ALGEBRAIC STRUCTURES

Chapter 1

Group Theory

1.1 Basics of Group

Definition 1.1.1. A *group* is an ordered pair $(G, *)$, where G is a nonempty set and $*$ is a binary operation on G such that the following properties hold:
(G1) For all $a, b, c \in G$, $a * (b * c) = (a * b) * c$ (associative law).
(G2) There exists $e \in G$ such that for all $a \in G$, $a * e = a = e * a$ (existence of an identity).
(G3) For all $a \in G$, there exists $a' \in G$ such that $a * a' = e = a' * a$ (existence of an inverse).

Definition 1.1.2. A group G is said to be *abelian* if $ab = ba$ for all $a, b \in G$. A group which is not abelian is called a *non-abelian* group.

Examples 1.1.3.

1. Let $G = \{e\}$ and $e * e = e$. Obviously G is a trivial group.
2. $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ are groups under usual addition.

3. The set of all 2×2 matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ where $a, b, c, d \in \mathbb{R}$ is a group under matrix addition. $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ is the identity element and $\begin{pmatrix} -a & -b \\ -c & -d \end{pmatrix}$ is the inverse of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

4. The set of all 2×2 non-singular matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ where $a, b, c, d \in \mathbb{R}$ is a group under matrix multiplication. $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is the identity element. The inverse of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is $\frac{1}{|A|}\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ where $|A| = ad - bc \neq 0$.

5. $\mathbb{N}$ is not a group under usual addition since there is no element $e \in \mathbb{N}$ such that $x + e = x$.

6. The set $\mathbb{E}$ of all even integers under usual addition is a group.

7. $\mathbb{Q}^*$ and $\mathbb{R}^*$ under usual multiplication are groups. 1 is the identity element and the inverse of a non-zero element a is $1/a$.

8. $\mathbb{Q}^+$ is a group under usual multiplication. For $a, b \in \mathbb{Q}^+ \Rightarrow ab \in \mathbb{Q}^+$. Therefore usual multiplication is a binary operation in $\mathbb{Q}^+$.
$1 \in \mathbb{Q}^+$ is the identity element. If $a \in \mathbb{Q}^+$, $(1/a) \in \mathbb{Q}^+$ is the inverse of a.

9. $\mathbb{Z}$ under the usual multiplication is not a group.

10. $G = \{1, i, -1, -i\}$. G is a group under usual multiplication. The identity element is 1. The inverse of $1, i, -1$ and $-i$ are $1, -i, -1$ and i respectively.
The Cayley table for this group is given by

*	1	i	-1	-i
1	1	i	-1	i
i	i	-1	-i	1
-1	-1	-i	1	i
-i	-i	1	i	-1

11. Let $G = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \right\}$
G is a group under matrix multiplication. [Construct the Cayley table for this group]

12. $\mathbb{C}^*$ is a group under usual multiplication given by $(a+ib)(c+id) = (ac-bd) + i(ad+bc)$.

13. Let $G = \{z : \ z \in \mathbb{C} \text{ and } |z| = 1\}$. Then G is a under usual multiplication.

14. The set of all n^{th} roots of unity with usual multiplication is a group.

15. Let $G = \{a + b\sqrt{2} : \ a, b \in \mathbb{Z}\}$. Then G is a group under addition.

Definition 1.1.4. Let $\mathbb{Z}_n = \{0, 1, 2, \ldots, n-1\}$. Let $a, b \in \mathbb{Z}_n$. Then $a + b = qn + r$ where $0 \leq r < n$. We define $a \oplus b = r$. Let $ab = q'n + s$ where $0 \leq s < n$. We define $a \odot b = s$. The binary operations $\oplus$ and $\odot$ are called *addition modulo* n and *multiplication modulo* n respectively. Then $(\mathbb{Z}_n, \oplus)$ is an abelian group.

Let n be a prime. Then $\mathbb{Z}_n - \{0\}$ is a group under multiplication modulo n.

1.2 Elementary properties of group

Theorem 1.2.1. *Let G be a group. Then*

(i) There exists a unique identity element $e \in G$ such that $e * a = a = a * e$ for all $a \in G$.

(ii) For all $a \in G$, there exists a unique inverse $a' \in G$ such that $a * a' = e = a' * a$.

Proof. (i) Now G is group. Therefore, by (G2), there exists $e \in G$ such that $e * a = a = a * e$ for all $a \in G$. Suppose, let e and e' be two identity elements of G. Then $ee' = e'$ (since e is an identity element). Also $ee' = e$(since e' is an identity element). Hence $e = e'$.

(ii) Let $a \in G$. By (G3), there exists $a' \in G$ such that $a * a' = e = a' * a$. Suppose there exists $a'' \in G$ such that $a * a'' = e = a'' * a$. We show that $a' = a''$. Now

$$\begin{aligned} a' &= a' * e = a' * (a * a'')(\text{substituting } e = a * a'') \\ &= (a' * a) * a'' = e * a''(\text{because } a' * a = e) = a''. \end{aligned}$$

Thus, a' is unique. □

We denote the inverse of a by a^{-1}.

Theorem 1.2.2. *In a group, the left and right cancellation laws hold (i.e,) $ab = ac \Rightarrow b = c$ and $ba = ca \Rightarrow b = c$.*

Proof. Suppose $ab = ac \Rightarrow a^{-1}(ab) = a^{-1}(ac) \Rightarrow (a^{-1}a)b = (a^{-1}a)c \Rightarrow eb = ec \Rightarrow b = c$. Similarly, we can prove that $ba = ca \Rightarrow b = c$. □

Theorem 1.2.3. *Let G be a group and $a, b \in G$. Then the equation $ax = b$ and $ya = b$ have unique solutions for x and y in G.*

Proof. Consider $a^{-1}b \in G$. Then $a(a^{-1}b) = (aa^{-1})b = eb = b$. Hence $a^{-1}b$ is a solution of $ax = b$. Now, to prove the uniqueness, let x_1 and x_2 be two solutions of $ax = b$. Then $ax_1 = b$ and $ax_2 = b$. Therefore $ax_1 = ax_2$ which implies $x_1 = x_2$. Hence $x = a^{-1}b$ is the unique solution for $ax = b$. Similarly we can prove that $y = ba^{-1}$ is the solution of the equation $ya = b$. □

Theorem 1.2.4. *Let G be a group. Let $a, b \in G$. Then $(ab)^{-1} = b^{-1}a^{-1}$ and $(a^{-1})^{-1} = a$.*

Proof. Now $(ab)(b^{-1}a^{-1}) = a(bb^{-1})a^{-1} = aea^{-1} = aa^{-1} = e$. Similarly $(b^{-1}a^{-1})(ab) = e$. Hence $(ab)^{-1} = b^{-1}a^{-1}$. Proof of the second part is obvious. $\square$

Corollary 1.2.5. *If $a_1, a_2, \ldots, a_n \in G$ then $(a_1a_2\cdots a_n)^{-1} = a_n^{-1}a_{n-1}^{-1}\cdots a_1^{-1}$.*

Definition 1.2.6. Let G be a group and $a \in G$. For any positive integer n, we define $a^n = aa\cdots a$(a written n times). Clearly $(a^n)^{-1} = (aa\cdots a)^{-1} = (a^{-1}a^{-1}\cdots a^{-1}) = (a^n)^{-1}$. Now we define $a^{-n} = (a^{-1})^n = (a^n)^{-1}$. Finally we define $a^0 = e$. Thus a^n is defined for all integers n.

When the binary operation on G is "+", we denote $a + a + \cdots + a$ (a written n times) as na.

Theorem 1.2.7. *Let G be a group and $a \in G$. Then*
(i) $a^m a^n = a^{m+n}$, $m, n \in \mathbb{Z}$.
(ii) $(a^m)^n = a^{mn}$, $m, n \in \mathbb{Z}$.

1.3 Permutation Groups

Definition 1.3.1. Let A be a finite set. A bijection from A to itself is called a permutation of A.

For example, if $A = \{1, 2, 3, 4\}$ $f : A \to A$ given by $f(1) = 2, f(2) = 1, f(3) = 4$ and $f(4) = 3$ is a permutation of A. We shall write this permutation as $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$. An element in the bottom row is the image of the element just above it in the upper row.

Definition 1.3.2. Let A be a finite set containing n elements. The set of all permutations of A is clearly a group under the composition of functions. This group is called the *symmetric group* of degree n and is denoted by S_n.

Example 1.3.3. Let $A = \{1, 2, 3\}$. Then S_3 consists of $e = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$;
$p_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$; $p_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$; $p_3 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$; $p_4 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$;
$p_5 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$. In this group, e is the identity element. We now compute the product p_1p_2.

$$\begin{array}{cccc} & 1 & 2 & 3 \\ p_1 : & \downarrow & \downarrow & \downarrow \\ & 2 & 3 & 1 \\ p_2 : & \downarrow & \downarrow & \downarrow \\ & 1 & 2 & 3 \end{array} \quad \text{Hence } p_1p_2 : \begin{array}{ccc} 1 & 2 & 3 \\ \downarrow & \downarrow & \downarrow \\ 1 & 2 & 3 \end{array}$$

So that $p_1p_2 = e$. Now, $p_1p_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = p_5$.
Similarly we can compute all other products and Cayley table for this group is given by

	e	p_1	p_2	p_3	p_4	p_5
e	e	p_1	p_2	p_3	p_4	p_5
p_1	p_1	p_2	e	p_4	p_5	p_3
p_2	p_2	e	p_1	p_5	p_3	p_4
p_3	p_3	p_5	p_4	e	p_2	p_1
p_4	p_4	p_3	p_5	p_1	e	p_2
p_5	p_5	p_4	p_3	p_2	p_1	e

Thus S_3 is a group containing $3! = 6$ elements.

In S_3, $p_1p_2 = p_2p_1 = e$ so that the inverse of p_1 is p_2. In general the inverse of a permutation can be obtained by interchanging the rows of the permutation.

For example, if $p = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 2 & 5 & 1 \end{pmatrix}$ then the inverse of p is the permutation given by $p^{-1} = \begin{pmatrix} 3 & 4 & 2 & 5 & 1 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 1 & 2 & 4 \end{pmatrix}$.

In S_3, $p_1p_4 = p_5$ and $p_4p_1 = p_3$. Hence $p_1p_4 \neq p_4p_1$ so that S_3 is non-abelian.

The symmetric group S_n containing $n!$ elements, for, let $A = \{1, 2, \ldots, n\}$. Any permutation on A is given by specifying the image of each element. The image of 1 can be chosen in n different ways. Since the image of two is different from the image of 1, it can be chosen in $(n-1)$ different ways and so on. Hence the number of permutations of A is $n(n-1)\cdots 2 \cdot 1 = n!$ so that the number of elements in S_n is $n!$.

Definition 1.3.4. Let G be a finite group. Then the number of elements in G is called the order of G and is denoted by $|G|$ or $o(G)$.

Definition 1.3.5. Let p be a permutation on $A = \{1, 2, \ldots, n\}$. p is called *a cycle* of length r if there exist distinct symbols $a_1, a_2, \ldots, a_r$ such that $p(a_1) = a_2, p(a_2) = a_3, \ldots, p(a_{r-1}) = a_r$, and $p(a_r) = a_1$, and $p(b) = b$ for all $b \in A - \{a_1, a_2, \ldots, a_r\}$. This cycle is represented by the symbol $(a_1, a_2, \cdots, a_r)$.

Thus under the cycle $(a_1, a_2, \cdots, a_r)$ each symbol is mapped onto the following symbol except the last one which is mapped onto the first symbol and all the other symbols not in the cycle are fixed.

Example 1.3.6. Let $A = \{1, 2, 3, 4, 5\}$. Consider the cycle of length 4 given by $p = (2451)$. Then $p = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 3 & 5 & 1 \end{pmatrix}$ and so $(2451) = (4521) = (5124) = (1245)$.

Remark 1.3.7. Since cycles are special types of permutations, they can be multiplied in the usual way. The product of cycles need not be a cycle.

For example, let $p_1 = (234)$ and $p_2 = (1,5)$. Then

$p_1p_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 4 & 2 & 5 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 2 & 3 & 4 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 4 & 2 & 1 \end{pmatrix}$ which is not a cycle.

Definition 1.3.8. Two cycles are said to be disjoint if they have any no symbols in common.

For example (2 1 5) and (3 4) are disjoint cycles.

Remark 1.3.9. If p_1 and p_2 are disjoint cycles the symbols which are moved by p_1 are fixed by p_2 and vice versa. Hence multiplication of disjoint cycles is commutative.

Theorem 1.3.10. *Any permutation can be expressed as a product of disjoint cycles.*

The decomposition of a permutation into disjoint cycles is unique except for the order of the factors.

Definition 1.3.11. A cycle of length two is called a *transposition* . Thus a transposition (a_1a_2) interchanges the symbols a_1 and a_2 and leaves all the other elements fixed.

Theorem 1.3.12. *Any permutation can be expressed as a product of transpositions.*

Proof. Since any permutation is a product of disjoint cycles it is enough to prove that each cycle is a product of transpositions. Let $c = (a_1a_2 \cdots a_1)$ be a cycle. Then $(a_1a_2 \cdots a_1) = (a_1a_2)(a_2a_3) \cdots (a_1a_r)$. This proves the theorem. $\square$

Theorem 1.3.13. *If a permutation $p \in S_n$ is a product of r transpositions and also a product of s transpositions then either r and s are both even or both odd.*

Definition 1.3.14. A permutation $p \in S_n$ is called *even* or *odd* according as p can be expressed as a product of an even number of transpositions or an odd number of transpositions respectively.

Theorem 1.3.15. *(i) The product of two even permutations is an even permutation.*
(ii) The product of two odd permutations is an even permutation.
(iii) The product of an even permutation and an odd permutation is an odd permutation.
(iv) The inverse of an even permutation is an even permutation.
(v) The inverse of an odd permutation is an odd permutation.
(vi) The identity permutation e is an even permutation.

Theorem 1.3.16. *Let A_n be the set of all even permutations in S_n. Then A_n is a group containing $\dfrac{n!}{2}$ permutations.*

Definition 1.3.17. The group A_n of all even permutations in S_n is called the *alternating group* on n symbols.

1.4 Subgroups

Definition 1.4.1. Let G be a set with binary operation $*$ defined on it. Let $S \subseteq G$. If for each $a, b \in S$, $a * b$ is in S, we say that S is *closed* with respect to the binary operation $*$.

Examples 1.4.2. (i) $(\mathbb{Z}, +)$ is a group. The set $\mathbb{E}$ of all even integers is closed under $+$ and further $(\mathbb{E}, +)$ is itself a group.

(ii) The set of G of all non-singular 2×2 matrices form a group under matrix multiplication. Let H be the set of all matrices of the form $\begin{pmatrix} cos\ \theta & -sin\ \theta \\ sin\ \theta & cos\ \theta \end{pmatrix}$. Then H is subset of G and H itself a group under matrix multiplication.

Definition 1.4.3. A subset H of group G is called *subgroup* of G if H forms a group with respect to the binary operation in G.

Examples 1.4.4. (i) Let G be any group. Then $\{e\}$ and G are trivial subgroups of G. They are called improper subgroups of G.

(ii) $(\mathbb{Q}, +)$ is a subgroup of $(\mathbb{R}, +)$ and $(\mathbb{R}, +)$ is a subgroup of $(\mathbb{C}, +)$.

(iii) In $(\mathbb{Z}_8, \oplus)$, let $H_1 = \{0, 4\}$ and $H_2 = \{0, 2, 4, 6\}$. The Cayley tables for H_1 and H_2 are given by

$\oplus$	0	4
0	0	4
4	4	0

$\oplus$	0	2	4	6
0	0	2	4	6
2	2	4	6	0
4	4	6	0	2
6	6	0	2	4

It is easily seen that H_1 and H_2 are closed under $\oplus$ and $(H_1, \oplus)$ and $(H_2, \oplus)$ are groups. Hence H_1 and H_2 are subgroups of $\mathbb{Z}_8$.

(iv) $\{1, -1\}$ is a subgroup of $(\mathbb{R}^*, \cdot)$.

(v) $\{1, i, -1, -i\}$ is a subgroup of $(\mathbb{C}^*, \cdot)$.

(vi) For any integer n we define $n\mathbb{Z} = \{nx : \ x \in \mathbb{Z}\}$. Then $(n\mathbb{Z}, +)$ is a subgroup of $(\mathbb{Z}, +)$. For, let $a, b \in n\mathbb{Z}$. Then $a = nx$ and $b = ny$ where $x, y \in \mathbb{Z}$. Hence $a + b = n(x + y) \in n\mathbb{Z}$ and so $n\mathbb{Z}$ is closed under $+$. Clearly $0 \in n\mathbb{Z}$ is the identity element. Inverse of nx is $-nx = n(-x) \in n\mathbb{Z}$. Hence $(n\mathbb{Z}, +)$ is a group.

(vii) In the symmetric group S_3, $H_1 = \{e, p_1, p_2\}; H_2 = \{e, p_3\}; H_3 = \{e, p_4\}$; and $H_4 = \{e, p_5\}$ are subgroups.

(viii) A_n is a subgroup of S_n.

In all the above examples we see that the identity element in the subgroup is the same as the identity element of the group.

Theorem 1.4.5. *Let H be a subgroup of G. Then*

(a) the identity element of H is the same as that of G.

(b) for each $a \in H$ the inverse of a in H is the same as the inverse of a in G.

Proof. (a) Let e and e' be the identity of G and H respectively. Let $a \in H$. Now,

$$e'a = a \text{(since e' is the identity of } H)$$
$$= ea \text{(since e' is the identity of } G \text{ and } a \in G)$$
$$\therefore \quad e'a = ea \Rightarrow e' = a \text{(by cancellation law)}$$

(b) Let a' and a'' be the inverse of a in G and H respectively. Since by (a), G and H have the same identity element e, we have $a'a = e = a''a$. Hence by cancellation law, $a' = a''$. □

Theorem 1.4.6. *A subset H of a group G is a subgroup of G if and only if*

(i) it is closed under the binary operation in G.

(ii) The identity e of G is in H. (iii) $a \in H \Rightarrow a^{-1} \in H$.

Proof. Let H be subgroup of G. The result follows immediately from Theorem 1.4.5.

Conversely, let H be a subset of G satisfying conditions (i), (ii) and (iii). Then, obviously H itself a group with respect to the binary operation in G. Therefore H is a subgroup of G. □

Theorem 1.4.7. *A non-empty subset H of a group G is a subgroup of G if and only if $a, b \in H \Rightarrow ab^{-1} \in H$.*

Proof. Let H be a subgroup of G. Then $a, b \in H \Rightarrow a, b^{-1} \in H \Rightarrow ab^{-1} \in H$. Conversely, suppose H is a non-empty subset of G such that $a, b \in H \Rightarrow ab^{-1} \in H$. Since $H \neq \emptyset$, there exists $a \in H$. Hence $a, a^{-1} \in H$. Therefore, $e = aa^{-1} \in H$, i.e., H contains the identity element e. Also, since $a, b \in H$. $ea^{-1} \in H$. Hence $a^{-1} \in H$. Now, let $a, b \in H$. Then $a, b^{-1} \in H$. Hence $a(b^{-1})^{-1} = ab \in H$ and so H is closed under the binary operation in G. Hence H is a subgroup of G. □

If the operation is $+$ then H is a subgroup of G if and only if $a, b \in H \Rightarrow a - b \in H$.

Theorem 1.4.8. *Let H be a non-empty finite subset subset of G. If H is closed under the operation in G then H is a subgroup of G.*

Proof. Let $a \in H$. Then $a, a^2, \ldots, a^n, \ldots$ are all elements of H. But since H is finite the elements $a, a^2, a^3 \ldots,$ cannot all be distinct. Hence let $a^r = a^s, r < s$. Then $a^{s-r} = e \in H$. Now, let $a \in H$. We have proved that $a^n = e$ for some n. Hence $aa^{n-1} = e$. Hence $a^{-1} = a^{n-1} \in H$. Thus H is a subgroup of G. □

Theorem 1.4.8 is not true if H is infinite. For example, $\mathbb{N}$ is an infinite subset of $(\mathbb{Z}, +)$ and $\mathbb{N}$ is closed under addition. However $\mathbb{N}$ is not a subgroup of $(\mathbb{Z}, +)$.

Theorem 1.4.9. *If H and K are subgroups of a group G then $H \cap K$ is also a subgroup of G.*

Proof. Clearly $e \in H \cap K$ and so $H \cap K$ is non-empty. Now let $a, b \in H \cap K$. Then $a, b \in H$ and $a, b \in K$. Since H and K are subgroups of G, $ab^{-1} \in H$ and $ab^{-1} \in K$. Therefore $ab^{-1} \in H \cap K$. Hence by Theorem 1.4.8, $H \cap K$ is a subgroup of G. □

It can be similarly proved that the intersection of any number of subgroups of G is again a subgroup of G.

The union of two subgroups of a group need not be a subgroup. For example, $2\mathbb{Z}$ and $3\mathbb{Z}$ are subgroups of $(\mathbb{Z}, +)$ but $2\mathbb{Z} \cup 3\mathbb{Z}$ is not a subgroup of $\mathbb{Z}$ since $3, 2 \in 2\mathbb{Z} \cup 3\mathbb{Z}$ but $3 + 3 = 5 \notin 2\mathbb{Z} \cup 3\mathbb{Z}$.

Theorem 1.4.10. *The union of two subgroups of a group G is a subgroup if and only if one is contained in the other.*

Proof. Let H and K be two subgroups of G such that one is contained in the other. Then either $H \subseteq K$ or $K \subseteq H$. Therefore $H \cup K = K$ or $H \cup K = H$. Hence $H \cup K$ is a subgroup of G.

Conversely, suppose H is not contained in K and K is not contained in H. Then there exist elements a, b such that $a \in H$, $a \notin K$, $b \in K$, and $b \notin H$.

Clearly $a, b \in H \cup K$. Since $H \cup K$ is a subgroup of G $ab \in H \cup K$. Hence $ab \in H$ or $ab \in K$. If $ab \in H$, then $a^{-1} \in H$ since $a \in H$. Hence $a^{-1}(ab) = b \in H$, a contradiction. If $ab \in K$, $b^{-1} \in K$ since $b \in K$. Hence $(ab)b^{-1} = a \in K$, a contradiction. Hence our assumption that H is not contained in K and K is not contained in H is false. Therefore $H \subseteq K$ or $K \subseteq H$. □

1.5 Cosets

Definition 1.5.1. Let H be a subgroup of a group G and $a \in G$. The sets $aH = \{ah : h \in H\}$ and $Ha = \{ha : h \in H\}$ are called the *left* and *right cosets* of H in G, respectively. The element a is called a representative of aH and Ha.

Examples 1.5.2.

1. Let us determine the left cosets of $(5\mathbb{Z}, +)$ in $(\mathbb{Z}, +)$. Here the operation is $+$.

$0 + 5\mathbb{Z} = 5\mathbb{Z}$ is itself a left coset. Another left coset is $1 + 5\mathbb{Z} = \{1 + 5n : n \in \mathbb{Z}\}$. We notice that this left coset contains all integers having remainder 1 when divided by 5. Similarly $2 + 5\mathbb{Z} = \{2 + 5n : n \in \mathbb{Z}\}$, $3 + 5\mathbb{Z} = \{3 + 5n : n \in \mathbb{Z}\}$ and $4 + 5\mathbb{Z} = \{4 + 5n : n \in \mathbb{Z}\}$.

These are all the left cosets of $(5\mathbb{Z}, +)$ in $\mathbb{Z}$. Here also we note that all the left cosets are mutually disjoint, and their union is $\mathbb{Z}$. In other words the collection of all left cosets forms a partition of the group.

2. Consider $(\mathbb{Z}_{12}, \oplus)$. Then $H = \{0, 4, 8\}$ is a subgroup of G. The left cosets of H are given by $0 + H = \{0, 4, 8\} = H$, $1 + H = \{1, 5, 9\}$, $2 + H = \{2, 6, 10\}$, and $3 + H = \{3, 7, 11\}$. We notice that $4 + H = \{4, 8, 0\} = H$, and $5 + H = \{5, 9, 1\}$ etc.

Theorem 1.5.3. *Let G be a group and H be a subgroup of G. Then*
(i) $a \in H \Rightarrow aH = H$.
(ii) $aH = bH \Rightarrow a^{-1}b \in H$. (iii) $a \in bH \Rightarrow a^{-1} \in Hb^{-1}$.
(iv) $a \in bH \Rightarrow aH = bH$.

Proof. (i) Let $a \in H$. We claim that $aH = H$. Let $x \in aH$. Then $x = ah$ for some $h \in H$. Now, $a \in H$ and $h \in H \Rightarrow ah = x \in H$(since H is a subgroup). Hence $aH \subseteq H$. Let $x \in H$. Then $x = a(a^{-1}x) \in aH$. Hence $H \subseteq aH$. Thus $H = aH$. Conversely, let $aH = H$. Now $a = ae \in aH$ and $a \in H$.

(ii) Let $aH = bH$. Then $a^{-1}(aH) = a^{-1}(bH)$ and $H = (a^{-1}b)H$. Hence $a^{-1}b \in H$(by (i)).

Conversely let $a^{-1}b \in H$. Then $a^{-1}bH = H$(by (i)), $aa^{-1}bH = aH$ and so $bH = aH$.

(iii) Let $a \in bH$. Then $a = bH$ for some $h \in H$ and so $a^{-1} = (bH)^{-1} = h^{-1}b^{-1} \in Hb^{-1}$. Converse can be similar proved.

(iv) Let $a \in bH$. We claim that $aH = bH$. Let $x \in aH$. Then $x = ah_1$ for some $h_1 \in H$. Also $a \in bH \Rightarrow a = bh_2$ for some $h_2 \in H$. Therefore $x = ((bh_2)h_1) = b(h_2h_1) \in bH$ and so $aH \subseteq bH$. Now, let $x \in bH$. Then $x = bh_3$ for some $h_3 \in H$ and so $b = ah_2^{-1}$. Therefore $x = ah_2^{-1}h_3 \in aH$ and so $bH \subseteq aH$. Hence $aH = bH$.

Conversely, let $aH = bH$. Then $a = ae \in aH$ and so $a \in bH$. □

Theorem 1.5.4. *Let H be a subgroup of G. Then*

(i) any two left cosets of H are either identical or disjoint.

(ii) union of all the left cosets of H is G.

(iii) the number of elements in any left coset aH is the same as the number of elements in H.

Proof. (i) Let aH and bH be two left cosets. Suppose aH and bH are not disjoint. We claim that $aH = bH$. Since aH and bH are not disjoint, $aH \cup bH \neq \emptyset$ and so there exists an element $c \in aH \cup bH$. Clearly $c \in aH$, $c \in bH$ and so $aH = cH$, $bH = cH$. Hence $aH = bH$.

(ii) Let $a \in G$. Then $a = ae \in aH$ and every element of G belongs to a left cosets of H.Thus the union of all the left cosets of H is G.

(iv) The map $f : H \to aH$ defined by $f(h) = ah$ is clearly a bijection. Hence every left coset has the same number of elements as H. □

This theorem shows that the collection of all left cosets forms a partition of the group. The above result is true if we replace left cosets by right cosets. In what follows, the result we prove for left cosets are also true for right cosets.

Remark 1.5.5. Let H be a subgroup of G. We define a relation in G as follows. Define $a \sim b \Leftrightarrow a^{-1}b \in H$. Then $\sim$ is an equivalence relation.

For, $a^{-1}a = e \in H$, $a \sim a$ and hence $\sim$ is reflexive.

Now , $a \sim b \Rightarrow a^{-1}b \in H \Rightarrow (a^{-1}b)^{-1} \in H \Rightarrow b^{-1}a \in H \Rightarrow b \sim a$.

Therefore $a \sim b \Rightarrow b \sim a$ and $\sim$ is symmetric.

Now, $a \sim b$ and $b \sim c \Rightarrow a^{-1}b \in H$ and $b^{-1}c \in H \Rightarrow (a^{-1}b)(b^{-1}c) \in H \Rightarrow a^{-1}c \in H \Rightarrow a \sim c$. Hence $\sim$ is transitive and so $\sim$ is an equivalence relation.

Now, we claim that equivalence class $[a] = aH$. Let $b \in [a]$. Then $b \sim a$.

$\therefore \quad a^{-1}b \in H$.

$\therefore \quad a^{-1}b = h$ for some $h \in H$.

$\therefore \quad b = ah$ Hence $b \in aH$.

$\therefore \quad [a] \subseteq aH$.

Also, $b \in aH \Rightarrow b = ah$ for some $h \in H$.

$$\Rightarrow a^{-1}b = h \in H \Rightarrow a \sim b \Rightarrow b \in [a].$$

Thus the left cosets of H in G are precisely the equivalence classes determined by $\sim$. Hence the left cosets form a partition of G.

Theorem 1.5.6. *Let H be a subgroup of G. The number of left cosets of H is the same as the number of right cosets of H.*

Proof. Let L and R respectively denote the set of left and right cosets of H. We define a map $f : L \to R$ by $f(aH) = Ha^{-1}$. f is well defined. For $aH = bH \Rightarrow a^{-1}b \in H \Rightarrow a^{-1} \in Hb^{-1} \Rightarrow Ha^{-1} = Hb^{-1}$ f is 1-1. For, $f(aH) = f(bH) \Rightarrow Ha^{-1} = Hb^{-1} \Rightarrow a^{-1} \in Hb^{-1} \Rightarrow a^{-1} = hb^{-1}$ for some $h \in H \Rightarrow a = bh^{-1} \Rightarrow a \in bH \Rightarrow aH = bH$. f is onto. For, every right coset Ha has a pre-image under f namely $a^{-1}H$. Hence f is a bijection from L to R. Hence the number of left cosets is the same as the number of right cosets. □

Definition 1.5.7. Let H be a subgroup of G. The number of distinct left (right) cosets of H in G is called the *index* of H in G and is denoted by $[G : H]$.

Example 1.5.8. In $(\mathbb{Z}_8, \oplus)$, $H = \{0, 4\}$ is a subgroup. The left cosets of H are given by

$$0 + H = \{0, 4\} = H$$

$$1 + H = \{1, 5\}$$

$$2 + H = \{2, 6\}$$

$$3 + H = \{3, 7\}$$

These are the four distinct left cosets of H. Hence the index of the subgroup H is 4. Note that $[\mathbb{Z}_8 : H] \times [H] = 4 \times 2 = 8 = |\mathbb{Z}_8|$.

Theorem 1.5.9 (Lagrange's theorem). *Let G be a finite group of order n and H be a subgroup of G. Then the order of H divides the order of G.*

Proof. Let $|H| = m$ and $[G : H] = r$. Then the number of distinct left cosets of H in G is r. By Theorem 1.5.6, these r left cosets are mutually disjoint, they have the same number of elements namely m and their union is G. Therefore $n = rm$. Hence m divides n. □

1.6 A counting principle

Definition 1.6.1. Let A and B be two subsets of a group G. We define

$$AB = \{ab : \ a \in A, b \in B\}.$$

If H and K are two subgroups of G, then HK need not be a subgroup of G.

For example, consider $G = S_3$. $H = \{e, p_3\}$ and $K = \{e, p_4\}$. Then H and K are subgroups of S_3. Also $HK = \{ee, ep_4, ep_3, p_3p_4\} = \{e, p_4, p_3, p_2\}$. Now, $p_4p_2 = p_5 \notin HK$. Hence HK is not a subgroup of S_3.

Theorem 1.6.2. *Let H and K be subgroups of a group G. Then HK is a subgroup of G if and only if $HK = KH$.*

Proof. Suppose HK is a subgroup of G. Let $kh \in KH$, where $h \in H$ and $k \in K$. Now $h = he \in HK$ and $k = ek \in HK$. Because HK is a subgroup, it follows that $kh \in HK$. Hence, $KH \subseteq HK$. On the other hand, let $hk \in HK$. Then $(hk)^{-1} \in HK$, so $(hk)^{-1} = h_1k_1$ for some $h_1 \in H$ and $k_1 \in K$. Thus, $hk = (h_1k_1)^{-1} = k_1^{-1}h_1^{-1} \in KH$. This implies that $HK \subseteq KH$. Hence, $HK = KH$.

Conversely, suppose $HK = KH$. Let $h_1k_1, h_2k_2 \in HK$, where $h_1, h_2 \in H$ and $k_1, k_2 \in K$. We show that $(h_1k_1)(h_2k_2)^{-1} \in HK$. Now $k_2 \in K$ and $h_2 \in H$. Therefore,

$k_2^{-1}h_2^{-1} \in KH = HK$. This implies that $k_2^{-1}h_2^{-1} = h_3k_3$ for some $h_3 \in H$ and $k_3 \in K$. Similarly, $k_1h_3 \in KH = HK$, so $k_1h_3 = h_4k_4$ for some $h_4 \in H$ and $k_4 \in K$. Thus,

$$\begin{aligned}
(h_1k_1)(h_2k_2)^{-1} &= h_1k_1k_2^{-1}h_2^{-1}(\text{because } (h_2k_2)^{-1} = k_2^{-1}h_2^{-1}) \\
&= h_1k_1h_3k_3(\text{substitute } k_2^{-1}h_2^{-1} = h_3k_3) \\
&= h_1h_4k_4k_3 \in HK(\text{substitute } k_1h_3 = h_4k_4)
\end{aligned}$$

Hence, HK is a subgroup of G. □

Corollary 1.6.3. *If H and K are subgroups of an abelian group G, then HK is a subgroup of G.*

Proof. Let $x \in HK$. Then $x = ab$ where $a \in H$ and $b \in K$. Since G is abelian, $ab = ba$ and so $x \in KH$. Hence $HK \subseteq KH$. Similarly $KH \subseteq HK$ and $HK = KH$. Hence HK is a subgroup of G. □

Theorem 1.6.4. *Let H and K be finite subgroups of a group G. Then*

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

Proof. Let us write $A = H \cap K$. Since H and K are subgroups of G, A is a subgroup of G and since $A \subseteq H$, A is also a subgroup of H. By Lagranges theorem,$|A|$ divides $|H|$. Let $n = \frac{|H|}{|A|}$. Then $[H : A] = n$ and so A has n distinct left cosets in H. Let $\{x_1A, x_2A, \ldots, x_nA\}$ be the set of all distinct left cosets of A in H. Then $H = \cup_{i=1}^n x_iA$. Since $A \subseteq K$, it follows that

$$HK = (\cup_{i=1}^n x_iA)K = \cup_{i=1}^n x_iK.$$

We now show that $x_iK \cap x_jK = \Phi$ if $i \neq j$. Suppose $x_iK \cap x_jK \neq \Phi$ for some $i \neq j$. Then $x_jK = x_iK$. Thus, $x_i^{-1}x_j \in K$. Since $x_i^{-1}x_j \in H$, $x_i^{-1}x_j \in A$ and so $x_jA = x_iA$.

This contradicts the assumption that $x_1A, \dots, x_nA$ are all distinct left cosets. Hence, $x_1K, \dots, x_nK$ are distinct left cosets of K. Also, $|K| = |x_iK|$ by Theorem 1.5.6 for all $i = 1, 2, \cdots, n$. Thus,

$$|HK| = |x_1K| + \cdots + |x_nK| = n|K| = \frac{|H||K|}{|A|} = \frac{|H||K|}{|H \cap K|}. \qquad \square$$

Corollary 1.6.5. *If H and K are subgroups of the finite group $G$and $o(H) > \sqrt{G}$, $o(G) > \sqrt{G}$, then $H \cap K \neq \{e\}$.*

Proof. Since HK is a subset of G, $o(HK) \leq o(G)$. Also $o(HK) = \frac{o(H)o(K)}{o(H \cap K)} > \frac{o(G)}{o(H \cap K)}$. This implies that $o(H \cap K) > 1$. $\square$

Corollary 1.6.6. *Suppose G is a finite group of order pq where p and q are prime numbers with $p > q$. Then that G can have at most one subgroup of order p.*

Proof. For suppose H, K are subgroups of order p. Clearly $H \cap K$ is a subgroup of G. By the Corollary 1.6.5, $H \cap K \neq (e)$, and by Lagrange's Theorem, $o(H \cap K) = p$ and so $H \cap K = K = H$. Hence there is at most one subgroup of order p. $\square$

Problem 1.6.7. Let H be a subgroup of G and $a \in G$. Then $aHa^{-1} = \{aga^{-1} : g \in H\}$ is a subgroup of G.

Solution. Clearly $e = aea^{-1} \in aHa^{-1}$ and so $aHa^{-1} \neq \emptyset$. Now, let $x, y \in aHa^{-1}$. Then $x = ah_1a^{-1}$ and $y = ah_2a^{-1}$ where $h_1, h_2 \in H$. Now, $xy^{-1} = (ah_1a^{-1})(ah_2a^{-1})^{-1} = (ah_1a^{-1})(ah_2^{-1}a^{-1}) = a(h_1h_2^{-1})a^{-1} \in aHa^{-1}$. Hence aHa^{-1} is a subgroup of G.

1.7 Cylic group

Definition 1.7.1. Let G be a group and $a \in G$. Then $H = \{a^n : n \in \mathbb{Z}\}$ is a subgroup of G.

H is called the cyclic subgroup of G generated by a and is denoted by $\langle a \rangle$.

Examples 1.7.2. 1. In $(\mathbb{Z},+)$, $\langle a\rangle = 2\mathbb{Z}$ which is the group of even integers.

2. In the group $G = (\mathbb{Z}_{12}, \oplus)$, $\langle 3\rangle = \{0,3,6,9\}$, $\langle 5\rangle = \{0,5,10,3,8,1,6,11,4,9,2,7\} = \mathbb{Z}_{12}$.

3. In the group $G = \{1,i,-1,-i\}$, $\langle i\rangle = \{i,i^2,i^3,\cdots\} = \{i,-1,-i,1\} = G$.

Definition 1.7.3. Let G be a group and let $a \in G$, a is called a **generator** of G if $\langle a\rangle = G$.

A group G is *cyclic* if there exists an element $a \in G$ such that $\langle a\rangle = G$.

Note 1.7.4. If G is cyclic group generated by an element a, then every element of G is of the form a^n for some $n \in \mathbb{Z}$.

Examples 1.7.5. 1. $(\mathbb{Z},+)$ is a cyclic group and 1 is the generator of this group. Clearly -1 is also a generator of this group. Thus a cyclic group can have more than one generator.

2. $(n\mathbb{Z},+)$ is a cyclic group and n and $-n$ are generators of this group.

3. $(\mathbb{Z}_8,\oplus)$ is a cyclic group and $1,3,5,7$ are all generators of this group.

4. $(\mathbb{Z}_n,\oplus)$ is a cyclic group for all $n \in \mathbb{N}$; 1 is a generator of this group. In fact if $m \in \mathbb{Z}_n$ and $(m,n) = 1$ then m is a generator of this group.

5. $G = \{1,i,-1,-i\}$ is a cyclic group under usual multiplication; i is a generator, $-i$ is also a generator of G. However -1 is not a generator of G since $\langle -1\rangle = \{1,-1\} \neq G$.

6. $G = \{1,\omega,\omega^2\}$ where $\omega \neq 1$ is a cube root of unity is a cyclic group, ω and ω^2 are both generators of this group.

7. In this group $G = (\mathbb{Z}_7 - \{0\}, \odot)$, 3 and 5 are both generators. Here 2 is not a generator of G since $\langle 2 \rangle = \{2, 4, 1\} \neq G$.

8. Let A be a set containing more than one element. Then $(\varrho(A), \triangle)$ is not cyclic; for let $B \in \varrho(A)$ be any element. Then $B \triangle B = \Phi$ so that $\langle B \rangle = \{B, \Phi\} \neq \varrho(A)$.

9. $(\mathbb{R}, +)$ is not a cyclic group since for any $x \in \mathbb{R}, \langle x \rangle = \{nx : \ n \in Z\} \neq \mathbb{R}$

Theorem 1.7.6. *Any cyclic group is abelian.*

Proof. Let $G = \langle a \rangle$ be a cyclic group. Let $x, y \in G$. Then $x = a^r$ and $y = a^s$ for some $r, s \in \mathbb{Z}$. Hence $xy = a^r a^s = a^{r+s} = a^{s+r} = a^s a^r = yx$. Hence G is abelian. □

Theorem 1.7.7. *A subgroup of cyclic group is cyclic.*

Proof. Let G be a cyclic group generated by a and let H be a subgroup of G. We claim that H is cyclic. Clearly every element of H is of the form a^n for some integer n. Let m be the smallest positive integer such that $a^m \in H$. We claim that a^m is the generator of H. Let $b \in H$. Then $b = a^n$ for some $n \in \mathbb{Z}$. Then $b = a^n = a^{mq+r} = a^{mq}a^r = (a^m)^q a^r$. Therefore $a^r = (a^m)^{-q}b$. Now, $a^m \in H$. Since H is a subgroup, $(a^m)^{-q} \in H$. Also, $b \in H$. Clearly $a^r \in H$ and $0 \leq r < m$. But m is the least positive integer such that $a^n \in H$. Therefore $r = 0$. Hence $b = a^n = a^{qm} = (a^m)^q$. Every element of H is a power of a^m. Thus $H = \langle a^m \rangle$ and so H is cyclic. □

Theorem 1.7.8. *Every group of prime order is cyclic.*

Proof. Let G be a group of order p where p is prime. Let $a \in G$ and $a \neq e$. By above theorem order of a divides p. The order of a is 1 or p. Since $a \neq e$ order of a is p. Hence $G = \langle a \rangle$ so that G is cyclic. □

Theorem 1.7.9. *Let G be a group of order n and $a \in G$. Then $a^n = e$.*

Proof. Let the order of a is m. Then m divides n and so $n = mq$. Thus, $a^n = a^{mq} = (a^m)^q = e^q = e$. $\square$

Definition 1.7.10. Let G be a group and let $a \in G$. The least positive integer n(if it exists) such that $a^n = e$ is called the **order** of a. If there is no positive integer n such that $a^n = e$, then the order of a is said to be infinite.

Examples 1.7.11.

1. Consider the group S_3, $p_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$, $p_1^2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = p_4$ and $p_1^3 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = e$.
   In this case, 3 is the least positive integer such that $p_1^3 = e$. Thus p_1 is of order 3.

2. Consider $(\mathbb{R}^*, \cdot)$, From this sequence of elements $2, 2^2, 2^3, \ldots, 2^n, \ldots$. In this case there is no positive integer n such that $2^n = 1$ and $\langle 2 \rangle$ contains infinite numbers of elements. Thus the order 2 is infinite.

Theorem 1.7.12. *Let G be a group and $a \in G$. Then the order of a is the same as the order of the cyclic group generated by a.*

Proof. Let a be an element of order n. Then $a^n = e$. We claim that $e, a, a^2, \ldots, a^{n-1}$ are all distinct. Suppose $a^r = a^s$ where $0 < r < s < n$. Then $a^{s-r} = e$ and $s - r < n$ which contradicts the definition of the order of a. Hence $e, a, a^2, \ldots, a^{n-1}$ are n distinct elements and $\langle a \rangle = \{e, a, a^2, \ldots, a^{n-1}\}$ which is of order n.

If a is of infinite order, the sequence of elements $e, a, a^2, \ldots, a^{n-1}, \ldots$ are all distinct and are in $\langle a \rangle$. Hence $\langle a \rangle$ is an infinite group. $\square$

Theorem 1.7.13. Let G be a group and a be an element of order n in G. Then $a^m = e$ if and only if n divides m.

Proof. Suppose $n|m$. Then $m = nq$ where $q \in \mathbb{Z}$ and $a^m = a^{nq} = (a^n)^q = e^q = e$.

Conversely, let $a^m = e$. Let $m = nq + r$ where $0 \leq r < n$. Now $a^m = a^{nq+r} = a^{nq}a^r = ea^r = a^r$. Thus $a^r = e$ and $0 \leq r < n$. Now, since n is the least positive integer such that $a^n = e$, we have $r = 0$. Hence $m = nq$ and so $n|m$. $\square$

1.8 Normal Subgroup

Definition 1.8.1. A subgroup H of G is called a **normal subgroup** of G if $ghg^{-1} \in H$ for all $g \in G$ and $h \in H$.

Example 1.8.2.

1. For any group G, $\{e\}$ and G are normal subgroups.
2. In S_3, the subgroup $\{e, p_1, p_2\}$ is normal.
3. In S_3, the subgroup $\{e, p_3\}$ is not a normal subgroup.

Example 1.8.3. The alternating group A_n is a subgroup of index 2 in S_n and hence is a normal subgroup of S_n.

Lemma 1.8.4. *Every subgroup of an abelian group is a normal subgroup.*

Proof. For any $g \in G$ and $h \in G$, $ghg^{-1} = h \in H$ and hence H is normal subgroup of G $\square$

Examples 1.8.5.

1. $n\mathbb{Z}$ is a normal subgroup of $(\mathbb{Z}, +)$.

2. Every subgroup of $(\mathbb{Z}_n, \oplus)$ is normal.

3. Since any cyclic group is abelian any subgroup of a cyclic is normal.

Lemma 1.8.6. The intersection two normal subgroups of a group G is a normal subgroup.

Proof. Let H and K be two normal subgroups of G. Then $H \cap K$ is a subgroup of G. Now, let $a \in G$ and $x \in H \cap K$. Then $x \in H$ and $x \in K$. Since H and K are normal $axa^{-1} \in H$ and $axa^{-1} \in K$. Hence $axa^{-1} \in H \cap K$. Thus $H \cap K$ is a normal subgroup of G. □

Lemma 1.8.7. The center $Z(G)$ of a group G is a normal subgroup of G.

Proof. Let $Z(G) = \{a : \ a \in G, ax = xa \text{ for all } x \in G\}$. Now let $x \in Z(G)$ and $a \in G$. Then $ax = xa$ and so $x = axa^{-1} \in Z(G)$. Hence $Z(G)$ is a normal subgroup of G. □

Theorem 1.8.8. *Let H be a subgroup of index 2 in a group G. Then H is a normal subgroup of G.*

Proof. If $a \in H$ then $H = aH = Ha$. If $a \notin H$, then aH is a left coset different from H. Hence $H \cap aH = \emptyset$. Further, since index of H in G is 2, $H \cup aH = G$. Hence $aH = G - H$. Similarly $Ha = G - H$ so that $aH = Ha$. Hence H is a normal subgroup of G. □

Theorem 1.8.9. *Let N be a subgroup of G. Then the following are equivalent.*
(ii) $aNa^{-1} = N$ *for all* $a \in G$.
(iii) $aNa^{-1} \subseteq N$ *for all* $a \in G$.
(iv) $ana^{-1} \in N$ *for all* $n \in N$ *and* $a \in G$.

Problem 1.8.10. Let H be a subgroup of G. Let $a \in G$. Then aHa^{-1} is a subgroup of G.

Solution. $e = aea^{-1} \in aHa^{-1}$ and hence $aHa^{-1} \neq \Phi$. Now, let $x, y \in aHa^{-1}$. Then $x = ah_1a^{-1}$ and $y = ah_2a^{-1}$ where $h_1, h_2 \in H$. Now, $xy^{-1} = (ah_1a^{-1})(ah_2a^{-1})^{-1} = (ah_1a^{-1})(ah_2^{-1}a^{-1}) = a(h_1h_2^{-1})a^{-1} \in aHa^{-1}$. $\therefore$ aHa^{-1} is a subgroup of G.

Problem 1.8.11. Show that if a group G has exactly one subgroup H of given order, then H is a normal subgroup of G.

Solution. Let the order of H be m. Let $a \in G$. Then by above problem, aHa^{-1} is also a subgroup of G. We claim that $|H| = |aHa^{-1}| = m$. Now, consider $f : H \to aHa^{-1}$ defined by $f(h) = aha^{-1}$. f is 1-1, for, $f(h_1) = f(h_2) \Rightarrow ah_1a^{-1} = ah_2a^{-1} \Rightarrow h_1 = h_2$. f is onto, for, let $x = aha^{-1} \in aHa^{-1}$. Then $f(h) = x$. Thus f is a bijection. $\therefore$ $|H| = |aHa^{-1}| = m$. But H is the only subgroup of G of order m. $\therefore$ $aHa^{-1} = H$. Hence $aH = Ha$. $\therefore$ H is a normal subgroup of G.

Problem 1.8.12. Show that if H and N are subgroups of a group G and N is normal in G, then $H \cap N$ is normal in H. Show by an example that $H \cap N$ need not be normal in G.

Solution. Let $x \in H \cap N$ and $a \in H$. We claim that $axa^{-1} \in H \cap N$. Now, $x \in N$ and $a \in H \Rightarrow axa^{-1} \in N$ (since N is a normal subgroup). Also $x \in H$ and $a \in H \Rightarrow axa^{-1} \in H$ (since H is a group). Hence $axa^{-1} \in H \cap N$. $\therefore$ $H \cap N$ is a normal subgroup of H.

The following example shows that $H \cap N$ need not be normal in G. Let $G = S_3$. Take $N = G$ and $H = \{e, p_3\}$. Now $H \cap N = H$ which is not normal in G.

Problem 1.8.13. If H is a subgroup of G and N is a normal subgroup of G then HN is a subgroup of G.

Solution. To prove that HN is a subgroup of G, it is enough if we prove that $HN = NH$(theorem 1.9.17).

Let $x \in HN$. Then $x = hn$ where $h \in H$ and $n \in N$. $\therefore \quad x \in hN$. But $hN = Nh$(since N is normal) $\therefore \quad x \in Nh$. Hence $x = n_1h$ where $n_1 \in N$. $\therefore \quad x \in Nh$. Hence $HN \subseteq NH$. Similarly $NH \subseteq HN$. $\therefore \quad HN = NH$. Hence HN is a subgroup of G.

Problem 1.8.14. M and N are normal subgroups of a group G such that $M \cap N = \{e\}$. Show that every element of M commutes with element of N.

Solution. Let $a \in M$ and $b \in N$. We claim that $ab = ba$.

Consider the element $aba^{-1}b^{-1}$. Since $a^{-1} \in M$ and M is normal, $ba^{-1}b^{-1} \in M$. Also, since $b \in M$, so that $aba^{-1}b^{-1} \in N$. Thus $aba^{-1}b^{-1} \in M \cap N = \{e\}$. $\therefore \quad aba^{-1}b^{-1} = e$, so that $ab = ba$.

Theorem 1.8.15. A subgroup N of G is normal if and only if the product of two right cosets of N is again a right coset of N.

Proof. Suppose N is a normal subgroup of G. Then

$$\begin{aligned} NaNb &= N(aN)b = N(Nab) \text{ (since } aN = Na) \\ &= NNab \ = Nab \text{ (since } NN = N). \end{aligned}$$

Conversely suppose that the product of any two right cosets of N is again a right coset of N. Then $NaNb$ is a right coset of N. Further $ab = (ea)(eb) \in NaNb$. Hence $NaNb$ is the right coset containing ab. $\therefore \quad NaNb = Nab$.

Now, we prove that N is a normal subgroup of G. Let $a \in G$ and $n \in N$. Then $ana^{-1} = eana^{-1} \in NaNa^{-1} = Naa^{-1} = N$. $\therefore \quad ana^{-1} \in N$. Hence N is a normal subgroup of G. $\square$

Chapter 2

Unit 1: Counting Principle

2.1 Class equation for finite group

Definition 2.1.1. Let G be a group. If $a, b \in G$, then b is said to be a conjugate of a in G if there exists an element $c \in G$ such that $b = c^{-1}ac$.

We shall write, for this, $a \sim b$ and shall refer to this relation as **conjugacy.**

Lemma 2.1.2. *Conjugacy is an equivalence relation on G.*

Proof. Define a relation $\sim$ on G by $a \sim b$ if a is conjugate to b

Clearly $a = e^{-1}ae$ and so $a \sim a$.

If $a \sim b$, then $b = x^{-1}ax$ for some $x \in G$, hence, $a = (x^{-1})^{-1}b(x^{-1})$ and since $y = x^{-1} \in G$ and $a = y^{-1}by$, and hence $b \sim a$.

Suppose that $a \sim b$ and $b \sim c$ where $a, b, c \in G$. Then $b = x^{-1}ax$, $c = y^{-1}by$ for some $x, y \in G$. Substituting for b in the expression for c we obtain, $c = y^{-1}(x^{-1}ax)y = (xy)^{-1}a(xy)$ and so $a \sim c$. Hence the conjugacy is an equivalence relation on G. □

For $a \in G$, let $C(a) = \{x \in G : a \sim x\}$. Then $C(a)$, the equivalence class of a in G under our relation, is usually called the conjugate class of a in G. From this, these conjugacy classes form a partition of G and hence $G = \bigcup_{a \in G} C(a)$.

Definition 2.1.3. If $a \in G$, then $N(a)$, the normalizer of a in G, is the set $N(a) = \{x \in G : xa = ax\}$.

$N(a)$ consists of precisely those elements in G which commute with a.

Lemma 2.1.4. *Let G be a group and $Z(G) = \{a : \ a \in G$ and $ax = xa$ for all $x \in G\}$. Then $Z(G)$ is a subgroup of G. Here $Z(G)$ is the center of G.*

Proof. Clearly $ex = xe = x$ for all $x \in G$. Hence $e \in Z(G)$, so that $Z(G)$ is non-empty. Now, let $a, b \in Z(G)$. Then $ax = xa$ and $bx = xb$ for all $x \in G$. Now, $bx = xb \Rightarrow b^{-1}(bx)b^{-1} = b^{-1}(xb)b^{-1} \Rightarrow (b^{-1}b)xb^{-1} = b^{-1}x(bb^{-1}) \Rightarrow exb^{-1} = b^{-1}xe \Rightarrow xb^{-1} = b^{-1}x$.

Now $(ab^{-1})x = a(b^{-1}x) = a(xb^{-1}) = (ax)b^{-1} = (xa)b^{-1} = x(ab^{-1})$. Thus ab^{-1} commutes with every element of G and so $ab^{-1} \in Z(G)$. Hence $Z(G)$ is a subgroup of G. $\square$

Lemma 2.1.5. *Let G be a group and $a \in G$. Let $C_G(a) = \{ \ x \in G : ax = xa\}$. Then $C_G(a)$ is a subgroup of G. Here $C_G(a) = N(a)$ is called the normalizer of a in G.*

Proof. Clearly $ea = ae = a$. Hence $e \in N(a)$ so that $N(a)$ is non-empty. Then $ax = xa$ and $ay = ya$. Now, $ay = ya \Rightarrow y^{-1}a = ay^{-1}$. Hence $a(xy^{-1}) = (ax)y^{-1} = (xa)y^{-1} = x(ay^{-1}) = x(y^{-1}a) = (xy^{-1})a$. Hence xy^{-1} commutes with a, $xy^{-1} \in N(a)$ and so $N(a)$ is a subgroup of G. $\square$

Lemma 2.1.6. *Let H be a subgroup of G. Then $N(H) = \{g \in G : gHg^{-1} = H\}$ is a subgroup of G*

Proof. Clearly $aea^{-1} = e \in H$ and so $e \in N(H)$. Hence $N(H)$ is non-empty. Let $x, y \in N(H)$. Then $xHx^{-1} = H$ and $yHy^{-1} = H$. This implies $(xy)H(xy)^{-1} = x(yHy^{-1})x^{-1} = xHx^{-1} = H$. Hence $N(H)$ is a subgroup of G. $\square$

Theorem 2.1.7. *If G is a finite group, then the number of elements conjugate C_a to a in G is the index of the normalizer of a in G.*

Proof. Let $H = N(a)$, where $a \in G$ and $\mathcal{L} = \{gH : g \in G\}$ be the set of all left cosets of $N(a)$ in G. Define $f : \mathcal{L} \to C(a)$ by $f(gH) = gag^{-1}$ for all $gH \in \mathcal{L}$. Let $xH, yH \in \mathcal{L}$. Suppose $xH = yH$. Then $xy^{-1} \in H$ implies $xy^{-1}a = axy^{-1}$. From this, we get $x^{-1}(xy^{-1}ay = x^{-1}axy^{-1}y$ implies $y^{-1}ay = x^{-1}ax$. Thus, $f(xH) = f(yH)$ and so f is well defined. Suppose $f(xH) = f(yH)$. Then $xax^{-1} = yay^{-1}$ implies $y^{-1}xax^{-1}x = y^{-1}yay^{-1}x$. From this, $y^{-1}xa = ay^{-1}x$ and so $y^{-1}x \in H = N(a)$. Thus $xH = yH$, since $y^{-1}x \in H \Leftrightarrow xH = yH$. Hence f is one to one.

For $z \in C(a)$, $z = cac^{-1}$ for some $c \in G$ and by definition of f, we have $z = cac^{-1} = f(cH)$ and f is onto. Hence $C_a = o(\mathcal{L}) = o(G)/o(N(a))$. □

Corollary 2.1.8. *(Class Equation for finite group) Let G be a finite group. Then $o(G) = \sum \frac{o(G)}{o(N(a))}$, where this sum runs over one element a in each conjugate class.*

Proof. By Lemma 2.1.2, for $a \in G$, let $C(a) = \{x \in G : a \sim x\}$. Then $C(a)$, the equivalence class of a in G under our relation, is usually called the conjugate class of a in G. From this, these conjugacy classes form a partition of G and hence $G = \bigcup\limits_{a \in G} C(a)$. By Theorem 2.1.7, $c_a = o(G)/o(N(a))$ and

$$o(G) = \sum o(C(a)) = \sum C_a = \sum o(G)/o(N(a)).$$

□

Lemma 2.1.9. *$a \in Z(G)$ if and only if $N(a) = G$. If G is finite, $a \in Z(G)$ if and only if $o(N(a)) = o(G)$.*

Proof. If $a \in Z(G)$, then $xa = ax$ for all $x \in G$, whence $N(a) = G$ and so $o(N(a)) = o(G)$. □

Corollary 2.1.10. *(Class Equation for finite group) Let G be a finite group. Then*

$$o(G) = o(Z(G)) + \sum_{a \notin Z(G)} \frac{o(G)}{o(N(a))},$$

where this sum runs over one element a in each conjugate class.

Proof. If $a \in Z(G)$, then $ax = xa$ for all $x \in G$, $C(a) = \{gag^{-1} : g \in G\} = \{a\}$ and hence $C_a = 1$. By Class equation,

$$o(G) = \sum_{a \in Z(G)} \frac{o(G)}{o(N(a))} + \sum_{a \notin Z(G)} \frac{o(G)}{o(N(a))} = o(Z(G)) + \sum_{a \notin Z(G)} \frac{o(G)}{o(N(a))}$$

□

Consider the group $S_3 = \{e, (1,2), (1,3), (2,3), (1,2,3), (1,3,2)\}$. We enumerate the conjugate classes: $C(e) = \{e\}$, $C(1,2) = \{g^{-1}(1,2)g : g \in S_3\} = \{(1,2), (1,3), (2,3)\}$ and $C(1,2,3) = \{(1,2,3), (1,3,2)\}$

Hence the class equation for S_3 is $C_e + C_{(1,2)} + C_{(1,2,3)} = 1 + 2 + 3$

Theorem 2.1.11. *If $o(G) = p^n$ where p is a prime number, then $Z(G) \neq (e)$.*

Proof. Since $N(a)$ is a subgroup of G, $o(N(a))$ divides $o(G) = p^n$ and so $o(N(a)) = p^{n_a}$. Also $a \in Z(G)$ if and only if $n_a = n$. Let $m = o(Z(G))$. Then by Corollary 2.1.10, $p^n = o(G) = m + \sum_{a \notin Z(G)} (p^n/p^{n_a})$. If $a \notin Z(G)$, then $n_a < n$, p divides $p^n - p^{n_a}$ and so p divides $\sum_{a \notin Z(G)} p^{n-n_a}$. Hence p divides $p^n - \sum_{a \notin Z(G)} p^{n-n_a} = m$ and so $Z(G) \neq \{e\}$. □

Corollary 2.1.12. *If $o(G) = p^2$ where p is a prime number, then G is abelian.*

Proof. Our aim is to show that $Z(G) = G$. By Theorem 2.1.11, $Z(G) \neq (e)$ is a subgroup of G so that $o(Z(G)) = p$ or p^2. Suppose that $o(Z(G)) = p$; let $a \in G$, $a \notin Z(G)$. Thus $Z(G) \subset N(a)$. Since $a \in N(a)$ and by Lagrange's Theorem, $o(N(a)) > p$, $o(N(a)) = p^2$ and so $a \in Z(G)$, a contradiction. □

Theorem 2.1.13. *(Cauchy's Theorem for abelian group) If G is a finite abelian group, p is a prime number and $p|o(G)$, then G has an element of order p.*

Theorem 2.1.14. *(Cauchy's Theorem) If G is any finite group, p is a prime number and $p|o(G)$, then G has an element of order p.*

Proof. To prove its existence we proceed by induction on $o(G)$. If $o(G) = 2$, then $G = \mathbb{Z}_2$ and so $o(1) = 2$. If $o(G) = \mathbb{Z}_3$, then $o(1) = o(2) = 3$. We assume the theorem to be true for all groups T such that $o(T) < o(G)$.

Let W be a proper subgroup of G. Then $o(W) < o(G)$ If p divides $o(W)$, then by our induction hypothesis, there exist $a \in W$ such that $a^p = e$ and $a \neq e$.

Suppose p doesnot divide $o(W)$ for any proper subgroups W of G. If $a \notin Z(G)$, then $N(a)$ is a proper subgroup of G, p doesnot divide $o(N(a))$ and so p divides $o(G)/o(N(a))$. From this, we get p divides $\sum\limits_{a \notin Z(G)} \frac{o(G)}{o(N(a))}$ so p divides $o(G) - \sum\limits_{a \notin Z(G)} \frac{o(G)}{o(N(a))}$. Hence p divides $o(Z(G))$. Since $Z(G)$ is abelian and by Cauchy's theorem for abelian group 2.1.13, there exist an element $x \in Z(G)$ such that $x^p = e$. $\square$

We conclude this section with a consideration of the conjugacy relation in a specific class of groups, namely, the symmetric groups S_n.

Given the integer n we say the sequence of positive integers $n_1, n_2, \ldots, n_r$ constitute a partition of n if n $= n_1 + n_2 + \cdots + n_r$. Let $p(n)$ denote the number of partitions of n. Let us determine $p(n)$ for small values of n:

$p(1) = 1$ since $1 = 1$ is the only partition of 1,

$p(2) = 2$ since $2 = 2$ and $2 = 1 + 1$,

$p(3) = 3$ since $3 = 3$, $3 = 1 + 2$, $3 = 1 + 1 + 1$,

$p(4) = 5$ since $4 = 4, 4 = 1 + 3, 4 = 1 + 1 + 2, 4 = 1 + 1 + 1 + 1, 4 = 2 + 2$

Some others are $p(5) = 7$, $p(6) = 11$, $p(61) = 1,121,505$. There is a large mathematical literature on $p(n)$.

Lemma 2.1.15. *The number of conjugate classes in S_n is $p(n)$, the number of partitions of n.*

2.2 Sylow's Theorems

Before entering the first proof of the theorem we digress slightly to a brief number-theoretic and combinatorial discussion. The number of ways of picking a subset of k elements from a set of n elements can easily be shown to be

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

If $n = p^{\alpha}m$ where p is a prime number and $(p, m) = 1$, and if $p^{\alpha}|n$ but $p^{\alpha+1} \nmid n$, consider

$$\begin{aligned}\binom{p^{\alpha}m}{p^{\alpha}} &= \frac{(p^{\alpha}m)!}{(p^{\alpha})!(p^{\alpha}m - p^{\alpha})!} \\ =& \frac{p^{\alpha}m(p^{\alpha}m-1)\cdots(p^{\alpha}m-i)\cdots(p^{\alpha}m-p^{\alpha}+1)}{p^{\alpha}(p^{\alpha}-1)\cdots(p^{\alpha}-i)\cdots(p^{\alpha}-p^{\alpha}+1)}.\end{aligned}$$

Theorem 2.2.1. *(Sylow's Theorem) If p is a prime number and $p^{\alpha}|o(G)$ where p is prime and α is non-negative integer, then G has a subgroup of order p^{α}.*

Proof. We prove, by induction on the order of the group G, that for every prime p dividing the order of G, G has a p-Sylow subgroup. If $o(G) = 2$, then $G = \mathbb{Z}_2$, then the group certainly has a subgroup of order 2, namely itself. So we suppose the result to be correct for all groups of order less than $o(G)$.

From this we want to show that the result is valid for G. Suppose, then, that $p^m|o(G)$, $p^{m+1}|o(G)$, where p is a prime, $m \geq 1$. If $p^m|o(H)$ for any proper subgroup H of G, then $o(H) < o(G)$ and by the induction hypothesis, H has a subgroup T of order p^m. However, since T is a subgroup of H, H is a subgroup of G, T is a subgroup of G.

We may assume that p^m doesnot divide $o(H)$ for any proper subgroup H of G. We restrict our attention to a limited set of such subgroups. If $a \notin Z(G)$, then $N(a) \neq G$ and so p^m does not divide $o(N(a))$, but p^m divides $o(G)/o(N(a))$. Thus, p^m divides $\sum_{a \notin Z(G)} o(G)/o(N(a))$. Since p divides $o(G)$, p divides $o(G) - \sum_{a \notin Z(G)} o(G)/o(N(a))$ and so p divides $o(Z(G))$. By Cauchy's Theorem, there exist an element $b \neq e$ in $Z(G)$ such that $b^p = e$.

Let $B = \langle b \rangle$, the subgroup of G generated by b. Then $o(B) = p$. Since $b \in Z(G)$, B is normal in G. Hence G/B is a group and $o(G/B) < o(G)$ and p^{m-1} divides $o(G)$. By the induction hypothesis, G/B has a subgroup P/B of order p^{m-1}, where P is a subgroup of G. Thus $p^{m-1} = o(P/B) = o(P)/o(B) = o(B)/p$ and so $o(P) = p^m$. □

In view of Sylow's Theorem, we have the following.

Corollary 2.2.2. *If $p^m | o(G)$, $p^{m+1} \nmid o(G)$, then G has a subgroup (p-Sylow subgroup) of order p^m.*

Let $n(k)$ be defined by $p^{n(k)} | (p^k)!$ but $p^{n(k)+1}$ does not divide $(p^k)!$.

Lemma 2.2.3. *$n(k) = 1 + p + \cdots + p^{k-1}$.*

Proof. If $k = 1$ then $p! = 1.2...(p-1).p$, it is clear that $p|p!$ but $p^2 \nmid p!$. Hence $n(1) = 1$. Clearly, only the multiples of p; that is, $p, 2p, \ldots, p^{k-1}p$. In other words $n(k)$ must be the power of p which divides $(2p)(3p)\cdots(p^{k-1}p) = p^{p^{k-1}}(p^{k-1})!$. But then $n(k) = p^{k-1} + n(k-1)$.

Similarly, $n(k-1) = n(k-2) + p^{k-2}$, and so on. Write these out as $n(k) - n(k-1) = p^{k-1}$, $n(k-1) - n(k-2) = p^{k-2}, \ldots, n(2) - n(1) = p$, $n(1) = 1$. Adding these up, with the cross-cancellation that we get, we obtain $n(k) = 1 + p + p^2 + \cdots + p^{k-1}$. □

We are now ready to show that S_{p^k} has a p-Sylow subgroup; that is, we shall show a subgroup of order $p^{n(k)}$ in S_{p^k}.

Lemma 2.2.4. *Let p be a prime number. Then S_{p^k} has a p-Sylow subgroup.*

Proof. We go by induction on k. If $k = 1$, then the element $(1\ 2\ \dots\ p)$, in S_p, is of order p, so generated a subgroup of order p. Since $n(1) = 1$, the result certainly checks out for $k = 1$.

Suppose that the result is correct for $k-1$; we want then must follow for k. Divide the integers $1, 2, \dots, p^k$ into p clumps each with p^{k-1} elements as follows: $\{1, 2, \dots, p^{k-1}\}$, $\{p^{k-1}+1, p^{k-1}+2, \dots, 2p^{k-1}\}, \dots, \{(p-1)p^{k-1}+1, \dots, p^k\}$.

The permutation σ defined by $\sigma = (1, p^{k-1}+1, 2p^{k-1}+1, \dots, (p-1)p^{k-1}+1) \cdots (j, p^{k-1}+j, 2p^{k-1}+j, \dots, (p-1)p^{k-1}+1+j) \cdots, (p^{k-1}, 2p^{k-1}, \cdots, (p-1)p^{k-1}, p^k)$ has the following properties: $\sigma^p = e$ and If τ is a permutation that leaves all i fixed for $i > p^{k-1}$(hence, affects only $1, 2, \dots, p^{k-1}$), then $\sigma^{-1}\tau\sigma$ moves only elements in $\{p^{k-1}+1, p^{k-1}+2, \dots, 2p^{k-1}\}$, and more generally, $\sigma(-j)\tau\sigma^j$ moves only elements in $\{jp^{k-1}+1, jp^{k-1}+2, \dots, (j+1)p^{k-1}\}$.

Consider $A = \{\tau \in S_{p^k} : \tau(i) = i \text{ if } i > p^{k-1}\}$. Then A is a subgroup of S_{p^k} and elements in a can carry out any permutation on $1, 2, \dots, p^{k-1}$. From this it follows easily that $A \cong S_{p^{k-1}}$. By induction hypothesis, A has a subgroup P_1 of order $p^{n(k-1)}$.

Let $T = P_1(\sigma^{-1}P_1\sigma)(\sigma^{-2}P_1\sigma^2)\cdots(\sigma^{-(p-1)}P_1\sigma^{p-1})$ where $P_i = \sigma^{-i}P_1\sigma^i$. Each P_i is isomorphic to P_1 so has order $p^{n(k-1)}$. Also elements in distinct $P_i{}^s$ influence non overlapping sets of integers, hence commute. Thus T is a subgroup of S_{p^k}. Since $P_i \cap P_j = (e)$ if $0 \le i \ne j \le p-1$, $o(T) = o(P_1)^p = p^{pn(k-1)}$.

Since $\sigma^p = e$ and $\sigma^{-i}P_1\sigma^i = P_i$, we have $\sigma^{-1}T\sigma = T$. Let $P = \{\sigma^j t : t \in T, 0 \le j \le p-1\}$. Since $\sigma \notin T$ and $\sigma^{-1}T\sigma = T$, T is a subgroup of S_{p^k} and $o(P) = po(T) = p\, p^{n(k-1)p} = p^{n(k-1)p+1}$. It is $p^{n(k-1)p+1}$. But $n(k-1) = 1+p+\cdots+p^{k-2}$, hence $pn(k-1)+1 = 1+p+\cdots+p^{k-1} = n(k)$. Since $o(P) = p^{n(k)}$, P is a p-Sylow subgroup of S_{p^k}. $\square$

Definition 2.2.5. Let G be a group, A, B subgroups of G. If $x, y \in G$ define $x \sim y$ if $y = axb$ for some $a \in A, b \in B$.

Lemma 2.2.6. *The relation defined above is an equivalence relation on G.*

Proof. Let $x, y \in G$. Then $x = exe$, since $e \in A \cap B$. Hence $x \sim x$. Suppose $x \sim y$. Then $y = axb$ for some $a \in A$ and $b \in B$. This implies $x = a^{-1}yb^{-1}$ and by definition, $y \sim x$. $\square$

For $x \in G$, the equivalence class of $x \in G$ is the set $AxB = \{axb | a \in A, b \in B\}$. These equivalence classes form a partition of G and so $G = \bigcup_{x \in G} AxB$. We call the set AxB a double coset of A, B in G.

Lemma 2.2.7. *If A, B are finite subgroups of G, then*

$$o(AxB) = \frac{o(A)o(B)}{o(A \cap xBx^{-1})}.$$

Proof. Define $T : AxB \to AxBx^{-1}$ given by $T(axb) = axbx^{-1}$ for all $axb \in AxB$. Let $axb, cxd \in AxB$. Suppose $T(axb) = T(cxd)$. Then $axbx^{-1} = cxdx^{-1}$ and by cancellation law, we have $axb = cxd$ and hence T is one-to-one. For any $y \in AxBx^{-1}$, $y = axbx^{-1} = T(axb)$ and hence T is onto. From this, we get $o(AxB) = o(AxBx^{-1})$. Since xBx^{-1} is a subgroup of G, of order $o(B)$, $o(AxB) = o(AxBx^{-1}) = \frac{o(A)\ o(xBx^{-1})}{o(A \cap xBx^{-1})} = \frac{o(A)\ o(B)}{o(A \cap xBx^{-1})}$. $\square$

Lemma 2.2.8. *Let G be a finite group and suppose that G is a subgroup of the finite group M. Suppose further that M has a p-Sylow subgroup Q. Then G has a p-Sylow subgroup P. In fact, $P = G \cap xQx^{-1}$ for some $x \in M$.*

Theorem 2.2.9. *(Second Part of Sylow's Theorem) If G is a finite group, p a prime and $p^n | o(G)$ but $p^{n+1} \nmid o(G)$, then any two subgroups of G of order p^n are conjugate.*

Proof. Let A and B be subgroups of G, each of order p^n. We want to show that $A = gBg^{-1}$ for some $g \in G$. Decompose G into double cosets of A and B; $G = \bigcup_{x \in G} AxB$.

Now, by Lemma 2.2.7,

$$o(AxB) = \frac{o(A)o(B)}{o(A \cap xBx^{-1})}.$$

If $A \neq xBx^{-1}$ for every $x \in G$, then $o(A \cap xBx^{-1}) = p^m$ where $m < n$. Thus

$$o(AxB) = \frac{o(A)o(B)}{p^m} = \frac{p^{2n}}{p^m} = p^{2n-m}$$

and $2n - m \geq n + 1$. Since $p^{n+1}|o(AxB)$ for every x and $o(G) = \sum_{x \in G} o(AxB)$, we would get the contradiction $p^{n+1}|o(G)$. Thus $A = gBg^{-1}$ for some $g \in G$. From this, we conclude that, for a given prime p, any two p-Sylow subgroups of G are conjugate. □

Lemma 2.2.10. *The number of p-Sylow subgroups in G equals $o(G)/o(N(P))$, where P is any p-Sylow subgroup of G. In particular, this number is a divisor of $o(G)$.*

Proof. Let P be a p-Sylow subgroup of G. Then $N(P) = \{g \in G : gPg^{-1} = P\}$ is a subgroup of G and by Theorem 2.1.7, we get the required result. □

Theorem 2.2.11. *(Third Part of Sylow's Theorem) Let G be a finite group and $p|o(G)$, where p is prime. Then the number of p-Sylow subgroups in G is of the form $1 + kp$.*

Proof. Let P be a p-Sylow subgroup of G. We decompose G into double cosets of P and P. Thus $G = \bigcup_{x \in G} PxP$. By Theorem 2.2.7,

$$o(PxP) = \frac{o(P)^2}{o(P \cap xPx^{-1})}.$$

Thus, if $P \cap xPx^{-1} \neq P$, then $p^{n+1}|o(PxP)$, where $p^n = o(P)$. If $x \notin N(P)$, then $p^{n+1}|o(PxP)$. Also, if $x \in N(P)$, then $PxP = P(Px) = P^2x = Px$, so $o(PxP) = p^n$ in this case.

Now

$$o(G) = \sum_{x \in N(P)} o(PxP) + \sum_{x \notin N(P)} o(PxP),$$

where each sum runs over one element from each double coset. However, if $x \in N(P)$, since $PxP = Px$, the first sum is merely $\sum_{x \in N(P)} o(Px)$ over the distinct cosets of P in $N(P)$. Thus this first sum is just $o(N(P))$. We saw that each of its constituent terms is divisible by p^{n+1}, hence

$$p^{n+1} | \sum_{x \notin N(P)} o(PxP).$$

We can thus write this second sum as

$$\sum_{x \notin N(P)} o(PxP) = p^{n+1}u.$$

Therefore $o(G) = o(N(P)) + p^{n+1}u$, so

$$\frac{o(G)}{o(N(P))} = 1 + \frac{p^{n+1}u}{o(N(P))}.$$

Now $o(N(P))|o(G)$ since $N(P)$ is a subgroup of G, hence $p^{n+1}u|o(N(P))$ is an integer. Also, since $p^{n+1} \nmid$ o(G), p^{n+1} can't divide $o(N(P))$. But then $p^{n+1}u|o(N(P))$ must be divisible by p, so we can write $p^{n+1}u|o(N(P))$ as kp, where k is an integer. Hence, the number of p-Sylow subgroups of G is

$$\frac{o(G)}{o(N(P))} = 1 + kp.$$

and by Lagrange's Theorem, $1 + kp$ divides $o(G)$. $\square$

Problem 2.2.12. *Let G be a group of order pqr, where $p < q < r$ are primes. Then some Sylow subgroup of G is normal.*

Proof. Suppose that no Sylow subgroup of G is normal. Then the number of p-Sylow subgroup of G is $1+kp$ and $1+kp \neq 1$ divides qr. Since q and r are distinct, $1+kp = q$, $1+kp = r$ or $1+kp = qr$. From this, we get G has at least $q(p-1)$ elements of order $q(p-1)$ elements of order p.

Also the number of q-Sylow subgroups of G is $1+kq = p$, $1+kq = r$ or $1+kq = pr$ and so G has at least $r(q-1)$ elements of q. Simillarly, G has at least $pq(r-1)$ elements of order r. Therefore, $o(G) \geq q(p-1)+r(q-1)+pq(r-1)+1 = pq-q+rq-r+pqr-pq > pqr$, a contradiction. Hence some Sylow subgroup in G is normal. $\square$

Chapter 3

Unit 2

3.1 Solvable group

Definition 3.1.1. A group G is said to be solvable(or soluble) if there exists a chain of subgroups

$$\{e\} = H_0 \subseteq \cdots \subseteq H_n = G$$

such that each H_i is a normal subgroup of H_{i+1} and the factor groups H_{i+1}/H_i is abelian for every $i = 0, \ldots n-1$.

The above series is referred to as solvable series of G.

Example 3.1.2. Any abelian group is solvable.

Example 3.1.3. Any non-abelian simple group is not solvable.

Definition 3.1.4. Let G be a group and $a, b \in G$. Then $aba^{-1}b^{-1}$ is called the *commutator* of a and b and is denoted by $[a, b]$. Let $A = \{aba^{-1}b^{-1} : a, b \in G\} = \{[a, b] : a, b \in G\}$ be the set of all commutators of elements in G.

Definition 3.1.5. Let G be a group. The subgroup of G generated by the commutators of elements of G is called the *commutator subgroup* of G. The commutator subgroup of a group G is denoted by G' or $G^{(1)}$ or $[G,G]$. Note that commutator subgroup is also called derived subgroup of G.

Theorem 3.1.6. *Let G be a group. Then $G' = \{e\}$ if and only if G is abelian.*

Proof. Let G' be the commutator subgroup of G. Assume that $G' = \{e\}$. Then by Definition 3.1.5, $aba^{-1}b^{-1} = e$ for all $a, b \in G$ and hence $ab = ba$ for all $a, b \in G$. Hence G is abelian.

Conversely, assume that G is abelian. Then $ab = ba$ for all $a, b \in G$ which implies $ab\,(ba)^{-1} = aba^{-1}b^{-1} = e$ for all $a, b \in G$ and hence $G' = \{e\}$. □

Theorem 3.1.7. *Let G be a group. Then*
(i) G' is a normal subgroup of G.
(ii) G/G' is abelian.
(iii) If H is a subgroup of G, then G/H is abelian and H is a normal subgroup of G if and only if $G' \subseteq H$.

Proof. (*i*) Let $g \in G$ and $x \in G'$. Then $x = c_1 \dots c_n$ where $c_i^{'}$ s are commutators of elements in G and hence $c_i = a_i b_i a_i^{-1} b_i^{-1}$ for some $a_i, b_i \in G$ for all $i = 1, \dots, n$. Now

$$\begin{aligned} gxg^{-1} &= g\,(c_1 \dots c_n)\,g^{-1} \\ &= g\left(a_1b_1a_1^{-1}b_1^{-1} \cdots a_nb_na_n^{-1}b_n^{-1}\right)g^{-1} \\ &= \left(ga_1g^{-1}\right)\left(gb_1g^{-1}\right)\left(ga_1^{-1}g^{-1}\right)\left(gb_1^{-1}g^{-1}\right)\cdots\left(ga_ng^{-1}\right) \\ &\quad \left(gb_ng^{-1}\right)\left(ga_n^{-1}g^{-1}\right)\left(gb_n^{-1}g^{-1}\right) \end{aligned}$$

Hence $gxg^{-1} \in G'$ and so G' is normal subgroup of G.

(*ii*) By (i), G/G' is a group and also $aba^{-1}b^{-1} \in G'$ for all $a, b \in G$. From this, we get $abG' = baG'$ for all $a, b \in G$ and so $aG'bG' = bG'aG'$ for all $a, b \in G$. Hence G/G' is abelian.

(iii) Assume that G/H is abelian and H is a normal subgroup of G. Then $xH\ yH = yH\ xH$ for all $x, y \in G$ and so $(xy)(yx)^{-1} \in H$ for all $x, y \in G$. Thus $xyx^{-1}y^{-1} \in H$ for all $x, y \in G$ and so $G' \subseteq H$.

Conversely, assume that $G' \subseteq H$. For any $g \in G$ and $x \in H$,
$gxg^{-1} = gxg^{-1}x^{-1}x \in H$, which shows that H is a normal subgroup of G. Since $G' \subseteq H$, $aba^{-1}b^{-1} \in H$ for all $a, b \in G$ and so $aH\ bH = bH\ aH$ for all $a, b \in G$. Hence G/H is abelian. $\square$

Example 3.1.8. *For $n \geq 3$,*

$$D'_{2n} = \begin{cases} \mathbb{Z}_n & \text{if } n \text{ is odd,} \\ \mathbb{Z}_{\frac{n}{2}} & \text{if } n \text{ is even} \end{cases}$$

Proof. Let $D_{2n} = \{1, r, r^2, \ldots, r^{n-1}, s, sr, \ldots, sr^{n-1}\}$. Then

$$\langle r^2 \rangle = \begin{cases} \mathbb{Z}_n & \text{if n is odd,} \\ \mathbb{Z}_{\frac{n}{2}} & \text{if n is even.} \end{cases}$$

Hence it is enough to prove that $D'_{2n} = \langle r^2 \rangle$.

As $[r, s] = rsr^{-1}s^{-1} = r^2 \in D'_{2n}$ and so $\langle r^2 \rangle \subseteq D'_{2n}$ is clear. Also $D'_{2n}/\langle r^2 \rangle$ is abelian and $\langle r^2 \rangle$ is a normal subgroup of D_{2n}. By Theorem 3.1.7(iii), $D'_{2n} \subseteq \langle r^2 \rangle$ and hence $D'_{2n} = \langle r^2 \rangle$. $\square$

Example 3.1.9. $\mathbb{Q}'_8 = \{\pm 1\}$

Proof. Let $\mathbb{Q}_8 = \{\pm 1, \pm i, \pm j, \pm k\}$ be a non-abelian group of order 8. Then by Theorem 3.1.6, $\{1\}$ is not a commutator subgroup of $\mathbb{Q}_8$. Note that $\{\pm 1\}$, $\{\pm, \pm i\}$, $\{\pm 1, \pm j\}$ and $\{\pm 1, \pm k\}$ are nontrivial normal subgroup of $\mathbb{Q}_8$. By Remark **??**, $\{\pm 1\}$ is the commutator subgroup of $\mathbb{Q}_8$. $\square$

Example 3.1.10. $S_n' = A_n,\ n \geq 3$

Proof. A_n is a normal subgroup of S_n and $|A_n| = \frac{n!}{2}$. Then $[S_n : A_n] = 2$ and so S_n/A_n is abelian. By Theorem 3.1.7(iii), $S_n' \subseteq A_n$. Since A_n is generated by 3-cycles for $n \geq 3$, it is enough to prove that every 3-cycle in A_n is the commutator of some element in S_n. Let $(a\ b\ c)$ be a 3-cycle in A_n. Then $(a\ b\ c) = (a\ b)(a\ c)(a\ b)^{-1}(a\ c)^{-1} \in S_n'$. Hence $A_n \subseteq S_n'$ and so $S_n' = A_n$. □

Theorem 3.1.11. If G is a non-abelian simple group, then G is $G' = G$.

Proof. Since G is simple, $\{e\}$ and G are only normal subgroup of G. Since G is non-abelian, by theorem 3.1.6, $G' \neq \{e\}$ and so $G' = G$. □

Example 3.1.12. $A_n' = A_n,\ n \geq 5$.

Proof. Clearly A_n is simple non-abelian group for $n \geq 5$. By Theorem 3.1.11, $A_n' = A_n,\ n \geq 5$. □

Example 3.1.13. $A_4' = \mathbb{V}_4$

Proof. Let $A_4 = \{e, (1\ 2\ 3), (1\ 2\ 4), (1\ 3\ 4), (2\ 3\ 4), (1\ 3\ 2), (1\ 4\ 2), (1\ 4\ 3), (2\ 4\ 3), (1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3)\}$. Let $H = \{e, (1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3)\}$ be a subgroup of A_4. Then $[A_4 : H] = 2$, H is a normal subgroup of A_4 and so A_4/H is abelian. By Theorem 3.1.7(iii), $A_4' \subseteq H$. For any $(a\ b)(c\ d) \in H$, $(a\ b)(c\ d) = (a\ b\ c)(a\ b\ d)(a\ b\ c)^{-1}(a\ b\ d)^{-1} \in A_4'$. Hence $A_4' = H$. Since every element in H other than identity is of order 2, H is isomorphic to $\mathbb{V}_4$. Hence $A_4' = \mathbb{V}_4$. □

Remark 3.1.14. Let G be a group. G' is the commutator subgroup of G, which is also denoted by $G^{(1)}$. $G^{(2)}$, the commutator subgroup of $G^{(1)}$ is the 2^{nd} commutator subgroup of G. In general $G^{(n)}$ is the n^{th} commutator subgroup of the group G. If $G^{(n)} = \{e\}$ for some positive integer n, the smallest such positive integer n is the commutator length or derived length of the group G.

Theorem 3.1.15. *Let G be a group. Then G is solvable if and only if $G^{(m)} = \{e\}$ for some positive integer m.*

Proof. Assume that G is solvable. Then there exists a series $G_0 = \{e\} \subseteq \ldots \subseteq G_n = G$ such that $G_i \lhd G_{i+1}$ and $\dfrac{G_{i+1}}{G_i}$ is abelian for every $i = 0, \ldots, n-1$. By Theorem 3.1.7(iii), $G'_{i+1} \subseteq G_i$ for every $i = 0, \ldots, n-1$. Thus $G' \subseteq G_{n-1}$. By Theorem **??**, $G^{(2)} \subseteq G'_{n-1}$. Again by Theorem 3.1.7(iii), $G'_{n-1} \subseteq G_{n-2}$ and so $G^{(2)} \subseteq G_{n-2}$ and then by Theorem **??**, $G^{(3)} \subseteq G_{n-3}$. Proceeding like this, a stage is reached where $G^{(n)} \subseteq G_0 = \{e\}$. Thus $G^{(m)} = \{e\}$ for some positive integer $m \leq n$.

Conversely, assume that $G^{(m)} = \{e\}$ for some positive integer m. Consider the series $G^{(m)} = \{e\} \subseteq G^{(m-1)} \subseteq \ldots \subseteq G = G^{(0)}$. $G^{(i+1)}$ is the commutator subgroup of $G^{(i)}$ for every $i = 0, \ldots, m-1$. Hence by Theorem 3.1.7(i) and (ii), $G^{(i+1)} \lhd G^{(i)}$ and $\dfrac{G^{(i)}}{G^{(i+1)}}$ is abelian for every $i = 0, \ldots, m-1$. Thus the series is a solvable series of G and G is solvable. $\square$

Example 3.1.16. $\mathbb{Q}_8$ is solvable.

Proof. Let $\mathbb{Q}_8 = \{\pm 1, \pm i, \pm j, \pm k\}$. Then by Example 3.1.9, $\mathbb{Q}'_8 = \{\pm 1\}$, which is abelian. Hence by Theorem 3.1.6, $\mathbb{Q}_8^{(2)} = \{e\}$ and by Theorem 3.1.15, $\mathbb{Q}_8$ is solvable. $\square$

Example 3.1.17. D_{2n} is solvable.

Proof. By Example 3.1.8, $D_{2n}' = \begin{cases} \mathbb{Z}_n & \text{if n is odd,} \\ \mathbb{Z}_{n/2} & \text{if n is even.} \end{cases}$

Then D_{2n}' is abelian. By Theorem 3.1.6, $D_{2n}^{(2)} = \{e\}$. Hence by Theorem 3.1.15, D_{2n} is solvable. □

Example 3.1.18. For $n \geq 5$, A_n is not solvable.

Example 3.1.19. For $n \geq 5$, S_n is not solvable.

Proof. By Example 3.1.10, $S_n' = A_n$. But by Example 3.1.12, $A_n' = A_n$. Hence $S_n^{(m)} = A_n$ for every positive integer m. Hence by Theorem 3.1.15, S_n is not solvable. □

Example 3.1.20. A_4 is solvable.

Proof. Clearly $\{e\} \subseteq \mathbb{V}_4 \subseteq A_4$ is a solvable series for A_4, hence is solvable. □

Example 3.1.21. S_3 and S_4 are solvable.

Proof. From Example 3.1.10, $S_3' = A_3$ and so S_3' is abelian. By Theorem 3.1.6, $S_3^{(2)} = \{e\}$. Thus by theorem 3.1.15, S_3 is solvable.

$$\{e\} \subseteq \mathbb{V}_4 \subseteq A_4 \subseteq S_4$$

is a solvable series for S_4. Hence, S_4 is solvable. □

Theorem 3.1.22. *Subgroup of a solvable group is solvable*

Proof. Let G be a solvable group and H be a subgroup of G. Since G is solvable and by Theorem 3.1.15, $G^{(n)} = \{e\}$ for some positive integer n and so $H' \subseteq G'$, $H^{(2)} \subseteq G^{(2)}$ and so on. In particular, $H^{(n)} \subseteq G^{(n)} = \{e\}$. Thus $H^{(m)} = \{e\}$ for some positive integer $m \leq n$. Hence by Theorem 3.1.15, H is solvable. □

Theorem 3.1.23. *Homomorphic image of a solvable group is solvable.*

Proof. Let G be a solvable group and let $f : G \longrightarrow K$ be a homomorphism. Let $a, b \in G$. Then $aba^{-1}b^{-1} \in G'$, $f(a), f(b) \in f(G)$, $f(aba^{-1}b^{-1}) \in f(G')$ and so $f(a) f(b) f(a)^{-1} f(b)^{-1} \in (f(G))'$. Since f is a homomorphism, for every $a, b \in G$,

$$f\left(aba^{-1}b^{-1}\right) = f(a) f(b) f(a)^{-1} f(b)^{-1}$$

. Hence $(f(G))' = f(G')$. Since G is solvable and by Theorem 3.1.15, there exists a positive integer n, such that $G^{(n)} = \{e_G\}$. $(f(G))' = f(G')$ implies that $(f(G))^{(n)} = f\left(G^{(n)}\right) = f(e_G) = e_K$. Hence by Theorem 3.1.15, $f(G)$ is solvable. □

Theorem 3.1.24. *Quotient group of a solvable group is solvable.*

Proof. Let G be a solvable group and N be a normal subgroup of G. Then G/N is a group. Define $f : G \to G/N$ by $f(g) = gN$. Then f is a natural homomorphism and $f(G) = G/N$. By Theorem 3.1.23, G/N is solvable. □

Remark 3.1.25. Let G be a solvable group. Suppose H is a subgroup of G with $H \neq \{e\}$. Then $H \neq H'$.

Proof. Suppose $H = H'$, $H^{(2)} = H' = H$. Then $H^{(n)} = H$ for any positive integer n and also by Theorem 3.1.15, H is not solvable, which gives a contradiction to Theorem 3.1.22. Hence $H \neq H'$. □

Theorem 3.1.26. *Let G be a group and N be a normal subgroup of G. Then G is solvable if and only if N and G/N are solvable.*

Proof. Assume that G is solvable. Then by Theorem 3.1.22 and Theorem 3.1.24, N and G/N are solvable.

Conversely, assume that N and G/N are solvable. Then there exists two series,

$$N_0 = \{e\} \subseteq \cdots \subseteq N_m = N$$

and

$$N = \frac{G_0}{N} = \frac{N}{N} \subseteq \cdots \subseteq \frac{G_k}{N} = \frac{G}{N}$$

such that $N_i \lhd N_{i+1}$, $\dfrac{N_{i+1}}{N_i}$ is abelian for every $i = 0, \ldots, m-1$ and $\dfrac{G_i}{N} \lhd \dfrac{G_{i+1}}{N}$, $\dfrac{G_{i+1}/N}{G_i/N}$ is abelian for every $i = 0, \ldots, k-1$. Since $\dfrac{G_i}{N} \lhd \dfrac{G_{i+1}}{N}$, $gNhNg^{-1}N \in \dfrac{G_i}{N}$ which implies that $ghg^{-1} \in G_i$ for every $g \in G_{i+1}$ and $h \in G_i$. Hence $G_i \lhd G_{i+1}$ for every $i = 0, \ldots, n-1$.

Now, G_i, $N \lhd G_{i+1}$ and $N \lhd G_i$ and by third theorem of isomorphism $\dfrac{G_{i+1}}{G_i} \cong \dfrac{G_{i+1}/N}{G_i/N}$. Since $\dfrac{G_{i+1}/N}{G_i/N}$ is abelian, $\dfrac{G_{i+1}}{G_i}$ is abelian. Thus

$$N = G_0 \subseteq \cdots \subseteq G_k = G$$

is a series such that $G_i \lhd G_{i+1}$ and $\dfrac{G_{i+1}}{G_i}$ is abelian for every $i = 0, \ldots n-1$. Hence

$$\{e\} = N_0 \subseteq \cdots \subseteq N_m = N = G_0 \subseteq \cdots \subseteq G_k$$

is a solvable series of G and so G is solvable. $\square$

3.2 Direct Product

Definition 3.2.1. Let $n > 1$ be any positive integer and let $(G_1, *_1), \ldots, (G_n, *_n)$ be any n groups. Let

$$G = G_1 \times G_2 \times \cdots \times G_n = \{(x_1, \ldots, x_n) : x_i \in G_i\}$$

Define $*$ on G by $(x_1, \ldots, x_n) * (y_1, \ldots, y_n) = (x_1 *_1 y_1, x_2 *_2 y_2, \ldots, x_n *_n y_n)$. Then $(e_1, e_2, \ldots, e_n)$ is an identity element of G, where each e_i is identity element of G_i. Also $(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1})$ is an inverse of $(x_1, \ldots, x_n)$ in G. Hence $(G, *)$ is a group.

We call this group G the external direct product of $G_1, \ldots, G_n$

Definition 3.2.2. Let G be a group and $N_1, N_2, \ldots, N_n$ normal subgroups of G such that

(i) $G = N_1 N_2 \ldots N_n$.

(ii) Given $g \in G$ then $g = m_1 m_2 \ldots m_n$, $m_i \in N_i$ in a unique way.

We then say that G is the internal direct product of $N_1,\ N_2, \ldots, N_n$.

Theorem 3.2.3. *Let G be a group and suppose that G is the internal direct product of $N_1, \ldots, N_n$. Let $T = N_1 \times N_2 \times \cdots \times N_n$. Then G and T are isomorphic.*

Proof. Define the mapping $\Psi : T \to G$ by

$$\Psi((b_1, b_2, \ldots, b_n)) = b_1 b_2 \cdots b_n,$$

where each $b_i \in N_i$, $i = 1, \ldots, n$. We claim that Ψ is an isomorphism of T onto G. If $x \in G$ then $x = a_1 a_2 \ldots a_n$ for some $a_1 \in N_1, \ldots, a_n \in N_n$. But then $\Psi((a_1, a_2, \ldots, a_n)) = a_1 a_2 \ldots a_n = x$ and hence Ψ is onto.

The mapping Ψ is one-to-one by the uniqueness of the representation of every element as a product of elements from $N_1, \ldots, N_n$. For, if $\Psi((a_1, \ldots, a_n)) = \Psi((c_1, \ldots, c_n))$, where $a_i \in N_i$, $c_i \in N_i$, for $i = 1, 2, \ldots, n$, then, by definition, $a_1a_2 \ldots a_n = c_1c_2 \ldots c_n$. The uniqueness in the definition of internal direct product forces $a_1 = c_1, a_2 = c_2, \ldots, a_n = c_n$. Thus Ψ is one-to-one.

If $X = (a_1, \ldots, a_n)$, $Y = (b_1, \ldots, b_n)$ are elements of T then $\Psi(XY) = \Psi((a_1, \ldots, a_n)(b_1, \ldots, b_n)) = \Psi(a_lb_l, a_2b_2, \ldots, a_nb_n) = a_1b_1a_2b_2 \ldots a_nb_n$. Thus However, by Lemma 3.2.4, $a_ib_i = b_ia_i$ if $i \neq j$. This implies that $a_1b_1 \ldots a_nb_n = a_1a_2 \ldots a_nb_1b_2 \ldots b_n$. Thus $\Psi(XY) = a_1a_2 \ldots a_nb_1b_2 \ldots b_n$. But we can recognize $a_1a_2 \ldots a_n$ as $\Psi((a_1, a_2, \ldots, a_n)) = \Psi(X)$ and $b_1b_2 \ldots b_n$ as $\Psi(Y)$. Hence $\Psi(XY) = \Psi(X)\Psi(Y)$. □

Lemma 3.2.4. *Suppose that G is the internal direct product of $N_1, \ldots, N_n$. Then for $i \neq j$, $N_i \cap N_j = \{e\}$, and if $a \in N_i$, $b \in N_j$ then $ab = ba$.*

Proof. Suppose that $x \in N_i \cap N_j$. Then we can write x as $x = e_1 \ldots e_{i-1}xe_{i+1} \ldots e_j \ldots e_n$ where $e_t = e$, viewing x as an element in N_i. Similarly, we can write x as $x = e_1 \ldots e_i \ldots e_{i-1}xe_{i+1} \ldots e_m$ where $e_t = e$, viewing x as an element of N_j. But every element and so, in particular x has a unique representation in the form $m_1m_2 \ldots m_n$, where $m_i \in N_1, \ldots, m_n \in N_n$. Since the two decompositions in this form for x must coincide, the entry from N_i in each must be equal. In our first decomposition this entry is x, in the other it is e; hence $x = e$. Thus $N_i \cup N_j = \{e\}$ for $i \neq j$.

Suppose $a \in N_i$, $b \in N_j$, and $i \neq j$. Then $aba^{-1} \in N_j$ since N_j is normal; thus $aba^{-1}b^{-1} \in N_j$. Similarly, since $a^{-1} \in N_i$, $ba^{-1}b^{-1} \in N_i$, whence $aba^{-1}b^{-1} \in N_i$. But then $aba^{-1}b^{-1} \in N_i \cap N_j = \{e\}$. Thus $aba^{-1}b^{-1} = e$; this gives the desired result $ab = ba$. □

Remark 3.2.5. If $G = G_1 \times \cdots \times G_n$ is the external direct product of $G_1, \ldots, G_n$, then $H_i = \{(e_1, \ldots, e_{i-1}, x_i, e_{i+1}, \ldots, e_n) \in G : x \in G_i\}$ is a normal subgroup of G and by definition 3.2.2 and Lemma 3.2.4, G is internal direct product of $H_1, \ldots, H_n$.

Theorem 3.2.6. *Let G be a finite abelian group. Then G is isomorphic to the direct product of its Sylow subgroups.*

Proof. Let $o(G) = p_1^{k_1} \cdots p_r^{k_r} > 1$, where $p_1, \ldots, p_r$ are distinct primes. Since G is abelian, all p-Sylow subgroups are normal and so G has unique p-Sylow subgroup for all prime p divides $o(G)$. Let H_i be p_i-Sylow subgroup of G and $o(H_i) = p_i^{k_i}$ for $i = 1, 2, \ldots, r$. Then H_i is normal subgroup of G, $H_i \cap H_j = \{e\}$ for all $i \neq j$ and $o(H_i H_j) p_i^{k_i} p_j^{k_j}$. By Theorem 1.6.4,

$$o(H_1 \cdots H_r) = o((H_1 \cdots H_{r-1})H_r) = \frac{o(H_1 \cdots H_{r-1})o(H_r)}{o((H_1 \cdots H_{r-1}) \cap H_r)} = o(G).$$

. Since each H_i is normal, $H_1 \cdots H_r$ is subgroup of G and so $G = H_1 \cdots H_r$. Hence, by Theorem 3.2.3, G is the external direct product of $H_1, \ldots, H_r$. □

Example 3.2.7. Let $G = \{e, a, b, c\}$ be the Klein 4-group. Then $H = \{e, a\}$ and $K = \{e, b\}$ are normal subgroups of G, $H \cap K = \{e\}$ and $HK = G$. Hence G is the internal direct product of H and K and so Theorem 3.2.3, $G = \mathbb{Z}_2 \times \mathbb{Z}_2$.

Example 3.2.8. Let $S_3 = \{e, (1\ 2\ 3), (1\ 3\ 2), (1\ 2), (1\ 3), (2\ 3)\}$. Then $H = \{e, (1\ 2\ 3), (1\ 3\ 2)\}$ is unique nontrial proper normal subgroup of S_3 and so S_3 is not the internal direct product of its normal subgroups.

3.3 Finite abelian groups

Our first step is to reduce the problem to a slightly easier one. If we knew that each such Sylow subgroup was a direct product of cyclic goups we could put the results together for these Sylow subgroups to realize G as a direct product of cyclic groups. **Thus it suffices to prove the following theorem for abelian groups of order p^n where p is a prime.**

Theorem 3.3.1. *Let G be an abelian group of order p^n, where p is prime. Then G is the direct product cyclic groups.*

Proof. Let a_1 be an element in G of highest possible order, p^{n_1}, and let $A_1 = (a_1)$. Pick b_2 in G such that $\bar{b}_2$, the image of b_2 in $\bar{G} = G/A_1$, has maximal order p^{n_2}. Since the order of $\bar{b}_2$ divides that of b_2, and since the order of a_1 is maximal, we must have that $n_1 \geq n_2$. In order to get a direct product of A_1 with (b_2) we would need $A_1 \cap (b_2) = (e)$; this might not be true for the initial choice of b_2, so we may have to adapt the element b_2. Suppose that $A_1 \cap (b_2) \neq (e)$; then, since $b_2^{p_{n_2}} \in A_1$ and is the first power of b_2 to fall in A_1 we have that $b_2^{p_{n_2}} = a_1^i$. Therefore $(a_1^i)^{p^{n_1-n_2}} = (b_2^{p^{n_2}})^{p^{n_1-n_2}} = b_2^{p_{n_1}} = e$, whence $(a_1^i)^{p^{n_1-n_2}} = e$. Since a_1 is of order p^{n_1} we must have that $p^{n_1}|ip^{n_1-n_2}$, and so $p_{n_2}|i$. Thus, re-calling what i is, we have $b_2^{p^{n_2}} = a_1^i = a_1^{jp^{n_2}}$. This tells us that if $a_2 = a_1^{-j}b_2$ then $a_2^{p^{n_2}} = e$. The element a_2 is indeed the element we seek. Let $A_2 = (a_2)$. We claim that $A_1 \cap A_2 = (e)$. For, suppose that $a_2^t \in A_1$; since $a_2 = a_1^{-j}b_2$, we get $(a_1^{-j}b_2)^t \in A_1$ and so $b_2^t \in A_1$. By choice of b_2, this last relation forces $p^{n_2}|t$, and since $a_2^{p^{n_2}} = e$ we must have that $a_2^t = e$. Hence $A_1 \cap A_2 = (e)$.

We continue one more step in the program we have outlined. Let $b_3 \in G$ map into an element of maximal order in $G/(A_1A_2)$. If the order of the image of b_3 in $G/(A_1A_2)$ is p^{n_3}, we claim that $n_3 \leq n_2 \leq n_1$. By the choice of n_2 , $b_3^{p^{n_2}} \in A_1$ so is certainly in A_1A_2. Thus $n_3 \leq n_2$. Since $b_3^{p^{n_2}} \in A_1A_2$, $b_3^{p^{n_2}} = a_1^{i_1}a_2^{i_2}$. We claim that $p^{n_3}|i_1$ and $p^{n_3}|i_2$. For, $b_3^{p^{n_2}} \in A_1$ hence $(a_1^{i_1}a_2^{i_2})p^{n_2-n_3} = (b_3^{p^{n_3}})^{p^{n_3-n_2}} = b_3^{p^{n_2}} \in A_1$. This tells us that $a_2^{i_2p^{n_2-n_3}} \in A_1$ and so $p^{n_2}|i_2p^{n_2-n_3}$, which is to say, $p^{n_3}|i_2$. Also $b_3^{p^{n_1}} = e$, hence $(a_1^{i_1}a_2^{i_2})p^{n_1-n_3} = b_3^{p^{n_1}} = e$; this says that $a_1^{i_1})p^{n_1-n_3} \in A_1 \cap A_2 = (e)$, that is, $a_1^{i_1p^{n_1-n_3}} = (e)$. This yields that $p^{n_3}|i_1$. Let $i_1 = j_1p^{n_3}$, $i_2 = j_2p^{n_3}$; thus $b_3p^{n_3} = a_1^{j_1p^{n_3}}a_2^{j_2p^{n_3}}$. Let $a_3 = a_1^{-j_1}a_2^{-j_2}b_3$, $A_3 = (a_3)$; note that $a_3^{p^{n_3}}a = e$. We claim that $A_3 \cap (A_1A_2) = (e)$. For if $a_3^t \in A_1A_2$ then $(a_1^{-j_1}a_2^{-j_2}b_3)^t \in A_1A_2$, giving us $b_3^t \in A_1A_2$. But then $p^{n_3}|t$, whence, since $a_3^{p^{n_3}} = e$, we have $a_3^t = e$. Thus, $A_3 \cap (A_1A_2) = (e)$.

Continuing this way we get cyclic subgroups $A_1 = (a_1), A_2 = (a_2), \ldots, A_k = (a_k)$ of order $p^{n_1}, p^{n_2}, \ldots, p^{n_k}$ respectively, with $n_1 \geq n_2 \geq \cdots \geq n_k$ such that $G = A_1A_2 \ldots A_k$

and such that, for each i, $A_i \cap (A_1 A_2 \ldots A_{i-1}) = (e)$. This tells us that every $x \in G$ has a unique representation as $x = a_1' a_2' \ldots a_k'$ where $a_1' \in A_1, \ldots, a_k' \in A_k$. Hence, G is the direct product of the cyclic subgroups $A_1, A_2, \ldots, A_k$. $\square$

Definition 3.3.2. If G is an abelian group of order p^n, p a prime, and $G = A_1 \times A_2 \times \cdots \times A_k$ where each A_i is cyclic of order p^{n_i}; with $n_1 \geq n_2 \geq \ldots n_k > 0$, then the integers $n_1, n_2, \ldots, n_k$ are called the invariants of G.

Theorem 3.3.3. *The number of non-isomorphic abelian groups of order p^n, p a prime, equals the number of partitions of n.*

Corollary 3.3.4. *The number of non-isomorphic abelian groups of order $p_1^{\alpha_1} \ldots p_r^{\alpha_r}$, where the p_i are distinct primes and where each $\alpha_i > 0$, is $p(\alpha_1)p(\alpha_2)\ldots p(\alpha_r)$, where $p(u)$ denotes the number of partitions of u.*

Example 3.3.5. Let G be an abelian group of order p^n, where p is a prime number.

n=1 $G = \mathbb{Z}_p$

n=2 $G = \mathbb{Z}_p \times \mathbb{Z}_p$ or $\mathbb{Z}_{p^2}$;

n=3 $G = \mathbb{Z}_p \times \mathbb{Z}_p \times \mathbb{Z}_p$, $\mathbb{Z}_{p^2} \times \mathbb{Z}_p$ or $\mathbb{Z}_{p^3}$

n=4 $G = \mathbb{Z}_p \times \mathbb{Z}_p \times \mathbb{Z}_p \times \mathbb{Z}_p$, $\mathbb{Z}_{p^2} \times \mathbb{Z}_{p^2}$, $\mathbb{Z}_{p^2} \times \mathbb{Z}_p \times \mathbb{Z}_p$, $\mathbb{Z}_{p^3} \times \mathbb{Z}_p$ or $\mathbb{Z}_{p^4}$

Example 3.3.6. Let G be an abelian group of order $100 = 2^2 5^2$.

$G = G_1 \times G_2$, where G_1 is 2-Sylow subgroup of G and G_2 is a 5-Sylow subgroups of G

$G_1 = \mathbb{Z}_2 \times \mathbb{Z}_2$ or $\mathbb{Z}_4$;

$G_2 = \mathbb{Z}_5 \times \mathbb{Z}_5$ or $\mathbb{Z}_{25}$;

$G = \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_5 \times \mathbb{Z}_5$, $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_{25}$, $\mathbb{Z}_4 \times \mathbb{Z}_5 \times \mathbb{Z}_5$ or $\mathbb{Z}_4 \times \mathbb{Z}_{25}$.

Theorem 3.3.7. *Let G be a group and A and B be subgroups of G. If*
(i) $G = AB$
(ii) $ab = ba$ for all $a \in A$, $b \in B$, and
(iii) $A \cap B = \{e\}$,
prove that G is an internal direct product of A and B.

Proof. let us first show that A and B are normal subgroup of G. For this, let $a \in A$, $g \in G$. There exist $c \in A$ and $b \in B$ such that $g = cb$ by(i). Now $gag^{-1} = (cb)a(cb)^{-1}$ $= cbab^{-1}c^{-1} = cabb^{-1}c^{-1} = cac^{-1} \in A$. Hence, A is a normal subgroup of G. Similarly, B is a normal subgroup of G.Let $g \in G$. Then $g = ab$ for some $a \in A$, $b \in$ B. Suppose $g = a_1b_1$, where $a_1 \in A$, $b_1 \in B$. Then $ab = a_1b_1$, which implies that $a_1^{-1}a = b_1b^{-1} \in A \cap B = \{e\}$. Thus $a = a_1$ and $b = b_1$. Therefore, we find that every element g of G can be expressed uniquely as $g = ab$, $a \in A$, $b \in B$. Consequently, G is an internal direct product of A, B. □

Theorem 3.3.8. *Let A and B be two cyclic groups of order m and n, respectively. Show that $A \times B$ is a cyclic group if and only if $gcd(m, n) = 1$.*

Proof. Let $A = \langle a \rangle$ for some $a \in A$ and $B = \langle b \rangle$ for some $b \in B$. Suppose $gcd(m, n) = 1$. Let $g = (a, b)$. Then $g^{mn} = (a, b)^{mn} = (a^{mn}, b^{mn}) = (e_A, e_B)$, where e_A denotes the identity of A and e_B denotes the identity of B. Suppose $o(g) = t$. Then $(a, b)^t = (e_A, e_B)$. This implies that $a^t = e_A$ and $b^t = e_B$. Thus, $m|t$ and $n|t$. Since $gcd(m, n) = 1$, $mn|t$. Hence, mn is the smallest positive integer such that $g^{mn} = e$. Thus, $o(g) = mn$. Now $|A \times B| = mn$ and $A \times B$ contains an element g of order mn. As a result, $A \times B$ is cyclic. Conversely, assume that $A \times B$ is a cyclic and

$gcd(m, n) = d \neq 1$. Let $(a, b) \in A \times B$. Then $o(a)|m$ and $o(b)|n$. Now $\frac{mn}{d} = \frac{m}{d}n = m\frac{n}{d}$ is and integer and $\frac{mn}{d} < mn$. Also,

$$(a, b)^{\frac{mn}{d}} = (a^{m\frac{n}{d}}, b^{n\frac{m}{d}}) = (e_A, e_B)$$

.

Hence, $A \times B$ does not contain any element of order mn. This implies that $A \times B$ is not cyclic, a contradiction. Therefore, $gcd(m, n) = 1$. $\square$

Chapter 4

Unit 3: Canonical Form

4.1 Basics of Linear Transformation

Definition 4.1.1. A nonempty set V is said to be vector space over field F if

(i) (V,+) is a abelin group.

(ii) $\alpha \cdot (v_1 + v_2) = \alpha \cdot v_1 + \alpha \cdot v_2$

(iii) $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$

(iv) $\alpha(\beta \cdot v) = (\alpha\beta) \cdot v$

(v) $1.v = v$ for all $v \in V$.

Example 4.1.2.

1. Every field is a vector space over itself
2. Every field is a vector space over its subfield
3. If F is a field, then $F[x]$ is a vector space over F
4. If F is a fiel, then $M_{n\times m}(F)$ is a vector space over a field F
5. $C[0,1]$ is a vector space over $\mathbb{R}$
6. Let $V_n = \{f(x) \in F[x] : deg(f(x)) \leq n\}$. Then V_n is vector space over a field F.

Definition 4.1.3. Let V be vector space over F. A subset B of V is a basis for V over F if B span V and B is linearly independent.

Example 4.1.4.

1. If F is a vector space over itself, then $\{1\}$ is a basis for F over F

2. If $F[x]$ is a vector space over F, then $\{1, x, x^2, \ldots\}$ is a basis for $F[x]$ over F

3. If $M_{n\times m}(F)$ is a vector space over a field F, then $B = \{E_{ij} : ij^{th}$ entry is 1 other entries are $0\}$ is a basis for $M_{n\times m}(F)$.

4. Let $V_n = \{f(x) \in F[x] : deg(f(x)) \leq n\}$ be a vector space over F. Then $\{1, x, x^2, x^3, \ldots, x^n\}$ is a basis for V_n over F.

Definition 4.1.5. Let V and W be vector space over the same field F. A function $T : V \to W$ is a linear transformtion if

$$T(\alpha u + v) = \alpha T(u) + T(v)$$

for all $\alpha \in F$ and $u, v \in V$.

Example 4.1.6. Define $O : V \to W$ by $O(v) = 0_w$ for all $v \in V$. Then $O(\alpha u + v) = 0_w = \alpha O(u) + O(v)$ and so O is Zero transformation

Example 4.1.7. Define $D : F[x] \to F[x]$ by $D(f(x)) = f'(x)$ for all $f(x) \in F[x]$. Then $D(\alpha f(x) + g(x)) = (\alpha f(x) + g(x))' = \alpha f'(x) + g'(x) = \alpha D(f(x)) + D(g(x))$ and so D is linear transformation.

Definition 4.1.8. Let $T \in A(V)$. A subspace W of V is invariant under T if $T(W) \subseteq W$. Clearly (0) and V are invariant subspace under T.

Example 4.1.9. Let $T \in A(V)$. Then $T(V)$ is invariant subspace of V under T and $Ker(T)$ is subspace of V under T.

Definition 4.1.10. Let F be a field and $p(x) \in F[x]$. Then $p(x)$ is the minimal polynomial for $T \in A(V)$ if $p(x)$ is monic, $p(T) = 0$ and $g(T) \neq 0$ for all $g(x) \in F[x]$.

Example 4.1.11. Let $I : V \to V$ by $I(v) = v$ for all $v \in V$. Then the minimal polynomial for I is $(x-1)^n$.

Example 4.1.12. Let $O : V \to W$ by $O(v) = O_W$ for all $v \in V$. Then the minimal polynomial for O is x.

Example 4.1.13. Define $D : V_n \to V_n$ by $D(f(x)) = f'(x)$ for all $f(x) \in F[x]$. Then the minimal polynomial for D is x^{n+1}.

Definition 4.1.14. A linear operator T on V is called nilpotent if $T^n = 0$ for some positive integer n.

Example 4.1.15. Let $O : V \to W$ by $O(v) = 0_W$ for all $v \in V$. Then O is nilpotent transformation.

Example 4.1.16. Define $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T(x, y) = (0, x)$. Then $T^2(x, y) = T(T(x, y)) = T(0, x) = T(0, 0) = (0, 0)$ and hence T is nilpotent transformtion.

4.2 Triangular Form

Definition 4.2.1. The linear transformations $S, T \in A(V)$ are said to be similar if there exists an invertible element $C \in A(V)$ such that $T = CSC^{-1}$.

Definition 4.2.2. The subspace W of V is invariant under $T \in A(V)$ if $WT \subset W$.

Lemma 4.2.3. *If $W \subset V$ is invariant under T, then T induces a linear transformation $\bar{T}$ on a vector space V/W, defined by $(v+W)\bar{T} = vT+W$. If T satisfies the polynomial $q(x) \in F[x]$, then so does $\bar{T}$. If $p_1(x)$ is the minimal polynomial for $\bar{T}$ over F and if $p(x)$ is that for T, then $p_1(x)|p(x)$.*

Proof. Let $\bar{V} = V|W = \{u + W : u \in V\}$. Given $\bar{v} = v + W \in \bar{V}$ define $\bar{T} : V/W \to V/W$ by $\bar{v}\bar{T} = vT + W$. Then $(\alpha(\bar{v}) + \bar{u})\bar{T} = (\alpha v + u)T + W = \alpha(vT) + uT + W = \alpha(vT + W) + uT + W = \alpha\bar{v}\bar{T} + \bar{u}\bar{T}$ and hence T is a linear operator on V/W.

Suppose that $\bar{v} = v_1 + W = v_2 + W$ where $v_1, v_2 \in V$. We must show that $v_1T + W = v_2T + W$. Since $v_1 + W = v_2 + W$, $v_1 - v_2$ must be in W, and since W is invariant under T, $(v_1 - v_2)T$ must also be in W. Consequently $v_1T - v_2T \in W$, from which it follows that $v_1T + W = v_2T + W$, as desired. We now know that $\bar{T}$ defines a linear transformation on $\bar{V} = V|W$.

If $\bar{v} = v + W \in \bar{V}$, then $\bar{v}(\bar{T^2}) = vT^2 + W = (vT)T + w = (vT + W)\bar{T} = ((v + W)\bar{T})\bar{T} = \bar{v}(\bar{T})^2$; thus $()\bar{T^2}) = (\bar{T})^2$ Similarly, $(\bar{T^k}) = (\bar{T})^k$ for any $k \geq 0$. Consequently, for any polynomial $q(x) \in F[x]$, $\bar{q(T)} = q(\bar{T})$. For any $q(x) \in F[x]$ with $q(T) = 0$, since $\bar{0}$ is the zero transformation on $\bar{V}$, $0 = \bar{q(T)} = q(\bar{T})$.

Let $p_1(x)$ be the minimal polynomial over F satisfied by $\bar{T}$. If $q(T) = 0$ for $q(x) \in F[x]$, then $P_i(x)Iq(x)$. If $p(x)$ is the minimal polynomial for T over F, then $p(T) = 0$, whence $p(T) = 0$; in consequence, $p_1(x)|p(x)$. ☐

Note that all the characteristic roots of $\bar{T}$ which lie in F are roots of the minimal polynomial of T over F. We say that all the characteristic roots of T are in F if all the roots of the minimal polynomial of T over F lie in F.

We defined a matrix as being triangular if all its entries above the main diagonal were 0. Equivalently, if T is a linear transformation on V over F, the matrix of T in the basis $v_1, \ldots, v_n$ is triangular if

$v_1T = \alpha_{1,1}v_1$

$v_2T = \alpha_{2,1}v_1 + \alpha_{2,2}v_2$

$\ldots$

$v_nT = \alpha_{n,1}v_1 + \cdots + \alpha_{m,n}v_n.$

Theorem 4.2.4. *If $T \in A(V)$ has all its characteristic roots in F, then there is a basis of V in which the matrix of T is triangular*

Proof. The proof by induction on the dimension of V over F. If $dim_F(V) = 1$, then every element in $A(V)$ is a scalar, and so the theorem is true here. Suppose that the theorem is true for all vector spaces over F of dimension $n-1$, and let V be of dimension n over F.

Note that the linear transformation T on V has all its characteristic roots in F. Let $\lambda_i \in F$ be a characteristic root of T. Then there exists a nonzero vector v_1 in V such that $v_1T = \lambda_1 v_1$. Let $W = \{\alpha v_1 : \alpha \in F\}$; W is a one-dimensional subspace of V, and is invariant under T. Let $\bar{V} = V/W$. Then $dim\bar{V} = dimV - dimW = n-1$. By Lemma 4.2.3, T induces a linear transformation $\bar{T}$ on $\bar{V}$ whose minimal polynomial over F divides the minimal polynomial of T over F. Thus all the roots of the minimal polynomial of $\bar{T}$, being roots of the minimal polynomial of T, must lie in F. Hence the linear transformation $\bar{T}$ in its action on V satisfies the hypothesis of the theorem; since $\bar{V}$ is $(n-1)$-dimensional over F, by our induction hypothesis, there is a basis $\bar{v}_2, \bar{v}_3, \ldots, \bar{v}_n$ of $\bar{V}$ over F such that $\bar{v}_1\bar{T} = \alpha_{1,1}\bar{v}_1$

$\bar{v}_2\bar{T} = \alpha_{2,1}\bar{v}_1 + \alpha_{2,2}\bar{v}_2$

$\ldots$

$\bar{v}_n\bar{T} = \alpha_{n,1}\bar{v}_1 + \cdots + \alpha_{m,n}\bar{v}_n$

Let $v_2, \ldots, v_n$ be elements of V mapping into $\bar{v}_2, \bar{v}_3, \ldots, \bar{v}_n$ of $\bar{V}$ respectively. Then $v_1, \ldots, v_n$ form a basis of V. Since $\bar{v}_2\bar{T} = \alpha_{2,2}\bar{v}_2$, $\bar{v}_2\bar{T} = \alpha_{2,2}\bar{v}_2 = 0$, whence $v_2T - \alpha_{2,2}v_2$ must be in W. Thus $v_2T - \alpha_{2,2}v_2$ is a multiple of v_1, say $\alpha_{2,1}v_1$, yielding, after transposing, $v_2T = \alpha_{2,1}v_1 + \alpha_{2,2}v_2$.

Similarly, $v_iT - \alpha_{i,2}v_2 - \alpha_{i,3}v_3 - \cdots - \alpha_{i,i}v_i \in W$, whence $v_iT = \alpha_{i,1}v_1 + \alpha_{i,2}v_2 + \alpha_{i,3}v_3 + \cdots + \alpha_{i,i}v_i$. The basis $v_1, \ldots, v_n$ of V over F provides us with a basis where every v_iT is a linear combination of v_i and its predecessors in the basis. Therefore, the matrix of T in this basis is triangular. □

Theorem 4.2.5. *If V is n-dimensional over F and if $T \in A(V)$ has all its characteristic roots in F, then T satisfies a polynomial of degree n over F.*

Proof. By Theorem 4.2.4, we can find a basis $v_1, \ldots, v_n$ of V over F such that: $v_1T = \lambda_1v_1$, $v_2T = \alpha_{2,1}v_1 + \lambda_2v_2$, $\ldots$, $v_iT = \alpha_{i,1}v_1 + \cdots + \alpha_{i,i-1}v_{i-1} + \lambda_iv_i$, for $i = 1, 2, \ldots, n$. Equivalently $v_1(T - \lambda_1) = 0$, $v_2(T - \lambda_2) = \alpha_{2,1}v_1$, $\ldots$, $vi(T - At) = \alpha_{i,1}v_1 + \cdots + \alpha_{i,i-1}v_{i-1}$, for $i = 1, 2, \ldots, n$.

As a result of $v_2(T-\lambda_2) = \alpha_{2,1}v_1$ and $v_1(T-\lambda_1) = 0$, we obtain $v_2(T-\lambda_2)(T-\lambda_1) = 0$. Since $(T - \lambda_2)(T - \lambda_1) = (T - \lambda_1)(T - \lambda_2)$,

$$v_1(T - \lambda_2)(T - \lambda_1) = v_1(T - \lambda_1)(T - \lambda_2) = 0.$$

Continuing this type of computation yields
$v_1(T - \lambda_i)(T - \lambda_{i-1}) \ldots (T - \lambda_1) = 0$,
$v_2(T - \lambda_i)(T - \lambda_{i-1}) \ldots (T - \lambda_1) = 0$,
$\ldots$
$v_i(T - \lambda_i)(T - \lambda_{i-1}) \ldots (T - \lambda_1) = 0$.

For $i = n$, the matrix $S = (T - \lambda_n)(T - \lambda_{n-1}) \cdots (T - \lambda_1)$ satisfies $v_1S = v_2 = \cdots = v_n = 0$. Then, since S annihilates a basis of V, S must annihilate all of V. Therefore,

$S = 0$. Consequently, T satisfies the polynomial $(x - \lambda_1)(x - \lambda_2) \dots (x - \lambda_n)$ in $F[x]$ of degree n. $\square$

4.3 Nilpotent Transformations

Definition 4.3.1. Let V be a vector space over F and $T \in A(V)$. If $T^m = 0$ for some m, then T is nilpotent linear transformation on V.

The smallest positive integer k such that $T^k = 0$ is called nilpotent index of T.

If T is nilpotent operator with nilpotent index k, then $T^s \neq 0$ for all $s < k$.

Lemma 4.3.2. *All characteristic roots of the nilpotent linear transformation are zero.*

Proof. Let T be a nilpotent lineartransformation of nilpotent index m. Then $T^m = 0$. Let α be a characteristic root of T. Then there exist $u \neq 0$ in B such that $uT = \alpha u$. Since $uT = \alpha u$, $uT^2 = \alpha(uT) = \alpha\alpha u = \alpha^2 u$. From this, we get $uT^\ell = \alpha^\ell$. Since $T^m = 0$, $uT^m = \alpha^m u = 0$. Since $u \neq 0$, $\alpha^m = 0$ and hence $\alpha = 0$. $\square$

Lemma 4.3.3. *If $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$, where each subspace V_i is of dimension n_i and is invariant under T, an element of $A(V)$, then a basis of V can be found so that the matrix of T in this basis is of the form*

$$\begin{bmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & A_k \end{bmatrix}$$

where each A_i is an $n_i \times n_i$ matrix and is the matrix of the linear transformation induced by T on V_i.

Proof. Choose a basis of V as follows: $v_1^{(1)}, \dots, v_n^{(1)}$ is a basis of V_1 ,$v_1^{(2)}, \dots, v_n^{(2)}$ is a basis of V_2, and so on. Since each V_i is invariant under T, $v_j^{(i)}T \in V_i$ so is a linear combination of $v_1^{(i)}, v_2^{(i)}, \dots, v_n^{(i)}$, and of only these. Thus the matrix of T in the basis so chosen is of the desired form. That each A_i is the matrix of T_i, the linear transformation induced on V_i by T, is clear from the very definition of the matrix of a linear transformation. □

Definition 4.3.4. If $T \in A(V)$ is nilpotent, then k is called the index of nilpotence of T if $T^k = 0$ but $T^{k-1} \neq 0$.

In a ring, sum of unit element and nilpotent element is unit.

Lemma 4.3.5. *If $T \in A(V)$ is nilpotent, then $\alpha_0 + \alpha_1 T + \cdots + \alpha_m T^m$ is invertible, where $\alpha_i \in F$, if $\alpha_0 \neq 0$.*

Proof. Since T is nilpotent, $T^r = 0$ for some r. Let $S = \alpha_1 T + = \alpha_2 T^2 + \cdots + \alpha_m T^m$. Then S^r is the linear combination of $T^r, \dots, T^{rm}$. Since $T^r = 0$, $S^r = 0$. Since $A(V)$ is ring and $\alpha_0 \neq 0$, $\alpha_0 I$ is unit and so $\alpha_0 I + S = \alpha_0 + S$ is unit. □

Notation: M_t will denote the $t \times t$ matrix all of whose entries are 0 except on the superdiagonal, where they are all $1's$.

$$M_t = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

Theorem 4.3.6. *If $T \in A(V)$ is nilpotent, of index of nilpotence n_1, then a basis of V can be found such that the matrix of T in this basis has the form*

$$\begin{bmatrix} M_{n_1} & 0 & \dots & 0 \\ 0 & M_{n_2} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & M_{n_r} \end{bmatrix}$$

where $n_1 \geq n_2 \geq \cdots \geq n_r$, and where $n_1 + n_2 + \cdots + n_r = dim_F V$.

Proof. The proof will be a little detailed, so as we proceed we shall separate parts of it out as lemmas. Since $T^{n_1} = 0$ but $T^{n_1-1} \neq 0$.

Claim 1: We can find a vector $v \in V$ such that $vT^{n_1-1} \neq 0$. We claim that the vectors $v, vT, \ldots, vT^{n_1-1}$ are linearly independent over F.

For, suppose that $\alpha_1 v + \alpha_2 vT + \cdots + \alpha_{n_1} vT^{n_1-1} = 0$ where the $\alpha_i \in F$; let α_s be the first nonzero α, hence

$$vT^{s-1}(\alpha_s + \alpha_{s+1}T + \cdots + \alpha_{n_1}T^{n_1-s}) = 0$$

Since $\alpha_s \neq 0$, by Lemma 4.3.5, $\alpha_s + \alpha_{s+1}T + \cdots + \alpha_{n_1}T^{n_1-s}$ is invertible, and therefore $vT^{s-1} = 0$. However, $s < n_1$, thus this contradicts that $vT^{n_1-1} \neq 0$. Thus no such nonzero α_s exists and $v, vT, \ldots, vT^{n_1-1}$ have been shown to be linearly independent over F.

Let V_1 be the subspace of V spanned by $v_1 = v, v_2 = vT, \ldots, v_{n_1} = vT^{n_1-1}$; V_1 is invariant under T, and, in the basis above, the linear transformation induced by T on V_1 has as matrix M_{n_1}

Claim 2: If $u \in V_1$ is such that $uT^{n_1-k} = 0$, where $0 < k \leq n_1$, then $u = u_0T^k$ for some $u_0 \in V_1$.

Since $u \in V_1$, $u = \alpha_1 v + \alpha_2 vT + \cdots + \alpha_k vT^{k-1} + a_{k+1}vT^k + \cdots + \alpha_{n_1}vT^{n_1-1}$. Thus $0 = uT^{n_1-k} = \alpha_1 vT^{n_1-1} + + \alpha_k vT^{n_1-1}$. However, $vT^{n_1-k}, \ldots, vT^{n_1-1}$ are linearly independent over F, whence $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$, and so, $u = \alpha_{k+1}vT^k + \cdots + \alpha_{n_1}vT^{n_1-1} = u_0T^k$ where $U_o = \alpha_{k+l}v + \cdots + \alpha_{n_1}vT^{n_1-k-1} \in V_1$.

Claim 3: There exists a subspace W of V, invariant under T, such that $V = V_1 \bigoplus W$.

Let W be a subspace of V, of largest possible dimension, such that

1. $V_l \cap W = (0)$;

2. W is invariant under T

We want to show that $V = V_1 + W$. Suppose not; then there exists an element $z \in V$ such that $z \notin V_1 + W$. Since $T^{n_1} = 0$, there exists an integer k, $0 < k \leq n_1$, such that $zT^k \in V_1 + W$ and such that $zT^i \notin V_1 + W$ for $i < k$. Thus $zT^k = u + w$, where $u \in V_l$ and where $w \in W$. But then $0 = zT^{n_1} = (zT^k)T^{n_1-k} = uT^{n_1-k} + wT^{n_1-k}$; however, since both V_1 and W are invariant under T, $uT^{n_1-k} \in V_l$ and $wT^{n_1-k} \in W$. Now, since $V_1 \cap W = (0)$, this leads to $uT^{n_1-k} = -wT^{n_1-k} \in V_l \cap W = (0)$, resulting in $uT^{n_1-k} = 0$. By Claim 2, $u = u_0T^k$ for some $u_0 \in V_1$; therefore, $zT^k = u + w = u_0T^k + w$. Let $z_1 = z - u_0$; then $z_1T^k = zT^k - u_0T^k = w \in W$, and since W is invariant under T this yields $z_1T^m \in W$ for all $m \geq k$. On the other hand, if $i < k$, $Z_1T^i = zT^i - U_oT^i \ni v_1 + w$, for otherwise zT^i must fall in $V_1 + W$, contradicting the choice of k.

Let W_1 be the subspace of V spanned by W and $Z_1, Z_1T, \ldots, Z_1T^{k-1}$. Since $z_1 \notin W$, and since $W_l \supset W$, the dimension of W_1 must be larger than that of W. Moreover, since $z_1T^k \in W$ and since W is invariant under T, W_1 must be invariant under T. By the maximal nature of W, there must be an element of the form $w0 + \alpha_1 Z_1 + \alpha_2 z_1 T + \cdots + \alpha_k z_1 T^{k-1} \neq 0$ in $W_1 \cap V_1$ where $w_o \in W$. Not all of $\alpha_l, \ldots, \alpha_k$ can be 0; otherwise we would have $0 \neq w_o \in W \cup V_1 = (0)$ a contradiction.

Let α_s be the first nonzero α; then $w_0 + z_1T^{s-1}(\alpha_s + \alpha_{s+1}T + \cdots + \alpha_k T^{k-s}) \in V_1$. Since $\alpha_s \neq 0$, by Lemma 4.2.4 , $\alpha_s + \alpha_{s+l}T + \cdots + \alpha_k T^{k-s}$ is invertible and its inverse, R, is a polynomial in T. Thus W and V_1 are invariant under R; however, from the above, $w_oR + z_1T^{s-l} \in V_1R \subset V_1$, forcing $z_1T^{s-1} \in V_1 + WR \subset V_1 + W$. Since $s - 1 < k$ this is impossible; therefore $V_1 + W = V$. Because $V_1 \cap W = (0), V = V_1 \bigoplus W$.

By Claim 3, $V = V_1 + W$, where W is invairant under R. Using the basis $v_1, \ldots, v_{n_1}$ of V_1 and any basis ov W as a basis of V. By Lemma 4.2.3, the matrix of T in this basis haas the form

$$\begin{bmatrix} M_{n_1} & 0 \\ 0 & A_2 \end{bmatrix},$$

where A_2 is the matrix of T_2, the linear transformation induced on W by T.

Since $T^{n_1} = 0$, $T_2^{n_2} = 0$ for some $n_2 \leq n_1$. Repeating the argument used for T on V for T_2 on W we can decompose W. Continuing this way, we get a basis of V in which the matrix of T is of the form

$$\begin{bmatrix} M_{n_1} & 0 & \dots & 0 \\ 0 & M_{n_2} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & M_{n_r} \end{bmatrix}.$$

From this, we get $n_1 + n_2 + \cdots + n_r = dim_F V$. □

Definition 4.3.7. The integers $n_1, n_2, \ldots, n_r$ are called the invariants of T.

Definition 4.3.8. If $T \in A(V)$ is nilpotent, the subspace M of V, of dimension m, which is invariant under T, is called cyclic with respect to T if

1. $MT^m = (0)$, $MT^{m-1} \neq (0)$;

2. there is an element $z \in M$ such that $z, zT, \ldots, zT^{m-1}$ form a basis of M

Lemma 4.3.9. *lf M, of dimension m, is cyclic with respect to T, then the dimension of MT^k is $m - k$ for all $k \leq m$.*

Proof. A basis of MT^k is provided us by taking the image of any basis of M under T^k. Using the basis $z, zT, \ldots, zT^{m-1}$ of M leads to a basis zT^k,$zT^{k+1}, \ldots, zT^{m-1}$ of MT^k. Since this basis has $m - k$ elements, the dimension of MT^k is $m - k$. □

Lemma 4.3.10. *If T is nilpotent operator on V, then the invaiant of T are unique.*

Proof. Let if possible there are two sets of invariants $n_1, n_2, \dots, n_r$ and $m_1, m_2, \dots, m_s$ of T. Then $V = V_1 \oplus \cdots \oplus V_r$ and $V = U_1 \oplus \cdots \oplus U_s$, where V_i and U_i are cyclic subspace of V of dimension n_i and m_i, respectively. Now we show that $r = s$ and $n_i = m_i$.

Suppose that k be the first integer such that $n_k \neq m_k$. Then $n_i = m_i$ for $i < k$. Without loss of generality, $n_k > m_k$. Consider

$$T^{m_k}(V) = T^{m_k}(V_1) \oplus \cdots \oplus T^{m_k}(V_r)$$

and

$$\dim T^{m_k}(V) = \dim T^{m_k}(V_1) \oplus \cdots \oplus \dim T^{m_k}(V_r)$$

By the above Lemma, $\dim T^{m_k}(V_i) = n_i - m_k$. Therefore $\dim T^{m_k}(V) > (n_1 - m_k) + \cdots + (n_{k-1} - n_k)$.

Simillarly,

$$\dim T^{m_k}(V) = \dim T^{m_k}(U_1) \oplus \cdots \oplus \dim T^{m_k}(U_s).$$

As $m_j \leq m_k$ for $j > k$, we have $T^{m_k}(U_j) = \{0\}$. Therefore, $\dim T^{m_k}(U_j) = 0$ for $j > k$. Hence,

$$\dim T^{m_k}(V) = (m_1 - m_k) + \cdots + (m_{k-1} - n_k)$$

. By assumption,

$$\dim T^{m_k}(V) = (n_1 - m_k) + \cdots + (n_{k-1} - n_k)$$

, a contradiction. Hence $n_i = m_i$. Since $\dim V = \sum_{i=1}^{r} n_i = \sum_{j=1}^{s} m_j$, $r = s$. □

Theorem 4.3.11. *Two nilpotent linear transformations are similar if and only if they have the same invariants.*

Proof. Suppose S and T are similar. Then there exist a regular mapping A such that $A^{-1}TA = S$.

Let $n_1, n_2, \ldots, n_r$ be invariants of S and $m_1, m_2, \ldots, m_s$ be invariants of T. Then $V = V_1 \oplus \cdots \oplus V_r$ and $V = U_1 \oplus \cdots \oplus U_s$, where V_i and U_j are cyclic and invariant subspaces of V of dimension n_i and m_j, respectively.

As $S(V_i) \subset V_i$, $(A^{-1}TA)(V_i) \subset V_i$ implies $(A^{-1}T)A(V_i) \subset V_i$. Put $A(V_i) = U_i$, (since A is regular). Thus, $\dim V_i = \dim U_i = n_i$. Further $T(U_i) = TA(V_i) = AS(V_i)$. As $S(V_i) \subset V_i$, therefore $T(U_i) \subset U_i$. Equivalently, we have to show that U_i is invariant under T.

Moreover,

$$V = A(V) = A(V_1) \oplus \cdots \oplus A(V_r) = U_1 \oplus \cdots \oplus U_s.$$

By the above theorem, the invariants of nilpotent transformations are unique. Therefore $n_i = m_i$ and $r = s$. Conversely, suppose that two nilpotent transformations S and T have same invariants. Then there exists two bases say, $\{v_1, v_2, \ldots, v_n\}$ and $\{u_1, u_2, , u_n\}$ of V such that the matrix of S under $\{v_1, v_2, \ldots, v_n\}$ is equal to the matrix of T under $\{u_1, u_2, \ldots, u_n\}$.

Let it be

$$m(S) = m(T) = \begin{bmatrix} M_{n_1} & \ldots & 0 \\ \vdots & \ldots & \vdots \\ 0 & \ldots & M_{n_r} \end{bmatrix}$$

where $m(S) = [a_{ij}]$ and $m(T) = [b_{ij}]$

Define a linear transformation $A : V \to V$ by $A(v_i) = u_i$. Then $A^{-1}TA(v_i) = A^{-1}T(u - i) = A^{-1}(\sum\limits_{j=1}^{n} a_{ij}u_j) = \sum\limits_{j=1}^{n} a_{ij}A^{-1}(u_j) = \sum\limits_{j=1}^{n} a_{ij}v_j = S(v_i)$. Hence $A^{-1}TA = S$ and so S and T are similar. $\square$

Chapter 5

Unit 4: Canonical Forms: Jordan Form

Let V be a finite-dimensional vector space over F and let T be an arbitrary element in $A_F(V)$. Suppose that V_1 is a subspace of V invariant under T. Therefore T induces a linear transformation T_1 on V_1 defined by $uT_1 = uT$ for every $u \in V_1$. Given any polynomial $q(x) \in F[x]$, we claim that the linear transformation induced by $q(T)$ on V_1 is precisely $q(T_1)$. In particular, if $q(T) = 0$ then $q(T_1) = 0$. Thus T_1 satisfies any polynomial satisfied by T over F. What can be said in the opposite direction?

Lemma 5.0.12. *Suppose that $V = V_1 \oplus V_2$ where V_1 and V_2 are subspaces of V invariant under T. Let T_1 and T_2 are the linear transformations induced by T on V_1 and V_2 respectively. If the minimal polynomial of T_1 over F is $p_1(x)$ while that of T_2 is $p_2(x)$, then the minimal polynomial for T over F is the $l.c.m\{p_1(x), p_2(x)\}$.*

Proof. Let $q(x)$ be the $l.c.m\{p_1(x), p_2(x)\}$ and let $p(x)$ be the minimal polynomial of T.

Since $p(x)$ is the minimal polynomial of T. Then $p(T) = 0 \Rightarrow p(T_1) = 0$ and $p(T_2) = 0$. Since $p_1(x)$ and $p_2(x)$ are the minimal polynomial of T_1 and T_2 respectively,

$p_1(x)|p(x)$ and $p_2(x)|p(x)$. From this we get $p(x)$ is one among all the multiples of $p_1(x)$ and $p_2(x)$ and so $q(x)|p(x)$.

On the other hand, if $q(x)$ is the least common multiple of $p_1(x)$ and $p_2(x)$, consider $q(T)$. For $v_1 \in V_1$, since $p_1(x)|q(x)$, $v_1q(T) = v_1q(T_1) = 0$; similarly, for $v_2 \in V_2$, $v_2q(T) = 0$. Given any $v \in V$, v can be written as $v = v_1 + v_2$, where $v_i \in V_i$, in consequence of which $vq(T) = (v1 + v2)q(T) = v1q(T) + v2q(T) = 0$. Thus $q(T) = 0$ and T satisfies $q(x)$. Since $p(x)$ is minimal polynomial for T, $p(x)|q(x)$. $\square$

Corollary 5.0.13. *If $V = V_1 \oplus \cdots \oplus V_k$ where each V_i is invariant under T and if $p_i(x)$ is the minimal polynomial over F of T_i the linear transformation induced by T on V_i, then the minimal polynomial over F is the l.c.m$\{p_1(x), \ldots, p_k(x)\}$.*

Lemma 5.0.14. *Any polynomial in $F[x]$ can be written in a unique manner as a product of irreducible polynomials in $F[x]$.*

Lemma 5.0.15. *Given two polynomials $f(x), g(x) \in F[x]$, they have g.c.d $d(x)$ which can be realized as $d(x) = \lambda(x)f(x) + \mu(x)g(x)$.*

Lemma 5.0.16 (Integers)**.** *If a and b are integers, not both 0 then we can find integers m_0 and n_0 such that $(a, b) = m_0a + n_0b$.*

Theorem 5.0.17. *Prove that for each $i = 1, \ldots, k, V_i \neq 0$ and $V = V_1 \oplus \cdots \oplus V_k$. The minimal polynomial of T_i is $(q_i(x))^{l_i}$, where q_i is irreducible and l_i is an integer.*

Proof. Let $T \in A_F(V)$ and $p(x)$ be the minimal polynomial over F. By Lemma 5.0.14, $p(x) \in F[x]$ is factorized in a unique way i.e, $p(x) = q_1(x)^{l_1}q_2(x)^{l_2} \ldots q_k(x)^{l_k}$ where q_i are distinct irreducible polynomial in $F[x]$ where $l_1, \ldots, l_k$ are positive integers.

Let $V_i = \{v \in V : vq_i(T)^{l_i} = 0\}$ for $i = 1, 2, \ldots, k$. Then each V_i is a subspace of V.

Claim 1: V_i is invariant under T

Let $u \in V_i$. It is enough to prove $(uT)(q_i(T))^{l_i} = 0$. Now $(uT)(q_i(T))^{l_i} = (uq_i(T)^{l_i})T = 0T = 0$ and so $uT \in V_i$. Hence each V_i is invariant under T.

If $k = 1$, there is nothing to prove, assume that $k > 1$.

Claim 2: $V_i \neq (0)$

Let $h_i(x) = \frac{p(x)}{q_i(x)^{l_i}}$ for $i = 1, 2, \ldots, k$. Then clearly $q_i(x)^{l_i} h_i(x) = p(x)$, for $i = 1, 2, \ldots, k$. Moreover $h_i(x) \neq p(x)$ and $h_i(T) \neq 0$. Then for any given i, there is a $w \in V$ such that $w = vh_i(T) \neq 0$. But $wq_i(T)^l_i = v[h_i(T)q_i(T)^l_i] = vp(T) = 0$ and so $w \in V_i$. Therefore, $V_i \neq (0)$. Moreover $Vh_i(T) \neq 0$ and $Vh_i(T) \subseteq V_i$.

Claim 3: $V = V_1 + V_2 + \cdots + V_k$

Suppose $v_i \in V_j$ for $j \neq i$.. Then $q_j(x)^{l_j} | h_i(x) \Longrightarrow h_i(x) = q_j(x)^{l_j} f(x)$ for some $f(x)$. Now $v_j h_i(T) = [v_j q_j(T)^{l_j}]f(T) = 0$ for all $j \neq i$. Clearly, the polynomial $h_1(x), h_2(x) \ldots, h_k(x)$ are relatively prime. By Lemma 5.0.15, we can find polynomials $a_1(x), \ldots, a_k(x)$ in $F[x]$ such that $a_1(x)h_1(x) + \cdots + a_k(x)h_k(x) = 1$ implies $a_1(T)h_1(T) + \cdots + a_k(T)h_k(T) = I$. For any $v \in V$, $v = vI = v[a_1(T)h_1(T) + \cdots + a_k(T)h_k(T)] = va_1(T)h_1(T) + \cdots + va_k(T)h_k(T)$. Now, each $va_i(T)h_i(T)$ is in $Vh_i(T)$, implies $Vh_i(T) \subset V_i$. From this, we get $v = v_1 + \cdots + v_k$, where $v_i = va_i(T)h_i(T)$ and hence $V = V_1 + V_2 + \cdots + V_k$

Claim 4: If $u_1 + \cdots + u_k = 0$, then $u_1 = u_2 = \cdots = u_k = 0$ where each $u_i \in V_i$

Suppose not for some i, $u_i \neq 0$. Without loss of generality, we may assume that $u_1 \neq 0$. Since $u_1 + u_2 + \cdots + u_k = 0$, $u_1h_1(T) + u_2h_1(T) + \cdots + u_kh_1(T) = 0 \Longrightarrow u_jh_1(T) = 0$ for all $j \neq 1$. Since $u)j \in V_j$, $u_1h_1(T) = 0$. This implies that $u_1q_1(T)^{l_1} = 0$. Since $h_1(x)$ and $q_1(x)^{l_1}$ are relatively prime, $u_1 = u_1 I = u_1[b_1(T)h_1(T) + b_2(T)q_1(T)^{l_1}] = u_1h_1(T)b_1(T) + u_1q_1(T)^{l_1}b_2(T) = 0$, a contradiction.

Claim 5: Minimal polynomial of T_i on V_i is $q(x)^l_i$

By the definition of V_i, $V_iq_i(T)^{l_i} = 0 \Rightarrow q_i(T)^{l_i} = 0$. This implies the minimal polynomial for T_i must be a divisor of $q_i(x)^{l_i}$ and so the minimal polynomial of T is $q_i(x)^{f_i}$ where $f_i \leq l_i$. By Lemma 5.0.12, the minimal polynomial of T is the l.c.m

$\{q_1(x)^{f_1}, \ldots, q_k(x)^{f_k}\} = q_1(x)^{f_1} \cdots q_k(x)^{f_k}$. Since this is the minimal polynomial each $f_i \geq l_i$, $f_i = l_i$. $\square$

If all the characteristic roots of T should happen to lie in F, then the minimal polynomial of T takes on the especially nice form $q(x) = (x - \lambda_1)^{\ell_1} \cdots (x - \lambda_k)^{\ell_k}$, where $\lambda_1, \ldots, \lambda_k$ are the distinct characteristic roots of T. The irreducible factors $q(x)$ above are merely $q_i(x) = x - \lambda_i$, Note that on V_i, T_i only has λ_i as a characteristic root.

Corollary 5.0.18. *If all the distinct characteristic roots $\lambda_1, \lambda_2, \ldots, \lambda_k$ of T lie in F then V can be written as $V = V_1 \oplus V_2 \cdots \oplus V_k$ where $V_i = \{v_i \in V : V(T - \lambda_i)^{l_i} = 0\}$ and T_i has only one characteristic root $\lambda_i \in V_i$*

Definition 5.0.19. The matrix

$$\begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & \ldots & 0 \\ 0 & 0 & \lambda & \ldots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & \lambda \end{pmatrix}$$

where λ_i's are on diagonal, 1's on the super diagnal and 0's elsewhere is a Jordan block belonging to λ.

Jordan form: The matrix

$$\begin{pmatrix} J_1\lambda_1 & \cdots & \cdots & \cdots \\ \cdots & J_2\lambda_2 & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & J_k\lambda_k \end{pmatrix}$$

where

$$J_i = \begin{pmatrix} B_{i1} & \cdots & \cdots & \cdots \\ \cdots & B_{i2} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & B_{ir_i} \end{pmatrix}$$

where $B_{i1}, B_{i2}, \cdots, B_{ir_i}$ are basic Jordan blocks belonging to λ_i.

Let $A \in F_n$ and suppose that K is the splitting field of the minimal polynomial of A over F, then an invertible matrix $C \in K_n$ can be formed so that CAC^{-1} is in Jordan form.

Remark 5.0.20. *Two linear transformation $A_F(V)$ which have all their characteristic roots in F are similar iff can be bought to the same Jordan form.*

Theorem 5.0.21. *Let $T \in A_k(V)$ have all its distinct characteristic roots $\lambda_1, \lambda_2, \cdots, \lambda_k$ in F. Then a basis of V can be found in which the matrix of T is of the form*

$$\begin{pmatrix} J_1 & 0 & \cdots & \cdots & 0 \\ 0 & J_2 & \cdots & \cdots & 0 \\ \vdots & & & & \\ 0 & \cdots & \cdots & \cdots & J_k \end{pmatrix}$$

where each

$$J_i = \begin{pmatrix} B_{i1} & \cdots & \cdots & \cdots \\ \cdots & B_{i2} & \cdots & \cdots \\ \cdots & & & \\ \cdots & \cdots & \cdots & B_{ir} \end{pmatrix}$$

where $B_{i1}, \cdots, B_{ir}$ are basic Jordan block belongs to λ_i.

Proof. Consider the case that T has only one characteristic root λ. Then by above corollary, $V = \{v \in V : T(T-\lambda)^l = 0\}$. $T - \lambda$ is nilpotent. Now $T = \lambda + T - \lambda$. Since $T - \lambda$ is nilpotent, there is a basis in which its matrix is of the form

$$\begin{pmatrix} M_{n1} & \cdots & \cdots \\ \cdots & M_{n2} & \cdots \\ \vdots & & \\ \cdots & \cdots & M_{nr} \end{pmatrix}.$$

Then the matrix of

$$T = \begin{pmatrix} \lambda & \cdots & \cdots \\ \vdots & & \\ \cdots & \cdots & \lambda \end{pmatrix} + \begin{pmatrix} M_{n1} & \cdots & \cdots \\ \vdots & & \\ \cdots & \cdots & M_{nr} \end{pmatrix} = \begin{pmatrix} B_{n1} & \cdots & \cdots \\ \vdots & & \\ \cdots & \cdots & B_{nr} \end{pmatrix}.$$

Hence the theorem is proved. □

5.1 Rational Canonical form

To obtain the Jordan form of $T \in A(V)$, T must have its characteristic roots in F. In rational canonical form the location of characteristic roots is not assumed.

Given $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in F[x]$ and $v \in V$, $f(x)v \in f(T)v$, then v is called an $F[x]$ module through T.

Remark 5.1.1. *1. If V is finite dimensional vector space then V becomes a finitely generated $F[x]$ module.*
2. By remark 1 and by Fundamental theorem of finitely generated modules $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$, where each V_i is cyclic submodules, $F[x]$ is Euclidean ring

Definition 5.1.2. *V is said to be cyclic relative to T^r if for every $w \in V$ there exist $v \in V$, $w = vf(T)$.*

Lemma 5.1.3. *Suppose that T in $A_F(V)$, has the minimal polynomial over F, the polynomial $p(x) = \gamma_o + \gamma_1 x + \cdots + \gamma_{r-1}x^{r-1} + x^r$. Suppose V is cyclic relative to T, then there is basis of V over F such that, in this basis, the matrix of T is*

$$\begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ -\gamma_0 & -\gamma_1 & . & . & \cdots & -\gamma_{r-1} \end{pmatrix}.$$

Proof. Since V is cyclic relative to T, there exists a vector v in V such that every element w, in V, is of the form $w = vf(T)$ for some $f(x)$ in $F[x]$.

Claim 1.

If $vs(T) = 0$, for some polynomial $s(x)$ in $F[x]$, then $s(T) = 0$. From this, $vs(T) = 0$ implies for any $w \in V$ such that $wS(T) = vf(T)s(T) = vs(T)f(T) = 0$. Therefore $S(T) = 0$. Hence the claim 1.

Claim 2

Note that $\{v, vT, VT^2, \cdots, VT^{r-1}\}$ is a basis of V. Since $p(x)$ is a minimal polynomial of T, $p(x)|s(x)$. First we have to prove $v, vT, VT^2, \cdots, VT^{r-1}$ are linearly independent. Suppose not, $\alpha_0 v + \alpha_1 vT + \alpha_2 vT^2 + \cdots \alpha_{r-1}vT^{r-1} = 0$ implies not $\alpha_i' s$ are zero. This implies $v(\alpha_0 + \alpha_1 T + \alpha_2 T^2 + \cdots \alpha_{r-1}T^{r-1}) = 0$ and so $vg(T) = 0$, where $g(T) = \alpha_0 + \alpha_1 T + \alpha_2 T^2 + \cdots \alpha_{r-1}T^{r-1}$. Thus $g(T) = 0$ (By claim 1) implies T satisfies $g(x)$. Hence $p(x)|g(x)$ implies $p(x)|\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots \alpha_{r-1}x^{r-1}$. This is possible only if $\alpha_0 = \alpha_1 = \cdots \alpha_{r-1} = 0$.

Next we will prove the vectors $v, vT, VT^2, \cdots, VT^{r-1}$ span V. So $vT^r = \gamma_0 v - \gamma_1 vT - \cdots - \gamma_{r-1}vT^{r-1}$ and

$$m(T) = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ -\gamma_0 & -\gamma_1 & . & . & \cdots & -\gamma_{r-1} \end{pmatrix}.$$

□

Definition 5.1.4. *If $f(x) = \gamma_o + \gamma_1 x + \cdots + \gamma_{r-1}x^{r-1} + x^r \in F[x]$ then the $r \times r$ matrix*

$$\begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ -\gamma_0 & -\gamma_1 & . & . & \cdots & -\gamma_{r-1} \end{pmatrix}$$

is called the companion matrix of $f(x)$. We write it as $C(f(x))$.

Example 5.1.5. *Let $f(x) = x^3 + 3x^2 + 4x - 7$. Then*

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 7 & -4 & -3 \end{pmatrix}.$$

Theorem 5.1.6. *If T in $A_F(V)$ has as minimal polynomial $p(x) = q(x)^e$, where $q(x)$ is a monic, irreducible polynomial in $F[x]$, then a basis of V over F can found in which* *the matrix of T is of the form*

$$\begin{pmatrix} C(q(x)^e) & & & \\ & C(q(x)^{e_2}) & & \\ & & \ddots & \\ & & & C(q(x)^{e_r}) \end{pmatrix}$$

where $e = e_1 \geq e_2 \geq e_2 \geq \cdots \geq e_r$.

Proof. Since V is finitely generated $F[x]$- module $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$, where $V_i = \{v \in V : v \in v(q(T))^{e_i} = 0\}$. Since $T^r = -\gamma_0 - \gamma_1 T - \cdots - \gamma_{r-1}T^{r-1}$, $T^{r+k}, k \geq 0$ is a linear combination of $1, T, T^2, \cdots, T^{r-1}$. This implies $f(T)$ is a linear combination of $1, T, T^2, \cdots, T^{r-1}$. over F. Since any w in V is of the form $w = vf(T)$, w is a linear combination of $v, vT, vT^2, \cdots, vT^{r-1}$. Let $V_1 = v, V_2 = vT, V_3 = vT^2 \cdots V_r = vT^{r-1}$. Thus we have to prove $V_1T = VT = V_2 = 0V_1 + 1V_2 + \cdots + 0V_r$ and so $V_2T = VT^2 = V_3 = 0V_1 + 0V_2 + 1V_3 + \cdots + 0V_r$. Note that each V_i is cyclic sub-module. Also each V_i is invariant under T and hence induces a linear transformation T_i on V_i.

Since the minimal polynomial of T_i divides the minimal polynomial of $T = q(x)^e$, the minimal polynomial of T_i is of the form $q(x)^{e_i}$, where $e_i \leq e$........(1) By suitably rearranging $V_i's$ we have $e_1 \geq e_2 \geq \cdots \geq e_i$.

Since V_i is a cyclic submodule relative to T_i, there is a basis of V_i in which $m(T_i) = c(q(x)^{e_i})$,. From this, we get

$$m(T)\begin{pmatrix} C(q(x)^e) & & & \\ & C(q(x)^{e_2}) & & \\ & & \ddots & \\ & & & C(q(x)^{e_r}) \end{pmatrix}.$$

Finally we have to prove $e = e_1$. For $v_1 \in V_i$ implies $v_i[q(T)]^{e_i} = 0$ for $i = 1, \cdots, r$. This implies $v[q(T)]^{e_1} = 0$ implies $[q(T)]^{e_1} = 0$. But $q(x)^e$ is the minimal polynomial of

T. $e \leq e_1$.....(2). From (1) and (2), hence $e = e_1$. $\square$

Chapter 6

Unit 5

6.1 Trace and Transpose

Definition 6.1.1. Let F_n be the set of all $n \times n$ matrices over a field F. The trace of $A \in F_n$ is the sum of the elements on the main diagonal of A.

We shall write the trace of A as trA, if $A = (a_{ij})$, then

$$trA = \sum_{i=1}^{n} a_{ii}$$

Lemma 6.1.2. *For $A, B \in F_n$ and $\lambda \in F$,*

1. $tr(\lambda A) = \lambda \; tr \; A.$
2. $tr(A + B) = trA + trB.$
3. $tr(AB) = tr(BA).$

Proof. (i) Let $A = [a_{ij}], B = [b_{ij}] \in F_n$. Then $\lambda A = [\lambda a_{ij}]$ and so $tr(\lambda A) = \sum_{i=1}^{n} \lambda a_{ii} = \lambda \sum_{i=1}^{n} a_{ii} = \lambda tr(A)$.

(ii) $tr(A+B) = \sum_{i=1}^{n}(a_{ii}+b_{ii}) = \sum_{i=1}^{n} a_{ii} + \sum_{i=1}^{n} b_{ii} = tr(A) + tr(B)$.

If $A = (\alpha_{ij})$ and $B = (\beta_{ij})$ then $AB = (\gamma_{ij})$ where

$$\gamma_{ij} = \sum_{k=1}^{n} \alpha_{ik}\beta_{kj}$$

and $BA = (\mu_{ij})$ where

$$\mu_{ij} = \sum_{k=1}^{n} \beta_{ik}\alpha_{kj}$$

Thus

$$tr(AB) = \sum_{i} \gamma_{ii} = \sum_{i} \left(\sum_{k} \alpha_{ik}\beta_{ki} \right);$$

if we interchange the order of summation in this last sum, we get

$$tr(AB) = \sum_{k=1}^{n}\sum_{i=1}^{n} \alpha_{ik}\beta_{ki} = \sum_{k=1} \left(\sum_{i=1} \beta_{ki}\alpha_{ik} \right) = \sum_{k=1}^{n} \mu_{kk} = tr(BA).$$

□

Corollary 6.1.3. *If A is invertible then $tr(ACA^{-1}) = tr(C)$.*

Proof. Let $B = CA^{-1}$. Then $tr(ACA^{-1}) = tr(AB) = tr(BA) = tr(CA^{-1}A) = tr(C)$. □

Definition 6.1.4. If $T \in A(V)$ then $tr\ T$, the trace of T, is the trace of $m_1(T)$ where $m_1(T)$ is the matrix of T in some basis of V. We claim that the definition is meaningful and depends only on T and not on any particular basis of V. For if $m_1(T)$ and $m_2(T)$ are the matrices of T in two different bases of V, then $m_1(T)$ and $m_2(T)$ are similar matrices, so they have the same trace.

Lemma 6.1.5. *If $T \in A(V)$ then $tr(T)$ is the sum of the characteristic roots of T.*

Proof. We can assume that T is a matrix in F_n. If K is the splitting field for the minimal polynomial of T over F, then in K_n, T can be brought to its Jordan form, J. From this, J is a matrix on whose diagonal appear the characteristic roots of T, each root appearing as often as its multiplicity. Thus $tr(J)$ is the sum of the characteristic roots of T. However, since J is of the form ATA^{-1}, $tr(J) = tr(T)$. □

Lemma 6.1.6. *If F is a field of characteristic 0, and if $T \in A_F(V)$ is such that $tr(T^i) = 0$ for all $i \geq 1$ then T is nilpotent.*

Proof. Since $T \in A_F(V)$, T satisfies some minimal polynomial $p(x) = x^m + \alpha_1 x^{m-1} + \cdots + \alpha_m$ from $T^m + \alpha_1 T^{m-1} + \cdots + \alpha_{m-1}T + \alpha_m = 0$, taking traces of both sides yields

$$trT^m + \alpha_1 trT^{m-1} + \cdots + \alpha_{m-1} trT + tr\alpha_m = 0$$

However, by assumption, $tr(T^i) = 0$ for $i \geq 1$, thus we get $\alpha_m = 0$. If $dimV = n$, $tr(\alpha_m I) = n\alpha_m$ whence $n\alpha_m = 0$. But the characteristic of F is 0, therefore, $n \neq 0$, hence it follows that $\alpha_m = 0$. Since the constant term of the minimal polynomial of T is 0, T is singular and so 0 is a characteristic root of T.

We can consider T as a matrix in F_n and therefore also as a matrix in K_n, where K is an extension of F which in turn contains all the characteristic roots of T. In K_n, we can bring T to triangular form, and since 0 is a characteristic root of T, we can actually bring it to the form.

$$\left(\begin{array}{c|cccc} 0 & 0 & \cdots & 0 \\ \beta_2 & \alpha_2 & 0. & 0 \\ \vdots & & \ddots & \vdots \\ \beta_n & * & & \alpha_n \end{array}\right) = \left(\begin{array}{c|c} 0 & 0 \\ * & T_n \end{array}\right),$$

where,

$$T_2 = \begin{pmatrix} \alpha_2 & 0 & 0 \\ & \ddots & \vdots \\ & * & \alpha_n \end{pmatrix}$$

is an $(n-1) \times (n-1)$ matrix (the $*$'s indicate parts in which we are not interested in the explicit entries). Now

$$T^k = \left(\begin{array}{c|c} 0 & 0 \\ * & T_2^k \end{array} \right)$$

hence $0 = tr(T^k) = tr(T_2^k$. Thus T_2 is an $(n-1) \times (n-1)$ matrix with the property that $tr(T_2^k) = 0$ for all $k \geq 1$. Either using induction on n,or repeating the argument on T_2 used for T, we get, since $\alpha_2, \ldots \alpha_n$ are the characteristic roots of T_2, that $\alpha_2 = \cdots = \alpha_n = 0$. Thus when T is brought to triangular form, all its entries on the main diagonal are 0 and hence T is nilpotent. □

Lemma 6.1.7. *If F is of characteristic 0 and if S and T, in $A_F(V)$, are such that $ST - TS$ commutes with S, then $ST - TS$ is nilpotent.*

Proof. For any $k \geq 1$, we compute $(ST-TS)^k$. Now $(ST-TS)^k = (ST-TS)^{-1}(ST-TS) = (ST-TS)^{k-1}ST - (ST-TS)^{k-1}TS$. Since $ST-TS$ commutes with S, the term $(ST-TS)^{k-1}ST$ can be written in the form $S((ST-TS)^{k-1}T)$. If we let $B = (ST-TS)^{-1}T$, we see that $(ST-TS)^k = SB - BS$; hence $tr((ST-TS)^k) = tr(SB-BS) = tr(SB) - tr(BS) = 0$. By previous lemma, $ST-TS$ must be nilpotent.

□

Definition 6.1.8. If $A = [\alpha_{ij}] \in F_n$, then the transpose of A, written as A', is the matrix $A' = [\gamma_{ij}]$ where $\gamma_{ji} = \alpha_{ji}$ for each i and j.

Lemma 6.1.9. *For $A, B \in F_n$*

1. $(A')' = A$.

2. $(A+B)' = A' + B'$.

3. $(AB)' = B'A'$.

Proof. Let $A = [a_{ij}], B = [b_{ij}] \in F_n$.

(i) Let $A' = [c_{ij}]$. Then $c_{ij} = a_{ji}$. In $(A')' = [d_{ij}]$, $d_{ij} = c_{ji} = a_{ij}$ and hence $(A')' = A$.

(ii) Clearly $A + B = [a_{ij} + b_{ij}]$. Also $(A+B)' = [a_{ij} + b_{ij}]' = [x_{ij}]$. From this $x_{ij} = a_{ji} + b_{ji}$ and so $(A+B)' = A' + B'$.

Suppose that $A = [\alpha_{ij}]$ and $B = [\beta_{ij}]$. Then $AB = [\lambda_{ij}]$ where

$$\lambda_{ij} = \sum_{k=1}^{n} \alpha_{ik}\beta_{kj}.$$

Therefore, by definition, $(AB)' = [\mu_{ij}]$, where

$$\mu_{ij} = \lambda_{ji} = \sum_{k=1}^{n} \alpha_{jk}\beta_{ki}$$

On the other hand $A' = [\gamma_{ij}]$ where $\gamma_{ij} = \alpha_{ji}$ and $B' = [\xi_{ij}]$ where $\xi_{ij} = \beta_{ji}$, whence the (i,j) element of $B'A'$ is

$$\sum_{k=1}^{n} \xi_{ik}\gamma_{kj} = \sum_{k=1}^{n} \beta_{ki}\alpha_{jk} = \sum_{k=1}^{n} \alpha_{jk}\beta_{ki} = \mu_{ij}$$

That is, $(AB)' = B'A'$. □

Definition 6.1.10. The matrix A is said to be a symmetric matrix if $A' = A$.

Definition 6.1.11. *The matrix A is said to be a skew-symmetric matrix if $A' = -A$.*

Definition 6.1.12. *A mapping $*$ from F_n into F_n is called an adjoint on F_n if*

1. $(A^*)^* = A$.

2. $(A+B)^* = A^* + B^*$.

3. $(AB)^* = B^*A^*$.

for all $A, B \in F_n$.

6.2 Hermitian, Unitary and Normal Transformations

Lemma 6.2.1. *If* $T \in A(V)$ *is such that* $(vT, v) = 0$ *for all* $v \in V$*, then* $T = 0$*.*

Proof. Since $(vT, v) = 0$ for $v \in V$, given $u, w \in V$, $((u+w)T, u+w) = 0$. Expanding this out and making use of $(uT, u) = (wT, w) = 0$, we obtain

$$(uT, w) + (wT, u) = 0 \textit{ for all } u, w \in V \tag{6.1}$$

Since equation (6.1) holds for arbitrary w in V, it still must hold if we replace in it w by iw where $i^2 = -1$; but $(uT, iw) = -i(uT, w)$ whereas $((iw)T, u) = i(wT, u)$. Substituting these values in (6.1) and cancelling out i leads us to

$$-(uT, w) + (wT, u) = 0. \tag{6.2}$$

Adding (6.1) and (6.2) we get $(wT, u) = 0$ for all $u, w \in V$, whence, in particular, $(wT, wT) = 0$. By the defining properties of an inner-product space, this forces $wT = 0$ for all $w \in V$, hence $T = 0$. $\square$

Definition 6.2.2. The linear transformation $T \in A(V)$ is said to be unitary if $(uT, vT) = (u, v)$ for all $u, v \in V$.

Lemma 6.2.3. *If* $(vT, vT) = (v, v)$ *for all* $v \in V$ *then* T *is unitary.*

Proof. Let $u, v \in V$. Then by assumption $((u+v)T, (u+v)T) = (u+v, u+v)$. Expanding this out and simplifying, we obtain

$$(uT, vT) + (vT, uT) = (u, v) + (v, u) \tag{6.3}$$

for $u, v \in V$. In (6.3) replace v by iv; computing the necessary parts, this yields

$$-(uT, vT) + (vT, uT) = -(u, v) + (v, u). \tag{6.4}$$

Adding (6.3) and (6.4) results in $(uT, vT) = (u, v)$ for all $u, v \in V$, hence T is unitary. □

Theorem 6.2.4. *The linear transformation T on V is unitary if and only if it takes an orthonormal basis of V into an orthonormal basis of V.*

Proof. Suppose that $\{v_1, \ldots, v_n\}$ is an orthonormal basis of V. Then $(v_i, v_j) = 0$ for $i \neq j$ while $(v_i, v_i) = 1$. We wish to show that if T is unitary, then $\{v_1T, \ldots, v_nT\}$ is also an orthonormal basis of V. But $(v_iT, v_jT) = (v_i, v_j) = 0$ for $i \neq j$ and $(v_iT, v_iT) = (v_i, v_i) = 1$, thus indeed $\{v_1T, \ldots, v_nT\}$ is an orthonormal basis of V.

On the other hand, if $T \in A(V)$ is such that both $\{v_1, \ldots, v_n\}$ and $\{v_1T, \ldots, v_nT\}$ are orthonormal bases of V, if $u, w \in V$ then

$$u = \sum_{i=1}^{n} \alpha_i v_i, w = \sum_{i=1}^{n} \beta_i v_i.$$

whence by the orthonormality of the v_i's,

$$(u, w) = \sum_{i=1}^{n} \alpha_i \beta_i.$$

However,

$$uT = \sum_{i=1}^{n} \alpha_i v_i T \ \textit{and}\ wT = \sum_{i=1}^{n} \beta_i v_i T$$

whence by the orthonormality of the v_iT's,

$$(uT, wT) = \sum_{i=1}^{n} \alpha_i \beta_i = (u, w).$$

Hence T is unitary. $\square$

Lemma 6.2.5. *If $T \in A(V)$ then given any $v \in V$ there exists an element $w \in V$, depending on v and T, such that $(uT, v) = (u, w)$ for all $u \in V$. This element w is uniquely determined by v and T.*

Proof. To prove the lemma, it is sufficient to exhibit a $w \in V$ which works for all the elements of a basis of V.

Let $\{u_1, \ldots, u_n\}$ be an orthonormal basis of V; we define

$$w = \sum_{i=1}^{n} \overline{(u_iT, v)} u_i.$$

An easy computation shows that $(u_i, w) = (u_iT, v)$, hence the element w has the desired property. That w is unique can be seen as follows: Suppose that $(uT, v) = (u, w_1) = (u, w_2)$; then $(u, w_1 - w_2) = 0$ for all $u \in V$ which forces, on putting $u = w_1 - w_2, w_1 = w_2$. $\square$

Definition 6.2.6. If $T \in A(V)$ then the Hermitian adjoint of T, written as T^*, is defined by $(uT, v) = (u, vT^*)$ for all $u, v \in V$.

Lemma 6.2.7. *If $T \in A(V)$ then $T^* \in A(V)$. Moreover,*

1. $(T^*)^* = T$;

2. $(S + T)^* = S^* + T^*$;

3. $(\lambda S)^* = \lambda S^*$;

4. $(ST)^* = T^*S^*$*;*

for all $S, T \in A(v)$ *and all* $\lambda \in F$*.*

Proof. We must first prove that T^* is a linear transformation on V. If u, v, w are in V, then $(u, (v+w)T^*) = (uT, v+w) = (uT, v) + (uT, w) = (u, vT^*) + (u, wT^*) = (u, vT^* + wT^*)$, in consequence of which $(v+w)T^* = vT^* + wT^*$.

Similarly, for $\lambda \in F, (u, (\lambda v)T^*) = (uT, \lambda v) = \lambda(uT, v) = \lambda(u, vT^*) = (u, \lambda(vT^*))$, whence $(\lambda v)T^* = \lambda(vT^*)$. Hence T^* is a linear transformation on V.

To see that $(T^*)^* = T$ notice that $(u, v(T^*)^*) = (uT^*, v) = \overline{(v, uT^*)} = \overline{(vT, u)} = (u, vT)$ for all $u, v \in V$ whence $v(T^*)^* = vT$ which implies that $(T^*)^* = T$. We leave the proofs of $(S+T)^* = S^* + T^*$ and of $(\lambda T)^* = \lambda T$ to the reader.

Finally, $(u, v(ST)^*) = (uST, v) = (uS, VT^*) = (u, vT^*S^*)$ for all $u, v \in V$; this forces $v(ST)^* = vT^*S^*$ for every $v \in V$ which results in $(ST)^* = T^*S^*$. $\square$

Lemma 6.2.8. $T \in A(V)$ *is unitary if and only if* $TT^* = 1$.

Proof. If T is unitary, then for all $u, v \in V$, $(u, vTT^*) = (uT, vT) = (u, v)$ hence $TT^* = 1$. On the other hand, if $TT^* = 1$, then $(u, v) = (u, vTT^*) = (uT, vT)$, which implies that T is unitary. $\square$

Note that a unitary transformation is nonsingular and its inverse is just its Hermitian adjoint. Note, too, that from $TT^* = 1$ we must have that $T^*T = 1$.

Theorem 6.2.9. *If* $\{v_1, \ldots, v_n\}$ *is an orthonormal basis of* V *and if the matrix of* $T \in A(V)$ *in this basis is* (α_{ij}) *then the matrix of* T^* *in this basis is* (β_{ij})*, where* $\beta_{ij} = \alpha_{ji}$

Proof. Since the matrices of T and T^* in this basis are, respectively, (α_{ij}) and (β_{ij}), then

$$v_iT = \sum_{i=1}^{n} \alpha_{ij}v_j \textit{ and } v_iT^* = \sum_{i=1}^{n} \beta_{ij}v_j$$

Now

$$\beta_{ij} = (v_i T^*, v_j) = (v_i, v_j T) = (v_i, \sum_{i=1}^{n} \alpha_{jk} v_k) = \overline{\alpha}_{ji}$$

by the orthonormality of the v_i's. This proves the theorem. □

Definition 6.2.10. $T \in A(V)$ is called self-adjoint or Hermitian if $T^* = T$.

If $T^* = -T$ we call skew-Hermitian. Given any $S \in A(V)$,

$$S = \frac{S + S^*}{2} + i(\frac{S - S^*}{2i})$$

and since $\frac{S+S^*}{2}$ and $\frac{S-S^*}{2i}$ are Hermitian, $S = A + iB$ where both A and B are Hermitian.

Theorem 6.2.11. *If $T \in A(V)$ is Hermitian, then all its characteristic roots are real.*

Proof. Let λ be a characteristic root of T. Then there is a $\neq 0$ in V such that $vT = \lambda v$. Now $\lambda(v, v) = (\lambda v, v) = (vT, v) = (v, vT^*) = (v, vT) = (v, \lambda v) = \lambda(v, v)$; since $(v, v) \neq 0$ we are left with $\lambda = \lambda$, hence λ is real. □

Lemma 6.2.12. *If $S \in A(V)$ and if $vSS^* = 0$, then $vS = 0$.*

Proof. Consider (uSS^*, v); since $USS^* = 0, 0 = (vSS^*, v) = (vS, v(S^*)^*) = (vS, vS)$. In an inner-product space, this implies that $vS = 0$. □

Corollary 6.2.13. *If T is Hermitian and $vT^k = 0$ for $k > 1$ then $vT = 0$.*

Proof. We show that if $vT^{2m} = 0$ then $vT = 0$; for if $S = T^{2m-1}$, then $S^* = S$ and $SS^* = T^{2m}$, whence $(vSS^*, v) = 0$ implies that $0 = vS = vT^{2m-1}$. Continuing down in this way, we obtain $T = 0$. If $vT^k = 0$, then $vT^{2m} = 0$ for $2m > k$, hence $vT = 0$. □

Definition 6.2.14. $T \in A(V)$ is said to be normal if $TT^* = T^*T$.

Lemma 6.2.15. *If N is a normal linear transformation and if $vN = 0$ for $v \in V$, then $vN^* = 0$.*

Proof. Consider (vN^*, N^*); by definition, $(vN^*, vN^*) = (vN^*N, v) = (vNN^*, v)$, since $NN^* = N^*N$. However, $vN = 0$, whence, certainly, $vNN^* = 0$. In this way we obtain that $(vN^*, vN^*) = 0$, forcing $vN^* = 0$. □

Corollary 6.2.16. *If λ is a characteristic root of the normal transformation N and if $vN = \lambda v$ then $vN^* = \lambda v$.*

Proof. Since N is normal, $NN^* = N^*N$, therefore, $(N-\lambda)(N-\lambda)^* = (N-\lambda)(N^* - \lambda) = NN^* - \lambda N^* - \lambda N + \lambda = N^*N - \lambda N^* - \lambda N + \lambda\lambda = (N^* - \lambda)(N^* - \lambda)(N - \lambda) = (N-\lambda)^*(N-\lambda)$, that is to say $n - \lambda$ is normal. Since $v(N-\lambda) = 0$ by the normality of $N - \lambda$, from the lemma, $v(N-\lambda)^* = 0$, hence $vN^* = \lambda v$. □

Corollary 6.2.17. *If T is unitary and if λ is a characteristic root of T, then $|\lambda| = 1$.*

Proof. Since T is unitary it is normal. Let λ be a characteristic root of T and suppose that $vT = \lambda v$ with $v \neq$ in V. By above Corollary, $vT^* = \lambda v$, thus $v = vTT^* = \lambda T^* = \lambda\lambda v$ since $TT^* = 1$. Thus we get $\lambda\lambda = 1$, which, of course, says that $|\lambda| = 1$. □

Lemma 6.2.18. *If N is normal and if $vN^k = 0$, then $vN = 0$.*

Proof. Let $S = NN^*$; S is Hermitian, and by the normality of $N, vS^k = v(NN^*)^k = vN^k(N^*)^k = 0$. By the corollary to Lemma 6.10.6, we deduce that $vS = 0$, that is to say, $vNN^* = 0$. From this, we get $vN = 0$. □

Corollary 6.2.19. *If N is normal and if for $\lambda \in F, v(N-\lambda)^k = 0$, then $vN = \lambda v$.*

Proof. From the normality of N it follows that N is normal, whence by applying the lemma just proved to $N - \lambda$ we obtain the corollary. $\square$

Lemma 6.2.20. *Let N be a normal transformation and suppose that λ and μ are two distinct characteristic roots of N. If v, w are in V and are such that $vN = \lambda v, wN = \mu w$, then $(v, w) = 0$.*

Proof. We compute (vN, w) in two different ways. As a consequence of $vN = \lambda v, (vN, w) = (\lambda v, w) = \lambda(v, w)$. From $wN = \mu w$, using above Lemma, we obtain that $wN^* = \overline{\mu} w$, whence $(vN, w) = (v, wN^*) = (v, \overline{\mu} w) = \mu(v, w)$. Comparing the two computations gives us $\lambda(v, w) = \mu(v, w)$ and since $\lambda \neq \mu$, this results in $(v, w) = 0$. $\square$

Theorem 6.2.21. *If N is a normal linear transformation on V, then there exists an orthonormal basis, consisting of characteristic vectors of N, in which the matrix of N is diagonal.*

Proof. Let N be normal and let $\lambda_1, \ldots, \lambda_n$ be the distinct characteristic roots of N. By the above corollary, we can decompose $V = V_1 \oplus \cdot \oplus V_k$ where every $v_i \in V_i$, is annihilated by $(N - \lambda_i)^{n_i}$. From this, we get, V_i consists only of characteristic vectors of N belonging to the characteristic root λ_i. The inner product of V induces an inner product on V_i and hence we can find a basis of V_i orthonormal relative to this inner product. By above Lemma, elements lying in distinct V_i's are orthogonal. Thus putting together the orthonormal bases of the V_i's provides us with an orthonormal basis of V. This basis consists of characteristic vectors of N, hence in this basis the matrix of N is diagonal. $\square$

1. A change of basis from one orthonormal basis to another is accomplished by a unitary transformation.

2. In a change of basis the matrix of a linear transformation is changed by conjugating by the matrix of the change of basis.

Corollary 6.2.22. *If T is a unitary transformation, then there is an orthonormal basis in which the matrix of T is diagonal.*

Corollary 6.2.23. *If T is a Hermitian linear transformation, then there exists an orthonormal basis in which the matrix of T is diagonal.*

Lemma 6.2.24. *The normal transformation N is*

1. *Hermitian if and only if its characteristic roots are real.*

2. *Unitary if and only if its characteristic roots are all of absolute value 1.*

Proof. We argue using matrices. If N is Hermitian, then it is normal and all its characteristic roots are real. If N is normal and has only real characteristic roots, then for some unitary matrix U, UNU^{-1} $UNU^* = D$, where D is a diagonal matrix with real entries on the diagonal. Thus $D^* = D$; since $D^* = (UNU^*)^* = UN^*U^*$, the relation D^* D implies $UN^*U^* = UNU^*$, and since U is invertible we obtain N^* N. Thus N is Hermitian.

If A is any linear transformation on V, then tr (AA^*) can be computed by using the matrix representation of A in any basis of V. We pick an orthonormal basis of V; in this basis, if the matrix of A is $[\alpha_{ij}]$ then that of A^* is (βij) where $\beta_{ij} = \overline{\alpha}_{ji}$ A simple computation then shows that tr $(AA^*) = \sum_{i,j} |\alpha_{ij}|^2$ and this is 0 if and only if each $\alpha_{ij} = 0$, that is, if and only if $A = 0$. In a word, tr $(AA^*) = 0$ if and only if $A = 0$. $\square$

Lemma 6.2.25. *If N is normal and $AN = NA$, then $AN^* = N^*A$.*

Proof. We want to show that $X = AN^* - N^*A$ is 0; what we shall do is prove that $tr\ XX^* = 0$, and deduce from this that $X = 0$. Since N commutes with A and with N^*, it must commute with $AN^* - N^*A$, thus $XX^* = (AN^* - N^*A)(NA^* - A^*N) = (AN^* - N^*A)NA^* - (AN^* - N^*A)A^*N = N\{(AN^* - N^*A)A^*\} - \{(AN^* - N^*A)A^*\}N$. Being of the form $NB - BN$, the trace of XX^* is 0. Thus $X = 0$, and $AN^* = N^*A$.

□

Lemma 6.2.26. *The Hermitian linear transformation T is nonnegative. (positive) if and only if all of its characteristic roots are nonnegative (positive).*

Proof. Suppose that $T \geq 0$; if λ is a characteristic root of T, then $vT = \lambda v$ for some $v \neq 0$. Thus $0 \leq (vT, v) = (\lambda v, v) = \lambda(v, v)$; since $(v, v) > 0$ we deduce that $\lambda \geq 0$.

Conversely, if T is Hermitian with nonnegative characteristic roots, then we can find an orthonormal basis $\{v_1, \ldots, v_n\}$ consisting of characteristic vectors of T. For each v_i, $v_iT = \lambda_i v_i$, where $\lambda_i \geq 0$. Given $v \in V, v = \sum \alpha_i v_i$ hence $vT = \sum \alpha_i v_i T = \sum \lambda_i \alpha_i v_i$. But $(vT, v) = (\sum \lambda_i v_i, \sum \alpha_i v_i) = \sum \lambda_i \alpha_i \overline{\alpha_i}$ by the orthonormality of v_i's. Since $\lambda_i \geq 0$ and $\alpha_i \overline{\alpha_i} \geq 0$. We get $(vT, v) \geq 0$ hence $T \geq 0$. □

Lemma 6.2.27. *$T \geq 0$ if and only if $T = AA^*$ for some A.*

Proof. We first show that $AA^* \geq 0$, Given $v \in V, (vAA^*, V) = (vA, vA) \geq 0$, hence $AA^* \geq 0$.

On the other hand, if $T \geq 0$ we can find a unitary matrix U such that

$$UTU^* = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

where each λ_i is a characteristic root of T, hence each $\lambda_i \geq 0$. Let

$$S = \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix}$$

since each $\lambda_i \geq 0$, each $\sqrt{\lambda_i}$ is real, whence S is Hermitian. Therefore, U^*SU is Hermitian, but

$$(U^*SU)^2 = U^*S^2U = U^* \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix} U = T$$

We have represented T in the form AA^*, where $A = U^*SU$. Notice that we have actually proved a little more; namely, if in constructing S above, we had chosen the nonnegative λ_i for each λ_i, then S, and U^*SU, would have been nonnegative. Thus $T \geq 0$ is the square of a non- negative linear transformation; that is, every $T \geq 0$ has a nonnegative square root. This nonnegative square root can be shown to be unique □

6.3 Real Quadratic Forms

Definition 6.3.1. *Two real symmetric matrices A and B are congruent if there is a nonsingular real matrix T such that $B = TAT'$.*

Lemma 6.3.2. *Congruence is an equivalence relation.*

Proof. Let us write, when A is congruent to B, $A \cong B$.

1. $A \cong A$ for $A = |A|'$.
2. If $A \cong B$ then $B = TAT'$ where T is nonsingular, hence $A = SBS'$ where $S = T^{-1}$. Thus $B \cong A$.

3. If $A \cong B$ and $B \cong C$ then $B = TAT'$ while $C = RBR'$, hence $C = RTAT'R' = (RT)A(RT)'$, and so $A \cong C$.

Since the relation satisfies the defining conditions for an equivalence relation, the lemma is proved. □

Theorem 6.3.3. *Given the real symmetric matrix A there is an invertible matrix T such that*

$$UTU^* = \begin{pmatrix} I_r & & \\ & -I_s & \\ & & 0_t \end{pmatrix}$$

where I_r and I_s are respectively the $r \times r$ and $s \times s$ unit matrices and where 0, is the $t \times t$ zero-matrix. The integers $r + s$, which is the rank of A, and $r - s$, which is the signature of A, characterize the congruence class of A. That is, two real symmetric matrices are congruent if and only if they have the same rank and signature.

Proof. Since A is real symmetric its characteristic roots are all real; let $\lambda_1, \cdots, \lambda_r$ be its positive characteristic roots, $-\lambda_{r+1}, \cdots, -\lambda_{r+s}$ its negative. We can find a real orthogonal matrix C such that

$$CAC^{-1} = CAC' = \begin{pmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & & \\ & & \lambda_r & & & & \\ & & & \frac{1}{\sqrt{-\lambda_{r+1}}} & & & \\ & & & & \ddots & & \\ & & & & & \frac{1}{\sqrt{-\lambda_{r+s}}} & \\ & & & & & & 0_t \end{pmatrix}$$

where $t = n - r - s$. Let D be the real diagonal matrix shown above.

$$D = \begin{pmatrix} \frac{1}{\sqrt{\lambda_1}} & & & & & & \\ & \ddots & & & & & \\ & & \frac{1}{\sqrt{\lambda_r}} & & & & \\ & & & \frac{1}{\sqrt{-\lambda_{r+1}}} & & & \\ & & & & \ddots & & \\ & & & & & \frac{1}{\sqrt{-\lambda_{r+s}}} & \\ & & & & & & I_t \end{pmatrix}$$

A single composition shows that

$$DCAC'D^{'} = \begin{pmatrix} I_r & & \\ & -I_s & \\ & & 0_t \end{pmatrix}$$

Thus there is a matrix of the required form in the congruence class of A. Our task is now to show that this is the only matrix in the congruence class of A of this form, or, equivalently, that

$$L = \begin{pmatrix} I_r & & \\ & -I_s & \\ & & 0_t \end{pmatrix} \text{ and } M = \begin{pmatrix} I_{r'} & & \\ & -I_{s'} & \\ & & 0_{t'} \end{pmatrix}$$

are congruent only if $r = r', s = s^{'}$ and $t = t^{'}$. Suppose that $M = TLT^{'}$ where T is invertible and so the rank of M equals that of L; since the rank of M is $n - t^{'}$ while that of L is $n - t$ we get $t = t^{'}$.

Suppose that $r < r^{'}$; since $n = r + s + t = r^{'} + s^{'} + t^{'}$, and since $t = t'$, we must have $s > s^{'}$. Let U be the subspace of $F^{(n)}$ of all vectors having the first r and last t coordinates 0; U is s-dimensional and for $u \neq 0$ in U, $(uL, u) < 0$.

Let W be the subspace of $F^{(n)}$ for which the $r'+1, \cdots, r'+s'$ components are all 0; on $W, (wM, w) \geq 0$ for any $w \in W$. Since T is invertible, and since W is $(n-s')$-dimensional, WT is $(n-s')$-dimensional. For $w \in W$, $(wM, w) \geq 0$; hence $(wTLT', w) \geq 0$; that is, $(wTL, wT) \geq 0$. Therefore, on $WT, (wTL, wT) \geq 0$ for all elements.

Now $dim(WT) + dimU = (n-s') + r = n + s - s' > n$ and so $WT \cap U \neq 0$. This, however, is nonsense, for if $x \neq 0 \in WT \cap U$, on one hand, being in U, $(xL, x) < 0$, while on the other, being in WT, $(xL, x) \geq 0$. Thus $r = r'$ and so $s = s'$. The rank, $r+s$, and signature, rs, of course, determine r, s and so $t = (n-r-s)$, whence they determine the congruence class. $\square$