma(\eta(\alpha))$, $\eta(\alpha) \in F$

$= \bar{\sigma}(\eta(\alpha))$

$= \bar{\sigma}(\bar{\eta}(\alpha))$

$= (\bar{\sigma}\ \bar{\eta})(\alpha)$, for all $\alpha \in F$

$\Rightarrow \quad \overline{\sigma\eta} = \bar{\sigma}\ \bar{\eta}$

$\Rightarrow \quad \psi(\sigma\eta) = \psi(\sigma)\psi(\eta)$

$\Rightarrow \psi$ is a homomorphism

Let $\theta \in N$. Then θ can be extended to k-automorphism σ of $E \Rightarrow \sigma \in G$

$\Rightarrow \psi(\sigma) = \bar{\sigma} = \theta$. So, ψ is onto. Now $\sigma \in \text{Ker}\ \psi \Leftrightarrow \psi(\sigma) =$ Identity of $N \Leftrightarrow \bar{\sigma} =$ Identity on $F \Leftrightarrow \bar{\sigma}(\alpha) = \alpha$, for all $\alpha \in F$.

The result now follows by using fundamental theorem of homomorphism.

Case 9: (*i*) Let E be a minimal splitting field of $f(x) = x^3 - 2$ over $\mathbf{Q}$. Let α be the real cube root of 2.

NOTES

Then $E = \mathbf{Q}(\alpha, \alpha w, \alpha w^2) = \mathbf{Q}(\alpha, \sqrt{3}\ i) = \mathbf{Q}(\alpha, \alpha w) = \mathbf{Q}(\alpha, w)$

Also, $[E : \mathbf{Q}] = 6$. Since char $\mathbf{Q} = 0$, $E/\mathbf{Q}$ is separable (as $\mathbf{Q}$ is perfect $\Rightarrow$ every algebraic extension of $\mathbf{Q}$ is separable.)

Also, E is a minimal splitting field of $f(x)$ over $\mathbf{Q} \Rightarrow E/\mathbf{Q}$ is finite normal.

So, $E/\mathbf{Q}$ is Galois.

Let $G = G(E/\mathbf{Q})$ be the group of all $\mathbf{Q}$-automorphisms of E.

Then $\mathbf{Q}$ is the fixed field of G. By Artin's theorem $o(G) = [E : \mathbf{Q}] = 6$.

Since α, αw are roots of $f(x)$, there exists $\mathbf{Q}$-isomorphism

$\sigma_0 : \mathbf{Q}(\alpha) \to \mathbf{Q}(\alpha w)$ such that,

$$\sigma_0(\alpha) = \alpha w$$

Let $g(x) = x^2 + x + 1$, then $g(x)$ is irreducible over $\mathbf{Q}(\alpha) \subseteq \mathbf{R}$

and $\sigma_0(g(x)) = g(x)$ is irreducible over $\mathbf{Q}(\alpha w)$

Since w, w are roots of $g(x)$, there exists an isomorphism

$\sigma : \mathbf{Q}(\alpha, w) = E \to \mathbf{Q}(\alpha w, w) = E$ such that,

$$\sigma(w) = w$$
$$\sigma(\alpha) = \sigma_0(\alpha) = \alpha w$$
$$\sigma(a) = \sigma_0(a) = a \quad \forall\ a \in \mathbf{Q}$$

Thus σ is $\mathbf{Q}$-automorphism of E, $\sigma \neq I$.

Also w, w^2 are roots of $g(x)$ which is irreducible over $\mathbf{Q}(\alpha)$ and $\exists$ $\mathbf{Q}(\alpha)$ isomorphism

$\tau : \mathbf{Q}(\alpha, w) = E \to \mathbf{Q}(\alpha, w^2) = E$ such that,

$$\tau(w) = w^2,\ \tau(\alpha) = \alpha$$

and so τ is $\mathbf{Q}$-automorphism of E, $\tau \neq \sigma$, $\tau \neq I$

Now $\sigma^2(\alpha) = \alpha w^2,\ \sigma^2(w) = w$

$(\sigma\ \tau)\ (\alpha) = \alpha w,\ (\sigma\tau)\ (w^2) = w^2$

$(\sigma^2\tau)\ (\alpha) = \alpha w^2,\ (\sigma^2\tau)\ (w^2) = w^2.$

Since $o(G) = 6,\ G = \{I, \sigma, \sigma^2, \tau, \sigma\tau, \sigma^2\tau\}$

Also $(\tau\sigma)\ (\alpha) = \tau(\alpha w) = \alpha w^2,\ \tau\sigma \neq \sigma\tau$

So G is a non abelian group of order 6 and so $G \cong S_3$.

Denote αw by 1, αw^2 by 2 and αw^3 by 3 and we get

$$\tau = (12),\ \sigma T = (13),\ \sigma^2 T = (23),$$
$$\sigma = (123),\ \sigma^2 = (132)$$

Write $\tau = \sigma_2,\ \sigma = \sigma_3,\ \sigma T = \sigma_4,\ \sigma^2 = \sigma_5$ and $\sigma^2\ T = \sigma_6$

Then $G = \{I, \sigma_2, \sigma_3, \sigma_4, \sigma_5, \sigma_6\}$

Subgroups of G are:

$$H_1 = \{I, \sigma_2\},\ H_2 = \{I, \sigma_4\},$$
$$H_3 = \{I, \sigma_6\},\ H_4 = \{I, \sigma_3, \sigma_5\},\ H_5 = G,\ H_6 = \{I\}.$$

Let $F_1 = H_1^*$, the fixed field of H_1.

NOTES

Now H_1 fixes $\alpha \Rightarrow \mathbf{Q} \subseteq \mathbf{Q}(\alpha) \subseteq F_1 \subseteq E$.

But $[\mathbf{Q}(\alpha) : \mathbf{Q}] = 3, [E : F_1] = [E : H_1^*] = o(H_1) = 2$

and $[E : \mathbf{Q}] = 6 \Rightarrow F_1 = \mathbf{Q}(\alpha)$.

Let $F_2 = H_2^*$, the fixed field of H_2.

Then $F_2 = \mathbf{Q}(\alpha w^2)$ and F_3, the fixed field of H_3 is $\mathbf{Q}(\alpha\omega)$

Let $F_4 = H_4^*$, the fixed field of H_4. Now H_4 fixes $\sqrt{3i}$

$\Rightarrow \mathbf{Q} \subseteq \mathbf{Q}(\sqrt{3i}) \subseteq F_4 \subseteq E$.

Since $[E : F_4] = 3$. $[\mathbf{Q}(\sqrt{3}\, i) : \mathbf{Q}] = 2$, $[E : \mathbf{Q}] = 6$, $F_4 = \mathbf{Q}(\sqrt{3}\, i)$.

Clearly, $F_5 =$ Fixed field of $G = \mathbf{Q}$

and $F_6 =$ Fixed field of $H_6 = E$.

So, we have 6 intermediate fields between $\mathbf{Q}$ and E corresponding to 6 subgroups of G.

Since H_1, H_2, H_3 are not normal, $F_1/\mathbf{Q}, F_2/\mathbf{Q}, F_3/\mathbf{Q}$ are also not normal. Also H_4, H_5, H_6 are normal subgroup of G, and thus $F_4/\mathbf{Q}, F_5/\mathbf{Q}, F_6/\mathbf{Q}$ are normal subgroups of G.

(*ii*) Let E be a minimal splitting field of $f(x) = x^4 + 1$ over $\mathbf{Q}$.

Then $\alpha, \alpha^3, \alpha^5, \alpha^7$ are roots of $f(x)$, where $\alpha = \cos\frac{\pi}{4} + i\sin\frac{\pi}{4}$

and $E = \mathbf{Q}(\alpha) = \mathbf{Q}(\alpha^3) = \mathbf{Q}(\alpha^5) = \mathbf{Q}(\alpha^7)$

Then $[E : \mathbf{Q}] = [\mathbf{Q}(\alpha) : \mathbf{Q}] = \deg Irr\ (\mathbf{Q}, \alpha) = \deg f(x) = 4$.

Char $\mathbf{Q} = 0 \Rightarrow E/\mathbf{Q}$ is separable.

Also E is a minimal splitting field of $f(x)$ over $\mathbf{Q}$ implies $E/\mathbf{Q}$ is normal.

Hence $E/\mathbf{Q}$ is Galois.

Let $G = G(E/\mathbf{Q})$ be the Galois group of $E/\mathbf{Q}$.

By Artin's theorem, $o(G) = [E : \mathbf{Q}] = 4$

Since α and α^3 are roots of an irreducible polynomial $f(x)$ over $\mathbf{Q}$, there exists $\mathbf{Q}$-automorphism

$\sigma_3 : \mathbf{Q}(\alpha) = E \to \mathbf{Q}(\alpha^3) = E$, such that,

$\sigma_3(\alpha) = \alpha^3$

Similarly, there exists $\mathbf{Q}$-automorphisms

$\sigma_5 : \mathbf{Q}(\alpha) = E \to \mathbf{Q}(\alpha^5) = E$ such that,

$\sigma_5\ (\alpha) = \alpha^5$

$\sigma_7 : \mathbf{Q}(\alpha) = E \to \mathbf{Q}(\alpha^7) = E$ such that,

$\sigma_7(\alpha) = \alpha^7$

So $G = \{I, \sigma_3, \sigma_5, \sigma_7\}$

Also $\sigma_3^2 = \sigma_5^2 = \sigma_7^2 = I$

Thus G is an abelian non cyclic group of order 4 and so it is the Klein's four group.

NOTES

Subgroups of G are

$$H_1 = \{I, \sigma_3\},\ H_2 = \{I, \sigma_5\},$$
$$H_3 = \{I, \sigma_7\},\ H_4 = G,\ H_5 = \{I\}.$$

Now $\sigma \in G$

$\Rightarrow \quad \sigma(\sqrt{2})^2 = \sigma(2) = 2$

$\Rightarrow \quad (\sigma(\sqrt{2}))^2 = 2 = 0$

$\Rightarrow$ $\sigma(\sqrt{2})$ is a zero of $x^2 + 2$ in $E \subseteq \mathbf{C}$

$\Rightarrow \quad \sigma(\sqrt{2}) = \pm\sqrt{2}$. Similarly $\sigma(i) = \pm i$.

So, $\quad \sigma_3(\alpha) = \alpha^3 \Rightarrow \sigma_3\left(\frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}}\right) = \frac{-1}{\sqrt{2}} + \frac{i}{\sqrt{2}}$

$\Rightarrow \sigma_3(\sqrt{2}) = -\sqrt{2}$, $\sigma_3(i) = -i$

$\Rightarrow \sigma_3(\sqrt{2}\, i) = \sqrt{2}\, i$

$\Rightarrow H_1$ fixes $\sqrt{2}\, i$

Let $F_1 = H_1^*$, the fixed field of H_1

Then $\quad \mathbf{Q} \subseteq \mathbf{Q}(\sqrt{2}\, i) \subseteq F_1 \subseteq E$

But $[\mathbf{Q}(\sqrt{2}\, i) : \mathbf{Q}] = 2$, $[E : F_1] = 2$, $[E : \mathbf{Q}] = 4$

So, $\quad F_1 = \mathbf{Q}(\sqrt{2}\, i)$

Also, $\quad \sigma_5(\alpha) = \sigma^5 \Rightarrow \sigma_5\left(\frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}}\right) = -\frac{1}{\sqrt{2}} - \frac{i}{\sqrt{2}}$.

$\sigma_5(\sqrt{2}) = -\sqrt{2}$ and $\sigma_5(i) = i \Rightarrow H_2$ fixes i.

Let $F_2 = H_2^*$, the fixed field of H_2.

Then $\quad \mathbf{Q} \subseteq \mathbf{Q}(i) \subseteq F_2 \subseteq E$

and $\quad [E : F_2] = 2$, $[\mathbf{Q}(i) : \mathbf{Q}] = 2$, $[E : \mathbf{Q}] = 4 \Rightarrow F_2 = \mathbf{Q}(i)$.

Now $\quad \sigma_7(\alpha) = \sigma^7 \Rightarrow \sigma_5\left(\frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}}\right) = \frac{1}{\sqrt{2}} - \frac{i}{\sqrt{2}}$

$\Rightarrow \quad \sigma_5(\sqrt{2}) = \sqrt{2}$

$\Rightarrow H_3$ fixes $\sqrt{2}$. Let $F_3 = H_3^*$, the fixed field of H_3.

Then $\quad \mathbf{Q} \subseteq \mathbf{Q}(\sqrt{2}) \subseteq F_3 \subseteq E$

and $\quad [E : F_3] = 2$, $[\mathbf{Q}(\sqrt{2}) : \mathbf{Q}] = 2$, $[E : \mathbf{Q}] = 4 \Rightarrow F_3 = \mathbf{Q}(\sqrt{2})$.

Clearly F_4 = fixed field of H_4 $(= G)$ is $\mathbf{Q}$ and F_5 = fixed field of $H_5 = E$.

So, F_1, F_2, F_3, F_4, F_5 are intermediate fields lying between $\mathbf{Q}$ and E.

Since F_1, F_2, F_3 are quadratic extensions of $\mathbf{Q}$, $F_1/\mathbf{Q}$, $F_2/\mathbf{Q}$, $F_3/\mathbf{Q}$ are normal. Also $F_4/\mathbf{Q}$, $F_5/\mathbf{Q}$ are normal. But G being abelian, all subgroup of G are normal subgroups of G.

3.7 SOLUTION OF POLYNOMIAL EQUATIONS BY RADICALS

NOTES

In this section you will learn how to establish the solvability through radicals of polynomials of different degrees. Additionally, for polynomials which are solvable through radicals we require the Galois theoretic derivation of the general solution to the polynomial. The solvability through radicals can be revealed using the Galois Theory and also the characteristics of Group and Field theory.

Polynomials of degree one and two are simply established to be solvable by radicals because of the existence of similar general formula for both. Complex formulas for cubic and quartic polynomials are solved by radicals. Though, general polynomials of degree five are not solvable and so there is no general formula for this.

Basically, the polynomials are functions of the type,

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

where $a_n \neq 0$. The root(s) of a polynomial are considered the value(s) of x which satisfy the condition $p(x) = 0$. To solve the polynomial roots using radicals does not mean to find a root, because as per the fundamental theorem of algebra any polynomial of degree n has n complex roots which should not be distinct. Solving a polynomial by radicals involves the expression of all roots of a polynomial including the four basic operations: addition, subtraction, multiplication and division, and also taking the radicals from the arithmetic grouping of coefficients of any given polynomial. Solving for polynomial roots through radicals includes obtaining the general solution to the general form of a polynomial of some specific degree. The following analysis explains how all polynomials can be solved through radicals and to prove the resultant of the solvability of polynomials.

Cubic Functions: Cubic functions can be solved with the help of Cardano's method in which the general cubic equation is transformed into a depressed cubic without the x^2 term.

Consider the general form of a polynomial of degree three.

$$ax^3 + bx^2 + cx + d = 0 \qquad \ldots(3.6)$$

It is easy to work using a polynomial of foremost coefficient one, hence we divide a outside the entire equation to get,

$$x^3 + \frac{b}{a}x^2 + \frac{c}{a}x + \frac{d}{a} = 0,$$

By substituting $x = y - \frac{b}{3a}$ into above equation the polynomial becomes,

$$\left(y - \frac{b}{3a}\right)^3 + \frac{b}{a}\left(y - \frac{b}{3a}\right)^2 + \frac{c}{a}\left(y - \frac{b}{3a}\right) + \frac{d}{a}$$

$$= y^3 + y\left(\frac{b^2}{3a^2} - \frac{2b^2}{3a^2} + \frac{c}{a}\right) + \left(-\frac{b^3}{27a^3} + \frac{b^3}{9a^3} - \frac{cb}{3a^2} + \frac{d}{a}\right) = 0$$

This has been reduced to the cubic polynomial of the form,

$$y^3 + py + q = 0 \qquad \ldots(3.7)$$

Where,

$$p = \frac{b^2}{3a^2} - \frac{2b^2}{3a^2} + \frac{c}{d} \text{ and } q = -\frac{b^3}{27a^3} + \frac{b^3}{9a^3} - \frac{cb}{3a^2} + \frac{d}{a}$$

Such that,

$$(u+v)^3 - 3uv(u+v) - (u^3 + v^3) = 0 \qquad \ldots(3.8)$$

Equation (3.7) corresponds to Equation (3.8) so that,

$(u+v) = y, 3uv = -p, u^3 + v^3 = -q$

Equation (3.8) can be solved for y as follows,

$$y = w_i\left(\sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}\right),$$

where $i \in \{1,2,3\}$ and w_i is one of the 3rd roots of unity.

The general solutions for this equation is,

$$x = -\frac{b}{3a} + \frac{w_i}{3a}\left(\sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}\right)$$

Let us consider the Galois group of the irreducible depressed cubic equation. The Galois group of the splitting field of a general cubic equation is S_3 and also the possible Galois group of any cubic is isomorphic to either S_3 or A_3.

Let $f(x) = x^3 + px + q$ be an irreducible cubic in the polynomial ring $F[x]$ over a field F of characteristic zero with roots y_1, y_2 and y_3.

We include the relations, $y_1 + y_2 + y_3 = 0$, $y_1 y_2 + y_2 y_3 + y_3 y_1 = p$ and $y_1y_2y_3 = -q$.

Hence we have the chain of fields $F \subset F(y_1) \subseteq K$, where $K = F(y_1,y_2) = F(y_1,y_2,y_3)$. Hence, if two roots are in the field then the third root is automatically there.

Also, either $F(y_1) = K$ or $F(y_1) < K$.

Case (i): $F(y_1) = K$.

We know that $K = F(y_1)$ for any $i = \{1,2,3\}$ or $[K:F] = 3$.

Hence, $Gal(K/F) = A_3$. The composition series of $Gal(K/F)$ is thus $A_3 \triangleright 1$.

NOTES

NOTES

Case (ii): $F(y_1) < K$.

We know that $G = Gal(K/F)$ is a subgroup of S_3. Since $f(x)$ factors over K and $F(y_1)$ does not contain y_2, consider $h(x) = (x - y_2)(x - y_3)$. Also, $h(x)$ is irreducible over $F(y_1)$, hence $[K: F] = 6$.

Since $[K: F] = 6$ and $G = S_3$, so S_3 has only one degree 3 subgroup A_3. This implies that there exists a field L such that $[K: L] = |A_3| = 3$ and $[L: F] = 2$. L is thus acquired by adjoining a square root of the discriminant D where,

$$D = \prod_{1 \le i < j \le 3} (y_j - y_i)^2$$

We comprehend that $\sqrt{D}$ is fixed by any even permutation of the roots and $\sigma\left(\sqrt{D}\right) = -\sqrt{D}$ for any odd permutation σ where σ acts naturally on the subscripts in the above expression of D. Thus D is fixed by all of S_4, so if D is not a square $\sqrt{D} \notin F$, hence $\left[F\left(\sqrt{D}\right): F\right] = 2$ or is a radical extension. Since $Gal(K/F) = S_3$ it can be shown that $L = F(\sqrt{D})$.

Thus, $K = F(y_1, y_2) = F\left(\sqrt{D}, y_1\right)$ and the composition series of $Gal(K/F)$:

$S_3 \triangleright A_3 \triangleright 1$

This is so because,

$$\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3 = \left(-\frac{b^3}{27a^3} + \frac{b^3}{9a^3} - \frac{cb}{3a^2} + \frac{d}{c}\right)^2 + \left(\frac{b^2}{9a^2} - \frac{2b^2}{9a^2} + \frac{c}{3a}\right)$$

$$= -\frac{1}{108}(b^2c^2 - 4db^3 - 4ac^3 + 18abcd - 27d^2a^2)$$

$$= -\frac{1}{108}(y_1 - y_2)^2 (y_2 - y_3)^2 (y_1 - y_3)^2$$

Therefore, the adjoining of the square root of the discriminant gives rise to the field L which contains the term,

$$\sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}$$

Quartic Functions: Quartic polynomials can be solved using Ferrari's method which transforms a quartic polynomial into a depressed quartic which has no x^3 term.

We start with the general form of a quartic equation,

$$x^4 + ax^3 + bx^2 + cx + d = 0 \qquad \ldots.(3.9)$$

NOTES

In fact, all quartic polynomials can be reduced to the above monic polynomials by dividing throughout with the leading coefficient and replacing the coefficients of the other terms with a, b, c and d.

Substitute $x = y - \frac{a}{4}$ into Equation (3.9) to get an equation of the form,

$$y^4+py^2+ qy + r = 0 \quad \ldots(3.10)$$

We can add $2zy^2 + z^2$ to the above equation to obtain,

$y^4 + 2zy^2 + z^2 = (2z - p)y^2 - qy + (z^2 - r)$

Since we would like the right hand side to be a square so we should let the discriminant of the quadratic on the RHS be 0. Specifically, we assume that,

$q^2–4(z^2–r)(2z–p) = 0$

Rearranging the terms we get a cubic in z as,

$$8z^3 - 4pz^2 - 8rz + 4rp - q^2 = 0 \quad \ldots(3.11)$$

Thus we find the root z of this equation and solve for y by substituting that value into Equation (3.10) to get a quadratic in y^2. Solving the resultant quadratic in y^2 gives the roots of the depressed quartic from which we can derive x.

Thus we get the solutions for the quartic Equation (3.9). One root of Equation (3.9) is fixed in this formula,

$$x = \frac{1}{2}\sqrt{2z - p} \pm \sqrt{\frac{1}{2}z - \frac{1}{2}p \pm \sqrt{z^2 - r}} - \frac{a}{4}.$$

The Galois theoretic derivation of the formula is as follows.

Solving for the roots of a quartic involves solving of the cubic Equation (3.11) in z:

$$8x^3 - 4pz^2 - 8rz + 4rp - q^2 = 0$$

For a *general irreducible* quartic equation f in $F[x]$, the Galois group $G = Gal(E/F)$ is S_4.

G = S_4 has the composition series as follows:

$1 \triangleleft < \sigma > \triangleleft V \triangleleft A_4 \triangleleft S_4$

where V is the Klein 4-group. σ is any of the 3 order 2 involutions in V.

The corresponding field extension is,

$$E \supset E_\sigma \supset E_V \supset E_{A4} \supset F.$$

The part $E_{A_4} \supset$ F(corresponding to $A_4 \triangleleft S_4$) is of degree two and corresponds to the degree two extension in solving z. The element z is solved via taking a degree two extension, i.e., square root of the discriminant and followed by a cubic root (as explained eralier for cubic equations). Note that $Gal(E/F) = S_4/V$, which is isomorphic to S_3. In fact, $S_4 = VS_3 = gh$ {g in V, h in S_3}. The group V acts on E_v trivially and hence S_4/V (identified with S_3) acts on E_v

NOTES

which fixes exactly elements in F. The extension $E_\sigma \supset E_v$ is of degree 2 and corresponds to the taking of either $\sqrt{2z-p}$ or $\sqrt{z^2-r}$. These are equivalent since we have, $(2z-p)(z^2-r)=\frac{q^2}{4}$ which is a square. There are 3 possible groups $<\sigma>$ which correspond to the adjoining of the 3 possible values of z as solutions of the Equation (3.11). The last radical extension $(\boldsymbol{E \supset E_\sigma})$ corresponds to,

$$\sqrt{-\frac{1}{2}z-\frac{1}{4}p+\sqrt{z^2-r}} \text{ or } \sqrt{-\frac{1}{2}z-\frac{1}{4}p-\sqrt{z^2-r}}$$

Adjoining either of these two to $E\sigma$ will give rise to the same field E since the degree $[E: E_\sigma] = 2$.

3.7.1 Insolvability of the General Equation of Degree 5

In algebra, the Abel-Ruffini theorem also known as Abel's impossibility theorem states that there is no general algebraic solution, i.e., solution in radicals to polynomial equations of degree five or higher. This theorem states that every non-constant polynomial equation in one unknown, with real or complex coefficients, has at least one complex number as solution.

The theorem defines the *form* that a solution must take. It also states that *not all* solutions of higher degree equations can be obtained by starting with the equation's coefficients and rational constants, and repeatedly forming sums, differences, products, quotients and radicals (nth roots for some integer n) of previously obtained numbers. In fact if the roots happen to be rational numbers, they can trivially be expressed as constants. The simplest nontrivial example is the monomial equation $ax^n = b$, whose solutions are,

$$\sqrt[n]{\frac{b}{a}} \cdot e^{i2\pi k/n} \quad k = 0, 1, ..., n-1$$

Here the expression $e^{i2\pi k/n}$ appears to involve the use of the exponential function that gives the possible values of $\sqrt[n]{1}$ (the nth roots of unity), so it involves only extraction of radicals.

The Abel-Ruffini theorem states that there are *some* fifth-degree equations whose solution cannot be so expressed, for example the equation $x^5 - x + 1 = 0$. Some other fifth degree equations can be solved by radicals, for example $x^5 - x^4 - x + 1 = 0$, which factorizes to $(x - 1)(x - 1)(x + 1)(x + i)(x - i) = 0$. The precise criterion that distinguishes between those equations that can be solved by radicals and those that cannot be solved was given by É variste Galois and is termed as Galois Theory. A polynomial equation can be solved by radicals if and only if its Galois group (over the rational numbers or more generally over the base field of admitted constants) is a solvable group.

NOTES

In modern algebraic context, second, third and fourth degree polynomial equations can always be solved by radicals because the symmetric groups S_2, S_3 and S_4 are solvable groups, whereas S_n is not solvable for $n \geq 5$. This is so because for a polynomial of degree n with indeterminate coefficients (i.e., given by symbolic parameters), the Galois group is the full symmetric group S_n and is called the 'general equation of the nth degree'. This remains true if the coefficients are concrete but algebraically independent values over the base field.

The following proof is based on Galois Theory. Historically, Ruffini and Abel's proofs precede Galois Theory. One of the fundamental theorems of Galois theory states that an equation is solvable in radicals if and only if it has a solvable Galois group, so the proof of the Abel-Ruffini theorem is based on the Galois group of the general polynomial of the fifth degree.

Let y_1 be a real number transcendental over the field of rational numbers Q and let y_2 be a real number transcendental over $Q(y_1)$ and so on to y_5 which is transcendental over $Q(y_1,y_2,y_3,y_4)$. These numbers are called independent transcendental elements over Q.

Let $E = Q(y_1,y_2,y_3,y_4,y_5)$ and let,

$f(x) = (x - y_1)\ (x - y_2)\ (x - y_3)\ (x - y_4)\ (x - y_5) \in \mathrm{E}[x]$

Multiplying $f(x)$ yields the elementary symmetric functions of the y_n:

$S_1 = y_1 + y_2 + y_3 + y_4 + y_5$

$S_2 = y_1y_2 + y_1y_3 + ... + y_4y_5$

and so for,

$S_5 = y_1y_2y_3y_4y_5$.

The coefficient of x^n in $f(x)$ is thus $(-1)^{5-n}S_{5-n}$. Because our independent transcendental y_n act as indeterminate over Q, so every permutation σ in the symmetric group on 5 letters S_5 induces an automorphism σ' on E that leaves Q fixed and permutes the elements y_n. An arbitrary rearrangement of the roots of the product form produces the same polynomial of the form,

$(y - y_3)(y - y_1)(y - y_2)(y - y_5)(y - y_4)$

This is same polynomial as,

$(y - y_1)(y - y_2)(y - y_3)(y - y_4)(y - y_5)$

The automorphism σ' also leave E fixed, so they are elements of the Galois group $G(E/Q)$. Now, since $|S_5| = 5!$ so it must be $\boldsymbol{|G(E/Q)| \geq 5!}$, as there could possibly be automorphism there that is not in S_5. However, since the splitting field of a quintic polynomial has at most 5! elements because $\boldsymbol{|G(E/Q)| \geq 5!}$ and so $G(E/Q)$ must be isomorphic to S_5. Generalizing this argument shows that the Galois group of every general polynomial of degree n is isomorphic to S_n.

The only composition series of S_5 is $S_5 \geq A_5 \geq \{e\}$, where A_5 is the alternating group on five letters also known as the icosahedral group. However, the quotient

NOTES

group $A_5/\{e\}$ which is isomorphic to A_5 itself is not an abelian group and so S_5 is not solvable. Hence, the general polynomial of the fifth degree has no solution in radicals. Since the first nontrivial normal subgroup of the symmetric group on n letters is always the alternating group on n letters and since the alternating groups on n letters for $n \geq 5$ are always simple and non-abelian hence not solvable. It also says that the general polynomials of all degrees higher than the fifth also have no solution in radicals. Note that the above construction of the Galois group for a fifth degree polynomial only applies to the *general polynomial*. Specific polynomials of the fifth degree may have different Galois groups with quite different properties, for example $x^5 - 1$ has a splitting field generated by a primitive 5th root of unity and hence its Galois group is abelian and the equation itself solvable by radicals.

Check Your Progress

7. When is an extension called a simple extension?
8. Write about the quartic function.
9. State the Abel's theorem.

3.8 ANSWERS TO 'CHECK YOUR PROGRESS'

1. Let K be a field and suppose F is a subfield of K then K is called the extension of F.
2. A complex number is said to be an algebraic number if it is algebraic over the field of rational numbers.
3. Let F be a field. The intersection of all subfields of F is the smallest subfield of F and is called the prime subfield of F.
4. Let E be an extension of K. E is called normal extension of K if E/K is algebraic and $\alpha \in E \Rightarrow p(x) = Irr(K, \alpha)$ splits in $E[x]$ or E.
5. A field having a finite number of elements is called a finite field or a Galois field.
6. Let S be a set of polynomials over k. Suppose each $f \in S$ splits in a field E containing k. Then E is called a splitting field of S over k.
7. An extension K/F is called a simple extension if $K=F(a)$ for some $a \in K$.
8. Quartic polynomials can be solved using Ferrari's method which transforms a quartic polynomial into a depressed quartic which has no x^3 term.
9. Abel's theorem states that the generic algebraic equation of degree higher than four is not solvable by radicals.

3.9 SUMMARY

- Ift K is a field and suppose F is a subfield of K, then K is called an extension of F.
- If K is an extension of F. $a \in K$ is said to be algebraic over F if $\exists$ non-zero polynomial $f(x) \in F[x]$ such that $f(a) = 0$.
- An element $a \in K$ is said to be algebraic of degree n over F if it satisfies a polynomial of degree n over F and does not satisfy any polynomial of lesser degree (than n).
- A field K is called perfect field if every algebraic extension of K is separable.
- If E is an extension of K. E is called normal extension of K if
  (i) E/K is algebraic (ii) $\alpha \in E \Rightarrow p(x) = Irr\ (K, \alpha)$ splits in $E[x]$ or E.
- A field k is called algebraically closed if every polynomial f over k splits in k.
- Let $\sigma_1, \sigma_2, ..., \sigma_n$ be homomorphisms from a field E into a field E'. Then, $\sigma_i s$ are called linearly independent over E' if $\alpha_1\sigma_1 + ... + \alpha_n\sigma_n = 0, \Rightarrow \alpha_i = 0\ \forall\ i$ where $\alpha_i \in E'$.
- An extension E of F is called a Galois extension if
  (i) E/F is finite
  (ii) F is the fixed field of a group of automorphisms of E.

- Cubic functions can be solved with the help of Cardano's method in which the general cubic equation is transformed into a depressed cubic without the x^2 term.
- Quartic polynomials can be solved using Ferrari's method which transforms a quartic polynomial into a depressed quartic which has no x^3 term.
- The Abel-Ruffini theorem states that there are some fifth-degree equations whose solution cannot be so expressed, for example the equation $x^5 - x + 1 = 0$.

3.10 KEY TERMS

- **Algebraic number:** A complex number is said to be an algebraic number if it is algebraic over the field of rational numbers.
- **Separable polynomial:** A polynomial is said to be separable if all its roots are simple.
- **Finite field**: A field having finite number of elements is called a finite field or a Galois field.

NOTES

NOTES

- **Cubic functions:** Cubic functions can be solved with the help of Cardano's method in which the general cubic equation is transformed into a depressed cubic without the x^2 term.
- **Quartic functions:** Quartic polynomials can be solved using Ferrari's method which transforms a quartic polynomial into a depressed quartic which has no x^3 term.

3.11 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. Define a field.
2. What is algebraic extension?
3. What is the difference between separable and inseparable extensions?
4. When is a field said to be perfect?
5. What do you mean by the term normal closure?
6. Define product of fields.
7. Define linear independence.
8. What is primitive element?
9. What is Galois group?
10. What do you understand the solvability of a quadratic equation.

Long-Answer Questions

1. If $a, b \in K$ are algebraic over F of degrees m and n respectively and if m and n are relatively prime, prove that $F(a, b)$ is of degree mn over F.
2. If $a \in K$ is algebraic over F of odd degree, show that $F(a) = F(a^2)$.
3. Show that degree of $\sqrt{2} + \sqrt{3}$ over Q is 4 and degree of $\sqrt{2} + \sqrt[3]{5}$ over Q is 6.
4. If a is an algebraic integer and m is an ordinary integer, prove
   (*i*) $a + m$ is an algebraic integer.
   (*ii*) ma is an algebraic integer.
5. Prove that sum and product of two algebraic integers is an algebraic integer.
6. Find a basis of $Q(\sqrt{2}, \sqrt{3})$ over Q. $[1, \sqrt{2}, \sqrt{3}, \sqrt{6}]$
7. Let K be an extension of F. Suppose E_1, E_2 are contained in K and are extensions of F. If $[E_1 : F]$ and $[E_2 : F]$ are primes, show that either $E_1 \cap E_2 = F$ or $E_1 = E_2$.
8. If K is an extension of F, $c \in K$, $a, b \in F$, $a \neq 0$ then show that $F(c) = F(ac + b)$.

9. Suppose that a field F has finite number of elements q. Show that
   (*i*) $q = p^n$ for some prime p and integer n.
   (*ii*) $a^q = a$ for all $a \in F$.
   (*iii*) If $b \in K$ is algebraic over F, then $b^{q^m} = b$ for some $m > 0$.
10. Let K be a finite extension of F. Suppose if F_1 and F_2 are any two subfields of K such that, $F \subseteq F_1$ and $F \subseteq F_2$ then either $F_1 \subseteq F_2$ or $F_2 \subseteq F_1$. Show that K will be a simple extension of F.

NOTES

3.12 FURTHER READING

Herstein, I.N. 1975. *Topics in Algebra*, 3rd Edition. New Delhi: Wiley Eastern Ltd.

Khanna, V.K. and S.K. Bhambri. 2008. *A Course in Abstract Algebra*, 3rd Edition. New Delhi: Vikas Publishing Hous Pvt. Ltd.

Bhattacharya, P.B., S.K. Jain and S.R. Nagpaul. 1997. *Basic Abstract Algebra*, 2nd Edition. New Delhi: Cambridge University Press.

Artin, M.1991. *Algebra*. New Delhi: Prentice-Hall of India.

Lang, S. 1993. *Algebra*, 3rd Edition. New York: Addison-Wesley.

Datta, K.B. 2000. *Matrix and Linear Algebra*. New Delhi: Prentice-Hall of India.

UNIT 4 NOETHERIAN AND ARTINIAN MODULES AND RINGS

Structure

4.0 INTRODUCTION

In mathematics, more specifically in the area of abstract algebra known as ring theory, a Noetherian ring is a ring that satisfies the 'Ascending Chain Condition' on left and right ideals. In other hand the notion of a Noetherian ring is of fundamental importance in both commutative and non-commutative ring theory, due to the role it plays in simplifying the ideal structure of a ring. In abstract algebra, a Noetherian module is a module that satisfies the ascending chain condition on its submodules, where the submodules are partially ordered by inclusion.

In abstract algebra, an Artinian module is a module that satisfies the descending chain condition on its poset of submodules. They are for modules what Artinian rings are for rings, and a ring is Artinian iff it is an Artinian module over itself (with left or right multiplication). Both concepts are named for Emil Artin.

In algebra, the Wedderburn–Artin theorem is a classification theorem for semisimple rings and semisimple algebras. The Wedderburn–Artin theorem reduces the problem of classifying finite-dimensional central simple algebras over a field K to the problem of classifying finite-dimensional central division algebras over K.

In abstract algebra, a module is called a uniform module if the intersection of any two non-zero submodules is non-zero. Alfred Goldie used the notion of uniform modules to construct a measure of dimension for modules, now known as the uniform dimension (or Goldie dimension) of a module. Uniformly primary ideals in a commutative ring with non-zero identity have been introduced and studied by J. A. Cox and A. J. Hetzel.

In mathematics, the Lasker–Noether theorem states that every Noetherian ring is a Lasker ring, which means that every ideal can be decomposed as an intersection, called primary decomposition, of finitely many primary ideals.

In this unit, you will study about the rings and modules, simple modules, Schur's Lemma, free modules fundamental structure theorem, Noetherian and Artinian module or ring, Hilbert's basis theorem, Wedderburn–Artin theorem, uniform module, primary module, Lasker–Noether theorem.

4.1 OBJECTIVES

After going through this unit, you will be able to:

- Define simple modules, uniform modules and Schur's lemma
- Understand the fundamental structure theorem for modules
- Describe the Neotherian and Artinian rings as well as modules
- State the Hilbert basis and Wedderburn Artin theorem
- Elaborate on the primary modules and Noether-Lasker theorem

4.2 RINGS AND MODULES: INTRODUCTION

A group we noticed is a system with a non-empty set and a binary composition. One can of course talk about non-empty sets with two binary compositions also, the set of integers under usual addition and multiplication being an example. Though this set forms a group under addition and not under multiplication, it does have certain specific properties satisfied with respect to multiplication as well. We single out some of these and generalize the concept in the form of a ring. We start with the formal definition.

Definition 1: A non-empty set R, together with two binary compositions + and . is said to form a *Ring* if the following axioms are satisfied:

(*i*) $a + (b + c) = (a + b) + c$ for all $a, b, c \in R$

(*ii*) $a + b = b + a$ for $a, b \in R$

(*iii*) $\exists$ some element 0 (called zero) in R, such that, $a + 0 = 0 + a = a$ for all $a \in R$

(*iv*) for each $a \in R$, $\exists$ an element $(-a) \in R$, such that, $a + (-a) = (-a) + a = 0$

(*v*) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ for all $a, b, c \in R$

(*vi*) $a \cdot (b + c) = a \cdot b + a \cdot c$

$(b + c) \cdot a = b \cdot a + c \cdot a$ for all $a, b, c \in R$

NOTES

Notes: 1. Since we say that + and . are binary compositions on *R*, it is understood that the closure properties with respect to these hold in *R*. In other words, for all $a, b \in R$, $a + b$ and $a \cdot b$ are unique in *R*.

2. One can use any other symbol instead of + and ., but for obvious reasons, we use these two symbols (the properties look so natural with these). In fact, in future, the statement that *R* is a ring would mean that *R* has two binary compositions + and . defined on it and satisfies the above axioms.
3. Axiom (*v*) is named associativity with respect to . and axiom (*vi*) is referred to as distributivity (left and right) with respect to . and +.
4. Axioms (*i*) to (*iv*) could be restated by simply saying that $< R, + >$ forms an abelian group.
5. Since 0 in axiom (*iii*) is identity with respect to +, it is clear that this element is unique (see groups).

Definitions 2: A ring *R* is called a *commutative* ring if $ab = ba$ for all $a, b \in R$. Again if $\exists$ an element $e \in R$ such that,

$$ae = ea = a \quad \text{for all } a \in R$$

we say, *R* is a ring with *unity*. Unity is generally denoted by 1. (It is also called unit element or multiplicative identity).

It would be easy to see that if unity exists in a ring then it must be unique.

Note: We recall that in a group by a^2 we meant $a \cdot a$ where '.' was the binary composition of the group. We continue with the same notation in rings as well. In fact, we also introduce similar notation for addition, and shall write *na* to mean $a + a + ...+ a$ (*n* times), *n* being an integer.

Case 1: Sets of real numbers, rational numbers, integers form rings with respect to usual addition and multiplication. These are all commutative rings with unity.

Case 2: Set **E** of all even integers forms a commutative ring, without unity (under usual addition and multiplication).

Case 3: (*a*) Let *M* be the set of all 2×2 matrices over integers under matrix addition and matrix multiplication. It is easy to see that *M* forms a ring with unity $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, but is not commutative.

(*b*) Let *M* be set of all matrices of the type $\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}$ over integers under matrix addition and multiplication. Then *M* forms a non commutative ring without unity.

NOTES

Case 4: The set $\mathbf{Z}_7 = \{0, 1, 2, 3, 4, 5, 6\}$ forms a ring under addition and multiplication modulo 7. (In fact, we could take n in place of 7).

Case 5: Let F be the set of all continuous functions $f : \mathbf{R} \to \mathbf{R}$, where $\mathbf{R}$ = set of real numbers. Then F forms a ring under addition and multiplication defined by:

for any $f, g \in F$

$$(f + g)x = f(x) \qquad \text{for all } x \in \mathbf{R}$$

$$(f\ g)x = f(x)g(x) \qquad \text{for all } x \in \mathbf{R}$$

zero of this ring is the mapping $O : \mathbf{R} \to \mathbf{R}$, such that,

$$O(x) = 0 \text{ for all } x \in \mathbf{R}$$

Also additive inverse of any $f \in F$ is the function $(-f) : \mathbf{R} \to \mathbf{R}$ such that, $(-f)x = -f(x)$

In fact, F would have unity also, namely the function $i : \mathbf{R} \to \mathbf{R}$ defined by $i(x) = 1$ for all $x \in \mathbf{R}$.

Note: Although the same notation fg has been used for product here it should not be mixed up with fog defined earlier.

Case 6: Let $\mathbf{Z}$ be the set of integers, then $\mathbf{Z}[i] = \{a + ib \mid a, b \in \mathbf{Z}\}$ forms a ring under usual addition and multiplication of complex numbers. $a + ib$ where a, $b \in \mathbf{Z}$ is called a Gaussian integer and $\mathbf{Z}[i]$ is called the ring of Guassian integers.

We can similarly get $\mathbf{Z}_n[i]$ the ring of Gaussian integers modulo n. For instance,

$$\mathbf{Z}_3[i] = \{a + ib \mid a, b \in \mathbf{Z}_3 = \{0, 1, 2\} \bmod 3\}$$
$$= \{0, 1, 2, i, 1 + i, 2 + i, 2i, 1 + 2i, 2 + 2i\}$$

Case 7: Let X be a non-empty set. Then $\mathscr{P}(X)$ the power set of X (*i.e.*, set of all subsets of X) forms a ring under + and $\cdot$ defined by

$$A + B = (A \cup B) - (A \cap B)$$
$$A \,.\, B = A \cap B$$

In fact, this is a commutative ring with unity and also satisfies the property $A^2 = A$ for all $A \in \mathscr{P}(X)$.

Case 8: Let M = set of all 2×2 matrices over members from the ring of integers modulo 2. *It would be a finite non-commutative ring*. M would have $2^4 = 16$ members as each element a, b, c, d in matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ can be chosen in 2 ways. Compositions in M are given by

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} x & y \\ z & u \end{bmatrix} = \begin{bmatrix} a \oplus x & b \oplus y \\ c \oplus z & d \oplus u \end{bmatrix}$$

where $\oplus$ denotes addition modulo 2 and

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x & y \\ z & u \end{bmatrix} = \begin{bmatrix} a\otimes x \oplus b\otimes z & a\otimes y \oplus b\otimes u \\ c\otimes x \oplus d\otimes z & c\otimes y \oplus d\otimes u \end{bmatrix}$$

$\otimes$ being multiplication modulo 2.

That M is non-commutative follows as $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

But $\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

Case 9: Let $R = \{0, a, b, c\}$. Define + and . on R by

+	0	a	b	c
0	0	a	b	c
a	a	0	c	b
b	b	c	0	a
c	c	b	a	0

.	0	a	b	c
0	0	0	0	0
a	0	a	b	c
b	0	a	b	c
c	0	0	0	0

Then one can check that R forms a non-commutative ring without unity. In fact it is an example of the smallest non-commutative ring.

Theorem 4.1: *In a ring R, the following results hold*

(*i*) $a \,.\, 0 = 0 \,.\, a = 0$ *for all* $a \in R$

(*ii*) $a(-b) = (-a)b = -ab$ *for all* $a, b \in R$

(*iii*) $(-a)\,(-b) = ab$. $\forall\ a, b \square R$

(*iv*) $a(b - c) = ab - ac$. $\forall\ a, b, c \in R$

Proof: (*i*) $a \,.\, 0 = a \,.\, (0 + 0)$

$\Rightarrow \quad a \,.\, 0 = a \,.\, 0 + a \,.\, 0$

$\Rightarrow a \,.\, 0 + 0 = a \,.\, 0 + a \,.\, 0$

$\Rightarrow \quad 0 = a \,.\, 0$

using cancellation w.r.t + in the group $< R, + >$.

(*ii*) $a \,.\, 0 = 0$

$\Rightarrow \quad a\,(-b + b) = 0$

$\Rightarrow \quad a\,(-b) + ab = 0$

$\Rightarrow \quad a\,(-b) = -(ab)$

similarly $(-a)\,b = -ab$.

(*iii*) $(-a)\,(-b) = -[a\,(-b)] = -[-ab] = ab$

(*iv*) $a\,(b - c) = a\,(b + (-c))$

$= ab + a\,(-c)$

$= ab - ac$.

NOTES

NOTES

Notes: 1. If R is a ring with unity and $1 = 0$, then since for any $a \in R$, $a = a.1 = a.0 = 0$, we find $R = \{0\}$ which is called the *trivial* ring. We generally exclude this case and thus whenever, we say R is a ring with unity, it will be understood that $1 \neq 0$ in R.

2. If n, m are integers and a, b elements of a ring, then it is easy to see that

$$n(a + b) = na + nb$$
$$(n + m)a = na + ma$$
$$(nm)a = n(ma)$$
$$a^m \; a^n = a^{m+n}$$
$$(a^m)^n = a^{mn}$$

We are so much used to the property that whenever $ab = 0$ then either $a = 0$ or $b = 0$ that it may need more than a bit of convincing that the result may not always be true. Indeed in the ring of integers (or reals or rationals) this property holds. But if we consider the ring of 2×2 matrices over integers, we notice, we can have two non-zero elements A, B s.t, $AB = 0$, but $A \neq 0$ $B \neq 0$. In fact, take $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ then $A \neq 0$, $B \neq 0$. But $AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. We formalise this notion through

Definition 1: Let R be a ring. An element $0 \neq a \in R$ is called a *zero-divisor*, if $\exists$ an element $0 \neq b \in R$ such that, $ab = 0$ or $ba = 0$.

Definition 2: A commutative ring R is called an *Integral domain* if $ab = 0$ in R $\Rightarrow$ either $a = 0$ or $b = 0$. In other words, a commutative ring R is called an integral domain if R has no zero divisors.

An obvious example of an integral domain is $< \mathbf{Z}, +, \cdot >$ the ring of integers whereas the ring of matrices, talked about above is an example of a ring which is not an integral domain.

Note: Some authors do not insist upon the condition of commutativity as a part of the definition of an integral domain. One can have non-commutative rings without zero divisors.

The following theorem gives us a necessary and sufficient condition for a commutative ring to be an integral domain.

Theorem 4.2: *A commutative ring R is an integral domain iff for all a, b, c* $\in R$ $(a \neq 0)$

$$ab = ac \Rightarrow b = c.$$

Proof: Let R be an integral domain

Let $\quad ab = ac \quad (a \neq 0)$

Then $\quad ab - ac = 0$

$\Rightarrow \quad a(b - c) = 0$

NOTES

$\Rightarrow$ $a = 0$ or $b - c = 0$

Since $a \neq 0$, we get $b = c$.

Conversely, let the given condition hold.

Let $a, b \in R$ be any elements with $a \neq 0$.

Suppose $ab = 0$

then $ab = a.0$

$\Rightarrow$ $b = 0$ using given condition

Hence $ab = 0 \Rightarrow b = 0$ whenever $a \neq 0$ or that R is an integral domain.

Note: A ring R is said to satisfy *left cancellation law* if for all $a, b, c \in R, a \neq 0$

$$ab = ac \Rightarrow b = c.$$

Similarly we can talk of *right cancellation law*. It might, of course, be noted that cancellation is of only non zero elements.

Definition 1: An element a in a ring R with unity, is called invertible (or a *unit*) with respect to multiplication if $\exists$ some $b \in R$ such that $ab = 1 = ba$.

Notice, unit and unit element (unity) are different concepts and should not be confused with each other.

Definition 2: A ring R with unity is called a *Division ring* or a *skew field* if non zero elements of R form a group with respect to multiplication.

In other words, a ring R with unity is a Division ring if non-zero elements of R have multiplicative inverse.

Definition 3: A commutative division ring is called a *field*.

Real numbers form a field, whereas integers do not, under usual addition and multiplication. Since a division ring (field) forms groups with respect to two binary compositions, it must contain two identity elements 0 and 1 (with respect to addition and multiplication) and thus a division ring (field) has at least two elements.

Case 10: A division ring which is not a field. Let M be the set of all 2×2 matrices of the type $\begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix}$ where a, b are complex numbers and $\bar{a}, \bar{b}$ are their conjugates, *i.e.*, if $a = x + iy$ then $\bar{a} = x - iy$. Then M is a ring with unity $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ under matrix addition and matrix multiplication.

Any non-zero element of M will be $\begin{bmatrix} x+iy & u+iv \\ -(u-iv) & x-iy \end{bmatrix}$

where x, y, u, v are not all zero.

NOTES

One can check that the matrix $\begin{bmatrix} \frac{x-iy}{k} & -\frac{u+iv}{k} \\ \frac{u-iv}{k} & \frac{x+iy}{k} \end{bmatrix}$

where $k = x^2 + y^2 + u^2 + v^2$, will be multiplicative inverse of the above non-zero matrix, showing that M is a division ring. But M will not be a field as it is not commutative as

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

But $$\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix}.$$

Case 11: Consider

$D = \{a + bi + cj + dk \mid a, b, c, d \in \mathbf{R}\}$ with $i^2 = j^2 = k^2 = -1$, then D forms a ring under multiplication.

Since $i = 0 + 1i + 0j + 0k$, $j = 0 + 0i + 1j + 0k$ gives $ij = k$, $ji = -k$, we find D is not commutative and hence is not a field. D has unity $1 = 1 + 0i + 0j + 0k$.

If $a + bi + cj + dk$ be any non-zero element of D (*i.e.*, at least one of a, b, c, d is non zero) then $(a + bi + cj + dk)\frac{(a - bi - cj - dk)}{a^2 + b^2\ c^2 + d^2} = 1$.

Hence D is a division ring but not a field.

Theorem 4.3: *A field is an integral domain.*

Proof: Let $< R, +, . >$ be a field, then R is a commutative ring.

Let $ab = 0$ in R. We want to show either $a = 0$ or $b = 0$. Suppose $a \neq 0$, then a^{-1} exists (definition of field)

thus $ab = 0$

$\Rightarrow$ $a^{-1}(ab) = a^{-1}0$

$\Rightarrow$ $b = 0.$

which shows that R is an integral domain.

A 'Partial Converse' of the above result also holds.

Theorem 4.4: *A non-zero finite integral domain is a field.*

Proof: Let R be a non-zero finite integral domain.

Let R' be the subset of R containing non-zero elements of R.

Since associativity holds in R, it will hold in R'. Thus R' is a finite semi group.

Again cancellation laws hold in R (for non zero elements) and therefore, these hold in R'.

NOTES

Hence R' is a finite semi group with respect to multiplication in which cancellation laws hold.

$\therefore < R', \cdot >$ forms a group.

In other words $< R, +, \cdot >$ is a field (it being commutative as it is an integral domain).

Aliter: Let $R = \{a_1, a_2, ..., a_n\}$ be a finite non-zero integral domain. Let $0 \neq a \in R$ be any element then $aa_1, aa_2, ..., aa_n$ are all in R and if $aa_i = aa_j$ for some $i \neq j$, then by cancellation we get $a_i = a_j$ which is not true. Hence $aa_1, aa_2, ..., aa_n$ are distinct members of R.

Since $a \in R$, $a = aa_i$ for some i

Let $x \in R$ be *any* element, then $x = aa_j$ for some j

Thus $ax = (aa_i)x = a(a_ix)$

i.e., $x = a_ix$

Hence using commutativity we find

$$x = a_ix = xa_i$$

or that a_i is unity of R. Let $a_i = 1$

Thus for $1 \in R$, since $1 = aa_k$ for some k

We find a_k is multiplicative inverse of a. Hence any non-zero element of R has multiplicative inverse or that R is a field.

Case 12: An infinite integral domain which is not a field is the ring of integers.

Definition: A ring R is called a *Boolean ring* if $x^2 = x$ for all $x \in R$.

Case 13: The ring $\{0, 1\}$ under addition and multiplication mod 2 forms a Boolean ring.

Example 4.1: *Show that a Boolean ring is commutative.*

Solution: Let $a, b \in R$ be any elements

Then $a + b \in R$ (closure)

By given condition

$$(a + b)^2 = a + b$$

$$\Rightarrow a^2 + b^2 + ab + ba = a + b$$

$$\Rightarrow a + b + ab + ba = a + b$$

$$\Rightarrow ab + ba = 0$$

$$\Rightarrow ab = -ba \quad ...(1)$$

$$\Rightarrow a(ab) = a(-ba)$$

$$\Rightarrow a^2b = -aba$$

$$\Rightarrow ab = -aba \quad ...(2)$$

NOTES

Again Equation (1) gives,

$$(ab)a = (-ba)a$$

$$\Rightarrow \quad aba = -ba^2 = -ba \qquad ...(3)$$

Equation (2) and (3) give,

$$ab = ba \ (= -aba)$$

or that R is commutative.

Example 4.2: (*a*) *Show that a non-zero element a in* $\mathbf{Z}_n$ *is a unit iff a and n are relatively prime.*

(*b*) *If a is not a unit then it is a zero divisor.*

Solution: (*a*) $\mathbf{Z}_n = \{0, 1, 2, ..., n-1\} \bmod n$

Let $a \in \mathbf{Z}_n$ be a unit, then $\exists\ b \in \mathbf{Z}_n$ such that,

$$a \otimes b = 1$$

i.e., when ab is divided by n, remainder is 1, in other words,

$$ab = nq + 1$$

or $\quad ab - nq = 1$

$\Rightarrow$ a and n are relatively prime.

Conversely, let $(a, n) = 1$, then $\exists$ integers u, v such that,

$$au + nv = 1$$

$$\Rightarrow \quad au = n(-v) + 1$$

Suppose, $\quad u = nq + r, \quad 0 \le r < n, \quad r \in \mathbf{Z}_n,$

Then $\quad au = anq + ar \ = n\ (-v) + 1$

$$\Rightarrow \quad ar \ = n\ (-v - aq) + 1, \quad r \in \mathbf{Z}_n$$

i.e., $\quad a \otimes r = 1, \quad r \in \mathbf{Z}_n$

i.e., $\quad a$ is a unit.

(*b*) Let a be not a unit and suppose g.c.d$(a, n) = d > 1$

Since $d \mid a$, $a = dk$ for some k. Also $d \mid n \Rightarrow n = dt$

$$\Rightarrow a.t = dk\ \frac{n}{d} = kn = 0 \bmod n$$

i.e., a is a zero divisor.

Example 4.3: *Show that* $\mathbf{Z}_p = \{0, 1, 2, ..., p-1\}$ *modulo p is a field iff p is a prime.*

Solution: Let $\mathbf{Z}_p$ be a field. Suppose p is not a prime, then $\exists\ a, b$, such that $p = ab$, $1 < a, b < p$

$\Rightarrow a \otimes b = 0$ where a, b are non zero $\Rightarrow \mathbf{Z}_p$ has zero divisors.

i.e. $\mathbf{Z}_p$ is not an integral domain, a contradiction as $\mathbf{Z}_p$ being a field is an integral domain.

Hence p is prime.

Conversely, let p be a prime. We need show that $\mathbf{Z}_p$ is an integral domain (it being finite will then be a field).

Let $a \otimes b = 0 \quad a, b \in \mathbf{Z}_p$

Then ab is a multiple of p

$\Rightarrow p \mid ab$

$\Rightarrow p \mid a$ or $p \mid b$ (p being prime)

$\Rightarrow a = 0$ or $b = 0$ (Notice $a, b \in \mathbf{Z}_p \Rightarrow a , b < p$)

$\Rightarrow \mathbf{Z}_p$ is an integral domain and hence a field.

Example 4.4: *If in a ring R, with unity,* $(xy)^2 = x^2y^2$ *for all* $x, y \in R$ *then show that R is commutative.*

Solution: Let $x, y \in R$ be any elements

then $y + 1 \in R$ as $1 \in R$

By given condition

$$(x(y + 1))^2 = x^2 (y +1)^2$$

$$\Rightarrow (xy + x)^2 = x^2 (y + 1)^2$$

$$\Rightarrow (xy)^2 + x^2 + xyx + xxy = x^2(y^2 + 1 + 2y)$$

$$\Rightarrow x^2y^2 + x^2 + xyx + xxy = x^2y^2 + x^2 + 2x^2y$$

$$\Rightarrow xyx = x^2y \qquad ...(1)$$

Since Equation (1) holds for all x, y in R, it holds for $x + 1, y$ also. Thus replacing x by $x + 1$, we get

$$(x + 1)\, y(x +1) = (x +1)^2y$$

$$\Rightarrow (xy + y)\, (x +1) = (x^2 +1 +2x)y$$

$$\Rightarrow xyx + xy + yx + y = x^2y + y + 2xy$$

$\Rightarrow yx = xy$ using Equation (1)

Hence R is commutative.

Example 4.5: *Show that the ring R of real valued continuous functions on* [0, 1] *has zero divisors*.

Solution: Consider the functions f and g defined on [0, 1] by

$$f(x) = \frac{1}{2} - x, \qquad 0 \le x \le \frac{1}{2}$$

$$= 0, \qquad \frac{1}{2} \le x \le 1$$

and $g(x) = 0, \qquad 0 \le x \le \frac{1}{2}$

$$= x - \frac{1}{2}, \qquad \frac{1}{2} \le x \le 1$$

NOTES

NOTES

then f and g are continuous functions and $f \neq 0, g \neq 0$

whereas $gf(x) = g(x)f(x) = 0 \cdot \left(\frac{1}{2} - x\right)$ if $0 \leq x \leq \frac{1}{2}$

$$= \left(x - \frac{1}{2}\right) \cdot 0 = 0 \text{ if } \frac{1}{2} \leq x \leq 1$$

i.e., $gf(x) = 0$ for all x

i.e., $gf = 0$ but $f \neq 0$, $g \neq 0$.

Definition: A non-empty subset S of a ring R is said to be a *subring* of R if S forms a ring under the binary compositions of R.

The ring $< \mathbf{Z}, +, \cdot >$ of integers is a subring of the ring $< \mathbf{R}, +, \cdot >$ of real numbers.

If R is a ring then $\{0\}$ and R are always subrings of R, called *trivial* subrings of R.

It is obvious that a subring of an integral domain will be an integral domain.

In practice it would be difficult and lengthy to check all axioms in the definition of a ring to find out whether a subset is a subring or not. The following theorem would make the job rather easy.

Theorem 4.5: *A non-empty subset S of a ring R is a subring of R iff $a, b \in S$*
$\Rightarrow ab, a - b \in S$.

Proof: Let S be a subring of R

then $a, b \in S \Rightarrow ab \in S$ (closure)

$a, b \in S \Rightarrow a - b \in S$

as $< S, + >$ is a subgroup of $< R, + >$.

Conversely, since $a, b \in S \Rightarrow a - b \in S$, we find $< S, + >$ forms a subgroup of $< R, + >$. Again for any $a, b \in S$, since $S \subseteq R$

$a, b \in R$

$\Rightarrow a + b = b + a$

and so we find S is abelian.

By a similar argument, we find that multiplicative associativity and distributivity hold in S.

In other words, S satisfies all the axioms in the definition of a ring.

Hence S is a subring of R.

Definition: A non-empty subset S of a field F is called a *subfield*, if S forms a field under the operations in F. Similarly, we can define a *subdivision ring* of a division ring.

The simple modules over a ring R are the (left or right) modules over R, which have no non-zero proper submodules.

NOTES

Module

A **left *R*-module** M over the ring R consists of an abelian group $(M, +)$ and an operation $R \times M \to M$ called scalar multiplication, such that for all $r, s \in R$ and $x, y \in M$, we have:

1. $r(x + y) = rx + ry$
2. $(r + s)x = rx + sx$
3. $(rs)x = r(sx)$
4. $1_R x = x$, if R has multiplicative identity 1_R.

A **right *R*-module** is defined in the similar way but the ring acts on the right, i.e., we have a scalar multiplication of the form $M \times R \to M$, and the axioms are written with scalars r and s on the right of x and y. If R is commutative, then left R-modules are the same as right R-modules and are called R-modules.

Submodule

Suppose M is a left R-module and N is a subgroup of M. Then N is a **submodule** or ***R*-submodule** if, for any $n \in N$ and any $r \in R$, the product $rn \in N$ or $nr \in N$ in the case of right R-module.

Quotient module

Given a module A over a ring R, and a submodule B of A, the quotient space A/B is defined by the equivalence relation

$a \sim b$ if and only if $b - a \in B$,

for any a and $b \in A$. The elements of A/B are the equivalence classes $[a] = \{ a + b : b \text{ in } B \}$.

The addition operation on A/B is defined for two equivalence classes as the equivalence class of the sum of two representatives from these classes as,

$[a] + [b] = [a + b]$ for $a, b \in A$ and $r \in R$

and the multiplication by elements of R as,

$r \cdot [a] = [r \cdot a]$, for all $a, b \in A$ and $r \in R$

In this way, A/B becomes itself a module over R, called the **quotient module**.

4.3 SIMPLE MODULES

Definition 1: A module is an algebraic object in which things can be added together commutatively by multiplying coefficients and in which most of the rules of manipulating vectors hold. If a module takes its coefficients in a ring R then it is called a module over R or an R-module. If a and b are two integers then the smallest module containing a and b is the module for their greatest common divisor.

Definition 2: The left R-module M is said to be finitely generated if there exist $m_1, m_2, \ldots, m_n \in M$ such that $M = \sum_{i=1}^{n} Rm_i$.

NOTES

In this case, we say that $\{ m_1, m_2, \ldots, m_n \}$ is a set of generators for M. The module M is called cyclic if there exists $m \in M$ such that $M = Rm$. The module M is called a free module if there exists a subset $X \leq M$ such that each element $m \in M$ can be expressed uniquely as a finite sum $m = \sum_i^n = 1\ a_i x_i$, with $a_1, \ldots, a_n \ R$ and $x_1, \ldots, x_n \in X$.

Definition 3: Let R be a ring and let M be a left R-module. For any element $m \in M$, the left ideal

$$\text{Ann}(m) = \{ r \in R \mid r\, m = 0 \}$$

is called the annihilator of m. The ideal

$$\text{Ann}\ (M) = \{ r \in R \mid r\, m = 0 \text{ for all } m \in M \}.$$

is called the **annihilator** of M.

The module M is called **faithful** if $\text{Ann}(M)=(0)$.

A module is *simple* if it is non-zero and does not admit a proper non-zero submodule. If a module M is simple then the following are equivalent:

- $Am = M$ for every m non-zero in M. simple module
- $M \sim A/m$ for some maximal left ideal of A.

In particular, simple modules are cyclic and the annihilator of any non-zero element of a simple module is a maximal left ideal.

The annihilator of a simple module is called a primitive ideal. The ring A is primitive if the zero ideal is primitive or equivalently, if A admits a faithful simple module.

- A module may have no simple submodules. Simple submodules of ${}_AA$ are minimal left ideals.
- The module ${}_AA$ is simple if and only if A is a division ring. In this case, any simple module is isomorphic to ${}_AA$.
- The Z-module Z/p^nZ where p is a prime is indecomposable. It is simple if and only if $n = 1$.
- Let $A = \text{End}_k V$ for a field k and a k-vector space V. The set a of finite rank endomorphisms is a two-sided ideal of A. Let B be the subring A generated by the identity endomorphism and a. Then V is a simple B-module, in particular a simple A-module and $B \ A$ if $\dim_k V$ is infinite. Let W be a codimension 1 subspace of V. The endomorphisms killing W form a minimal left ideal in A and in B. Thus A and B when $\dim_k V$ is infinite give examples of primitive rings that admit non-trivial proper two-sided ideals.

Definition 4: A *uniform* module is a non-zero module M such that the intersection of any two non-zero submodules of M is non-zero or equivalently such that every non-zero submodule of M is essential in M.

Note: An essential submodule of a module B is any submodule A which has non-zero intersection with every non zero submodule of B.

NOTES

4.4 SCHUR'S LEMMA

Schur's lemma is a fundamental result in representation theory, an elementary observation about irreducible modules, which is nonetheless noteworthy because of its profound applications.

Lemma 1: Let G be a finite group and let V and W be irreducible G-modules. Then, every G-module homomorphism $f: V \to W$ is either invertible or the trivial zero map.

Proof: Both the kernel, ker f and the image, im f are G-submodules of V and W, respectively. Since V is irreducible, ker f is either trivial or all of V. In the former case, im f is all of W also because W is irreducible and hence f is invertible. In the latter case, f is the zero map.

Given below is one of the most important consequences of Schur's lemma:

Corollary: Let V be a finite-dimensional, irreducible G-module taken over an algebraically closed field. Then, every G-module homomorphism $f: V \to V$ is equal to a scalar multiplication.

Proof: Since the ground field is algebraically closed, the linear transformation $f: V \to V$ has an eigenvalue λ, say. By definition, $f - \lambda$ is not invertible, and hence equal to zero by Schur's lemma. In other words, $f = \lambda$, i.e., a scalar.

4.5 FREE MODULES FUNDAMENTAL STRUCTURE THEOREM

In a principal ideal domain, the generators of an ideal is unique up to associates. If $a \in R$, then the generator of ann(a) (= $\{r \in R | ra = 0\}$) is called the order of a, denoted by $o(a)$. Now we attach a weight $P(a)$ to $a \in$ R. Since R is a unique factorization domain, we denote the number of prime factors (counting multiplicity) of a by $P(a)$. By convention, $P(0) = 1$. Thus, $a|b$ in R implies that $P(a) \leq P(b)$, where the equality holds if and only if a, b are associates.

Lemma 2: Let M be a finitely generated module over a principal ideal domain R, say $M = \{m_1, \ldots, m_n\}$. Suppose that there is a relation $a_1 m_1 + \ldots + a_n m_n = 0$, where not all the a_i are zero. Then there are elements $m'_1, \ldots, m'_n \in M$, such that $M = \{m'_1, \ldots, m'_n\}$, and the order of m'_1 divides every a_i.

Proof: If one of the a_i is a unit then the proof follows.

If a_1 is a unit, then m_1 is a linear combination of the other m_i. So take $m'_1 = 0$, $m'_i = m_i$, $i > 1$.

Let $s = \sum P(a_i)$ where $a_i \neq 0$. We will prove this by induction on s. If $s = 0$, every a_i is zero or a unit and at least one a_i is a unit.

NOTES

If only one a_i is non-zero, the result is easy to establish, so let us assume a_1, a_2 are non-zero and non-unit. Let b = g.c.d.(a_1, a_2), $a_1 = bc_1$, $a_2 = bc_2$, and $b_1c_1 + b_2c_2 = 1$. Now

$$M = \{m_1, m_2, \ldots, m_n\}$$

$$= \left\{(m_1, m_2)\begin{pmatrix} c_2 & b_1 \\ -c_1 & b_2 \end{pmatrix}, m_3, \ldots, m_n\right\}$$

$$0 = b(b_1m_1 + b_2m_2) + a_3m_3 + \ldots + a_nm_n$$

Now $P(b) \leq P(a_1 < P(a_1) + P(a_2)$. By induction, $M = \{m'_1, \ldots, m'_n\}$, with $o(m'_1) \mid b$, and $o(m'_1)|a_i$, for $i \geq 3$. But $b|a_1$, $b|a_2$, hence $o(m'_1)|a_i$, for all i.

Theorem 4.6: *Every n-generated module M over a principal ideal domain R is a direct sum of n cyclic modules* $M \simeq \oplus_{i=1}^{n} Rm_i$. *Equivalently,* $M = \{m_1, \ldots, m_n\}$, *and* $\sum a_i m_i = 0$ *implies* $a_i m_i = 0$, *for all i.*

Proof: If $n = 1$, this is true, as R is a principal ideal domain. Now let $n > 1$. We induct on n.

Amongst all possible set of generators of M having n elements choose one which has an element m with least $P(m)$. Let $M = \{m = m_1, m'_2, \ldots, m'_n\}$. If $M = R_m \oplus \sum_{i\geq 2} Rm'_i$, then by induction the submodule $\sum_{i\geq 2} Rm'_i$ has a basis $\{m_2, \ldots, m_n\}$. But then $\{m_1, \ldots, m_n\}$ is a basis of M.

We show that Rm is indeed a direct summand of M: If not, one has a relation $a_1m_1 + \ldots + a_nm_n = 0$, with $a_1m_1 \neq 0$. Let b =g.c.d.$(a_1, o(m_1)) = c_1a_1 + c_2o(m_1)$. Since $a_1m_1 \neq 0$, a_1 and $o(m_1)$ are not associates. Hence, $P(b) < P(o(m_1))$.

Note that $bm_1 + c_1a_2m_2 + \ldots + c_1a_nm_n = 0$. By above Lemma $M = \{m'_1, \ldots, m'_n\}$, with $o(m'_1)|b$, $o(m'_1)|c_1a_i$, for $i \geq 2$. Since $P(o(m'_1)) \leq P(b) < P(o(m_1))$, this contradicts the minimality of $\{m_1, \ldots, m_n\}$. Thus, Rm is a summand of M and the result follows.

4.6 NOETHERIAN AND ARTINIAN MODULES

A module is Artinian/Noetherian if it satisfies either of the following equivalent conditions:

- Every non-empty collection of submodules contains a minimal/maximal element with respect to inclusion.
- Any descending/ ascending chain of submodules stabilizes.

An infinite direct sum of non-zero modules is neither Artinian nor Noetherian. A vector space is Artinian/ Noetherian if and only if its dimension is finite. Submodules and quotient modules of Artinian modules are Artinian. If a submodule N of a module M and the quotient M/N by it are Artinian, then so is M.

NOTES

Theorem 4.7: *An R-module M is Noetherian if and only if each submodule of M is finitely generated.*

Proof:

Let N ≤ M be a submodule of *M* which is not finitely generated. Since no finite subset of *N* will generate *N*, we clearly can choose an infinite sequence of elements from $N, s_1, s_2, ... \in N$, and get a proper ascending sequence of submodules: $<\{s_1\}> \subseteq <\{s_1, s_2\}> \subseteq$ .... which contradicts with the fact that *M* is Noetherian.

Now, let $N_1 \subset N_0 \subset ...$ be an ascending sequence of submodules of *M*, $\bigcup_{i=1}^{\infty} N_i$ is again a submodule of *M* which, by assumption, is finitely generated. Let $\{v_1, v_2, ..., v_k\}$ be a set of generators for $\bigcup_{i=1}^{\infty} N_i$, and N_{ji} be some submodule which contains $\{v_i\}$. Let m = max $\{j_1, j_2, ..., jk\}$, since ...,

$$N_1 \subset N_0 \subset ..., N_m = N_{m+1} = ... = \bigcup_{i=1}^{\infty} N_i .$$

Theorem 4.8: *Given any short exact sequence* $0 \to Y \xrightarrow{\mu} X \xrightarrow{v} Z \to 0$, *X is Noetherian if and only if Y and Z are Noetherian.*

Proof: Let *N* be any submodule of the Noetherian module *X* and every submodule of *N* is also a submodule of *X*, hence, is finitely generated. Then from the above Theorem , *N* is Noetherian.

Now, since *Y* is *D*-isomorphic to a Noetherian submodule of *X*, *Y* is thus Noetherian. To conclude that *Z* is Noetherian, let us consider any ascending sequence of submodules of $Z: Z_1 \subset Z_2 \subset ...$. Clearly, $v^{-1}(Z_1) \subset v^{-1}(Z_2) \subset ...$ is a terminated ascending sequence, since X is Noetherian. This implies $Z_1 \subset Z_2 \subset ...$also terminates.

$\Leftarrow$: Let $N_1 \subset N_0 \subset ...$ be any ascending sequence of submodules of *X*, then $\mu^{-1}(N_1) \subset \mu^{-1}(N_2) \subset ...$ and $v(N_1) \subset v(N_2) \subset ...$ are ascending sequences in Noetherian modules *Y* and *Z*, respectively.

There is *m* so that $\mu^{-1}(N_m) = \mu^{-1}(N_{m+1}) = ...$ and $v(N_m) = v(N_{m+1}) = ...$ .We claim that $N_m = N_{m+1} = ...$.

For this claim to hold, we only need to show that $N_{m+1} \subset N_m$, i.e. for any $z \in N_{m+1}$, we show that $z \in N_m$.

For $z \in N_{m+1}$, $v(z) \in v(N_{m+1}) = v(N_m)$, there is $z' \in N_m$ so that $v(z) = v(z')$.

Since the given sequence is exact, $\exists y \in Y$ such that $\mu(y) = z - z'$.

This implies $y \in \mu^{-1}(z - z') \subset \mu^{-1}(N_{m+1}) = \mu^{-1}(N_m)$. We thus have $\mu(y) \in N_m$, and $z = \mu(y) + z' \in N_m$.

NOTES

4.7 NOETHERIAN AND ARTINIAN RINGS

We will, briefly, discuss noetherian rings here which are in fact a natural generalization of Principal Ideal Domain (PIDs). We begin with

Definition 1: A ring R is called a *noetherian* ring if every ideal of R is finitely generated.

Definition 2: A ring R is called *noetherain* ring if every ascending chain of ideals in R terminates after finite number of steps.

Before giving any examples let us first show the equivalence of the two definitions.

Definition 1 $\Rightarrow$ Definition 2

Let R be a ring in which every ideal is finitely generated. Let

$$A_1 \subseteq A_2 \subseteq A_3 \subseteq \ \ldots$$

be any ascending chain of ideals in R,

Let $A = \underset{i}{\cup} A_i$

then A is an ideal of R

Thus A is finitely generated.

Let $A = < a_1, a_2, \ldots\ldots, a_n >$

Consider any a_j, then $a_j \in A = \cup A_i$

$\Rightarrow \quad a_j \in A_i$ for some i

Suppose $a_1 \in A_{i_1}, a_2 \in A_{i_2} \ldots\ldots, a_n \in A_{i_n}$

Let k be such that $A_{i_j} \subseteq A_k \ \forall\, j = 1, 2, \ldots\ldots, n$

Then $a_1, a_2, \ldots\ldots, a_n \in A_k$

$\Rightarrow \quad A \subseteq A_k \subseteq A$

Hence $A_k = A$ or that the chain terminates at A_k which proves the result.

Definition 1 $\Rightarrow$ Definition 2

Let R be a ring satisfying the condition of Definition 2.

Let I be any ideal of R. We show I is finitely generated.

Let $a_1 \in I$ be any element.

If $I = < a_1 >$, we are done.

If $I \neq < a_1 >$ then $\exists$ same $a_2 \in I$ such that, $a_2 \notin < a_1 >$

Consider $< a_1, a_2 >$. If $I = < a_1, a_2 >$ then the result is proved.

If not then $\exists\ a_3 \in I$ such that, $a_3 \notin < a_1, a_2 >$ continuing like this we get an ascending chain of ideals

$$< a_1 > \subseteq < a_1, a_2 > \subseteq < a_1, a_2, a_3 > \subseteq \ \ldots$$

which must break off after a finite number of steps, say at $<a_1, a_2, ..., a_n>$. Then

$$I = < a_1, a_2, ..., a_n > \text{ and the result is proved.}$$

Case 14: A Principal Ideal Domain or PID is a noetherian ring. Thus in particular, $\mathbf{Z}$, $\mathbf{Z}[i]$, $F[x]$ where F is a field are all noetherian.

Case 15: A finite ring will be noetherian and so would be any field. Remeber a field F has only two ideal $\{0\}$ and F.

Remark: A ring R is defined to be *right noetherian* if every ascending chain of right ideals in R terminates after finite number of steps. Similarly one can talk of a *left noetherian ring* by considering left ideals.

Again the condition of termination of an ascending chain is also referred to as ACC (Ascending Chain Condition). A ring in which ACC holds for right as well as left ideals is called a noetherian ring.

One can have examples of right noetherian rings that are not left noetherian and vice versa.

Theorem 4.9: *Quotient ring of a noetherian ring is noetherian.*

Proof: Let R/I be any quotient ring of a noetherian ring R.

Let $f: R \to R/I$ be the the natural homomorphism, where $f(r) = r + I$

Let $\bar{J}$ be any ideal of R/I. We show $\bar{J}$ is finitely generated.

Let $J = \{r \in R \mid f(r) \in \bar{J}\}$

then it is easy to see that J is an ideal of R Since R is noetherian, J is finitely generated.

Let $J = < r_1, r_2, ..., r_n >$, then we can show that

$$\bar{J} = <f(r_1), f(r_2), ..., f(r_n)>$$

Let $f(r) \in \bar{J}$ be any element then $r \in J$ and as J is generated by $r_1, r_2, ..., r_n$, we get

$$r = \alpha_1 r_1 + \alpha_2 r_2 + ... + \alpha_n r_n \qquad \alpha_i \in R$$

$$\Rightarrow f(r) = f(\alpha_1) f(r_1) + f(\alpha_2) f(r_2) + ... + f(\alpha_n) f(r_n), \quad f(\alpha_i) \in R/I$$

Showing that $\bar{J} = <f(r_1), f(r_2), ..., f(r_n)>$

Hence R/I is noetherian.

Theorem 4.10: *Homomorphic image of a noetherian ring is noetherian*

Proof: Let $f: R \to R'$ be an onto homomorphism and suppose R is noetherian.

By Fundamental theorem of ring homomorphism

R' is isomorphic to a quotient ring of R, which will be noetherian by above theorem. Hence R' will be noetherian.

NOTES

NOTES

Example 4.6: *Let R be a noetherian ring. Show that any ideal $I \neq R$ is contained in a maximal ideal of R.*

Solution: If I itself is maximal we have nothing to prove. If I is not maximal then $\exists$ an ideal I_1, such that, $I \subseteq I_1$. If I_1 is maximal, we are done. If not then $\exists$ another ideal I_2 such that, $I \subseteq I_1 \subseteq I_2$ and continuing like this we get an ascending chain of ideals which must become stationary after a finite number of steps

i.e., $I \subseteq I_1 \subseteq I_2 \subseteq ..... \subseteq I_n = I_{n+1} = I_{n+2} .....$

and thus I_n will be maximal.

Example 4.7: *Let R be a commutative ring with unity. Let $R[x]$ be noetherian. Show that R is also noetherian.*

Solution: We know that

$$\frac{R[x]}{<x>} \cong R$$

Since $R[x]$ is noetherian, its quotient ring $\frac{R[x]}{<x>}$ is noetherian and therefore so is R.

We use the famous Hilbert Basis theorem which says that polynomial ring $R[x]$ of a noetherian ring R is noetherian in proving the following

Example 4.8: *Show by an example that subring of a noetherian ring may not be noetherian.*

Solution: Let $\mathbf{Q}$ be the field of rational numbers, then $\mathbf{Q}$ is a noetherian ring and thus $\mathbf{Q}[x]$ is noetherian.

Let $S = \{f(x) \in \mathbf{Q}[x] \mid f(x) = a_o + a_1x + a_2x^2 + ... + a_nx^n,\ a_o \in \mathbf{Z},\ a_i \in \mathbf{Q}\ \forall\ i \geq 1\}$

It is easy to see that S is a subring of $\mathbf{Q}[x]$.

We notice the chain

$$<x> \underset{\neq}{\subset} <\frac{x}{2}> \underset{\neq}{\subset} <\frac{x}{4}> \underset{\neq}{\subset} ...$$

is an ascending chain of ideals in S which does not terminate after finite number of steps. Suppose for instance, equality holds at $<x> = <\frac{x}{2}>$, then

$$\frac{x}{2} \in <x> \Rightarrow \frac{x}{2} = h(x)x$$

for some $h(x) = \alpha_0 + \alpha_1x + ... + \alpha_m x^m$ where $\alpha_0 \in \mathbf{Z}$

$$\Rightarrow \quad \frac{x}{2} = \alpha_0x + \alpha_1x^2 + ... + \alpha_m x^{m+1}$$

$$\Rightarrow \quad 0 + \frac{1}{2}x + 0x^2 + ... = 0 + \alpha_o x + \alpha_1x^2 + ... + \alpha_m x^{m+1}$$

$$\Rightarrow \quad \frac{1}{2} = \alpha_o \text{ But } \frac{1}{2} \notin \boldsymbol{Z}$$

Hence $< x > \underset{\neq}{\subset} < \frac{x}{2} >$. Similarly it follows that equality does not hold in the above chain at any step.

Definition: A ring R is called *artinian* ring if every decending chain of ideals

$$I_1 \supseteq I_2 \supseteq I_3 \supseteq .....$$

terminates after a finite number of steps.

It is clear that any finite ring is artinian and so would be a field. The ring **Z** of integers is not artinian as the decending chain

$$< n > \underset{\neq}{\supseteq} < 2n > \underset{\neq}{\supseteq} < 4n > \underset{\neq}{\supseteq} .....$$

of ideals (for any +ve integer n) is infinite.

This also shows that subring of an artinian ring may not be artinian. Notice **Q** the ring of rationals being a field is artinian. One can talk of left and right artinian rings also by considering chain of left (right) ideals.

NOTES

Check Your Progress

1. What is commutative ring?
2. What do you understand by submodule?
3. State the Schur's lemma.
4. Define the simple modules.
5. Give the statement of principal ideal domain.
6. Write the necessary and sufficient condition for a Noetherian and Artinian module.
7. Define Noetherian ring.

4.8 HILBERT BASIS THEOREM

In mathematics, specifically commutative algebra, Hilbert's basis theorem says that a polynomial ring over a Noetherian ring is Noetherian.

Theorem 4.11: *Let R be a right (left) Noetherian ring. Then R[x] is also right (left) Noetherian.*

Proof: Let R be a noetherian ring and let $f(x)=a_n x^n + a_{n-1}x^{n+1}+...+a_1 x+a_0 \in R[x]$ with $a_n \neq 0$. a_n is the initial coefficient of f.

Let I be an ideal in $R[x]$. We will show that I is finitely generated, so that $R[x]$ is noetherian. Now let f_0 be a polynomial of least degree in I and if $f_0, f_1, \ldots,$ f_k have been chosen then choose f_{k+1} from I\($f_0, f_1, \ldots, f_k$) of minimal degree. Continuing inductively gives a sequence (f_k) of elements of I.

NOTES

Let a_k be the initial coefficient of f_k and consider the ideal $J=(a_1, a_2, a_3,\ldots)$ of initial coefficients. Since R is noetherian, $J=(a_0,\ldots, a_N)$ for some N.

Then $I = (f_0, f_1,\ldots, f_N)$. Otherwise, $f_{N+1} \in I\,(f_0, f_1,\ldots, f_N)$ and $a_{N+1} = \sum_{k=0}^{N} u_k a_k$ for some $u_1, u_2,\ldots, u_N \in R$. Let $g(x) = \sum_{k=0}^{N} u_k f_k x^{v_k}$ where $v_k = \deg(f_{N+1}) - \deg(f_k)$. Then $\deg(f_{N+1} - g) < \deg(f_{N+1})$ and $f_{N+1} - g \in I$ and $f_{N+1} - g \notin (f_0, f_1,\ldots, f_N)$. But this contradicts minimality of $\deg(f_{N+1})$.

Hence, $R[x]$ is noetherian.

4.9 WEDDERBURN ARTIN THEOREM

Theorem 4.12: (Wedderburn): *A finite division ring is a field.*

Proof: Let R be a finite division ring.

Let $Z(R)$ be the centre of R. Then $Z(R)$ is a field and R can be regarded as a vector space over $Z(R)$. Since R is finite, R is finite dimensional over $Z(R)$. Let $\dim R = n$, $o(Z(R)) = q$ = power of a prime. Then $o(R) = q^n$. We show that $n = 1$. Because then $\dim R = 1$ would imply $R = Z(R) \Rightarrow R$ is a field. Let $n > 1$. Now $N(a) = \{x \in R \mid xa = ax\}$ is a subring of R containing $Z(R)$. So, $N(a)$ can also be regarded as a vector space over $Z(R)$. Let $o(N(a)) = q^{r_a}$ for some integer r_a.

Let $R^* = R - \{0\}$. Then R^* is a multiplicative group and $o(R^*) = q^n - 1$. Consider the class equation of R^*.

$$q^n - 1 = o(Z(R^*)) + \sum_{a \notin Z(R^*)} \frac{o(R^*)}{o(N(a))}$$

$$= q - 1 + \sum_{a \notin Z(R^*)} \frac{q^n - 1}{q^{r_a} - 1}$$

Now $\quad q^{r_a} - 1 \mid q^n - 1 \Rightarrow r_a \mid n, \quad 1 \le r_a < n$

By above Lemma then

$$|\Phi_n(q)| \left| \frac{q^n - 1}{q^{r_a} - 1} \right. \Rightarrow |\Phi_n(q)| \mid q - 1$$

$$\Rightarrow |\Phi_n(q)| \le q - 1$$

But $$|\Phi_n(q)| = \prod_{\substack{\alpha \\ o(\alpha)=m}} |q - \alpha| > \prod_{\alpha} |q - 1| > q - 1$$

So, we get a contradiction $\Rightarrow n = 1$

Hence R is a field.

4.10 PRIMARY MODULES AND NOETHER-LASKER THEOREM

NOTES

The Lasker-Noether theorem states that every Noetherian ring is a Lasker ring which specifies that every ideal can be written as an intersection of finitely many *primary ideals* which are related to but are not identical as powers of prime ideals. The theorem was first established by Emanuel Lasker for the special case of polynomial rings and convergent power series rings, and was verified by Emmy Noether. Basically, the Lasker-Noether theorem is an extension of the fundamental theorem of arithmetic and more specifically the fundamental theorem of finitely generated abelian groups to all Noetherian rings.

It has an extension to modules and states that every submodule of a set module over a Noetherian ring is a finite intersection of primary submodules. This refers to the situation for rings as a special case considering the ring as a module over itself such that ideals are submodules. This specifies the primary decomposition structure of the structure theorem for set modules over a principal ideal domain and for the special case of polynomial rings over a field.

Definitions

Write R for a commutative ring, and M and N for modules over it.

- A **zero divisor** of a module M is an element x of R such that $xm = 0$ for some non-zero m in M.
- An element x of R is called **nilpotent in M** if $x^n M = 0$ for some positive integer n.
- A module is called **coprimary** if every zero divisor of M is nilpotent in M. For example, groups of prime power order and free abelian groups are termed as coprimary modules over the ring of integers.
- A submodule M of a module N is called a **primary submodule** if N/M is coprimary.
- An ideal I is called **primary** if it is a primary submodule of R. This is equivalent to the statement that if ab is in I then either a is in I or b^n is in I for some n and to the condition that every zero-divisor of the ring R/I is nilpotent.
- A submodule M of a module N is called **irreducible** if it is not an intersection of two strictly larger submodules.
- An associated prime of a module M is a prime ideal that is the annihilator of some element of M.

Statement

The Lasker-Noether theorem for modules states that every submodule of a set module over a Noetherian ring is a finite intersection of primary submodules. For the special case of ideals it states that every ideal of a Noetherian ring is a finite

NOTES

intersection of primary ideals. An equivalent statement is that every finitely generated module over a Noetherian ring is contained in a finite product of coprimary modules.

The Lasker-Noether theorem results from the following three facts:

- Any submodule of a finitely generated module over a Noetherian ring is an intersection of a finite number of irreducible submodules.
- If M is an irreducible submodule of a finitely generated module N over a Noetherian ring then N/M has only one associated prime ideal.
- A finitely generated module over a Noetherian ring is coprimary if and only if it has at most one associated prime.

Irreducible Decomposition in Rings

The decomposition of ideals in rings was required when there was lack of unique factorization in number fields like $\mathbb{Z}[\sqrt{-5}]$, in which $6 = 2.3 = (1+\sqrt{-5})(1-\sqrt{-5})$.. If a number does not factor uniquely into primes, then the ideal generated by the number may however factor into the intersection of powers of prime ideals otherwise an ideal may at least factor into the intersection of primary ideals. Consider the example given below:

Let R be a Noetherian ring and I an ideal in R. Then I has a unique irredundant primary decomposition into primary ideals.

$$I = Q_1 \cap \cdots \cap Q_n$$

Irredundancy refers to:

- Removing any of the Q_i changes the intersection, i.e., $Q_1 \cap \cdots \cap \widehat{Q_i} \cap \cdots \cap Q_n \not\supseteq Q_i$ for all i, where the symbol hat denotes omission.
- The associated prime ideals $\sqrt{Q_i}$ are distinct.

Uniqueness refers to uniqueness for reordering the primary ideals. In the case of the ring of integers $\mathbb{Z}$,, the Lasker-Noether theorem is equivalent to the fundamental theorem of arithmetic. If an integer n has prime factorization $n = \pm p_1^{d_1} \cdots p_r^{d_r}$,, then the primary decomposition of the ideal generated by $n(n) \subset \mathbb{Z}$, is

$$(n) = (p_1^{n_1}) \cap \cdots \cap (p_r^{d_r})$$

Minimal Decompositions and Uniqueness

A primary decomposition of a submodule M of a module N is called **minimal** if it has the smallest possible number of primary modules. Consider the case where all modules will be finitely generated over a Noetherian ring R. For minimal decompositions, the primes of the primary modules are uniquely determined as

they are the associated primes of N/M. In addition the primary submodules associated to the **minimal** associated primes (those not containing any other associated primes) are also unique. Though the primary submodules associated to the non-minimal associated primes called **embedded primes** need not be unique. For example, let $N = R = k[x, y]$ for some field k and let M be the ideal (xy, y^2). Then M has two different minimal primary decompositions $M = (y)) \cap (x, y^2) = (y)) \cap (x + y, y^2)$. The minimal prime is (y) and the embedded prime is (x, y).

4.11 UNIFORM MODULES

In abstract algebra, a module is called a uniform module if the intersection of any two non-zero submodules is non-zero. This is equivalent to saying that every non-zero submodule of M is an essential submodule. A ring may be called a right (left) uniform ring if it is uniform as a right (left) module over itself.

Alfred Goldie used the notion of uniform modules to construct a measure of dimension for modules, now known as the uniform dimension (or **Goldie dimension**) of a module. Uniform dimension generalizes some, but not all, aspects of the notion of the dimension of a vector space. Finite uniform dimension was a key assumption for several theorems by Goldie, including Goldie's theorem, which characterizes which rings are right orders in a semi simple ring. Modules of finite uniform dimension generalize both '**Artinian Modules and Noetherian Modules'**. Uniform dimension is also referred to as simply the dimension of a module or the rank of a module. Uniform dimension should not be confused with the related notion, also due to Goldie, of the reduced rank of a module.

Properties and Examples of Uniform Modules

Being a uniform module is not usually preserved by direct products or quotient modules. The direct sum of two non-zero uniform modules always contains two submodules with intersection zero, namely the two original summand modules. If N_1 and N_2 are proper submodules of a uniform module M and neither submodule contains the other, then $M / (N_1 \cap N_2)$ fails to be uniform, as

$N_1 / (N_1 \cap N_2) \cap N_2 / (N_1 \cap N_2) = \{0\}$.

Uniserial modules are uniform, and uniform modules are necessarily directly indecomposable. Any commutative domain is a uniform ring, since if a and b are non-zero elements of two ideals, then the product ab is a non-zero element in the intersection of the ideals.

Uniform Dimension of a Module

The following theorem makes it possible to define a dimension on modules using uniform submodules. It is a module version of a vector space Refer Theorem 4.13.

Theorem4.13: If U_i and V_j are members of a finite collection of uniform submodules of a module M such that $\oplus_{i=1}^{n} U_i$ and $\oplus_{i=1}^{m} V_i$ are both essential submodules of M, then $n = m$.

The uniform dimension of a module M, denoted **u.dim(M),** is defined to be n if there exists a finite set of uniform submodules U_i such that is an essential submodule of M**.** The preceding theorem ensures that this n is well defined. If no such finite set of submodules exists, then **u.dim(M)** is defined to be **".** When speaking of the uniform dimension of a ring, it is necessary to specify whether **u.dim(R_R)** or rather **u.dim($_RR$)** is being measured. It is possible to have two different uniform dimensions on the opposite sides of a ring.

If N is a submodule of M**,** then **u.dim(N)** $\leq$ **u.dim(M)** with equality exactly when N is an essential submodule of M. In particular, M and its injective hull **E(M)** always have the same uniform dimension. It is also true that **u.dim(M) = n** if and only if **E(M)** is a direct sum of n indecomposable injective modules.

It can be shown that **u.dim(M) = ∞** if and only if M contains an infinite direct sum of non-zero submodules. Thus if M is either Noetherian or Artinian, M has finite uniform dimension. If M has finite composition length k**,** then **u.dim(M) $\leq$ k** with equality exactly when M is a semi simple module. (Lam 1999)

A standard result is that a right Noetherian domain is a right Ore domain. In fact, we can recover this result from another theorem attributed to Goldie, which states that the following three conditions are equivalent for a domain D:

- D is right Ore
- **u.dim(D_D) = 1**
- **u.dim(D_D) < ∞**

Check Your Progress

8. State the Hilbert basis theorem.
9. Define the Lasker-Noether theorem.

4.12 ANSWERS TO 'CHECK YOUR PROGRESS'

1. A ring R is called a commutative ring if $ab = ba$ for all $a, b \in R$. Again if $\exists$ an element $e \in R$ such that,

   $ae = ea = a$ for all $a \in R$

   we say, R is a ring with unity. Unity is generally denoted by 1. (It is also called unit element or multiplicative identity).
2. Suppose M is a left R-module and N is a subgroup of M. Then N is a submodule or R-submodule if, for any $n \in N$ and any $r \in R$, the product $rn \in N$ or $nr \in N$ in the case of right R-module.
3. A module is an algebraic object in which things can be added together commutatively by multiplying coefficients and in which most of the rules of manipulating vectors hold. If a module takes its coefficients in a ring R then

it is called a module over R or an R-module. If a and b are two integers then the smallest module containing a and b is the module for their greatest common divisor.

4. Let G be a finite group and let V and W be irreducible G-modules. Then, every G-module homomorphism $f: V \to W$ is either invertible or the trivial zero map.
5. In a principal ideal domain, the generators of an ideal is unique up to associates. If $a \in R$, then the generator of ann(a) ($= \{r \in R | ra = 0\}$) is called the order of a, denoted by $o(a)$.
6. A module is Artinian/Noetherian if it satisfies either of the following equivalent conditions:

   • Every non-empty collection of submodules contains a minimal/maximal element with respect to inclusion.

   • Any descending/ ascending chain of submodules stabilizes.

   An infinite direct sum of non-zero modules is neither Artinian nor Noetherian. A vector space is Artinian/ Noetherian if and only if its dimension is finite. Submodules and quotient modules of Artinian modules are Artinian. If a submodule N of a module M and the quotient M/N by it are Artinian, then so is M.
7. A ring R is called a northerian ring if every ideal of R is finitelty generated.
8. Let R be a right (left) Noetherian ring. Then $R[x]$ is also right (left) Noetherian.
9. The Lasker-Noether theorem states that every Noetherian ring is a Lasker ring which specifies that every ideal can be written as an intersection of finitely many primary ideals which are related to but are not identical as powers of prime ideals. The theorem was first established by Emanuel Lasker for the special case of polynomial rings and convergent power series rings, and was verified by Emmy Noether.

4.13 SUMMARY

- Sets of real numbers, rational numbers, integers form rings with respect to usual addition and multiplication. These are all commutative rings with unity.
- A commutative ring R is called an Integral domain if $ab = 0$ in $R \Rightarrow$ either $a = 0$ or $b = 0$. In other words, a commutative ring R is called an integral domain if R has no zero divisors.
- An element a in a ring R with unity, is called invertible (or a unit) with respect to multiplication if $\exists$ some $b \in R$ such that $ab = 1 = ba$.
- Real numbers form a field, whereas integers do not, under usual addition and multiplication. Since a division ring (field) forms groups with respect to two binary compositions, it must contain two identity elements 0 and 1

NOTES

(with respect to addition and multiplication) and thus a division ring (field) has at least two elements.

- A non-empty subset S of a ring R is said to be a *subring* of R if S forms a ring under the binary compositions of R.
- A non-empty subset S of a field F is called a subfield, if S forms a field under the operations in F. Similarly, we can define a subdivision ring of a division ring.
- A right R-module is defined in the similar way but the ring acts on the right, i.e., we have a scalar multiplication of the form $M \times R \to M$, and the axioms are written with scalars r and s on the right of x and y. If R is commutative, then left R-modules are the same as right R-modules and are called R-modules.
- A module is an algebraic object in which things can be added together commutatively by multiplying coefficients and in which most of the rules of manipulating vectors hold. If a module takes its coefficients in a ring R then it is called a module over R or an R-module. If a and b are two integers then the smallest module containing a and b is the module for their greatest common divisor.
- A module is simple if it is non-zero and does not admit a proper non-zero submodule.
- Schur's lemma is a fundamental result in representation theory, an elementary observation about irreducible modules, which is nonetheless noteworthy because of its profound applications.
- An infinite direct sum of non-zero modules is neither Artinian nor Noetherian. A vector space is Artinian/ Noetherian if and only if its dimension is finite. Submodules and quotient modules of Artinian modules are Artinian. If a submodule N of a module M and the quotient M/N by it are Artinian, then so is M.
- Homomorphic image of a noetherian ring is noetherian
- The Lasker-Noether theorem states that every Noetherian ring is a Lasker ring which specifies that every ideal can be written as an intersection of finitely many primary ideals which are related to but are not identical as powers of prime ideals.

4.14 KEY TERMS

- **Submodule:** Suppose M is a left R-module and N is a subgroup of M. Then N is a submodule or R-submodule if, for any $n \in N$ and any $r \in R$, the product $rn \in N$ or $nr \in N$ in the case of right R-module.

- **Schur's lemma:** Schur's lemma is a fundamental result in representation theory, an elementary observation about irreducible modules, which is nonetheless noteworthy because of its profound applications.
- **Module:** A module is an algebraic object in which things can be added together commutatively by multiplying coefficients and in which most of the rules of manipulating vectors hold.
- **Noetherian ring:** A ring is called noetherian ring if every ideal of the ring is finitely generated.
- **Lasker-Noether theorem:** The Lasker-Noether theorem states that every Noetherian ring is a Lasker ring which specifies that every ideal can be written as an intersection of finitely many primary ideals which are related to but are not identical as powers of prime ideals.

4.15 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. Give the axioms which are satisfied of a ring.
2. What are simple modules?
3. What is the significance of Schur's lemma?
4. State the fundamental structure theorem for modules.
5. What is the difference between noetherian rings and modules?
6. Write the applications of Hilbert Basis theorem.
7. State Wedderburn Artin theorem.
8. Define Noether-Lasker theorem.

Long-Answer Questions

1. Show that a ring R is commutative iff
   $(a + b)^2 = a^2 + b^2 + 2ab$ for all $a, b \in R$.
2. If in a ring R, $x^2 = x$ for all x then show that $2x = 0$ and $x + y = 0 \Rightarrow x = y$.
3. If R is a ring with unity and $(ab)^2 = (ba)^2$ for all $a, b \in R$ and $2x = 0 \Rightarrow x = 0$ then show that R is commutative.
4. Let $\mathbf{R}$ be the set of real numbers. Show that $\mathbf{R} \times \mathbf{R}$ forms a field under addition and multiplication defined by
   $(a, b) + (c, d) = (a + c, b + d)$
   $(a, b) \cdot (c, d) = (ac - bd, ad + bc)$.

NOTES

5. Let R be a commutative ring with unity. Show that
   (*i*) a is a unit iff a^{-1} is a unit.
   (*ii*) a, b are units iff ab is a unit.
6. Show that set of all units in a commutative ring with unity forms an abelian group.
7. Give an example of a non commutative ring R in which $(xy)^2 = x^2y^2$ for all $x, y \in R$.
8. If $< R, +, \cdot >$ be a system satisfying all conditions in the definition of a ring with unity except $a + b = b + a$, then show that this condition is also satisfied.
9. Show that if $1 - ab$ is invertible in a ring with 1 then so is $1 - ba$.
10. Show that a finite commutative ring R without zero divisors has unity. (See theorem 4 page 261).
11. Let R be the set of all 2×2 matrices $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ over **Q** such *that,* $a = d$ and $c = 0$. Let I be the set of all such matrices for which $a = d = 0$. Show that I is an ideal of R.

4.16 FURTHER READING

Herstein, I.N. 1975. *Topics in Algebra*, 3rd edition. New Delhi: Wiley Eastern Ltd.

Khanna, V.K. and S.K. Bhambri. 2008. *A Course in Abstract Algebra*, 3rd edition. New Delhi: Vikas Publishing House.

Bhattacharya, P.B., S.K. Jain and S.R. Nagpaul. 1997. *Basic Abstract Algebra*, 2nd edition. New Delhi: Cambridge University Press.

Artin, M.1991. *Algebra*. New Delhi: Prentice-Hall of India.

Lang, S. 1993. *Algebra*, 3rd edition. New York: Addison-Wesley.

Datta, K.B. 2000. *Matrix and Linear Algebra*. New Delhi: Prentice-Hall of India.

UNIT 5 ABELIAN GROUPS AND JORDAN FORM

NOTES

Structure

5.0 INTRODUCTION

In mathematics, finitely generated Abelian group is a non-empty set G, together with a binary composition * (star) is said to form a group. Specifically in the field of finite group theory. The rational canonical form of a square matrix A with entries in a field F is a canonical form for matrices formed by conjugation by invertible matrices over F in linear algebra. The shape represents a simple decomposition of the vector space into cyclic subspaces for A. (i.e., spanned by some vector and its repeated images under A). Because a given matrix can only have one normal form (thus the term 'Canonical'), matrix B is identical to A if and only if it has the same rational canonical form as A. This form can be determined without any operations that might change while extending the field F (thus the 'Rational'), such as factoring polynomials, demonstrating that whether two matrices are comparable does not change when the field is extended. Ferdinand Georg Frobenius, a German Mathematician, is the name of the form.

In this unit, you will learn about the finitely generated Abelian groups, rational canonical form and generalised Jordon form over any field.

5.1 OBJECTIVES

After going through this unit, you will be able to:

- Know about the finitely generated Abelian groups
- Define rational canonical form
- Learn about the generalised Jordon form over any field

NOTES

5.2 FINITELY GENERATED ABELIAN GROUPS

Definition: A non-empty set G, together with a binary composition $*$ (star) is said to form a group, if it satisfies the following postulates

(*i*) *Associativity*: $a * (b * c) = (a * b) * c$, for all $a, b, c \in G$

(*ii*) *Existence of Identity*: $\exists$ an element $e \in G$, such that,

$$a * e = e * a = a \quad \text{for all } a \in G$$

(e is then called *identity*)

(*iii*) *Existence of Inverse*: For every $a \in G$, $\exists\, a' \in G$ (depending upon a) such that,

$$a * a' = a' * a = e$$

(a' is then called inverse of a)

Remarks: (*i*) Since $*$ is a binary composition on G, it is understood that for all $a, b \in G$, $a * b$ is a unique member of G. This property is called *closure property*.

(*ii*) If, in addition to the above postulates, G also satisfies the *commutative law*

$$a * b = b * a \quad \text{for all } a, b \in G$$

then G is called an *abelian group* or *a commutative group*.

(*iii*) Generally, the binary composition for a group is denoted by '.' (dot) which is so convenient to write (and makes the axioms look so natural too).

This binary composition '.' is called product or multiplication (although it may have nothing to do with the usual multiplication, that we are so familiar with). In fact, we even drop '.' and simply write ab in place of $a \,.\, b$.

In future, whenever we say that G is a group it will be understood that there exists a binary composition '.' on G and it satisfies all the axioms in the definition of a group.

If the set G is finite (i.e., has finite number of elements) it is called a *finite group* otherwise, it is called an *infinite group*.

We shall always (unless stated otherwise) use the symbols e for identity of a group and a^{-1} for inverse of element a of the group.

Definition: By order of a group, we will mean the number of elements in the group and shall denote it by $o(G)$ or $|\, G \,|$.

We now consider a few examples of systems that form groups (or do not form groups).

Example 5.1: The set **Z** of integers forms an abelian group with respect to the usual addition of integers.

It is easy to verify the postulates in the definition of a group as sum of two integers is a unique integer (thus closure holds). Associativity of addition is known to us. 0 (zero) will be identity and negatives will be the respective inverse elements. Commutativity again being obvious.

Example 5.2: One can easily check, as in the previous example, that sets **Q** of rationals, **R** of real numbers would also form abelian groups with respect to addition.

Example 5.3: Set of integers, with respect to usual multiplication does not form a group, although closure, associativity, identity conditions hold.

Note 2 has no inverse with respect to multiplication as there does not exist any integer a such that, $2 \,.\, a = a \,.\, 2 = 1$.

Example 5.4: The set G of all +ve irrational numbers together with 1 under multiplication does not form a group as closure does not hold. Indeed $\sqrt{3} \,.\, \sqrt{3} = 3 \notin G$, although one would notice that other conditions in the definition of a group are satisfied here.

Example 5.5: Let G be the set $\{1, -1\}$. Then it forms an abelian group under multiplication. It is again easy to check the properties.

1 would be identity and each element is its own inverse.

Example 5.6: Set of all 2×2 matrices over integers under matrix addition would be another example of an abelian group.

Example 5.7: Set of all non-zero complex numbers forms a group under multiplication defined by

$$(a + ib)\ (c + id) = (ac - bd) + i\ (ad + bc)$$

$1 = 1 + i.0$ will be identity,

$\dfrac{a}{a^2 + b^2} - i \dfrac{b}{a^2 + b^2}$ will be inverse of $a + ib$.

Note $a + ib$ non-zero means that not both a & b are zero. Thus $a^2 + b^2 \neq 0$.

Example 5.8: The set G of all nth roots of unity, where n is a fixed positive integer forms an abelian group under usual multiplication of complex numbers.

We know that complex number z is an nth root of unity if $z^n = 1$ and also that there exist exactly n distinct roots of unity.

In fact the roots are given by $e^{2\pi ir/n}$

where $r = 1, 2, ..., n$ and $e^{ix} = \cos x + i \sin x$.

If $a, b \in G$ be any two members, then $a^n = 1$, $b^n = 1$ thus $(ab)^n = a^n\, b^n = 1$.

$\Rightarrow$ ab is an nth root of unity

$\Rightarrow$ $ab \in G \Rightarrow$ closure holds.

NOTES

Associativity of multiplication is true in complex numbers.

Again, since $1.\ a = a\ .\ 1 = a$, 1 will be identity.

Also for any $a \in G$, $\frac{1}{a}$ will be its inverse as $\left(\frac{1}{a}\right)^n = \frac{1}{a^n} = 1$.

So, inverse of $e^{2\pi ir/n}$ is $e^{2\pi i(n-r)/n}$ and identity is $e^{2\pi i0/n} = 1$

Commutativity being obvious, we find G is an abelian group.

As a particular case, if $n = 4$ then G is $\{1, -1, i, -i\}$

Example 5.9: (*i*) Let $G = \{\pm 1, \pm i, \pm j, \pm k\}$. Define product on G by usual multiplication together with

$$i^2 = j^2 = k^2 = -1, \quad ij = -ji = k$$
$$jk = -kj = i$$
$$ki = -ik = j$$

then G forms a group. G is not abelian as $ij \neq ji$.

This is called the ***Quaternion Group***.

(*ii*) If set G consists of the eight matrices

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

$$\begin{bmatrix}0 & i\\i & 0\end{bmatrix}, \begin{bmatrix}0 & -i\\-i & 0\end{bmatrix}, \text{ where } i = \sqrt{-1}$$

then G forms a non-abelian group under matrix multiplication.

Example 5.10: Let $G = \{(a, b) \mid a, b$ rationals, $a \neq 0\}$. Define $*$ on G by

$$(a, b) * (c, d) = (ac, ad + b)$$

Closure follows as $a, c \neq 0 \Rightarrow ac \neq 0$

$$[(a, b) * (c, d)] * (e, f) = (ac, ad + b) * (e, f)$$
$$= (ace, acf + ad + b)$$
$$(a, b) * [(c, d) * (e, f)] = (a, b) * (ce, cf + d)$$
$$= (ace, acf + ad + b)$$

proves associativity.

(1, 0) will be identity and $(1/a, -b/a)$ will be inverse of any element (a, b).

G is not abelian as

$$(1, 2) * (3, 4) = (3, 4 + 2) = (3, 6)$$
$$(3, 4) * (1, 2) = (3, 6 + 4) = (3, 10).$$

Example 5.11 (a): The set G of all 2×2 matrices of the form $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ over reals, where $ad - bc \neq 0$, i.e., with non-zero determinant forms a non-abelian group under matrix multiplication.

It is called the **general linear group** of 2×2 matrices over reals and is denoted by $GL(2, \mathbf{R})$.

The matrix $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ will act as identity and

the matrix $\begin{bmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix}$ will be inverse of $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

one can generalise and prove

(b) If G be the set of all $n \times n$ invertible matrices over reals, then G forms a group under matrix multiplication.

(c) The set of 2×2 matrices over **R** with determinant value 1 forms a non-abelian group under matrix multiplication and is called the **special linear group**, denoted by $SL(2, \mathbf{R})$.

One can take any field (e.g., **Q**, **C** or $\mathbf{Z}_p$) in place of **R** in the above examples.

Example 5.12: Let $G = \{2^r \mid r = 0, \pm 1, \pm 2, ...\}$

We show G forms a group under usual multiplication.

For any $2^r, 2^s \in G$, $2^r . 2^s = 2^{r+s} \in G$

Thus closure holds.

Associativity is obvious.

Again as $1 \in G$, and $x . 1 = 1 . x = x$ for all $x \in G$

1 is identity.

For any $2^r \in G$, as $2^{-r} \in G$ and $2^r . 2^{-r} = 2^0 = 1$,

we find each element of G has inverse. Commutativity is evidently true.

Example 5.13: *Group of Residues* : Let $G = \{0, 1, 2, 3, 4\}$. Define a composition $\oplus_5$ on G by $a \oplus_5 b = c$ where c is the least non –ve integer obtained as remainder when $a + b$ is divided by 5. For example. $3 \oplus_5 4 = 2$, $3 \oplus_5 1 = 4$, etc. Then $\oplus_5$ is a binary composition on G (called addition modulo 5). It is easy to verify that G forms a group under this.

One can generalise this result to

$$G = \{0, 1, 2, ..., n - 1\}$$

under addition modulo n where n is any positive integer.

NOTES

We thus notice

$$a \oplus_n b = \begin{cases} a+b & \text{if } a+b<n \\ a+b-n & \text{if } a+b \geq n \end{cases}$$

Also, in case there is no scope of confusion we drop the sub suffix n and simply write $\oplus$. This group is generally denoted by $\mathbf{Z}_n$.

Example 5.14: Let $G = \{x \in \mathbf{Z} \mid 1 \leq x < n, x, n \text{ being co-prime}\}$ where $\mathbf{Z}$ = set of integers and x, n being co-prime means H.C.F of x and n is 1.

We define a binary composition $\otimes$ on G by $a \otimes b = c$ where c is the least +ve remainder obtained when $a \,.\, b$ is divided by n. This composition $\otimes$ is called multiplication modulo n.

We show G forms a group under $\otimes$.

Closure: For $a, b \in G$, let $a \otimes b = c$. Then $c \neq 0$, because otherwise $n \mid ab$ which is not possible as a, n and b, n are co-prime.

Thus $c \neq 0$ and also then $1 \leq c < n$.

Now if c, n are not co-prime then $\exists$ some prime no. p such that, $p \mid c$ and $p \mid n$.

Again as $ab = nq + c$ for some q

We get $p \mid ab$ $\quad [p \mid n \Rightarrow p \mid nq, p \mid c \Rightarrow p \mid nq + c]$

$\Rightarrow p \mid a$ or $p \mid b$ (as p is prime)

If $p \mid a$ then as $p \mid n$ it means a, n are not co-prime.

But a, n are co-prime.

Similarly $p \mid b$ leads to a contradiction.

Hence c, n are co-prime and thus $c \in G$, showing that closure holds.

Associativity: Let $a, b, c \in G$ be any elements.

Let $a \otimes b = r_1$, $(a \otimes b) \otimes c = r_1 \otimes c = r_2$

then r_2 is given by $r_1 c = nq_2 + r_2$

Also $a \otimes b = r_1$ means

$$ab = q_1 n + r_1$$

thus $$ab - q_1 n = r_1$$

$$\Rightarrow (ab - q_1 n)c = r_1 c = nq_2 + r_2$$

$$\Rightarrow (ab)c = r_2 + nq_2 + nq_1 c = n(q_1 c + q_2) + r_2$$

or that r_2 is the least non-negative remainder got by dividing $(ab)c$ by n.

Similarly, if $a \otimes (b \otimes c) = r_3$ then we can show that r_3 is the least non –ve remainder got by dividing $a(bc)$ by n.

But since $a(bc) = (ab)c$, $r_2 = r_3$

Hence $a \otimes (b \otimes c) = (a \otimes b) \otimes c$.

NOTES

Existence of Identity: It is easy to see that

$$a \otimes 1 = 1 \otimes a = a \quad \text{for all } a \in G$$

or that 1 will act as identity.

Existence of Inverse: Let $a \in G$ be any element then a and n are co-prime and thus we can find integers x and y such that, $ax + ny = 1$

By division algorithm, we can write

$$x = qn + r, \quad \text{where } 0 \le r < n$$

$$\Rightarrow ax = aqn + ar$$

$$\Rightarrow ax + ny = aqn + ar + ny$$

$$\Rightarrow 1 = aqn + ar + ny$$

or that $ar = 1 + (-aq - y)n$

i.e., $a \otimes r = 1$. Similarly $r \otimes a = 1$. If r, n are co-prime, r will be inverse of a.

If r, n are not co-prime, we can find a prime number p such that, $p \mid r, p \mid n$

$$\Rightarrow p \mid qn \text{ and } p \mid r$$

$$\Rightarrow p \mid qn + r$$

$$\Rightarrow p \mid x$$

$$\Rightarrow p \mid ax \text{ also } p \mid ny$$

$$\Rightarrow p \mid ax + ny = 1$$

which is not possible. Thus r, n are co-prime and so $r \in G$ and is the required inverse of a.

It is easy to see that G will be abelian. We denote this group by U_n or $U(n)$ and call it the group of integers under multiplication modulo n.

Remark: Suppose $n = p$, a prime, then since all the integers 1, 2, 3, ..., $p - 1$ are co-prime to p, these will all be members of G. One can show that

$$G' = \{2, 4, 6, ..., 2(p - 1)\}$$

where $p > 2$ is a prime forms an abelian group under multiplication modulo $2p$.

Since for any $2n \in G$, $2n(p + 1) = 2np + 2n = 2n$

We noitce $p + 1$ will be identity of G'.

Again, for any $2n \in G'$, since $2n$ and p are co-prime $\exists$ x, y, s, t, $2nx + py = 1$

$$\Rightarrow py = 1 - 2nx = \text{odd}$$

$\Rightarrow y$ is odd as p is odd.

Let $y = 2k + 1$, then $2nx + p(2k + 1) = 1$

$$\Rightarrow 2nx + 2py + 2p = p + 1$$

$\Rightarrow 2nx + 2p(k + 1) = p + 1$

$\Rightarrow (2n)(x) = p + 1 =$ identity (under mod $2p$)

If x is even, x will be inverse of $2n$.

NOTES

If x is odd, $x + p$ will be inverse of $2n$.

Example 5.15: Let $G = \{0, 1, 2\}$ and define $*$ on G by

$$a * b = | a - b |$$

Then closure is established by taking a look at the composition table

$*$	0	1	2
0	0	1	2
1	1	0	1
2	2	1	0

Since $a * 0 = | a - 0 | = a = 0 * a$, 0 is identity

and $a * a = | a - a | = 0$ shows each element will be its own inverse.

But the system $(G, *)$ fails to be a group as associativity does not hold.

Indeed $1 * (1 * 2) = 1 * 1 = 0$

but $(1 * 1) * 2 = 0 * 2 = 2$

Example 5.16: Let $S = \{1, 2, 3\}$ and let $S_3 = A(S)$ = set all permutations of S. This set satisfies associativity, existence of identity and existence of inverse conditions in the definition of a group. Also clearly, since f, g permutations on S imply that fog is a permutation on S the closure property is ensured. Hence S_3 forms a group. That it is not abelian follows by the fact that $fog \neq gof$. This would, in fact, be the smallest non-abelian group and we shall have an occasion to talk about this group again under the section on permutation groups.

Remark: Let X be a non-empty set and let $M(X)$ = set of *all* maps from X to X, then $A(X) \subseteq M(X)$. $M(X)$ forms a semi group under composition of maps. Identity map also lies in $M(X)$ and as a map is invertible iff it is 1-1, onto *i.e.*, a permutation, we find $A(X)$ the subset of all permutations forms a group, denoted by S_X or Sym(X) and is called symmetric group of X. If X is finite with say, n elements then $o(M(X)) = n^n$ and $o(S_X) = \underline{|n}$ and in that case we use the notation S_n for S_X.

In the definition of a group, we only talked about the existence of identity and inverse of each element. We now show that these elements would also be unique, an elementary but exceedingly useful result. We prove it along with some other results in

NOTES

Lemma: *In a group G,*

(1) *Identity element is unique.*

(2) *Inverse of each $a \in G$ is unique.*

(3) $(a^{-1})^{-1} = a$, *for all* $a \in G$, *where* a^{-1} *stands for inverse of a.*

(4) $(ab)^{-1} = b^{-1} a^{-1}$ *for all* $a, b \in G$

(5) $ab = ac \Rightarrow b = c$

$ba = ca \Rightarrow b = c$ *for all* $a, b, c \in G$

(called the cancellation laws).

Proof: (1) Suppose e and e' are two elements of G which act as identity.

Then, since $e \in G$ and e' is identity,

$$e'e = ee' = e$$

and as $e' \in G$ and e is identity

$$e'e = ee' = e'$$

The two $\Rightarrow e = e'$

which establishes the uniqueness of identity in a group.

(2) Let $a \in G$ be any element and let a' and a'' be two inverse elements of a, then

$$aa' = a'a = e$$

$$aa'' = a''a = e$$

Now $a' = a'e = a'(aa'') = (a'a)a'' = ea'' = a''$.

Showing thereby that inverse of an element is unique. We shall denote inverse of a by a^{-1}.

(3) Since a^{-1} is inverse of a

$$aa^{-1} = a^{-1}a = e62$$

which also implies a is inverse of a^{-1}

Thus $(a^{-1})^{-1} = a$.

(4) We have to prove that ab is inverse of $b^{-1}a^{-1}$ for which we show

$$(ab)\,(b^{-1}a^{-1}) = (b^{-1}a^{-1})\,(ab) = e.$$

Now $(ab)\,(b^{-1}a^{-1}) = [(ab)\,b^{-1}]\,a^{-1}$

$= [(a(bb^{-1})]\,a^{-1}$

$= (ae)\,a^{-1} = aa^{-1} = e$

Similarly $(b^{-1}a^{-1})\,(ab) = e$

and thus the result follows.

NOTES

(5) Let $ab = ac$, then

$$\begin{aligned} b &= eb = (a^{-1}a)b \\ &= a^{-1}(ab) = a^{-1}(ac) \\ &= (a^{-1}a)c = ec = c \end{aligned}$$

Thus $ab = ac \Rightarrow b = c$

which is called the left cancellation law.

One can similarly, prove the right cancellation law.

Example 5.17 (*a*): Let $X = \{1, 2, 3\}$ and let $S_3 = A(X)$ be the group of all permutations on X. Consider $f, g, h \in A(X)$, defined by

$$f(1) = 2, \quad f(2) = 3, \quad f(3) = 1$$
$$g(1) = 2, \quad g(2) = 1, \quad g(3) = 3$$
$$h(1) = 3, \quad h(2) = 1, \quad h(3) = 2$$

It is easy then to verify that $fog = goh$

But $f \neq h$.

(*b*) If we consider the group in Example 5.10, we find

$$(1, 2) * (3, 4) = (3, 6) = (3, 0) * (1, 2)$$

But $(3, 4) \neq (3, 0)$

Hence we notice, cross cancellations *may not* hold in a group.

Theorem 5.1: *For elements a, b in a group G, the equations ax = b and ya = b have unique solutions for x and y in G.*

Proof: Now $ax = b$

$$\Rightarrow a^{-1}(ax) = a^{-1}b$$
$$\Rightarrow ex = a^{-1}b$$

or $x = a^{-1}b$

which is the required solution of the equation $ax = b$.

Suppose $x = x_1$ and $x = x_2$ are two solutions of this equation, then

$$ax_1 = b \text{ and } ax_2 = b$$
$$\Rightarrow ax_1 = ax_2$$
$$\Rightarrow x_1 = x_2 \quad \text{by left cancellation}$$

Showing that the solution is unique.

Similarly $y = ba^{-1}$ will be unique solution of the equation $ya = b$.

Theorem 5.2: *A non-empty set G together with a binary composition* '.' *is a group if and only if*

(1) $a(bc) = (ab)c$ *for all* $a, b, c \in G$

(2) *For any* $a, b \in G$, *the equations* $ax = b$ *and* $ya = b$ *have solutions in* G.

Proof: If G is a group, then (1) and (2) follow by definition and previous theorem.

Conversely, let (1) and (2) hold. To show G is a group, we need prove existence of identity and inverse (for each element).

NOTES

Let $a \in G$ be any element.

By (2) the equations $ax = a$

$ya = a$

have solutions in G.

Let $x = e$ and $y = f$ be the solutions.

Thus $\exists\ e, f \in G$, such that, $ae = a$

$fa = a$

Let now $b \in G$ be *any* element then again by (2) $\exists$ some x, y in G such that,

$ax = b$

$ya = b$.

Now $ax = b \Rightarrow f.(a\ .\ x) = f\ .\ b$

$\Rightarrow (f\ .\ a)\ .\ x = f\ .\ b$

$\Rightarrow a\ .\ x = f\ .\ b$

$\Rightarrow b = f\ .\ b$

Again $y\ .\ a = b \Rightarrow (y\ .\ a)\ .\ e = b\ .\ e$

$\Rightarrow y\ .\ (a\ .\ e) = b\ .\ e$

$\Rightarrow y\ .\ a = be$

$\Rightarrow b = be$

thus we have $b = fb$

$b = be$

for any $b \in G$

Putting $b = e$ in (i) and $b = f$ in (ii) we get

$e = fe$

$f = fe$

$\Rightarrow e = f.$

Hence $ae = a = fa = ea$

i.e., $\exists\ e \in G$, such that, $ae = ea = a$

$\Rightarrow e$ is identity.

Again, for any $a \in G$, and (the identity) $e \in G$, the equations $ax = e$ and $ya = e$ have solutions.

NOTES

Let the solutions be $x = a_1$, and $y = a_2$

then $aa_1 = e, \quad a_2a = e$

Now $a_1 = ea_1 = (a_2a)a_1 = a_2(aa_1) = a_2e = a_2.$

Hence $aa_1 = e = a_1a$ for any $a \in G$

i.e., for any $a \in G$, $\exists$ some $a_1 \in G$ satisfying the above relations $\Rightarrow a$ has an inverse. Thus each element has inverse and, by definition, G forms a group.

Remark: While proving the above theorem we have assumed that equations of the type $ax = b$ and $ya = b$ have solutions in G. The result may fail, if only one type of the above equations has solution. Consider for example:

G to be a set with at least two elements. Define '.' on G by $a \,.\, b = b$ for all $a, b \in G$.

then $a \,.\, (b \,.\, c) = a \,.\, c = c$

$(a \,.\, b) \,.\, c = b \,.\, c = c$

shows associativity holds.

Again as $ab = b$, the equation $ax = b$ has a solution for any $a, b \in G$.

We notice that G is not a group, as cancellation laws do not hold in G.

As let $a, b \in G$ be any two distinct members, then

$ab = b$

$bb = b \Rightarrow ab = bb$

But $a \neq b.$

Definition: A non-empty set G together with a binary composition '.' is called a *semi-group* if

$a \,.\, (b \,.\, c) = (a \,.\, b) \,.\, c$ for all $a, b, c \in G$

Obviously then every group is a semi-group. That the converse is not true follows by considering the set **N** of natural numbers under addition.

Theorem 5.3: *Cancellation laws may not hold in a semi-group.*

Proof: Consider M the set of all 2×2 matrices over integers under matrix multiplication, which forms a semi-group.

If we take $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}, C = \begin{bmatrix} 0 & 0 \\ 3 & 0 \end{bmatrix}$

then clearly $AB = AC = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

But $B \neq C.$

Set of natural numbers under addition is an example of a semi-group in which cancellation laws hold.

NOTES

Theorem 5.4: *A finite semi-group in which cancellation laws hold is a group.*

Proof: Let $G = \{a_1, a_2, ..., a_n\}$ be a finite semi-group in which cancellation laws hold.

Let $a \in G$ be any element, then by closure property

$$aa_1, aa_2, ..., aa_n$$

are all in G.

Suppose any two of these elements are equal

say, $aa_i = aa_j$ for some $i \neq j$

then $a_i = a_j$ by cancellation

But $a_i \neq a_j$ as $i \neq j$

Hence no two of $aa_1, aa_2, ..., aa_n$ can be equal.

These being n in number, will be distinct members of G (Note $o(G) = n$).

Thus if $b \in G$ be any element then

$$b = aa_i \quad \text{for some } i$$

i.e., for $a, b \in G$ the equation $ax = b$ has a solution ($x = a_i$) in G.

Similarly, the equation $ya = b$ will have a solution in G.

G being a semi-group, associativity holds in G.

Hence G is a group (by theorem 5.2).

Remark: The above theorem holds only in finite semi-groups. The semi-group of natural numbers under addition being an example where cancellation laws hold but which is not a group.

Theorem 5.5: *A finite semi-group is a group if and only if it satisfies cancellation laws.*

Proof: Follows by previous Theorem 5.4.

Definition: A non-empty set G together with a binary composition '.' is said to form a *monoid* if

(*i*) $a(bc) = (ab)c \quad \forall\ a, b, c \in G$

(*ii*) $\exists$ an element $e \in G$ such that, $ae = ea = a \quad \forall\ a \in G$

e is then called identity of G. It is easy to see that e is unique.

So all groups are monoids and all monoids are semi-groups.

When we defined a group, we insisted that $\exists$ an element e which acts both as a right as well as a left identity and each element has both sided inverse. We show now that it is not really essential and only one sided identity and the *same* sided inverse for each element could also make the system a group.

NOTES

Theorem 5.6: *A system* $< G, . >$ *forms a group if and only if*

(*i*) $a(bc) = (ab)c$ *for all* $a, b, c \in G$

(*ii*) $\exists\ e \in G$, *such* that, $ae = a$ *for all* $a \in G$

(*iii*) *for all* $a \in G$, $\exists\ a' \in G$, *such* that, $aa' = e$.

Proof: If G is a group, we have nothing to prove as the result follows by definition.

Conversely, let the given conditions hold.

All we need show is that $ea = a$ for all $a \in G$

and $a'a = a$ for any $a \in G$

Let $a \in G$ be any element.

By (*iii*) $\exists\ a' \in G$, such that, $aa' = e$

$\therefore$ For $a' \in G$, $\exists\ a'' \in G$, such that, $a'a'' = e$ (using (*iii*))

Now $a'a = a'(ae) = (a'a)e = (a'a)(a'a'')$

$= a'(aa')a'' = a'(e)a'' = (a'e)a'' = a'a'' = e.$

Thus for any $a \in G$, $\exists\ a' \in G$, such that, $aa' = a'a = e$

Again $ea = (aa')a = a(a'a) = ae = a$

$\therefore$ $ae = ea = a$ for all $a \in G$

i.e., e is identity of G.

Hence G is a group.

It would now be a routine exercise to prove

Theorem 5.7: *A system* $< G, . >$ *forms a group if and only if*

(*i*) $a(bc) = (ab)c$ *for all* $a, b, c \in G$

(*ii*) $\exists\ e \in G$, *such that,* $ea = a$ *for all* $a \in G$

(*iii*) *for all* $a \in G$, $\exists$ *some* $a' \in G$, *such* that, $a'a = e$.

A natural question to crop up at this stage would be what happens, when one sided identity and the other sided inverse exists. Would such a system also form a group? The answer to which is provided by

Example 5.18. Let G be a finite set having at least two elements. Define '.' on G by

$ab = b$ for all $a, b \in G$

then clearly associativity holds in G.

Let $e \in G$ by any fixed element.

Then as $ea = a$ for all $a \in G$

e will act as left identity.

Again $a \,.\, e = e$ for all $a \in G$

$\Rightarrow$ e is right inverse for any element $a \in G$.

NOTES

But we know G is not a group (cancellation laws do not hold in it).

Hence for a system $< G, . >$ to form a group it is essential that the same sided identity and inverse exist.

A Notation: Let G be a group with binary composition '.'. If $a \in G$ be any element then by closure property $a \,.\, a \in G$. Similarly $(a \,.\, a) \,.\, a \in G$ and so on.

It would be very convenient (and natural!) to denote $a \,.\, a$ by a^2 and $a \,.\, (a \,.\, a)$ or $(a \,.\, a) \,.\, a$ by a^3 and so on. Again $a^{-1} . a^{-1}$ would be denoted by a^{-2}. And since $a \,.\, a^{-1} = e$, it would not be wrong to denote $e = a^0$. It is now a simple matter to understand that under our notation

$$a^m \,.\, a^n = a^{m+n}$$
$$(a^m)^n = a^{mn}$$

where m, n are integers.

In case the binary composition of the group is denoted by +, we will talk of sums and multiples in place of products and powers. Thus here $2a = a + a$, and $na = a + a + ... + a$ (n times), if n is a +ve integer. In case n is –ve integer then $n = – m$, where m is +ve and we define $na = – ma = (– a) + (– a) + ... + (– a)$ m times.

Example 5.19. *If G is a finite group of order n then show that for any $a \in G$, $\exists$ some positive integer r, $1 \le r \le n$, such that, $a^r = e$.*

Solution: Since $o(G) = n$, G has n elements.

Let $a \in G$ be any element. By closure property $a^2, a^3, ...$ all belong to G.

Consider $\quad e, a, a^2, ..., a^n$

These are $n + 1$ elements (all in G). But G contains only n elements.

$\Rightarrow$ at least two of these elements are equal. If any of $a, a^2, ..., a^n$ equals e, our result is proved. If not, then $a^i = a^j$ for some i, j, $1 \le i, j \le n$. Without any loss of generality, we can take $i > j$

then $\quad a^i = a^j$

$$\Rightarrow a^i \,.\, a^{-j} = a^j \,.\, a^{-j}$$
$$\Rightarrow a^{i-j} = e \quad \text{where } 1 \le i - j \le n.$$

Putting $i – j = r$ gives us the required result.

Example 5.20. *Show that a finite semi-group in which cross cancellation holds is an abelian group.*

Solution: Let G be the given finite semi-group. Let $a, b \in G$ be any elements. Since G is a semi-group, by associativity

$$a(ba) = (ab)a$$

By cross cancellation then $ba = ab \Rightarrow G$ is abelian.

NOTES

Since G is abelian, cross cancellation laws become the cancellation laws. Hence G is a finite semi-group in which cancellation laws hold.

Thus G is a group.

Example 5.21. *If G is a group in which $(ab)^i = a^i b^i$ for three consecutive integers i and any a, b in G, then show that G is abelian.*

Solution: Let $n, n+1, n+2$ be three consecutive integers for which the given condition holds. Then for any $a, b \in G$,

$$(ab)^n = a^n b^n \quad ...(1)$$

$$(ab)^{n+1} = a^{n+1} b^{n+1} \quad ...(2)$$

$$(ab)^{n+2} = a^{n+2} b^{n+2} \quad ...(3)$$

Now $(ab)^{n+2} = a^{n+2} b^{n+2}$

$\Rightarrow (ab)(ab)^{n+1} = a^{n+2}\, b^{n+2}$

$\Rightarrow (ab)(a^{n+1} b^{n+1}) = a^{n+2}\, b^{n+2}$

$\Rightarrow ba^{n+1} = a^{n+1} b$ (using cancellation) ...(4)

Similarly $(ab)^{n+1} = a^{n+1} b^{n+1}$

gives $(ab)(ab)^n = a^{n+1} b^{n+1}$

i.e., $(ab)(a^n b^n) = a^{n+1} b^{n+1}$

$\Rightarrow ba^n = a^n b$

$\Rightarrow ba^{n+1} = a^n ba$

$\Rightarrow a^{n+1} b = a^n ba$ using Equation (4)

$\Rightarrow ab = ba.$

Hence G is abelian.

Remark: Conclusion of the above result may not follow if the given result holds only for two consecutive integers.

Consider, for example, the Quaternion group. One can check that $(ab)^i = a^i b^i$ for $i = 4, 5$ but the group is not abelian.

Example 5.22. *Suppose $(ab)^n = a^n b^n$ for all $a, b \in G$ where $n > 1$ is a fixed integer.*

Show that (*i*) $(ab)^{n-1} = b^{n-1} a^{n-1}$

(*ii*) $a^n\, b^{n-1} = b^{n-1} a^n$

(*iii*) $(aba^{-1}b^{-1})^{n(n-1)} = e$ *for all $a, b \in G$*

Solution: (*i*) We have

$$[b^{-1}(ba)b]^n = b^{-1}(ba)^n b$$

and $[b^{-1}(ba)b]^n = (ab)^n$

$$(ab)^n = b^{-1}(ba)^n b$$

NOTES

$\Rightarrow (ab)^{n-1}ab = b^{-1}(b^n a^n)b$

$\Rightarrow (ab)^{n-1} = b^{n-1}a^{n-1}$ for all $a, b \in G$

(*ii*) Now $(a^{-1}b^{-1}ab)^n = a^{-n}b^{-n}a^n b^n$

and $(a^{-1}b^{-1}ab)^n = a^{-n}(b^{-1}ab)^n$

$= a^{-n}b^{-1}a^n b$

$\therefore \quad a^{-n}b^{-n}a^n b^n = a^{-n}b^{-1}a^n b$

$\Rightarrow a^n b^{n-1} = b^{n-1}a^n$ for all $a, b \in G$

(*iii*) Consider $(aba^{-1}b^{-1})^{n(n-1)}$

$= [(aba^{-1}b^{-1})^{n-1}]^n$

$= [(ba^{-1}b^{-1})^{n-1}a^{n-1}]^n$ by (*i*)

$= [ba^{-(n-1)}b^{-1}a^{n-1}]^n = [b(a^{-(n-1)}b^{-1}a^{n-1})]^n$

$= b^n(a^{-(n-1)}b^{-1}a^{n-1})^n = b^n a^{-(n-1)}b^{-n}a^{n-1}$

$= a^{-(n-1)}b^n b^{-n}a^{n-1}$ by (*ii*)

$= e$ for all $a, b \in G$.

Example 5.23. *Let G be a group and suppose there exist two relatively prime positive integers m and n such that* $a^m b^m = b^m a^m$ *and* $a^n b^n = b^n a^n$ *for all a, b* $\in G$*. Show that G is abelian.*

Solution: Since m, n are relatively prime, there exist integers x and y such that $mx + ny = 1$.

For any a, b we have

$(a^m b^n)^{mx} = (a^m b^n)(a^m b^n).....(a^m b^n)$ mx times

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

$= a^m(b^n a^m)^{mx-1}b^n$

$= a^m(b^n a^m)^{mx}(b^n a^m)^{-1}b^n$

$= a^m c^m(b^n a^m)^{-1}b^n$ where $c = (b^n a^m)^x$

$= c^m a^m(b^n a^m)^{-1}b^n$

$= c^m \; a^m a^{-m}b^{-n}b^n = c^m = (b^n a^m)^{mx}$

Similarly $(a^m b^n)^{ny} = (b^n a^m)^{ny}$

giving $(a^m b^n)^{mx+ny} = (b^n a^m)^{mx+ny}$

$\Rightarrow a^m b^n = b^n a^m$ for all $a, b \in G$ ...(1)

Now $ab = a^{mx+ny} \; b^{mx+ny}$

$= a^{mx} \cdot (a^{ny} \; b^{mx})b^{ny}$

$= a^{mx}(a^m \; k^m)b^{ny}$ where $d = a^y$, $k = b^x$

$= a^{mx}(k^m \; d^m)b^{ny}$ by (1)

$= a^{mx} \cdot b^{mx} \cdot a^{ny} \cdot b^{ny}$

NOTES

$= (a^x)^m . (b^x)^m . (a^y)^n . (b^y)^n$

$= (b^x)^m . (a^x)^m . (b^y)^n . (a^y)^n$

$= b^{mx}(a^{mx} . b^{ny}) . a^{ny} = b^{mx} (b^{ny} . a^{mx}) . a^{ny}$

$= b^{mx + ny} . a^{mx + ny} = ba.$

Hence G is abelian.

Remark: In the following problem we give another proof to Theorem 5.6 done earlier.

Example 5.24. *Let G be a semi-group, Suppose $\exists\ e \in G$, such that, $ae = a$ for all $a \in G$ and for each $a \in G$, $\exists\ a' \in G$, such that, $aa' = e$. Show that G is a group.*

Solution: We first show that G satisfies the right cancellation law.

Let $ac = bc.$

As given $\exists\ c' \in G$, such that, $cc' = e$

$\therefore$ $(ac)c' = (bc)c'$

$\Rightarrow\ a(cc') = b(cc')$

$\Rightarrow\ ae = be \Rightarrow a = b.$

We now show that e is left identity.

Consider, $(ea)a' = e(aa') = e . e = e$

Also $aa' = e$

$\therefore$ $aa' = (ea) = a'$

By right cancellation law,

$a = ea$ for all $a \in G$

$\therefore$ e is also left identify of G.

Again $(a'a)a' = a'(aa') = a'e = a'$

and $ea' = a'$

$\Rightarrow\ (a'a)a' = ea'$

$\Rightarrow\ a'a = e$ by right cancellation law

$\Rightarrow\ a'$ is also left inverse of a

So, G is a group.

Example 5.25. *If in a semi-group S, $x^2y = y = yx^2 \quad \forall x, y$, then show that S is abelian.*

Solution: $x^2y = y \Rightarrow x^2y^2 = y^2$

$yx^2 = y \quad \forall\ x, y \in S$

$\Rightarrow\ xy^2 = x \quad \forall\ x, y \in S$

$\Rightarrow\ x^2y^2 = x^2$

So $x^2 = y^2 \quad \forall\ x, y \in S$

Now $x^2y = y \Rightarrow y^2y = y \Rightarrow y^3 = y \quad \forall\ y \in S$

Also $yx^2y = y^2$ ...(1)

Now $xy^2 = x \Rightarrow xy^2x = x^2$ ...(2)

By Equation (1) and (2), $xy^2x = yx^2y$

Since $y = y^3 \quad \forall\ y \in S$, we get

$$\begin{aligned} xy &= (xy)^3 = xy\ xy\ xy \\ &= xy\ xy\ x^3y = x(yx)^2x(xy) \\ &= (yx)x^2(yx)\ (xy) \\ &= yx^3\ yx^2y = yxy\ x^2y \\ &= (yx)xy^2x \\ &= yx^2y^2x \\ &= y(y^2x) \text{ (as } y = yx^2) \\ &= y^3x \\ &= yx \qquad (\text{as } y^3 = y) \end{aligned}$$

Thus $xy = yx\ \forall\ x, y \in S$

Hence S is abelian.

Example 5.26. *If G is a semi-group such that given $a \in G$, $\exists$ unique $a' \in G$ such that $aa'a = a$, then show that G is a group.*

Solution: Let e, f be idempotents in G, *i.e.*, $e^2 = e$, $f^2 = f$.

We show $(ef)^2 = ef$.

Now $ef \in G \Rightarrow \exists\ g \in G$, such that,

$(ef)\ g\ (ef) = ef$ ...(1)

Also $ef(gefg)\ ef = (efgef)\ gef = (ef)\ gef = ef$

$\Rightarrow\ g = gefg$ ...(2)

Again, $(ef)\ (ge)\ (ef) = efgef = ef$

$\Rightarrow\ ge = g$ ...(3)

Also, $ef(fg)\ ef = efgef = ef$

$\Rightarrow\ fg = g$ ...(4)

Now g^2 = $(ge)\ (fg)$ by (3) and (4)

= $g(ef) = g$ by (2)

i.e., g is an idempotent.

Also, $g^3 = g^2g = gg = g \Rightarrow ggg = g$

But $g(ef)\ g = g$ and so $g = ef$ and

NOTES

NOTES

Thus ef is an idempotent, *i.e.*, $(ef)^2 = ef$

Now $(ef)\, f\, (ef) = (ef)\, (ef) = ef$

and $(ef)\, e\, (ef) = ef$

$\Rightarrow f = e$ showing thereby that G has unique idempotent, say e.

Now $aa'a = a \Rightarrow (a'a)^2 = a'a \Rightarrow a'a$ is an idempotent.

$\Rightarrow a'a = e.$

Similarly $aa' = e$

Now $a = aa'a = ae$

$a = aa'a = ea$

$\Rightarrow ae = ea = a \quad \forall\ a \in G$

$\Rightarrow e$ is identity of G.

Also given $a \in G$, $aa' = e = a'a$ showing that a' is inverse of a.

Hence G is a group.

Check Your Progress

1. Define binary composition.
2. What do you understand by quaternion group?
3. Write the statement of general linear group.
4. What is special linear group?
5. When G is called semi-group?

5.3 RATIONAL NORMAL FORM

The rational canonical form of a square matrix A with entries in a field F is a canonical form for matrices formed by conjugation by invertible matrices over F in linear algebra. The shape represents a simple decomposition of the vector space into cyclic subspaces for A. (i.e., spanned by some vector and its repeated images under A). Because a given matrix can only have one normal form (thus the term 'Canonical'), matrix B is identical to A if and only if it has the same rational canonical form as A. This form can be determined without any operations that might change while extending the field F (thus the 'Rational'), such as factoring polynomials, demonstrating that whether two matrices are comparable does not change when the field is extended. Ferdinand Georg Frobenius, a German mathematician, is the name of the form.

Some authors use the phrase rational canonical form to refer to a somewhat different form, the **primary rational canonical form**. The fundamental form, rather of decomposing into a small number of cyclic subspaces, decomposes into a large number of them. It is similarly defined over F, but with major differences:

determining the form necessitates polynomial factorization, and as a result, the primary rational canonical form may change when the same matrix is evaluated over an extension field of F. This article focuses on the form that does not require factorization, and it uses the term 'Primary' when referring to the form that does.

Rational Normal Form Motivation

When determining whether two square matrices A and B are comparable, one way is to deconstruct the vector space as far as possible into a direct sum of stable subspaces for each of them and compare the actions on these subspaces. If both are diagonalizable, for example, one can decompose them into eigenspaces (for which the action is as basic as it gets, namely by a scalar), and then compare their eigenvalues and multiplicities to determine similarity. While this is typically a very illuminating strategy in practice, it has a number of limitations as a general method.

First, it necessitates the discovery of all eigenvalues, such as the roots of the characteristic polynomial, but an explicit statement for them may not be attainable. Second, a complete set of eigenvalues may exist only in a subset of the field under consideration, in which case there is no proof of similarity to the original field. Finally, even over this bigger field, A and B may not be diagonalizable, in which case a decomposition into generalized eigenspaces, and potentially Jordan blocks, must be used instead.

However, attaining such a detailed decomposition is not required to simply determine if two matrices are comparable. Instead, the rational canonical form relies on a direct sum decomposition into as many stable subspaces as possible, while yet permitting a fairly basic description of the action on each of them. These subspaces are called cyclic subspaces (by analogy with cyclic subgroups) and are clearly stable under the linear operator. They are formed by a single nonzero vector v and all of its images by repeated application of the linear operator associated with the matrix. Taking v and its consecutive images as long as they are linearly independent yields a basis for such a subspace. The companion matrix of a monic polynomial is the matrix of the linear operator with respect to such a basis; this polynomial (the minimal polynomial of the operator restricted to the subspace, which is analogous to the order of a cyclic subgroup) determines the action of the operator on the cyclic subspace up to isomorphism and is independent of the vector v generate.

There is always a direct sum decomposition into cyclic subspaces, and obtaining one does not necessitate factoring polynomials. However, cyclic subspaces may allow a decomposition as the direct sum of smaller cyclic subspaces (essentially by the Chinese remainder theorem). As a result, knowing the respective minimum polynomials and having some decomposition of the space into cyclic subspaces for both matrices is insufficient to determine their similarity. To verify that decompositions into cyclic subspaces for similar matrices are same, an extra requirement is imposed: in the list of associated minimum polynomials, each one must divide the next (and the constant polynomial 1 is forbidden to exclude trivial

cyclic subspaces of dimension 0). The invariant factors of (the K[X]-module formed by) the matrix are the resulting list of polynomials, and two matrices are equivalent if and only if they have identical lists of invariant factors. A matrix's rational canonical form is derived by expressing it on a basis adapted to a decomposition into cyclic subspaces whose associated minimum polynomials are the invariant factors of A; two matrices are identical if and only if their rational canonical forms are the same.

Example 5.27

Consider the following matrix A, which is centered on Q:

$$A=\begin{pmatrix} -1 & 3 & -1 & 0 & -2 & 0 & 0 & -2 \\ -1 & -1 & 1 & 1 & -2 & -1 & 0 & -1 \\ -2 & -6 & 4 & 3 & -8 & -4 & -2 & 1 \\ -1 & 8 & -3 & -1 & 5 & 2 & 3 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 & 1 & 0 \end{pmatrix}.$$

Solution: A has minimal polynomial $\mu = X^6 - 4X^4 - 2X^3 + 4X^2 + 4X + 1$, so that the dimension of a subspace generated by the repeated images of a single vector is at most 6. The characteristic polynomial is $\chi = X^8 - X^7 - 5X^6 + 2X^5 + 10X^4 + 2X^3 - 7X^2 - 5X - 1$, which is a multiple of the minimal polynomial by a factor $X^2 - X - 1$. There always exist vectors such that the cyclic subspace that they generate has the same minimal polynomial as the operator has on the whole space; indeed most vectors will have this property, and in this case the first standard basis vector e_1 does so: the vectors $A^k(e_1)$ for $k = 0, 1, \ldots, 5$ are linearly independent and span a cyclic subspace with minimal polynomial μ. There exist complementary stable subspaces (of dimension 2) to this cyclic subspace, and the space generated by vectors $v = (3, 4, 8, 0, -1, 0, 2, -1)^T$ and $w = (5, 4, 5, 9, -1, 1, 1, -2)^T$ is an example. In fact one has $A.v = w$, so the complementary subspace is a cyclic subspace generated by v; it has minimal polynomial $X^2 - X - 1$ must divided μ (and it is easily checked that it does), and we have found the invariant factors $X^2 - X - 1$ and $\mu = X^6 - 4X^4 - 2X^3 + 4X^2 + 4X + 1$ of A. Then the rational canonical form of A is the block diagonal matrix with the corresponding companion matrices as diagonal blocks, namely

$$C=\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

A basis on which this form is attained is formed by the vectors v, w above, followed by $A^k(e_1)$ for $k = 0, 1, \ldots, 5$; explicitly this means that for

$$P = \begin{pmatrix} 3 & 5 & 1 & -1 & 0 & 0 & -4 & 0 \\ 4 & 4 & 0 & -1 & -1 & -2 & -3 & -5 \\ 8 & 5 & 0 & -2 & -5 & -2 & -11 & -6 \\ 0 & 9 & 0 & -1 & 3 & -2 & 0 & 0 \\ -1 & -1 & 0 & 0 & 0 & 1 & -1 & 4 \\ 0 & 1 & 0 & 0 & 0 & 0 & -1 & 1 \\ 2 & 1 & 0 & 1 & -1 & 0 & 2 & -6 \\ -1 & -2 & 0 & 0 & 1 & -1 & 4 & -2 \end{pmatrix},$$

one has $A = PCP^{-1}$.

General Case and Theory of Rational Normal Form

Fix a base field F and a finite-dimensional vector space V over F. Given a polynomial $P \in F[X]$, there is associated to it a companion matrix C_P whose characteristic polynomial and minimal polynomial are both equal to P).

Theorem 5.8: Let V be a finite-dimensional vector space over a field F, and A a square matrix over F. Then V (viewed as an $F[X]$-module with the action of X given by A) admits a $F[X]$-module isomorphism

$$V \cong F[x]/f_1 \oplus \ldots \oplus F[X]/f_k$$

where the $f_i \in F[X]$ may be taken to be monic polynomials of positive degree (so they are non-units in $F[X]$) that satisfy the relations

$$f1 \mid f_2 \mid \ldots \mid f_k$$

where '$a \mid b$' is notation for 'a divides b'; with these conditions the list of polynomials f_i is unique.

Proof: Apply the structure theorem for finitely generated modules over a principal ideal domain to V, viewing it as an $F[X]$-module. The structure theorem provides a decomposition into cyclic factors, each of which is a quotient of $F[X]$ by a proper ideal; the zero ideal cannot be present since the resulting free module would be infinite-dimensional as F vector space, while V is finite-dimensional. For the polynomials f_i one then takes the unique monic generators of the respective ideals, and since the structure theorem ensures containment of every ideal in the preceding ideal, one obtains the divisibility conditions for the f_i.

Given an arbitrary square matrix, the elementary divisors used in the construction of the Jordan normal form do not exist over $F[X]$, so the invariant factors f_i as given above must be used instead. The last of these factors f_k is then minimal polynomial, which all the invariant factors therefore divide, and the product of the invariant factors gives the characteristic polynomial. Note that this implies that the minimal polynomial divides the characteristic polynomial (which is essentially the Cayley-Hamilton theorem), and that every irreducible factor of the characteristic polynomial also divides the minimal polynomial (possibly with lower multiplicity).

NOTES

For each invariant factor fi one takes its companion matrix C_{fi}, and the block diagonal matrix formed from these blocks yields the rational canonical form of A. When the minimal polynomial is identical to the characteristic polynomial (the case $k = 1$), the Frobenius normal form is the companion matrix of the characteristic polynomial. As the rational canonical form is uniquely determined by the unique invariant factors associated to A, and these invariant factors are independent of basis, it follows that two square matrices A and B are similar if and only if they have the same rational canonical form.

5.3.1 Generalised Jordon form over any Field

Even if it exists over the ground field F, the rational or Frobenius normal form does not reflect any sort of factorization of the characteristic polynomial. This means that when F is substituted by a different field, it remains invariant (as long as it contains the entries of the original matrix A). However, this distinguishes the Frobenius normal form from other normal forms that rely on factoring the characteristic polynomial, such as the diagonal form (if A is diagonalizable) or the Jordan normal form in general (if the characteristic polynomial splits into linear factors). A diagonal matrix with unique diagonal elements, for example, has a Frobenius normal form that is simply the partner matrix of its characteristic polynomial.

There is another way to define a normal form, that, like the Frobenius normal form, is always defined over the same field F as A, but that does reflect a possible factorization of the characteristic polynomial (or equivalently the minimal polynomial) into irreducible factors over F, and which reduces to the Jordan normal form when this factorization only contains linear factors (corresponding to eigenvalues). This form is sometimes called the **generalized Jordan normal** form, or **primary rational canonical form**. It is based on the fact that the vector space can be canonically decomposed into a direct sum of stable subspaces corresponding to the distinct irreducible factors P of the characteristic polynomial (as stated by the lemme des noyaux [fr]), where the characteristic polynomial of each summand is a power of the corresponding P. These summands can be further decomposed, non-canonically, as a direct sum of cyclic $F[x]$-modules (like is done for the Frobenius normal form above), where the characteristic polynomial of each summand is still a (generally smaller) power of P. The primary rational canonical form is a block diagonal matrix corresponding to such a decomposition into cyclic modules, with a particular form called generalized Jordan block in the diagonal blocks, corresponding to a particular choice of a basis for the cyclic modules. This generalized Jordan block is itself a block matrix of the form

$$\begin{pmatrix} C & 0 & \cdots & 0 \\ U & C & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & U & C \end{pmatrix}$$

where C is the companion matrix of the irreducible polynomial P, and U is a matrix whose sole non-zero entry is a 1 in the upper right hand corner. For the case of a linear irreducible factor $P = x - \lambda$, these blocks are reduced to single entries $C = \lambda$ and $U = 1$ and, one finds a (transposed) Jordan block. In any generalized Jordan block, all entries immediately below the main diagonal are 1. A basis of the cyclic module giving rise to this form is obtained by choosing a generating vector v (one that is not annihilated by $P^{k-1}(A)$ where the minimal polynomial of the cyclic module is P^k), and taking as basis

$v, A(v), A^2(v), \ldots, A^{d-1}(v), P(A)(v), A(P(A)(v)), \ldots, A^{d-1}(P(A)(v), \ldots, P^2(A)(v), \ldots, P^{k-1}(A)(v), \ldots, A^{d-1}(P^{k-1}(A)(v))$

where $d = \deg(P)$.

Check Your Progress

6. What is rational cononical form of a square matrix?
7. Define generalised Jordon block in diagonal block.

5.4 ANSWERS TO 'CHECK YOUR PROGRESS'

1. The binary composition for a group is denoted by '.' (dot) which is so convenient to write (and makes the axioms look so natural too).

   This binary composition '.' is called product or multiplication (although it may have nothing to do with the usual multiplication, that we are so familiar with). In fact, we even drop '.' and simply write ab in place of $a \cdot b$.

2. Let $G = \{\pm 1, \pm i, \pm j, \pm k\}$. Define product on G by usual multiplication together with

   $i^2 = j^2 = k^2 = -1, \qquad ij = -ji = k$

   $jk = -kj = i$

   $ki = -ik = j$

   then G forms a group. G is not abelian as $ij \neq ji$.

   This is called the Quaternion Group.

3. The set G of all 2×2 matrices of the form $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ over reals, where $ad - bc \neq 0$, i.e., with non zero determinant forms a non abelian group under matrix multiplication.

   It is called the general linear group of 2×2 matrices over reals and is denoted by $GL(2, \mathrm{R})$.

NOTES

4. The set of 2×2 matrices over **R** with determinant value 1 forms a non-abelian group under matrix multiplication and is called the special linear group, denoted by $SL(2, \mathbf{R})$.
5. A non-empty set G together with a binary composition '.' is called a *semi-group* if

   $a \,.\, (b \,.\, c) = (a \,.\, b) \,.\, c$ for all $a, b, c \in G$

   Obviously then every group is a semi-group. That the converse is not true follows by considering the set **N** of natural numbers under addition.
6. The rational canonical form of a square matrix A with entries in a field F is a canonical form for matrices formed by conjugation by invertible matrices over F in linear algebra. The shape represents a simple decomposition of the vector space into cyclic subspaces for A. (i.e., spanned by some vector and its repeated images under A).
7. The primary rational canonical form is a block diagonal matrix corresponding to such a decomposition into cyclic modules, with a particular form called generalized Jordan block in the diagonal blocks, corresponding to a particular choice of a basis for the cyclic modules.

5.5 SUMMARY

- Associativity: $a * (b * c) = (a * b) * c$, for all $a, b, c \in G$
- Existence of Identity: $\exists$ an element $e \in G$, such that,

  $a * e = e * a = a$ for all $a \in G$

  (e is then called identity)
- Existence of Inverse: For every $a \in G, \exists\, a' \in G$ (depending upon a) such that,

  $a * a' = a' * a = e$

  (a' is then called inverse of a)
- Since $*$ is a binary composition on G, it is understood that for all $a, b \in G, a * b$ is a unique member of G. This property is called *closure property*.
- This binary composition '.' is called product or multiplication (although it may have nothing to do with the usual multiplication, that we are so familiar with). In fact, we even drop '.' and simply write ab in place of $a \,.\, b$.
- Set of all non-zero complex numbers forms a group under multiplication defined by

  $(a + ib)\,(c + id) = (ac - bd) + i\,(ad + bc)$.
- The set G of all nth roots of unity, where n is a fixed positive integer forms an abelian group under usual multiplication of complex numbers.

- Let $G = \{\pm 1, \pm i, \pm j, \pm k\}$. Define product on G by usual multiplication together with

  $i^2 = j^2 = k^2 = -1, \qquad ij = -ji = k$

  $jk = -kj = i$

  $ki = -ik = j$

  then G forms a group. G is not abelian as $ij \neq ji$.

  This is called the Quaternion Group.
- For elements a, b in a group G, the equations $ax = b$ and $ya = b$ have unique solutions for x and y in G.
- Cancellation laws may not hold in a semi-group.
- A finite semi-group in which cancellation laws hold is a group.
- A finite semi-group is a group if and only if it satisfies cancellation laws.
- The rational canonical form of a square matrix A with entries in a field F is a canonical form for matrices formed by conjugation by invertible matrices over F in linear algebra. The shape represents a simple decomposition of the vector space into cyclic subspaces for A. (i.e., spanned by some vector and its repeated images under A).
- The primary rational canonical form is a block diagonal matrix corresponding to such a decomposition into cyclic modules, with a particular form called generalized Jordan block in the diagonal blocks, corresponding to a particular choice of a basis for the cyclic modules.

5.6 KEY TERMS

- **Closure property:** Since $*$ is a binary composition on G, it is understood that for all $a, b \in G$, $a * b$ is a unique member of G. This property is called closure property.
- **Existence of identity:** $\exists$ an element $e \in G$, such that,

  $a * e = e * a = a$ for all $a \in G$

  (e is then called identity)
- **Existence of inverse:** For every $a \in G$, $\exists\ a' \in G$ (depending upon a) such that,

  $a * a' = a' * a = e$

  (a' is then called inverse of a)
- **Rational normal form:** The rational canonical form of a square matrix A with entries in a field F is a canonical form for matrices formed by conjugation by invertible matrices over F in linear algebra. The shape represents a simple decomposition of the vector space into cyclic subspaces for A. (i.e., spanned by some vector and its repeated images under A).

NOTES

5.7 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. What do you mean by the finitely generated Abelian group?
2. What is rational normal form?
3. State the generalised Jordon form over any field.

Long-Answer Questions

1. Briefly discuss about the finitely generated Abelian group giving appropriate examples.
2. Elaborate on the is rational normal form give appropriate examples.
3. Discuss in detail about the generalised Jordon form over any field with the help of relevant examples.

5.8 FURTHER READING

Herstein, I.N. 1975. Topics in Algebra, 3rd Edition. New Delhi: Wiley Eastern Ltd.

Khanna, V.K. and S.K. Bhambri. 2008. A Course in Abstract Algebra, 3rd Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Bhattacharya, P.B., S.K. Jain and S.R. Nagpaul. 1997. Basic Abstract Algebra, 2nd Edition. New Delhi: Cambridge University Press.

Artin, M.1991. Algebra. New Delhi: Prentice-Hall of India.

Lang, S. 1993. Algebra, 3rd Edition. New York: Addison-Wesley.

Datta, K.B. 2000. Matrix and Linear Algebra. New Delhi: Prentice-Hall of India.