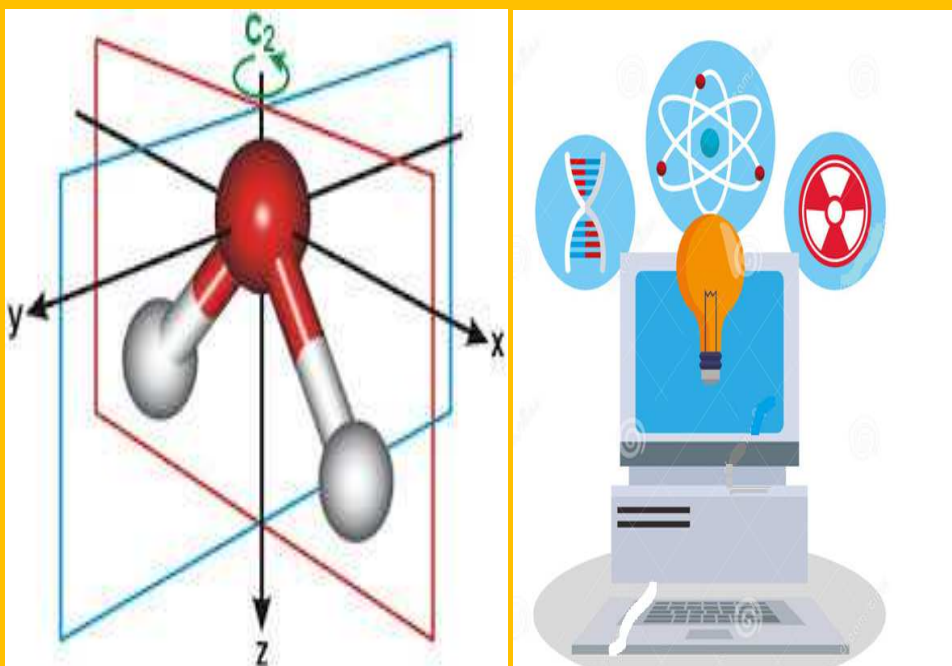




MSCCH-504

M. Sc. I Semester

**GROUP THEORY, INSTRUMENTATION
CHEMISTRY & COMPUTER FOR CHEMIST**



**SCHOOL OF SCIENCES
DEPARTMENT OF CHEMISTRY
UTTARAKHAND OPEN UNIVERSITY**

MSCCH-504

**GROUP THEORY,
INSTRUMENTATION CHEMISTRY &
COMPUTER FOR CHEMIST**



**SCHOOL OF SCIENCES
DEPARTMENT OF CHEMISTRY
UTTARAKHAND OPEN UNIVERSITY**

Phone No. 05946-261122, 261123

Toll free No. 18001804025

Fax No. 05946-264232, E. mail info@uou.ac.in

<http://uou.ac.in>

Expert Committee**Prof. B.S.Saraswat**

Department of Chemistry
Indira Gandhi National Open University
Maidan Garhi, New Delhi

Prof. A.K. Pant

Department of Chemistry
G.B.Pant Agriculture, University
Pantnagar

Prof. A. B. Melkani

Department of Chemistry
DSB Campus,
Kumaun University, Nainital

Prof. Diwan S Rawat

Department of Chemistry
Delhi University
Delhi

Dr. Hemant Kandpal

Assistant Professor
School of Health Science
Uttarakhand Open University, Haldwani

Dr. Charu C. Pant

Academic Consultant
Department of Chemistry
Uttarakhand Open University,

Board of Studies**Prof. A.B. Melkani**

Department of Chemistry
DSB Campus, Kumaun University
Nainital

Prof. G.C. Shah

Department of Chemistry
SSJ Campus, Kumaun University
Nainital

Prof. R.D.Kaushik

Department of Chemistry
Gurukul Kangri Vishwavidyalaya
Haridwar

Prof. P.D.Pant

Director I/C, School of Sciences
Uttarakhand Open University
Haldwani

Dr. Shalini Singh

Assistant Professor
Department of Chemistry
School of Sciences
Uttarakhand Open University, Haldwani

Dr. Charu C. Pant

Academic Consultant
Department of Chemistry
School of Science
Uttarakhand Open University,

Programme Coordinator**Dr. Shalini Singh**

Assistant Professor
Department of Chemistry
Uttarakhand Open University
Haldwani

Unit Written By	Unit No.
1. Dr. Priyanka Tiwari Department of Chemistry M.B. Govt. P.G. College, Haldwani	01
2. Dr. Mahesh Chandra Vishwakarma Department of Chemistry, L.S.M. Govt. PG. College, Pithoragarh	02
3. Dr. Renu Loshali Department of Chemistry, Govt. Degree College, Bazpur	03
4. Dr. Jyoti Singh Department of Applied Sciences Amarpali Group of Institute Haldwani, Nainital	04
5. Dr. Ajay Pali Department of Applied Sciences Amarpali Group of Institute Haldwani, Nainital	05
6. Dr. Devendra Singh Dhama Department of Chemistry SSJ Campus Kumaun University, Almaora	06
7. Mr. Gaurav Joshi Assistant Professor Department of Computer Science MB Govt. PG College, Haldwani	07 & 11
8. Mr. Shekhar Kumar Assistant Professor Department of Computer Science MB Govt. PG College, Haldwani	08, 09 & 10

Course Editors

- 1. Prof. (Dr.) Narendra Singh Bhandari**
Member,
Uttarakhand Public Service Commission
Gurukul Kangri, Haridwar
- 2. Dr. Neeta Joshi**
Associate Professor
Department of Chemistry,
PNG Govt. PG. College, Ramnagar
- 3. Mr. Harsh Vardhan Pant**
Department of Computer Science
Amrapali Group of Institute, Haldwani

Title : **Group Theory, Instrumentation Chemistry & Computer for Chemist**
ISBN No. :
Copyright : **Uttarakhand Open University**
Edition : **2021**

Published by : **Uttarakhand Open University, Haldwani, Nainital- 263139**

CONTENTS

BLOCK 1: GROUP THEORY IN CHEMISTRY AND X-RAY DIFFRACTION

Unit 1: Symmetry and Group Theory	01-28
Unit 2: Matrix	29-58
Unit 3: X-ray Diffraction Methods	59-82

BLOCK 2: CHROMATOGRAPHIC AND RADIO ANALYTICAL METHODS

Unit 4: Chromatographic methods	83-112
Unit 5: Gas Chromatography	113-143
Unit 6: Radio Analytical Methods	144-159

BLOCK 3: INTRODUCTION TO COMPUTERS

Unit 1 Data Analysis	160-188
Unit 2 History of development of computers	189-207
Unit 3 Introduction to Computers and Computing	208-255

BLOCK 4 : PROGRAMMING IN CHEMISTRY

Unit 4 Programming in Chemistry	256-270
Unit 5 Use of Computer Programmes	271- 304

UNIT 1- SYMMETRY AND GROUP THEORY

CONTENTS:

- 1.0 Objectives
- 1.1 Introduction
- 1.2 Element of symmetry
 - 1.1.1 Identity (E)
 - 1.1.2 Axis of symmetry (C_n)
 - 1.1.3 Plane of symmetry
 - 1.1.4 Point of symmetry
 - 1.1.5 Identity Operation
- 1.2 A complete Set of symmetry operations as mathematical group
 - 1.2.1 Abelian group
 - 1.2.2 Non-Abelian Group
- 1.3 Generator
- 1.4 Summary
- 1.5 Terminal questions

1.0 OBJECTIVES

In this chapter we will discuss about following:

- Symmetry and symmetry operations
- Basic definition of group theory
- A complete set of symmetry operations as mathematical group
- To discuss about Generator
- Flow chart for point group determination
- Characteristics of symmetry operations
- Determination of point groups

1.1 INTRODUCTION

Symmetry is a phenomenon of geometrical property of the world in which we live. In nature many type of flowers and plants, snowflakes, insects, certain fruits vegetables and various microscopic organism exhibit characteristic symmetry (Fig. 1.1). Symmetry concepts are extremely useful in Chemistry. On the basis of symmetry we can predict infrared spectra, type of orbital used in bonding, predict the optical activity and interpret electronic spectra and study of molecular properties.

Symmetry helps us understand molecular structure, some chemical properties, and characteristics of physical properties (spectroscopy) – used with group theory to predict vibrational spectra for the identification of molecular shape, and as a tool for understanding electronic structure and bonding.

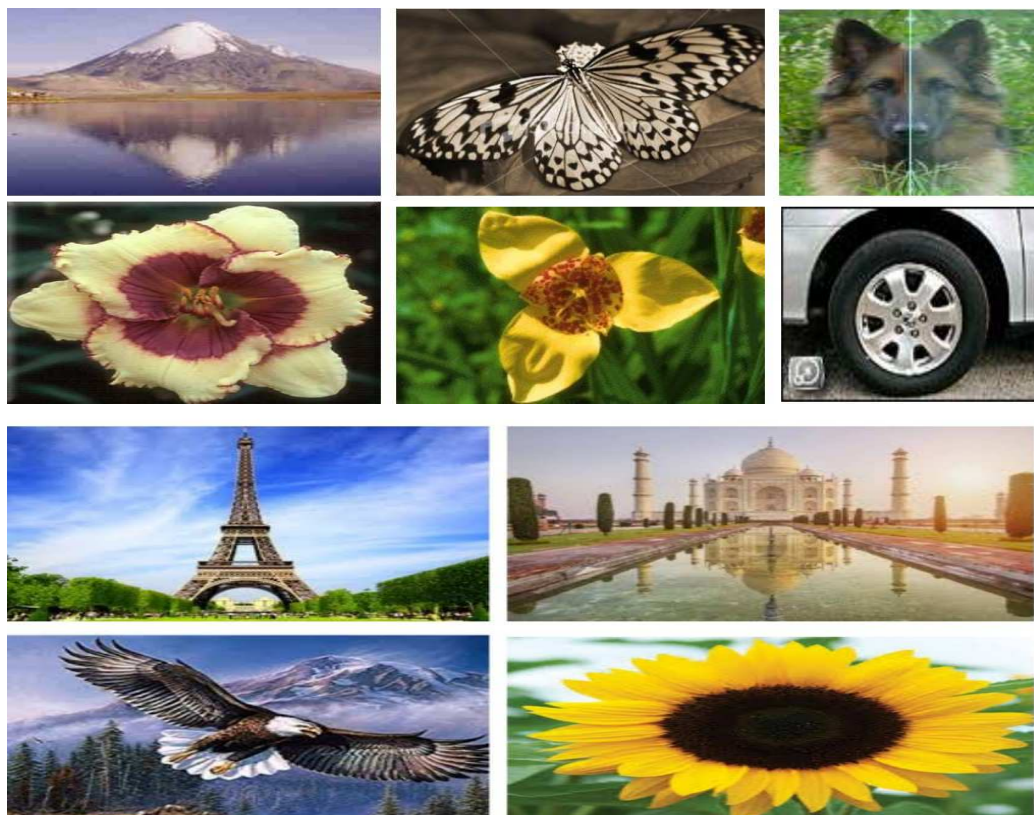


Figure-1.1

The term **symmetry** is derived from the Greek word “**symmetria**” which means “measured together”

A **group** consists of a set of symmetry elements (and associated symmetry operations) that completely describe the symmetry of an object.

Point groups have symmetry about a single point at the center of mass of the system.

1.2 ELEMENT OF SYMMETRY

Symmetry elements are geometric entities as point, line, plane of molecule about which a *symmetry operations* as rotations, reflections, inversions and improper rotations can be performed. In a point group, all symmetry elements must pass through the center of mass (the point). A **symmetry operation** is the action that produces an object identical (i.e. **position indistinguishable from the original position**) to the initial object (even though atoms and bonds may have been moved).

In generally we can say that the element of symmetry is geometrical tools of symmetry or entity of symmetry tools i.e. point, line, plane in the molecule. The actual reflection, rotation or inversion is called the symmetry operation. A molecule possesses a symmetrical element, which is unchanged in appearance after applying the symmetry operation, correspond to the element. The element of symmetry divided into five operations.

1. Identity symmetry (E)
2. Axis of symmetry (C_n where $n=360/\text{angle of rotation in } \theta$)
3. Plane of symmetry (σ)
4. Improper axis of symmetry (S_n and then $\perp C_n$)
5. Point of symmetry (i)

1.1.1 Identity (E):

The operation, which brings back the molecule to the original orientation, is called identity operation. It is represented as **E** from the German word Einhart meaning unity. The identity operation in effect means doing nothing on the molecule and hence does not seem to be of much consequence but it is important lies in considering the molecule as a group. Each molecule has this type of symmetry.

1.1.2 Axis of symmetry (C_n):

The symmetry element in which the molecule can represent the identical image after rotation of molecule by $360^\circ/n$ with respect to any imaginary axis called as axis of symmetry or rotational axis.

$$n = \frac{360}{\theta}$$

- Where n is always an integer. This axis defined is an n -fold rotation axis, C_n .
- In water there is a C_2 axis so we can perform a 2-fold (180°) rotation to get the identical arrangement of atoms Fig 1.2
- In BF_3 there is a C_3 axis so we can perform 3-fold (120°) rotations to get identical arrangement of atoms Fig 1.2

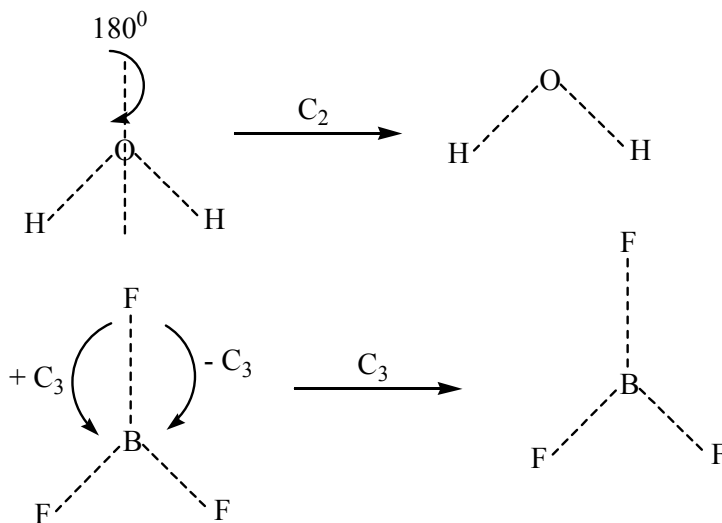


Figure 1.2

- Rotations are considered positive in the counter-clockwise direction.
- Each possible rotation operation is assigned using a superscript integer m of the form C_n^m .
- The rotation C_n^n is equivalent to the identity operation (nothing is moved)

Classification of axis of symmetry

Many molecules have more than one C_n axis

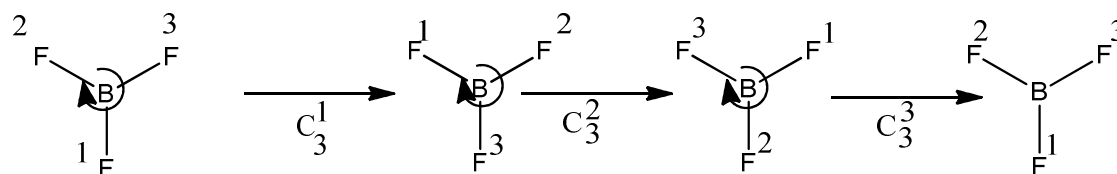
It can be divided into two different types:

- a) Principal axis of symmetry
- b) Secondary/ subsidiary axis of symmetry

(a) Principal axis of symmetry

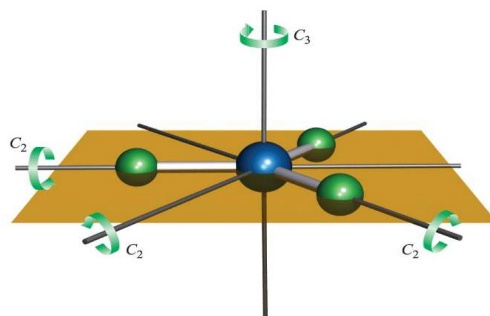
The *Principal axis* in an object is the highest order rotation axis i.e. having largest value of n. It is usually easy to identify the principle axis and this is typically assigned to the z-axis if we are using Cartesian coordinates. If there are more than one axis of same order, then the axis passing through maximum number of atoms is called Principle axis.

Example: In BF_3 molecule three fold rotation axis (C_3) and two fold rotational axis (C_2) are present. The three fold axis is coincident with the perpendicular to the plane and the rotation angle is $\frac{360}{3}=120^\circ$. After two C_3 operations molecule comes into identity operation. Three-fold axis have higher order than two fold axis, therefore C_3 is the principal axis of BF_3 molecule **Figure 1.3 and Figure 1.4**. In Benzene there is six fold axes of symmetry C_6 as principal axes and it has six axes of two fold symmetry three passing through centre of Benzene and two opposite carbon atoms and three passing through centre of symmetry and centre of two opposite edges **Figure 1.5**. In NH_3 molecule only three-fold axis of symmetry **Figure 1.6** and in water only two fold symmetry is present **Figure 1.7**.



$$C_3^1=120^\circ, C_3^2=240^\circ, C_3^3=E=360^\circ$$

Figure: 1.3



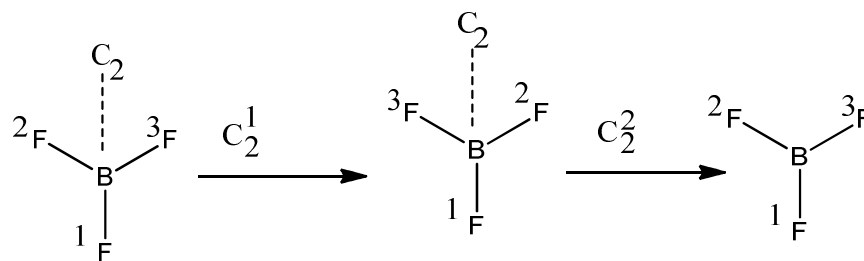


Figure: 1.4

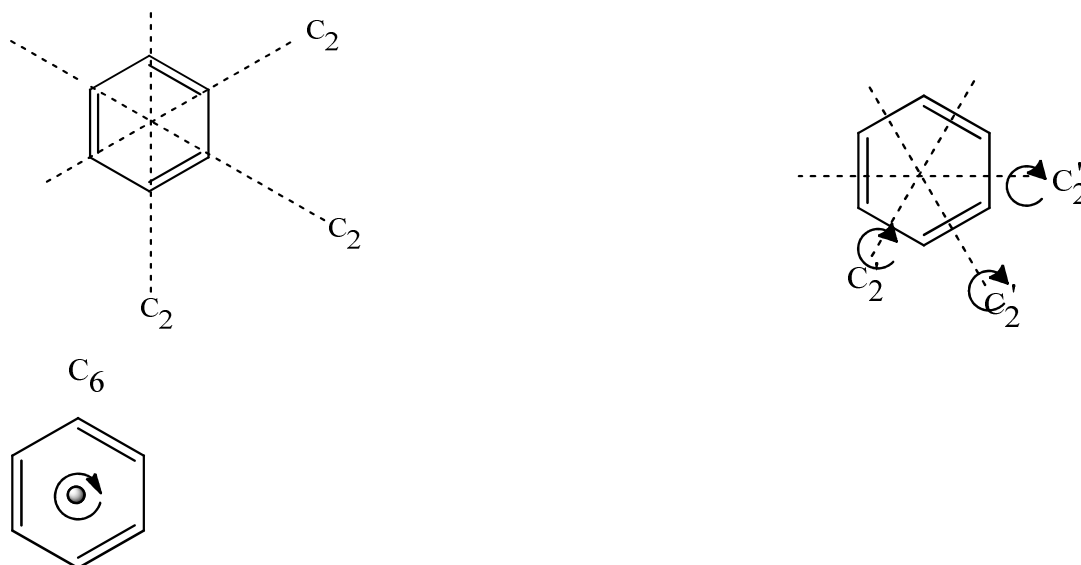


Figure: 1.5

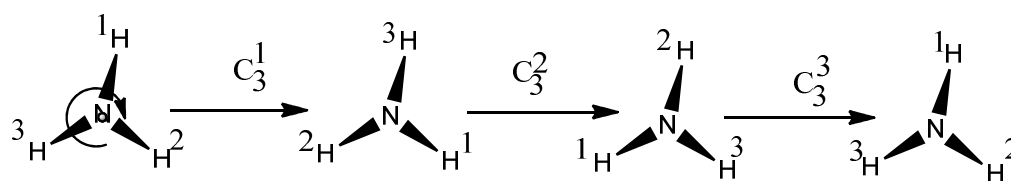


Figure: 1.6

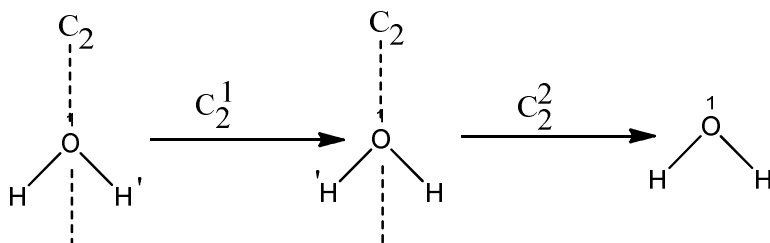


Figure: 1.7

If the order of rotation is similar, the principal axis is that in which the higher number of atoms present across in particular axis.

Example: In ethene molecule there are three C_2 axis are present shown in Figure C_2' is the principal axis of ethene molecule the two carbon atom present across this axis, while another two axis have no atom is present **Figure 1.8**. In naphthalene there are three C_2 axis are present where C_2' is the principal axis shown in **Figure: 1.8**

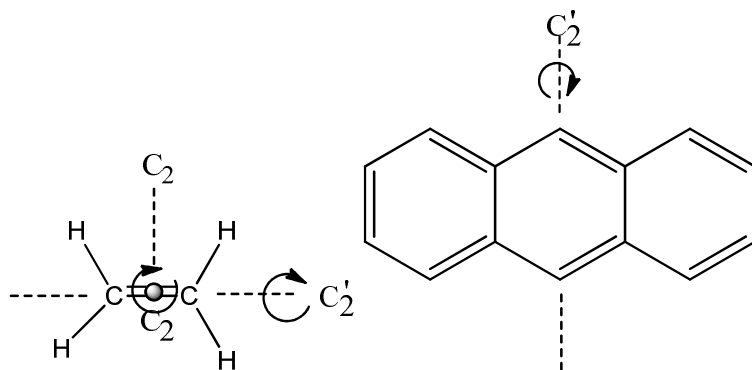


Figure: 1.8

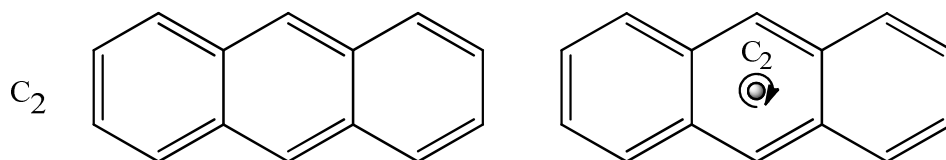
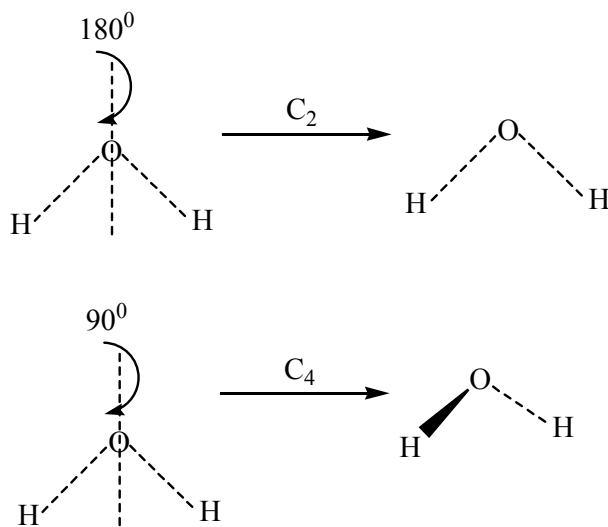


Figure: 1.9



However, rotation by 90° about the same axis does not give back the identical molecule. Therefore H_2O does not possess a C_4 symmetry axis.

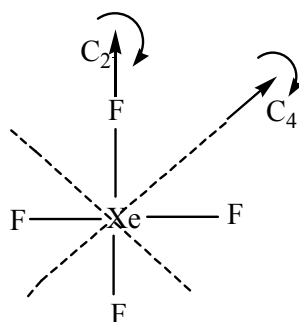


Figure: 1.10

XeF_4 is square planar. It has four different C_2 axes. A C_4 axis out of the page is called principal axis.

(b) Subsidiary or secondary axes of Symmetry:

Lower fold of rotation axis (lower order of axis of symmetry) is the secondary or subsidiary axis of the molecule.

Example: In BF_3 molecule C_3 axis is principal axis and C_2 axis is secondary axis. In ethane molecule C_2' is the principal axis and C_2 is subsidiary axis. In NH_3 molecule three-fold of axis present that is principal axis, here subsidiary axis is absent.

1.1.3 Plane of symmetry:

Plane of symmetry have generate only one operation, on repeating the reflection operation molecule comes into initial structure that is $\sigma^2 = E$. The imaginary plane bisects the molecule into two halves which are mirror image to each other.

$$\sigma^n = E \quad (\text{if } n \text{ is even})$$

$$\sigma^n = \sigma \quad (\text{if } n \text{ is odd})$$

The molecules possessing only plane of symmetry but no rotation axis other than $C_1 (=E)$, there can be no definition vertical and horizontal planes and the symbol for this plane is simply σ . The point group for this type of molecule is C_s .

Rotational axis along with plane, the planes can be classified into three types.

- a) Vertical Plane (σ_v : $\sigma \parallel$ principal axis)
- b) Horizontal plane (σ_h : $\sigma \perp$ principal axis)
- c) Dihedral plane (σ_d bisect plane with respect to two C_2 axis)

1.1.4 (a) Vertical plane:

The plane of operation undergoes parallel with respect to principal axis is vertical plane of symmetry and denoted by σ_v . The subscript “v” in σ_v , indicates a vertical plane of symmetry. This indicates that the mirror plane passing through the principle axis and one of the subsidiary axis (if present).

1.1.3 (b) Horizontal plane:

The plane reflection is perpendicular with respect to principal axis is called horizontal plane and denoted by σ_h .

Example: In bent molecule of H_2O have $\sigma_{v(yz)}$ and $\sigma_{v(xz)}$ C plane, in $\sigma_{v(xz)}$ all the atoms of H and O are bisect in the same plane and in $\sigma_{v(yz)}$, two hydrogen atoms are reflected to each other.

Example: In pyramidal NH_3 molecule, three plane of reflections are present. Which are parallel with respect to C_3 principal axis, i.e. there are three vertical planes are present.

Example: In trigonal planar BF_3 molecule have three reflection of plane which are parallel with respect to three fold (C_3) axis. One reflection of plane which is perpendicular with respect to principal axis.

Example: In square planar complex $[PtCl_4]^{2-}$ have a horizontal plane, two vertical plane and two dihedral plane (which bisect the two C_2 axis) shown in **figure: 1.11, and 1.12**

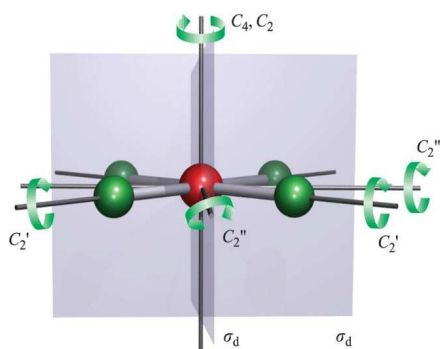


Figure: 1.11

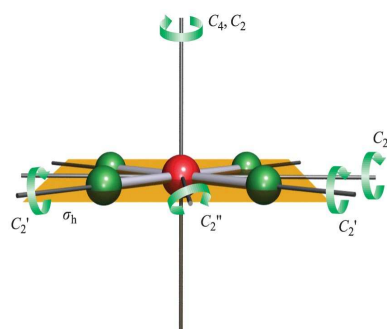


Figure 1.12

1.1.3 (c) Dihedral plane:

The plane passing through the principal axis but passing between two subsidiary axis is called dihedral plane and denoted by σ_d .

Example: In allene compound there are two C_2 axis which are different type one of them have three atoms are passed though the axis and principle axis. There are two dihedral angle which are bisecting to two C_2 axis.

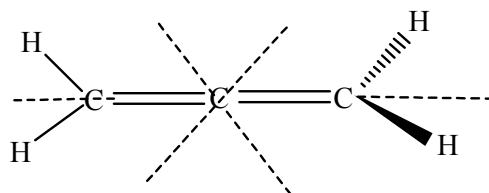


Figure: 1.13

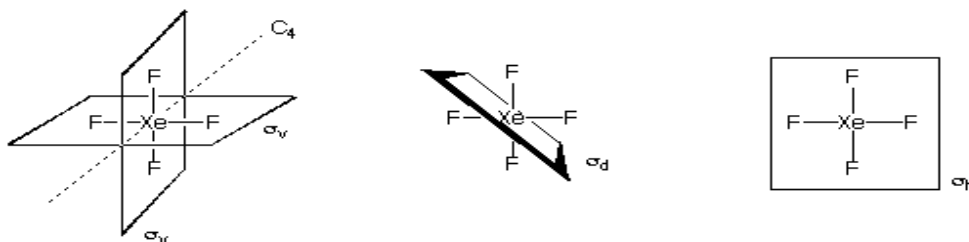


Figure: 1.14

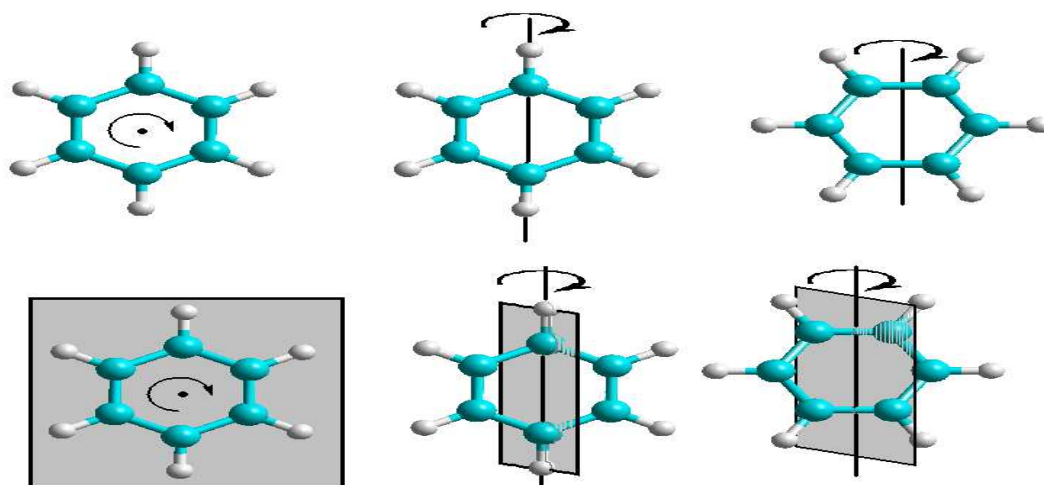


Figure 1.15 Rotational axes and mirror planes in Benzene

1.1.3 Improper axis of symmetry (S_n):

An improper rotation operation is one that comprises of a proper rotation operation around an axis followed by reflection through a plane perpendicular to it. It is denoted by S_n . n order of rotational axis. It can be also defined as, imaginary axis passing through molecule rotation in which perpendicular reflection to give equivalent orientation.

$$S_n = C_n \perp \sigma_n$$

Example: In CH_4 has tetrahedral geometry. If we place a tetrahedron in a cube then the four atoms touch the four corners of cubical box. CH_4 having S_4 it doesn't mean that C_4 symmetry is present.

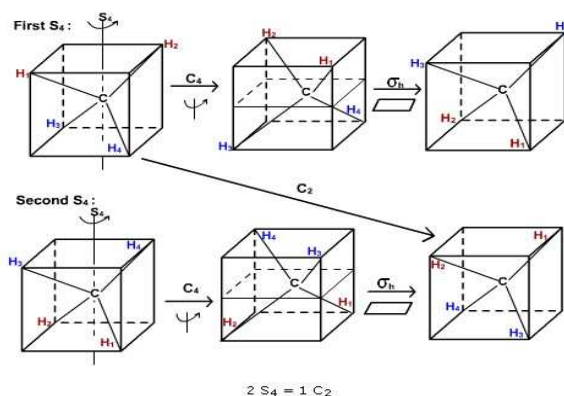
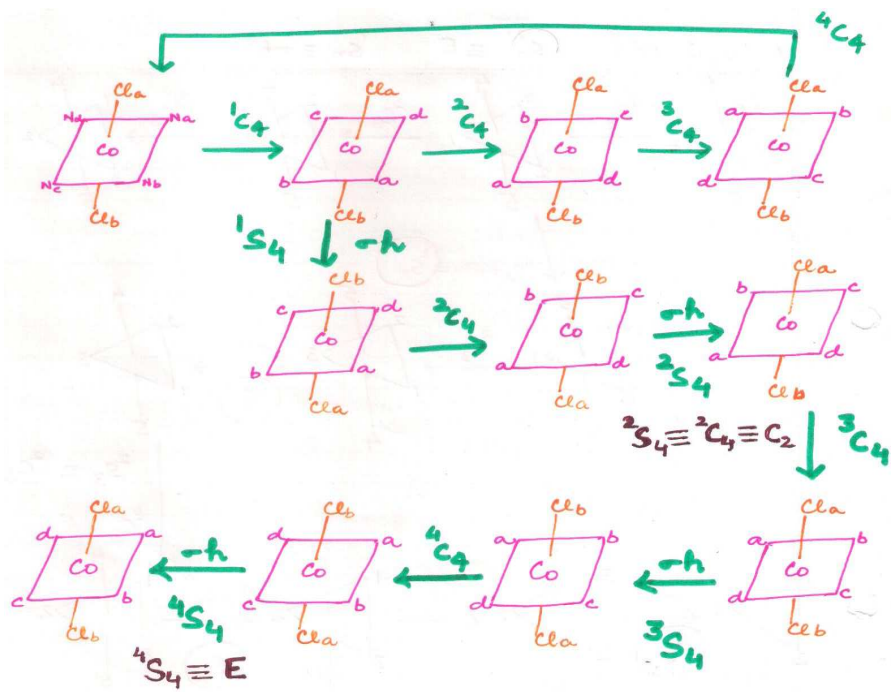


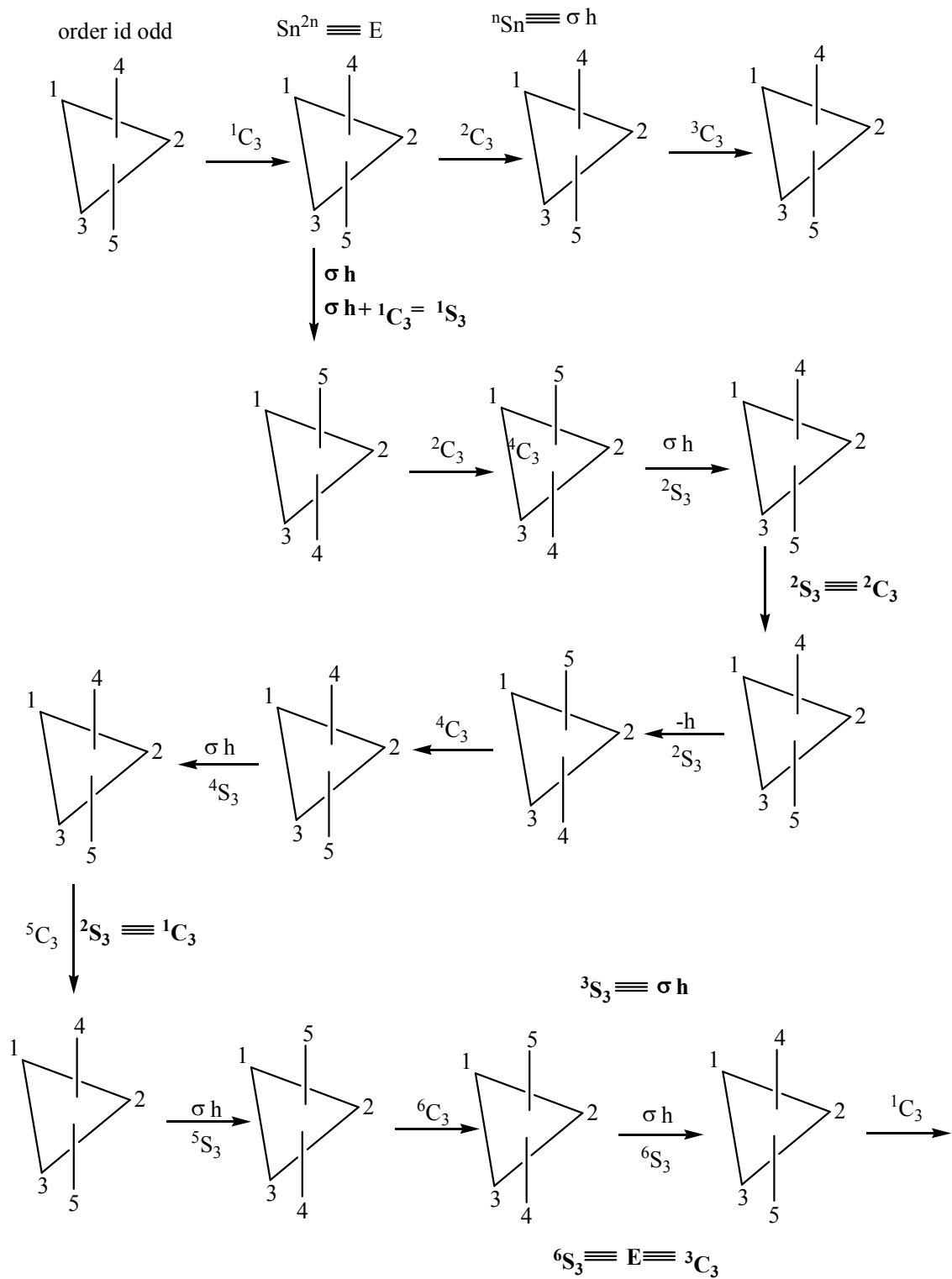
Figure: 1.16

IMPROPER AXIS IN OCTAHEDRAL STRUCTURE $S_n = 4$ even



Order of improper rotation is even $S_n^n \equiv E$

Improper axis in trigonal bipyramidal structure $s_n=3$ odd



Number of operation in S_n axis:

$$S_2 = C_2 \perp \sigma$$

$$S_3 = C_3 \perp \sigma$$

$$S_4 = C_4 \perp \sigma$$

$$\sigma(\text{odd}) = \sigma$$

$$\sigma(\text{even})$$

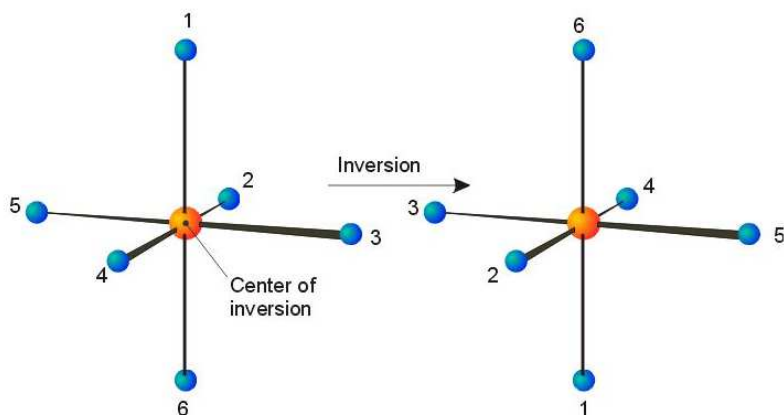
$$S_2^1 = C_2 \times \sigma^1 = S_2^1 = i$$

$$S_2^2 = C_2^2 \times \sigma^2 = E \times E = E$$

$$S_3 = C_3^1$$

1.1.4 Point of symmetry:

It is a point within the molecule through which all the atoms can be inverted within a point. The element corresponding to this operation is a center of symmetry. This is designated by i . In this operation, every part of the object is reflected through the inversion center, which must be at the center of mass of the object



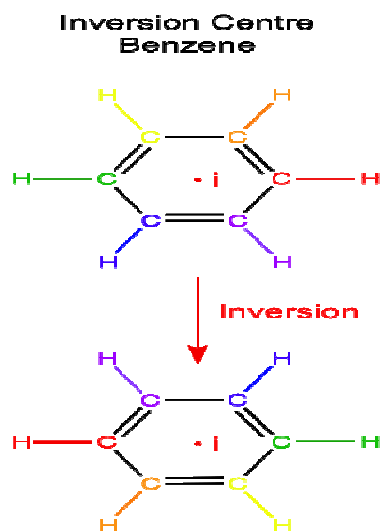


Figure: 1.17

Example:

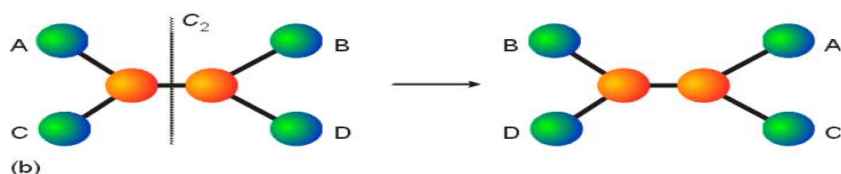
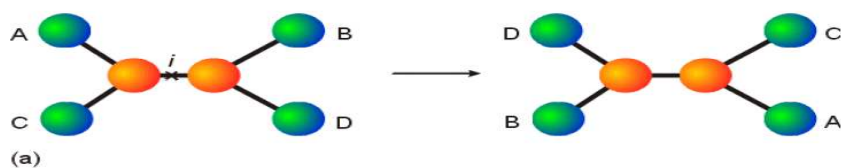
In ethene molecule each atoms are gives the mirror image at the center of the molecule. That center is called an inversion center.

Example: In benzene each carbon and hydrogen gives mirror image in a point.

We should remember that $C_2 \neq C_i$ **Figure:**

1.18

***i* Operation**



C_2 Operation

Figure 1.18

We should remember $C_2 \times \sigma_h = S_2 = C_i$

For example improper axis in Trans Dichloroethylene Figure :2.7

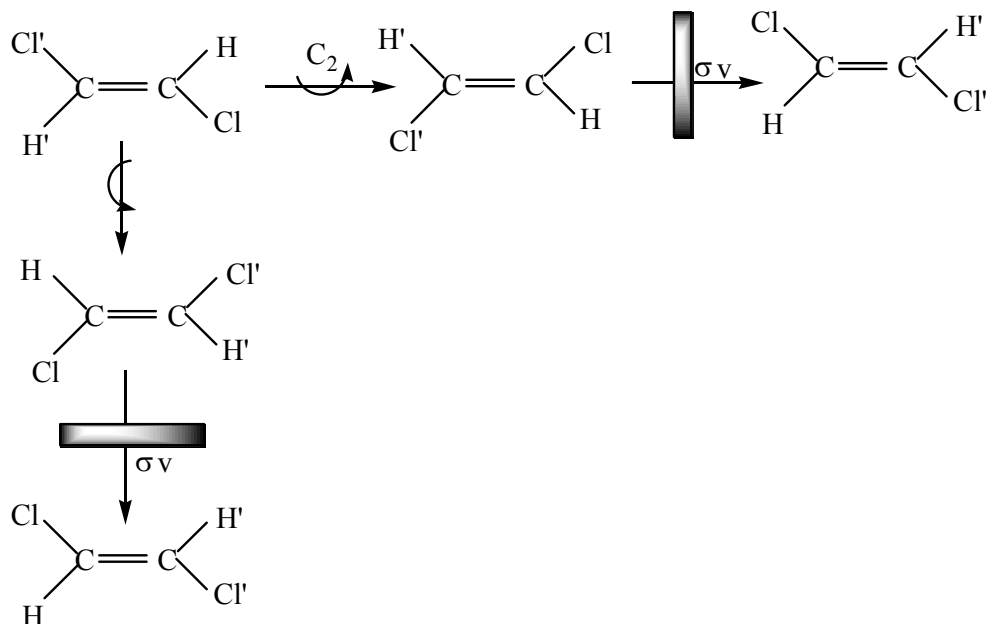


Figure: 1.19

1.1.5 Identity Operation:

This operation gives no change in the molecule. It is included mathematical completeness. An identity operation is characteristic of every molecule, even if it has no other symmetry.

Difference between symmetry operation and symmetry element:

S.N.	Symmetry operation	Symmetry element
1.	No change	Identity
2.	Rotation by $360/\theta$ about an axis of symmetry (n)	A n-fold axis of rotation
3.	The reflection in any	A plane of

	plane of symmetry parallel w.r.t. to principle axis of symmetry (σ_v).	symmetry containing the principle axis of symmetry (σ_v).
4.	The reflection of plane of symmetry perpendicular to the principle axis of symmetry (σ_h).	A plane of symmetry perpendicular to the principle axis of symmetry.
5.	Reflection of a plane of symmetry containing parallel w.r.t. principle axis of symmetry and bisecting between two C_2 axes of symmetry (σ_d).	A plane of symmetry containing the principle axes of symmetry and bisecting angle between two C_2 of symmetry (σ_d).
6.	Rotation by $360/\theta$ about axis followed by reflection in a plane perpendicular to that axis (S_n)	The n fold alternating axis (S_n).
7.	Inversion in a centre of symmetry (i)	The Centre of symmetry

1.2 A COMPLETE SET OF SYMMETRY OPERATIONS AS MATHEMATICAL GROUP

Group was defined by Arthur and Cayley in 1854. "Groups are a collection or set of element and obey the certain rules. It should denote number, matrix, vector, symmetry operation, element, addition-matrix and multiplication etc. are called group."

There are four postulate of the group:

- I) The product (multiplication of any two elements) in the group and the square of each element must be an element in the group.

Example: In H_2O , element of symmetry $C_2(z)$, E , $\sigma_{v_{xz}}$, $\sigma_{v_{yz}}$ are present.

$$C_2.C_2 = E$$

Since the product of $C_2.C_2 = E$ which is the element present in group.

$$C_2.\sigma_{v_{xy}} = C_2$$

In this, the product of $C_2.\sigma_{v_{xy}} = C_2$, which is also present in the group.

The product multiplication of elements means combination of elements of the group. The square of an element means multiplication an element with itself.

The product of two group of elements A and B of the group can be AB or BA these type of groups are known as Abelian group.

Case1- $A.B=B.A$

$$C_2.E = E.C_2$$

Case 2 $A.B \neq B.A$

$$C_3^1.\sigma_v \neq \sigma_v.C_3^1$$

Quite often the product A.B may give the element C, the group where B.A may give another element D of the same group.

- II) There is one element in each mathematical group which must commute ($A.B=B.A$) with all other elements of the group and leave them unchanged such element is known as identity element and this operation is called identity operation.

$$A.E = E.A = A$$

Example: In H_2O

$$C_2.E=E.C_2=C_2$$

Example: In NH_3

$$C_3.E=E.C_3=C_3$$

III) The element of mathematical group obeys the associative law of multiplication.

$$A.B(C.D.E)=A(B.C)(D.E)=(A.B)(C.D).E$$

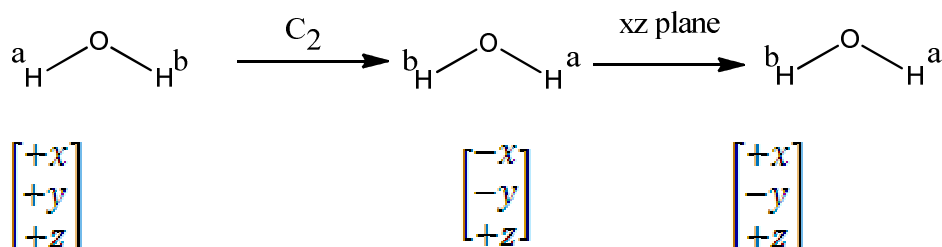
IV) Each element of mathematical group must have a reciprocal which is also an element of the group. When an element is multiplied by its reciprocal we get identity.

$$A.A^{-1}=A^{-1}.A=E$$

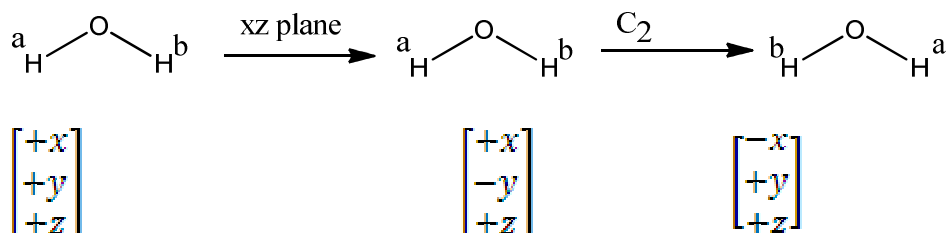
Type of group-

- 1) Abelian group
- 2) Non-abelian

1.2.1 Abelian group: “The groups in which all elements commute to each other are called Abelian group.”



(Matrix of plane of symmetry)

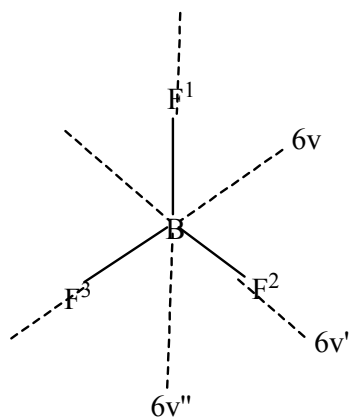


(Matrix of plane of symmetry)

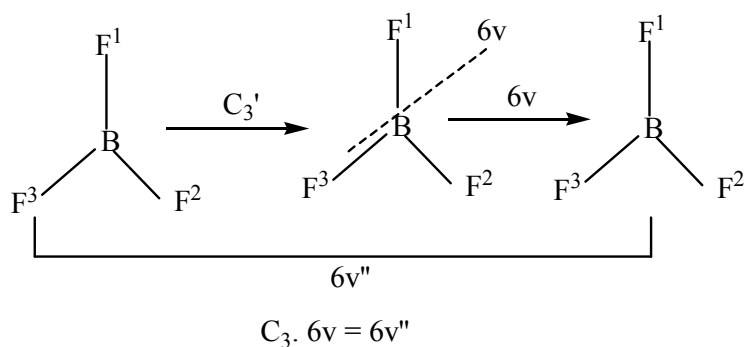
1.2.2 Non-Abelian Group: “If all the elements do not commute to each other known as Non-abelian group. Shown in **Figure: 2.1**

? $6v+$

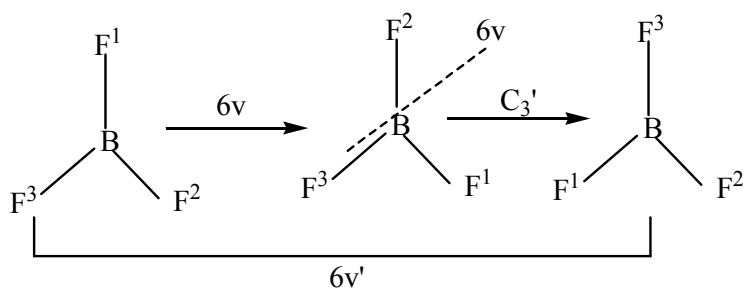
1. BF_3 - Element of symmetry - $^2C_3, 3C_2, 6v, 6v', 6v'', 6h$ for C_3' , $6v = 6v, C_3' = ?$



(a) For C_3' . $6v = ?$



(b) For $6v$. $C_3' = ?$



Since, $C_3' \cdot 6v = 6v''$
 $6v \cdot C_3' = 6v'$ } Non- abelian group

Figure: 2.1

1.3 GENERATOR

For arriving at the point group of a molecule we need to have knowledge of the generator of the symmetry operation. Generator can be defined as: “The minimum number of symmetry operation combined to each other be give the product of the group is known as **point group.**”

Classification of point group:

- 1) Non-axial point group
- 2) Axial point group

1.3 (a) Non-axial point group:

No axis of symmetry present in this type of point group.

- A) If only identity (E) present as generator. Only C_1 ($n=360/360$) or e operator is possible so that the point group is C_1 .
- B) Only plane of symmetry is present in a group, the point group is C_s .
- C) If only point of symmetry present the point group is C_i

1.3 (b) Axial point group:

Basic generator as well as additional generators are present in the molecule those are called axial group.

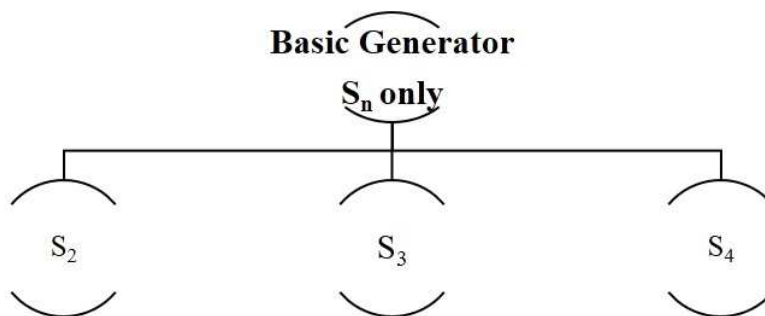


Figure: 2.2

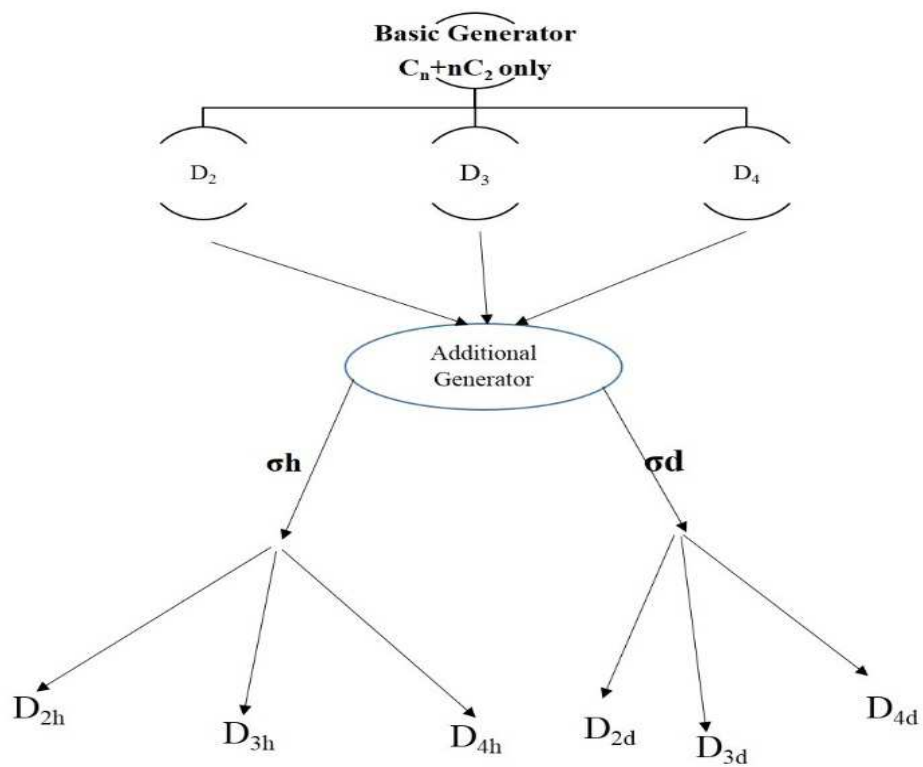
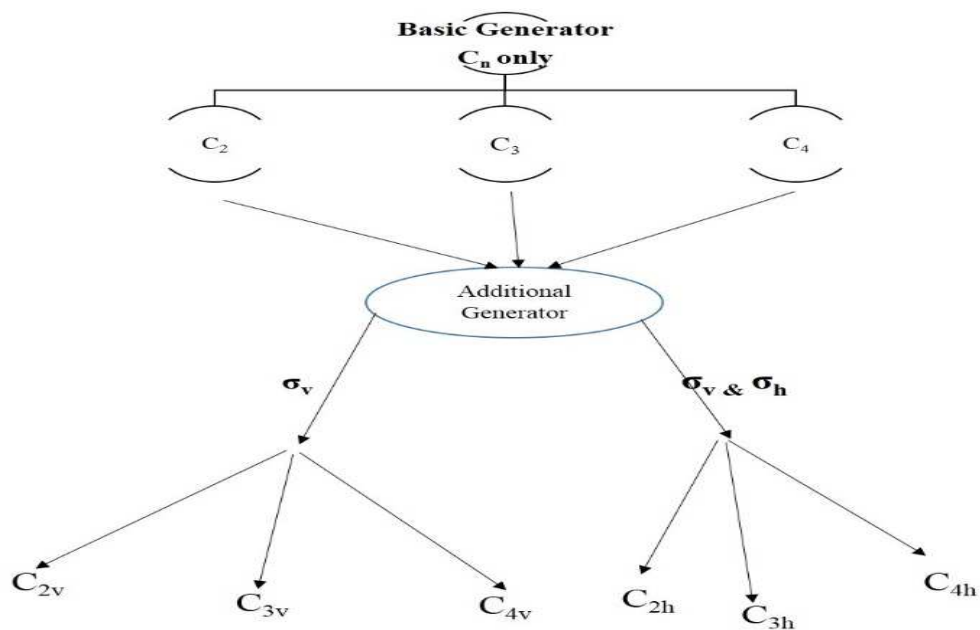


Figure: 2.3

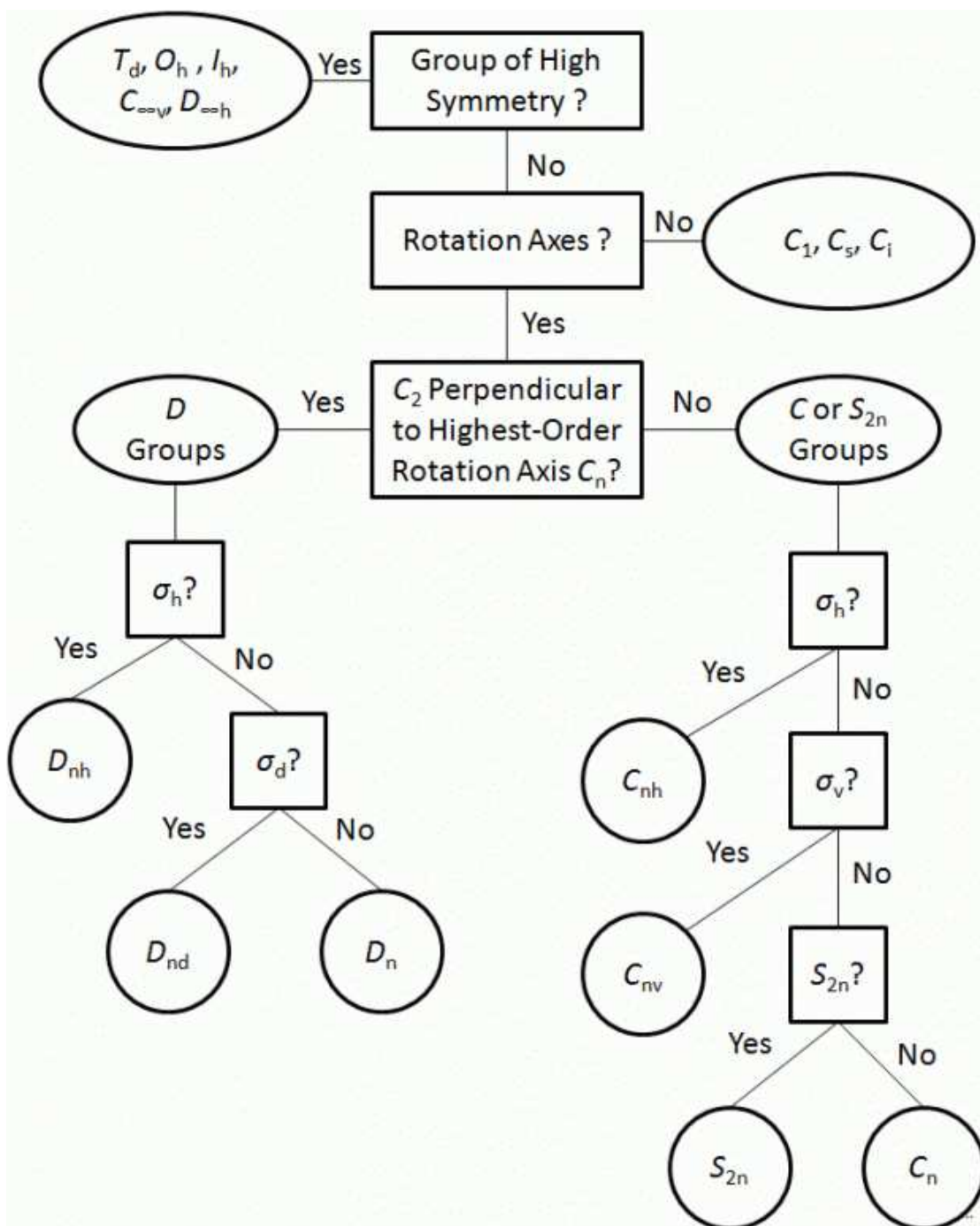
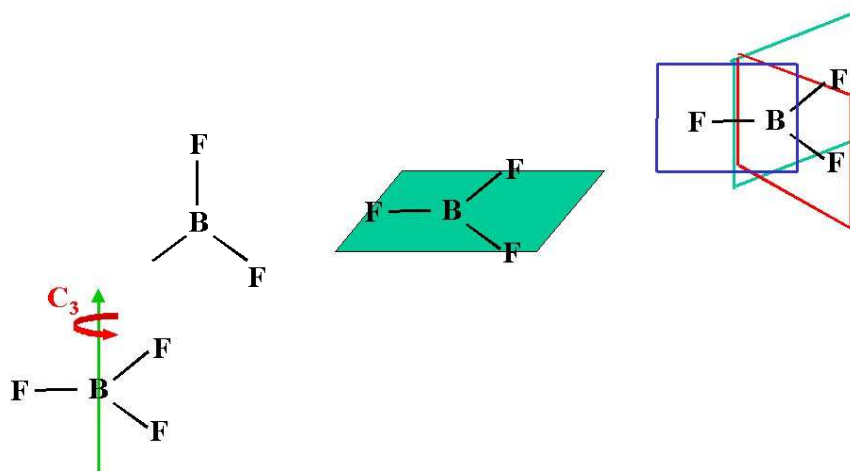


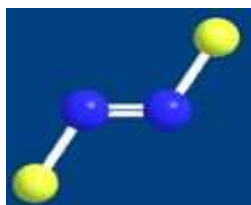
Figure: 2.4 Flow chart for point group determination



D_{3h}

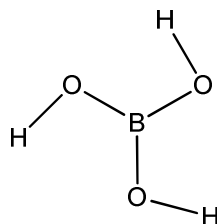
Figure: 2.5 BF_3 molecules the point group is D_{3h}

Example: C_{nh} Point Groups:



N_2F_2

C_{2h}

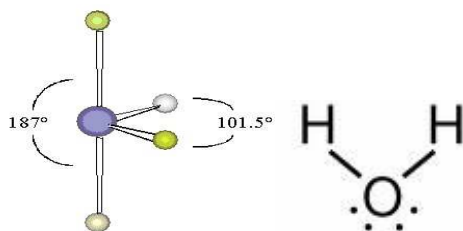


$B(OH)_3$

C_{3h}

Figure: 2.6 C_{nh} Point Groups

Example: C_{nv} Point Groups AX_2 bent type molecule. If a mirror plane contains the rotational axis, the group is called a C_{nv} group.

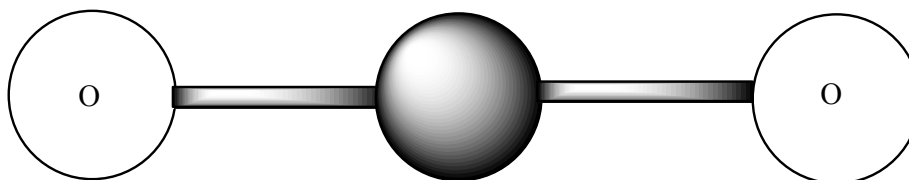


SF_4

C_{2v}

Figure: 2.7 C_{2v} point group

Example: D_{nh} type point group, in CO_2



There are an infinite number of possible C_n axes and σ_v mirror planes in addition to the σ_h .

Figure 2.8 D_{2h}

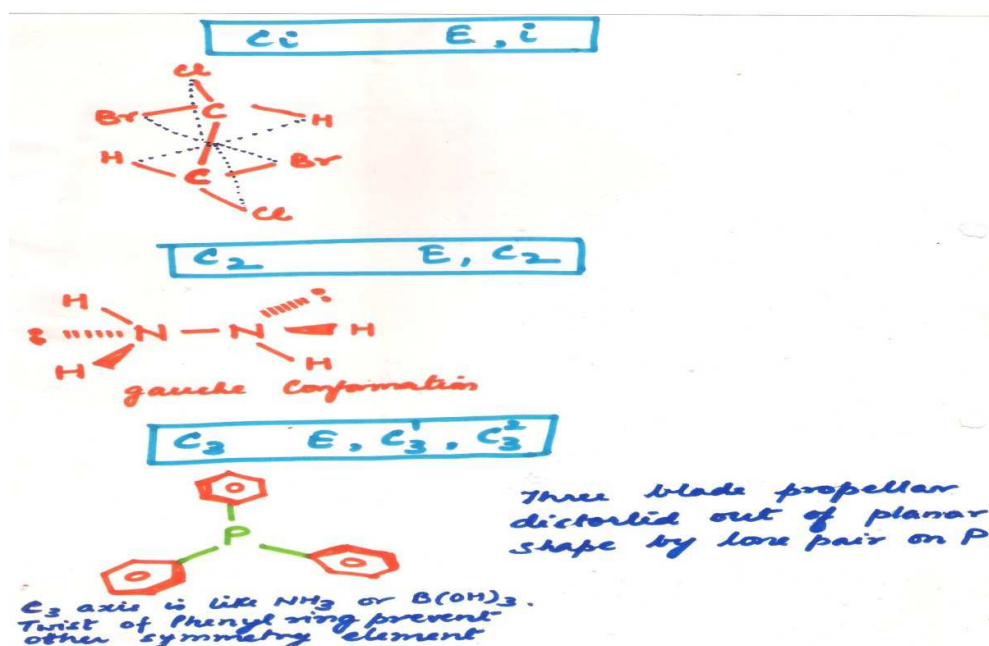


Figure 2.9 Examples of C_i , C_2 , and C_3

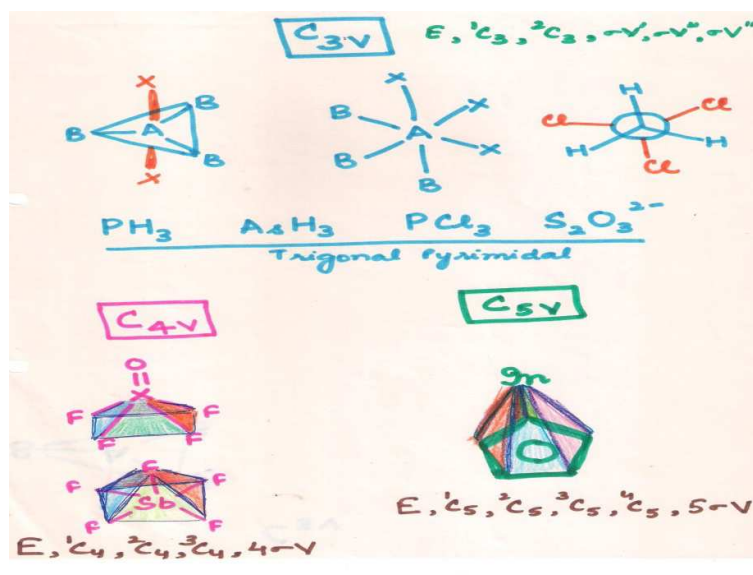


Figure 2.10 Examples of C_{nv}

Point Group	Characteristic symmetry operations
C_1	E
C_i	i
C_s	σ
C_n	C_n
C_{nv}	$C_n, n \sigma_v$
C_{nh}	$C_n, \sigma_h (S_n, i \text{ for even } n)$
D_n	$C_n, n \text{ perpendicular } C_2\text{'s, no } \sigma\text{'s}$
D_{nd}	$C_n, n \text{ perpendicular } C_2\text{'s, } n \sigma_d, S_{2n}, i \text{ for odd } n.$
D_{nh}	$C_n, n \text{ perpendicular } C_2\text{'s, } \sigma_h, S_n, (i, n/2 \sigma_d \text{'s, } n/2 \sigma_v\text{'s for even } n)$
S_n	$S_n \text{ only, } n \text{ is even. Possible } i, C_{n/2}$
$D_{\infty h}$	Linear, i
$C_{\infty v}$	Linear, no i
T_d	Tetrahedral
O_h	Octahedral
K_h	Sphere

1.4 SUMMARY

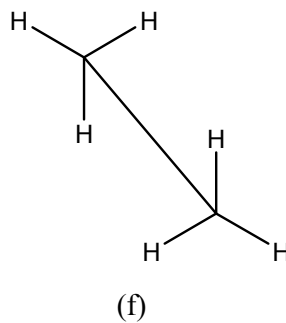
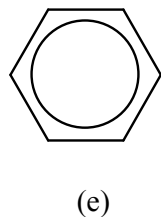
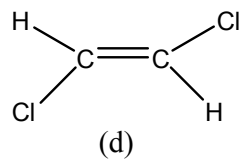
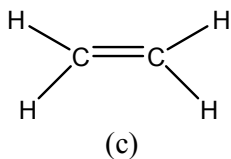
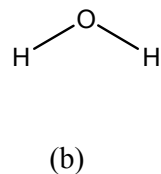
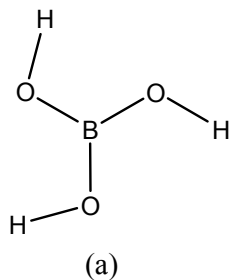
After reading this unit learner becomes familiar with group theory. The unit is summarized as -The experimental chemist in his daily work and thought is concerned with observing and, to as great an extent as possible, understanding and interpreting his observations on the nature of chemical compounds.

In this chapter we have discuss about absolute and non redundant list of the elements of a finite group and we know what all of the possible products are, then the group is absolutely and uniquely defined at least in an abstract sense.

1.5 TERMINAL QUESTIONS

1. What is the meant by a symmetry operation?
2. What is the relation between principal axis and horizontal plane?
3. Determine the symmetry elements in BF₃ molecule .
4. Determine the symmetry elements in BF₃ molecule .
5. What is abelian group. Explain with an example.

6. Give the example of molecules where their molecular plane are identified with
(a) vertical plane and (b) horizontal plane.
7. Define the dihedral plane of symmetry?
8. Write down the point group of the following molecules. How many and which symmetry operations are possible in each case?



9. What is symmetry?
10. What are symmetry elements?
11. What are symmetry operations?
12. How is axis of symmetry represented?
13. What is represented by point group ' C_n '?
14. What is plane of reflection?
15. What do you understand by S_4 ?

REFERENCE:

1. K Veera Reddy, Symmetry and Spectroscopy of molecule, New Age International publishers, 2014.
2. Gurdeep Raj, Ajay Bhagi and Vinod Jain, Group Theory and Symmetry in chemistry, Krishna publication, 2017.
3. B.R. Puri, L.R Sharma and K.C. Kalia, Principles of Inorganic Chemistry, Vishal publication, 2016
4. Group Theory and its Chemical Applications - P.K. Bhattacharya (Himalaya Publishing House) 2003

UNIT-2 MATRIX

CONTENTS:

- 2.1 Objectives
- 2.2 Introduction
- 2.3 Matrix
- 2.4 Order of group: denoted by 'h'
- 2.5 Character table of symmetry operation
- 2.6 Application of group theory in I.R. and Raman Spectroscopy
- 2.7 Terminal questions

2.1 OBJECTIVES

This chapter deals with the Matrix, order of group: denoted by 'h', Character table of symmetry operation and also discuss about the applications of group theory in I.R. and Raman Spectroscopy.

2.2 INTRODUCTION

Matrix can use for the study of symmetry. "A matrix is rectangular array of number or symbol for number." The symbol written down in rows and columns the square bracket is used.

2.3 MATRIX

The method of wave mechanics developed by Schrodinger can be applied to the solution of spectroscopic problems. Matrix can use for the study of symmetry. "A matrix is rectangular array of number or symbol for number." The symbol written down in rows and columns the square bracket is used.

Matrix representation of symmetry elements:

Element of symmetry- E, C_n, S_n, i and σ

2.3.1 Matrix for identity (E):

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{E} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ or } \begin{bmatrix} +1 \\ +1 \\ +1 \end{bmatrix}$$

Contribution of E=1+1+1= 3

2.3.2 Matrix for Rotation ($C_{n(z)}$):

$$\begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Anticlockwise Rotation

Clockwise Rotation

θ	0°	30°	45°	60°	90°	120°	180°	360°
Sin	0	$1/2$	$1/\sqrt{2}$	$\sqrt{3}/2$	1	$\sqrt{3}/2$	0	0
Cos	1	$\sqrt{3}/2$	$1/\sqrt{2}$	$1/2$	0	$-1/2$	-1	1

In $C_{4(z)}$ $\theta=90^\circ$

$$\begin{bmatrix} \cos 90 & -\sin 90 & 0 \\ \sin 90 & \cos 90 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -x_1 \\ y_1 \\ z_1 \end{bmatrix}$$

Contribution of $C_4=0+0+1=1$

For C_2 , $\theta=180^\circ$

$$\begin{bmatrix} \cos 180 & -\sin 180 & 0 \\ \sin 180 & \cos 180 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

Contribution= -1-1+1= -1

Contribution of C_n can be calculated with the help of above matrix chart.

2.3.3 Matrix for Reflection (σ):

This is the reflection plane σ_h , σ_v and σ_d

(a) For σ_{xz} operation:

$$[x_1 y_1 z_1] = [x_2 y_2 z_2]$$

$$x_2 = +x_1$$

$$y_2 = -y_1$$

$$z_2 = +z_1$$

$$\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ -y_1 \\ z_1 \end{bmatrix}$$

Contribution of $\sigma = +1-1+1=+1$

For each type of plane contribution is +1.

2.3.4 Matrix for inverse (C_i):

$$[x_1 y_1 z_1] = [x_2 y_2 z_2]$$

$$x_2 = -x_1$$

$$y_2 = -y_1$$

$$z_2 = -z_1$$

$$\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} -x_1 \\ -y_1 \\ -z_1 \end{bmatrix}$$

Contribution of $i = (-1)+(-1) +(-1) = -3$

2.3.5 Matrix for improper rotation ($S_{n(z)}$)

$$S_n = C_n \perp \sigma$$

$$\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ -z_1 \end{bmatrix}$$

Contribution for $S_2 = (-1) +(-1)+ (-1)= -3$

Table 1 Contribution of atoms in different symmetry

Symmetry element	Contribution
E	3
C_2	-1
C_3	0
C_4	+1
I	-3
S_3	-2
S_4	-1
Σ	1

2.4 ORDER OF GROUP: DENOTED BY 'h'

The total number of symmetry operations of point group is known as order of group.

Point group	C_s	C_n	C_{nv}	C_{nh}	D_n	D_{nh}	D_{nd}	S_n	Td	O_h	Ih
Order of group	2	n	2n	2n	2n	4n	4n	N	24	48	120

2.4.1 Class of operation in the group:

Class of Group has sorted elements of point group. Arrangement or sorting of elements require action of similarity transformation

If A, B AND X are elements of a group and

$X^{-1} A X = B$ 'B' is similarity transform of A by X. If this relationship is satisfied we say A & B as conjugate elements.

If elements conjugate then

1. Every element is conjugate to itself i.e. $X^{-1} A X = A$
2. If A is conjugate with B, then B must conjugate with A then $X^{-1} A X = B$ and then there must be some other element Y in the group that $Y^{-1} B Y = A$
3. If A conjugates with B and C and then B & C conjugate with each other then,

$$X^{-1} A X = B$$

$$Y^{-1} A Y = C$$

$$X^{-1} B X = C$$

$$Y^{-1} C Y = B$$

The set of symmetry operations of a group, which are conjugate to each other form conjugate (reciprocal) class or symmetry class.

- 1) Inversion (i) will be always in the separate class, if it is present.
- 2) σ_h will be always in the separate class.
- 3) In the case of proper axis C_n^m and C_n^{n-m} will be in the same class.

Example:



$$E C_3^1 E = C_3^1$$

$$C_3^2 C_3^1 C_3^1 = C_3^1$$

$$C_3^1 C_3^1 C_3^2 = C_3^1 \quad C_3^1 \text{ and } C_3^2 \text{ are inverse of each other.}$$

$$\sigma_v C_3^1 \sigma_v = C_3^2$$

$$\sigma_v' C_3^1 \sigma_v' = C_3^2$$

$$\sigma_v'' C_3^1 \sigma_v'' = C_3^2$$

So C_3^1 and C_3^2 are mutually conjugate, hence they can be in one class. Order is 2.

- 4) There will be separate different class for σ_d and σ_v
- 5) In the case of improper axis S_n^m and S_n^{n-m} will be in the same class.

Example:

(1) In H_2O (C_{2v})

Element of symmetry = $E, C_{2(z)}, \sigma_{v(xz)}, \sigma_{v(yz)}$

Number of class = 4

(2) In NH_3 (C_{3v})

Element of symmetry = $E, 2C_3, 3\sigma_v$. Here all three planes $\sigma_v', \sigma_v'', \sigma_v'''$ are in one class.

$$C_3^1 = C_3^2 = C_3^3 = E$$

C_3^1 and C_3^2 in the same class.

2.4.2 Representation of group:

Representation is defined as a set of matrix which corresponding to a single operation in the group that can be combine among themselves in a manner to the way in which the group elements.

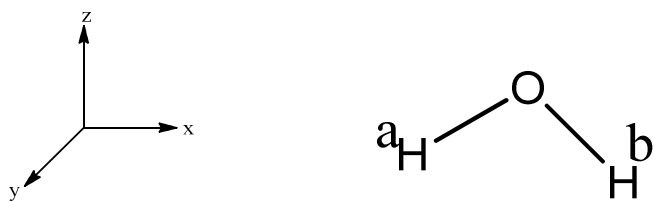
- 1) Reducible representation
- 2) Irreducible representation

2.4.2.1 Reducible Representation:

Representation of higher dimension which can reduce to a lower dimension representation is known as reducible representation. The character or identity (E) is known as dimension. To derive a reducible representation first of all we have the basis set according x, y and z direction to our need

Example: Reducible representation for H_2O i.e. C_{2v}

Basis set = 3 Cartesian coordinate (x, y, z)



Symmetry operation	No. Unshifted atom
\xrightarrow{E}	3
$\xrightarrow{C_2}$	1
$\xrightarrow{\text{plane } xz}$	3
$\xrightarrow{\text{plane } yz}$	1

C_{2v}	E	$C_{2(z)}$	$\sigma_{V(xz)}$	$\sigma_{V(yz)}$
No. of unshifted atom	3	1	1	3
Contribution of atom	3	-1	1	1
Reducible representation	3×3	1×-1	1×1	3×1
	9	-1	1	3

Dimension E=9

i.e. 9 dimension reducible representation

Example: Reducible representation for BF₃

Basis set= 3 Cartesian coordinate (x, y, z)

D_{3h}	E	$2C_3$	$3C_2$	$3\sigma_v$	σ_h	$2S_3$
----------	---	--------	--------	-------------	------------	--------

No. of un-shifted atom	4	1	2	2	4	1
Contribution of atom	3	0	-1	1	1	2
Reducible representation	4×3	1×0	2×-1	2×1	4×1	1×2
	12	0	-2	2	4	2

Dimension E=12

12 dimension reducible representation.

Example: Reducible representation for NH₃

Basis set= 3 Cartesian coordinate (x, y, z)

C_{3v}	E	2C₃	3σ_v
No. of un-shifted atom	4	1	2
Contribution of atom	3	0	1
Reducible representation	4×3	1×0	2×1
	12	0	2

2.4.3 Irreducible representation:

A representation of lower dimension which cannot be further reduced is known as irreducible representation. The irreducible representation for a group can derive by using great orthogonality theorem (GOT).

Postulates of Great Orthogonality Theorem (GOT) and derivation of Irreducible representation:

Ist Postulate: Number of irreducible representation is always equal to the number of class.

No. of IR = No. of class

Example: In H₂O molecule there are E, C_{2(z)}, σ_{v(xz)}, σ_{v(yz)} elements are present.

The number of order = 4

The number of class =4

C _{2v}	E	C _{2(z)}	σ _{v(xz)}	σ _{v(yz)}
I.R. 1	L ₁	M ₁	N ₁	O ₁
I.R. 2	L ₂	M ₂	N ₂	O ₂
I.R. 3	L ₃	M ₃	N ₃	O ₃
I.R. 4	L ₄	M ₄	N ₄	O ₄

Here four IR representations occur because C_{2v} has four classes.

IInd postulate: The sum of square of dimension of all the IR in the group will equal to the order of group.

$$L_1^2 + L_2^2 + L_3^2 + L_4^2 = 4$$

$$1^2 + 1^2 + 1^2 + 1^2 = 4$$

IIIrd postulate: The sum of square of the characters of IR will be equal to order of the group.

$$L_1^2 + M_2^2 + N_3^2 + O_4^2 = 4$$

C_{2v}	E	$C_{2(z)}$	$\sigma_{V(xz)}$	$\sigma_{V(yz)}$
I.R. 1	1	1	1	1
I.R. 2	1	M_2	N_2	O_2
I.R. 3	1	M_3	N_3	O_3
I.R. 4	1	M_4	N_4	O_4

IVth postulate: The characters of any two IR in the same group will be always orthogonal to each other (on multiplying any two IR the result should be zero).

$$IR_1 \times IR_2 = 0$$

$$IR_1 \times IR_3 = 0$$

$$IR_1 \times IR_4 = 0$$

$$IR_3 \times IR_2 = 0$$

$$IR_4 \times IR_2 = 0$$

$$IR_1 \times IR_2 = 0$$

$$1 \times 1 \times L_2 + 1 \times 1 \times M_2 + 1 \times 1 \times N_2 + 1 \times 1 \times O_2 = 0$$

$$1 \times 1 \times 1 + 1 \times 1 \times 1 + 1 \times 1 \times (-1) + 1 \times 1 \times (-1) = 0$$

C_{2v}	E	$C_{2(z)}$	$\sigma_{V(xz)}$	$\sigma_{V(yz)}$
I.R. 1	1	1	1	1
I.R. 2	1	1	-1	-1
I.R. 3	1	M_3	N_3	O_3
I.R. 4	1	M_4	N_4	O_4

$$IR_1 \times IR_3 = 0$$

$$1 \times 1 \times L_3 + 1 \times 1 \times M_3 + 1 \times 1 \times N_3 + 1 \times 1 \times O_3 = 0$$

$$1 \times 1 \times 1 + 1 \times 1 \times (-1) + 1 \times 1 \times 1 + 1 \times 1 \times (-1) = 0$$

C_{2v}	E	$C_{2(z)}$	$\sigma_{V(xz)}$	$\sigma_{V(yz)}$
I.R. 1	1	1	1	1
I.R. 2	1	1	-1	-1
I.R. 3	1	-1	1	-1
I.R. 4	1	M_4	N_4	O_4

$IR_1 \times IR_4 = 0$

$$1 \times 1 \times L_4 + 1 \times 1 \times M_4 + 1 \times 1 \times N_4 + 1 \times 1 \times O_4 = 0$$

$$1 \times 1 \times 1 + 1 \times 1 \times (-1) + 1 \times 1 \times (-1) + 1 \times 1 \times 1 = 0$$

C_{2v}	E	$C_{2(z)}$	$\sigma_{V(xz)}$	$\sigma_{V(yz)}$
I.R. 1	1	1	1	1
I.R. 2	1	1	-1	-1
I.R. 3	1	-1	1	-1
I.R. 4	1	-1	-1	1

Ist Postulate: Number of irreducible representation is always equal to the number of class.

No. of IR = No. of class

Example: In NH_3 molecule there are E, $2C_3$, $3\sigma_v$ elements are present.

The number of order = 6

The number of class = 3

C_{3v}	E	$2C_3$	$3\sigma_v$
I.R. 1	L_1	M_1	N_1
I.R. 2	L_2	M_2	N_2
I.R. 3	L_3	M_3	N_3

IInd postulate: The sum of square of dimension of all the IR in the group will equal to the order of group.

$$L_1^2 + L_2^2 + L_3^2 = 6$$

$$1^2 + 1^2 + 2^2 = 6$$

C_{3v}	E	$2C_3$	$3\sigma_v$
I.R. 1	1	1	2
I.R. 2	L_2	M_2	N_2
I.R. 3	L_3	M_3	N_3

IIIrd postulate: The sum of square of the characters of IR will be equal to order of the group.

$$\mathbf{I.R.1 \times I.R.2 = 0}$$

$$1 \times 1 \times L_2 + 1 \times 1 \times M_2 + 1 \times 1 \times N_2 + 1 \times 1 \times O_2 = 0$$

$$1 \times 1 \times 1 + 2 \times 1 \times 1 + 3 \times 1 \times (-1) = 0$$

C_{3v}	E	$2C_3$	$3\sigma_v$
I.R. 1	1	1	1
I.R. 2	1	1	-1
I.R. 3	L_3	M_3	N_3

$$\mathbf{IR_1 \times IR_3 = 0}$$

$$1 \times 1 \times L_3 + 1 \times 1 \times M_3 + 1 \times 1 \times N_3 + 1 \times 1 \times O_3 = 0$$

$$1 \times 1 \times 2 + 2 \times 1 \times (-1) + 3 \times 1 \times 0 = 0$$

C_{3v}	E	$2C_3$	$3\sigma_v$
I.R. 1	1	1	1
I.R. 2	1	1	-1
I.R. 3	2	1	0

2.5 CHARACTER TABLE OF SYMMETRY OPERATION

IR and their transformation properties which can be utilized for the chemical operation of group theory. Each character table has four columns.

- Character tables contain, in a highly symbolic form, information about how something of interest (a bond, an orbital, etc.) is affected by the operations of a given point group.
- Each point group has a unique character table, which is organized into a matrix.
- Column headings are the symmetry operations, which are grouped into classes.
- Horizontal rows are called irreducible representations of the point group.

- The main body consists of characters (numbers), and a section on the right side of the table provides information about vectors and atomic orbitals

For H₂O molecule the character table is given bellow.

Point group	Class of the symmetry elements	h=4	
I st Column	II nd Column	III rd Column	IV th Column

Column Ist: first column of character table shows Mulliken symbols for I.R. of the group.

(i) Dimension

- 1D (A or B)
- 2D (E)
- 3D (T)

(ii) Principal axis

(+) symmetry – A

(-) symmetry – B

(iii) If subsidiary axis is present

(+) gerade-1

(-) gerade- 2

(iv) If subsidiary axis absent then molecular plane

(+) subscript - 1

(-) subscript – 2

(v) If σ_h present

(+) single prime – ‘

(-) double prime- ‘’

(vi) If C_i present

(+) gerade- g

(-) ungerade- u

Example: For H₂O molecule

C_{2v}	E	$C_{2(z)}$	$\sigma_{V(xz)}$	$\sigma_{V(yz)}$
A_1	1	1	1	1
A_2	1	1	-1	-1
B_1	1	-1	1	-1
B_2	1	-1	-1	1

Column II has dimensions for each irreducible representation

Example: For H_2O molecule

C_{2v}	E	$C_{2(z)}$	$\sigma_{V(xz)}$	$\sigma_{V(yz)}$
I.R. 1	1	1	1	1
I.R. 2	1	1	-1	-1
I.R. 3	1	-1	1	-1
I.R. 4	1	-1	-1	1
C_{2v}	E	$C_{2(z)}$	$\sigma_{V(xz)}$	$\sigma_{V(yz)}$
R.R.	9	-1	3	1

Example: For NH_3 molecule

C_{3v}	E	$2C_3$	$3\sigma_v$
A_1	1	1	1
A_2	1	1	-1
E	2	1	0

C_{3v}	E	$2C_3$	$3\sigma_v$
R.R.	12	0	2

Column III:

Explain transformation properties of three Cartesian coordinate (x,y,z) and three rotational axis (R_x, R_y, R_z).

Column IV:

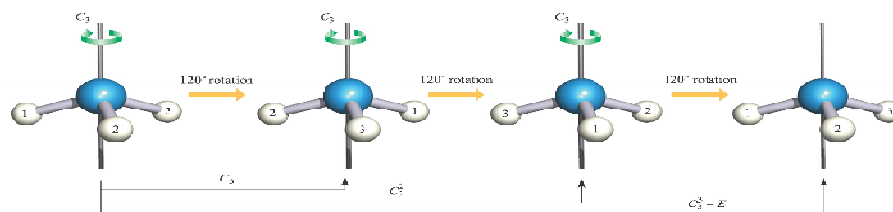
It explains transformation properties of quadratic function of coordinates like

transformational properties of x^2, y^2, z^2 .

Example: The character table for water molecule (for bent XY_2 type molecule) $h=4$

C_{2v}	E	$C_{2(z)}$	$\sigma v_{(xz)}$	$\sigma v_{(yz)}$	Basis functions or Symmetry of functions	
A_1	1	1	1	1	Z	x^2, y^2, z^2
A_2	1	1	-1	-1	Rz	Xy
B_1	1	-1	1	-1	(x, Ry)	Yz
B_2	1	-1	-1	1	(y, Rx)	Zx

The character table for ammonia molecule (for pyramidal XY_3 type molecule)



C_{3v}	E	$2C_3$	$3\sigma v$	$h=6$ Basis functions or Symmetry of functions	
A_1	1	1	1	z	$x^2 + y^2, z^2$
A_2	1	1	-1	Rz	
E	2	1	0	(x, y)(Rx, Ry)	$(x^2 - y^2, xy) (xz, yz)$

The reduction can be achieved using the reduction formula. It is a mathematical way of reducing that will always work when the answer cannot be spotted by inspection. It is particularly useful when there are large numbers of atoms and bonds involved.

The vibrational modes of the molecule are reduced to produce a reducible representation into the irreducible representations. This method uses the following formula reduction formula:

$$N = 1/h [\sum n^x X_r^x X_i^x]$$

N is the number of times a symmetry species occurs in the reducible representation,
 h is the 'order of the group': simply the total number of symmetry operations in the group.

The summation is over *all* of the symmetry operations.

For each symmetry operation, three numbers are multiplied together.

These are:

X_r is the character for a particular class of operation in the reducible representation

X_i is the character of the irreducible representation.

n is the number of symmetry operations in the class

The characters of the reducible representation can be determined by considering the combined effect of each symmetry operation on the atomic vectors

$$nA_1 = \frac{1}{\text{order}} [\sum \text{operation of group} \times R.R. \times I.R.]$$

$$nA_1 = \frac{1}{4} [1 \times 1 \times 9 + 1 \times 1 \times (-1) + 1 \times 1 \times 3 + 1 \times 1 \times 1]$$

$$nA_1 = 12/4$$

$$nA_1 = 3$$

The value of n is 3

3A₁

$$nA_2 = \frac{1}{4} [1 \times 1 \times 9 + 1 \times 1 \times (-1) + 1 \times (-1) \times 3 + 1 \times (-1) \times 1]$$

$$nA_2 = \frac{1}{4} [9 - 1 - 3 - 1]$$

$$nA_2 = 4/4$$

$$nA_2 = 1$$

The value of n is 1

A₂

$$nB_1 = \frac{1}{4} [1 \times 1 \times 9 + 1 \times (-1) \times (-1) + 1 \times 1 \times 3 + 1 \times (-1) \times 1]$$

$$nB_1 = 12/4$$

$$nB_1 = 3$$

The value of n is 3

3B₁

$$nB_2 = \frac{1}{4} [1 \times 1 \times 9 + 1 \times (-1) \times (-1) + 1 \times (-1) \times 3 + 1 \times 1 \times 1]$$

$$nB_2 = 8/4$$

$$nB_2 = 2$$

The value of n is 2

2B₂

Relation between I.R. and R.R. in H₂O is;

$$3A_1 + 1A_2 + 3B_1 + 2B_2$$

Example: For NH₃ molecule

C _{3v}	E	2C ₃	3σ _v
A ₁	1	1	1
A ₂	1	1	-1
E	2	1	0

C _{3v}	E	2C ₃	3σ _v
R.R.	12	0	2

$$nA_1 = \frac{1}{\text{order}} [\sum \text{operation of group} \times \text{R.R.} \times \text{I.R.}]$$

$$nA_1 = \frac{1}{3} [1 \times 1 \times 12 + 2 \times 1 \times 0 + 3 \times 1 \times 2]$$

$$nA_1 = 18/6$$

$$nA_1 = 3$$

The value of n is 3

3A₁

$$nA_2 = \frac{1}{3} [1 \times 1 \times 12 + 2 \times 1 \times 0 + 3 \times (-1) \times 2]$$

$$nA_2 = 6/6$$

$$nA_2 = 1$$

The value of n is 1

1A₂

$$nE = \frac{1}{6} [1 \times 2 \times 12 + 2 \times (-1) \times 0 + 3 \times 0 \times 2]$$

$$nE = 24/6$$

$$nE = 4$$

The value of n is 4

4E

Relation between I.R. and R.R. = $3A_1 + A_2 + 4E$

Column III:

Explain transformation properties of three Cartesian coordinate (x,y,z) and three rotational axis (R_x, R_y, R_z).

Column IV:

It explain transformation properties of quadratic function of coordinates like transformational properties of x^2, y^2, z^2 .

We know that

Example: The character table for water molecule (for bent XY_2 type molecule) $h=4$

C_{2v}	E	$C_{2(z)}$	$\sigma_{v(xz)}$	$\sigma_{v(yz)}$	Basis functions or Symmetry of functions	
A_1	1	1	1	1	Z	x^2, y^2, z^2
A_2	1	1	-1	-1	R_z	XY
B_1	1	-1	1	-1	(x, R_y)	YZ
B_2	1	-1	-1	1	(y, R_x)	ZX

The character table for ammonia molecule (for pyramidal XY_3 type molecule)

C_{3v}	E	$2C_3$	$3\sigma_v$	$h=6$	
A_1	1	1	1	z	$x^2 + y^2, z^2$
A_2	1	1	-1	R_z	
E	2	1	0	(x, y)(R_x, R_y)	$(x^2 - y^2, xy)(xz, yz)$

2.6 APPLICATION OF GROUP THEORY IN I.R. AND RAMAN SPECTROSCOPY

AB₂ type molecule can exist in two geometry:

S.N.	AB ₂ geometry	Degree of freedom=3N- (Rotational +Translational)
1.	Linear	3N- (2+3)
2.	Non-linear	3N- (3+3)

To obtain from this total set the representations for vibration only, it is necessary to subtract the representations for the other two form of motion: rotation and translation.

One of the most practical uses of point groups and group theory for the inorganic chemist in is predicting the number of infrared and Raman bands that may be expected from a molecule. Alternatively, given the IR or Raman spectrum, the symmetry of a molecule may be inferred. In both IR and Raman spectroscopy the molecule is viewed as containing moving vectors. How these vectors are affected by symmetry will provide a means to determine how many bands would be expected in these spectra.

Selection Rules: Infrared and Raman Spectroscopy:

- Infrared energy is absorbed for certain changes in vibrational energy levels of a molecule.
- For a vibration to be infrared active there must be a change in the molecular dipole moment vector associated with the vibration.
- For a vibration mode to be Raman active there must be a change in the net polarizability tensor.
- Polarizability is the ease in which the electron cloud associated with the molecule is distorted.
- For centro symmetric molecules, the rule of mutual exclusion states that vibrations that are IR active are Raman inactive, and vice versa.
- The transition from the vibrational ground state to the first excited state is the fundamental transition.

Quadratic function:

(1) $x^2, y^2, z^2, x^2-y^2, x^2-y^2-z^2, xy, yz, zx$ - **Raman active**

In which these quadratic function occurs in IV column are Raman active.

(2) R_x, R_y, R_z, x, y, z - **IR active**

In which these quadratic function occurs in III column

Character tables can be used in

Determining the orbitals involved in bonding in a molecule.

Determine the point group of the molecule.

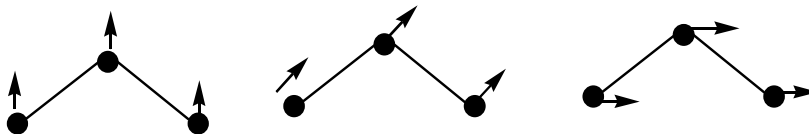
Determining the **Reducible Representation**, Γ , for the type of bonding you wish to describe (e.g. σ , π , π_{\perp} , π_{\parallel}). The Reducible Representation indicates how the bonds are affected by the symmetry elements present in the point group.

Identifying the **Irreducible Representation** that provides the Reducible Representation; there is a simple equation to do this. The Irreducible Representation (e.g. $2A_1 + B_1 + B_2$) is the combination of symmetry representations in the point group that sum to give the Reducible Representation.

Identifying orbitals that are involved from the Irreducible Representation and the character table.

Cartesian displacement vectors for a water molecule:

Translational Modes: A mode in which all atoms are moving in the same direction, equivalent to moving the molecule.



Rotational Modes: A mode in which atoms move to rotate (change the orientation of) the molecule. There are three rotational modes for nonlinear molecules, and two rotational modes for linear molecules.

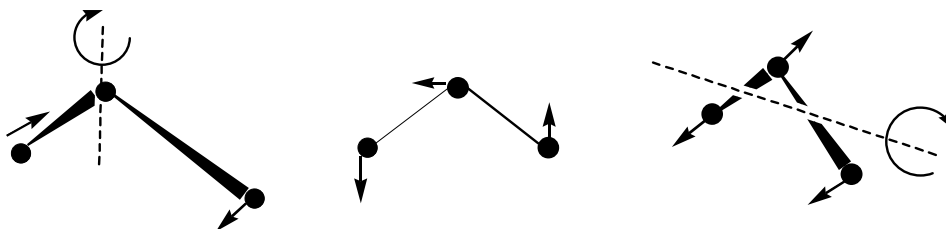
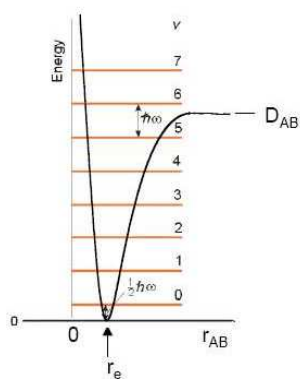


Figure: 2.12



r_e = equilibrium distance between A and B D_{AB} = energy required to dissociate into A and B atoms

For a diatomic molecule (A-B), the bond between the two atoms can be approximated by a spring that restores the distance between A and B to its equilibrium value. The bond can be assigned a force constant, k (in Nm^{-1}); the stronger the bond, the larger k and the relationship between the frequency of the vibration, ω , is given by the relationship:

$$\omega = \sqrt{\frac{k}{\mu}} \quad \text{or, more typically} \quad 2\pi c\bar{\nu} = \sqrt{\frac{k}{\mu}}$$

where, c is the speed of light, $\bar{\nu}$ is the frequency in "wave numbers" (cm^{-1}) and μ is the reduced mass (in amu) of A and B given by the equation:

$$\mu = \frac{m_A \times m_B}{m_A + m_B}$$

Figure: 2.13

Example: In H_2O IR and Raman active mode can explained with the help of group Theory. The relation between Irreducible and Reducible representation in H_2O is:

$$3N = 3A_1 + A_2 + 3B_1 + 2B_2$$

C_{2v}	E	$C_{2(z)}$	$\sigma_{v(xz)}$	$\sigma_{v(yz)}$	$h=4$	
A_1	1	1	1	1	Z	x^2, y^2, z^2
A_2	1	1	-1	-1	Rz	xy
B_1	1	-1	1	-1	(x, Ry)	yz
B_2	1	-1	-1	1	(y, Rx)	zx

Translational: $A_1 = z, B_1 = x, B_2 = y$

Rotational: $B_2 = R_x, B_1 = R_y, A_2 = R_z$

Degree of Freedom = $3N - (\text{Rotational} + \text{Translational})$

$$= 3A_1 + A_2 + 3B_1 + 2B_2 - (B_2 + B_1 + A_2 + A_1 + B_1 + B_2)$$

$$= 2A_1 + B_1$$

There are three modes of vibrations. Rotations have the same symmetry as R_x, R_y and R_z . So in C_{2v} The three vibrational modes remain. Two have A_1 symmetry, and one has B_1 symmetry

For a molecular vibration to be seen in the Raman spectrum (Raman active), it must change the polarizability of the molecule. The polarizability has the same symmetry properties as the quadratic functions: xy, yz, xz, x^2, y^2 and z^2

A_1 = Raman and IR active (due to presence of x^2 , y^2 and z present)

B_1 = Raman and IR active (due to presence of x , R_y and yz present)

Example: In NH_3 molecule the $3N$ is equal to $3A_1 + A_2 + 4E$

Translational: $A_1 = z$, $E = x, y$

Rotational: $A_2 = R_z$, $E = R_y, R_x$

Degree of Freedom = $3N - (\text{Rotational} + \text{Translational})$

$$= 3A_1 + A_2 + 4E - (A_1 + E + A_2 + E)$$

$$= 2A_1 + 2E$$

A_1 = Raman and IR active

E = Raman and IR active

Polarity of Molecules can be studied with group theory

A molecule cannot be polar if it has

1. A centre of inversion i.e. any group with C_i

2. Any electric dipole moment perpendicular to any mirror planes i.e. any of the groups D and their derivatives. A

3. Any electric dipole moment perpendicular to any axis of rotation i.e. the cubic groups T, O , the Icosohedral I , and their modifications. A

Group theory is useful in studying Chirality of molecules

A molecule is not chiral if

1. It possesses an improper rotation axis S_n .

2. It belongs to the group D_{nh} or D_{nd}

3. It belongs to T_d or O_h

4. It possesses a centre of symmetry

5. It possesses a mirror plane.

Examples

1. $C_s, C_{nv}, D_{nh}, D_{nd}$, and T_d possess a plane of symmetry

2. C_i has a centre of symmetry

3. O_h and I_h have a centre and a plane of symmetry

4 S_n has improper axis

Appendix A: Character Tables for selected point groups.

C_s	E	σ _h		
A' A''	1	1	x,y,R _z	x ² ,y ² ,z ² ,xy
	1	-1	z,R _x ,R _y	yz,xz

C₂	E	C ₂		
A B	1	1	z,R _z	x ² +y ² ,z ²
	1	-1	x,y,R _x ,R _y	yz,xz

D₂	E	C ₂ (z)	C ₂ (y)	C ₂ (x)		
A B ₁	1	1	1	1		x ² ,y ² ,z ² , xy
B ₂	1	1	-1	-1	z,R _z	Xy
	1	-1	1	-1	y, R _y	Xz
B ₃	1	-1	-1	1	x,R _x	Yz

D₃	E	2C ₃	3C ₂		
A ₁	1	1	1		x ² +y ² , z ²
A ₂	1	1	-1	z,R _z	
E	2	-1	0	(x,y);(R _x ,R _y)	(xz,yz);

C_{2v}	E	C ₂	σ _v (xz)	σ _v (yz)		
A ₁	1	1	1	1	z	x ² ,y ² ,z ²
A ₂	1	1	-1	-1	R _z	Xy
	1	-1	1	-1	x,R _x	Xz
B ₁	1	-1	-1	1	y,R _y	Yz

C_{3v}	E	2C ₃	3 σ _v		
A ₁	1	1	1	z	x ² +y ² , z ²
A ₂	1	1	-1	R _z	
	2	-1	0	(x,y), (R _x ,R _y)	(x ² -y ² ,xy), (xz,yz)

C_{4v}	E	2C ₄	C ₂	2 σ _v	2 σ _d		
A ₁	1	1	1	1	1	z	x ² +y ² , z ²
A ₂	1	1	1	-1	-1	R _z	
B ₁	1	-1	1	1	-1		x ² -y ²
B ₂	1	-1	1	-1	1		Xy
E	2	0	-2	0	0	(x,y) (R _x ,R _y)	(xz,yz)

C_i	E	i		
A _g	1	1	R _x ,R _y ,R _z	x ² ,y ² ,z ² ,xy,xz,yz
A _u	1	-1	x,y,z	

E	E	C ₂ (z)	C ₂ (y)	C ₂ (x)	I	(xy)	(xz)	(yz)		
A _g	1	1	1	1	1	1	1	1		x ² , y ² , z ²
B _{1g}	1	1	-1	-1	1	1	-1	-1	R _z	xy
B _{2g}	1	-1	1	-1	1	-1	1	-1	R _v	xz
B _{3g}	1	-1	-1	1	1	-1	-1	1	R _x	yz
A _u	1	1	1	1	-1	-1	-1	-1		
	1	1	-1	-1	-1	-1	1	1	z	
	1	-1	1	-1	-1	1	-1	1	y	
B _{1u}	1	-1	-1	1	-1	1	1	-1	x	

D_{3h}	E	2C ₃	3C ₂	σ _h	2S ₃	3 σ _v		
A ₁ '	1	1	1	1	1	1		x ² +y ² , z ²
A ₂ ' E'	1	1	-1	1	1	-1	R _z	
	2	-1	0	2	-1	0	(x,y)	(x ² -y ² , xy)
A ₁ ''	1	1	1	-1	-1	-1		
A ₂ ''	1	1	-1	-1	-1	1	z	
E''	2	-1	0	-2	1	0	(R _x , R _v)	(xz, yz)

D_{4h}	E	2C ₄	C ₂	2C ₂ '	2C ₂ ''	I	2S ₄	h	2 σ _v	2 σ _d		
A _{1g}	1	1	1	1	1	1	1	1	1	1		x ² +y ² , z ²
A _{2g}	1	1	1	-1	-1	1	1	1	-1	-1	R _z	
B _{1g}	1	-1	1	1	-1	1	-1	1	1	-1		x ² -y ²
	1	-1	1	-1	1	1	-1	1	-1	1		xy
	2	0	-2	0	0	2	0	-2	0	0	(R _x , R _v)	(xz, yz)
D_{5h}	1	1	1	1	1	-1	-1	-1	-1	-1		
A ₁ '	1	1	1	-1	-1	-1	-1	-1	1	1	z	
A ₂ ' E ₁ '	1	-1	1	1	-1	-1	1	-1	-1	1		
	1	-1	1	-1	1	-1	1	-1	1	-1		
E ₂ '	2	0	-2	0	0	-2	0	2	0	0	(x,y)	

D_{6h}	E	2C ₆	2C ₃	C ₂	3C ₂ '	3C ₂ "	i	2S ₃	2S ₆	h	3 d	3 v		
A _{1g}	1	1	1	1	1	1	1	1	1	1	1	1		x^2+y^2, z^2
A _{2g}	1	1	1	1	-1	-1	1	1	1	1	-1	-1	R _z	
	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1		x^2-y^2
B _{1g}	1	-1	1	-1	-1	1	1	-1	1	-1	-1	1		xy
B _{2g}	2	1	-1	-2	0	0	2	1	-1	-2	0	0	(R _x , R _y)	(xz, yz)
	2	-1	-1	2	0	0	2	-1	-1	2	0	0		
E _{1g}	1	1	1	1	1	1	-1	-1	-1	-1	-1	-1		
E _{2g}	1	1	1	1	-1	-1	-1	-1	-1	-1	1	1	z	
A _{1u}	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1		
A _{2u}	1	-1	1	-1	-1	1	-1	1	-1	1	1	-1		
	2	1	-1	-2	0	0	-2	-1	1	2	0	0	(x, y)	
B _{1u}	2	-1	-1	2	0	0	-2	1	1	-2	0	0		

D_{2d}	E	2S ₄	C ₂	2C ₂ '	2 d		
A ₁	1	1	1	1	1		x^2+y^2, z^2
A ₂	1	1	1	-1	-1	R _z	
	1	-1	1	1	-1		x^2-y^2
B ₁	1	-1	1	-1	1	z	xy
B ₂	2	0	-2	0	0	(x, y); (R _x , R _y)	(xz, yz)

A-2

D_{3d}	E	2C ₃	3C ₂	i	2S ₆	3 d		
A _{1g}	1	1	1	1	1	1		x^2+y^2, z^2
A _{2g} E _g	1	1	-1	1	1	-1	R _z	
	2	-1	0	2	-1	0	(R _x , R _y)	(x^2-y^2, xy); (xz, yz)
A _{1u}	1	1	1	-1	-1	-1		
A _{2u} E _u	1	1	-1	-1	-1	1	z	
	2	-1	0	-2	1	0	(x, y)	

S₄	E	S ₄	C ₂	S ₄ ³		
A B	1	1	1	1	R _z	x^2+y^2, z^2
E	1	-1	1	-1	z	x^2-y^2, xy
	1	i	-1	-(i)	(x, y); (R _x , R _y)	(xz, yz)

T_d	E	8C ₃	3C ₂	6S ₄	6 d		
A ₁	1	1	1	1	1		$x^2+y^2+z^2$
A ₂	1	1	1	-1	-1		
	2	-1	2	0	0		($2z^2-x^2-y^2, x^2-y^2$)
E T ₁	3	0	-1	1	-1	(R _x , R _y , R _z)	
T ₂	3	0	-1	-1	1	(x, y, z)	(xz, yz, xy)

Appendix B: Constants & Useful Energy Conversions:

Planck's constant, $h = 6.626 \times 10^{-34}$ J-s

Boltzman's Constant, $k = 1.381 \times 10^{-23}$ J/K = $0.6950 \text{ cm}^{-1}/\text{K}$

Speed of light, $c = 2.998 \times 10^8$ m/s

$1 \text{ eV} = 1.60219 \times 10^{-19} \text{ J} = 96.485 \text{ kJ/mol} = 22.58 \text{ kcal/mol} = 8065.5 \text{ cm}^{-1}$

$1 \text{ cm}^{-1} = 11.96 \text{ J/mol} = 2.859 \text{ cal/mol} = 0.1240 \text{ meV}$

Appendix C: Some Direct Products

Note that in some instances, g and u must be added ($g \times g = u \times u = g$; $g \times u = u$), some subscripts must be omitted and 'x' and 'y' must be added ('x' = "x" = 'x' = "y")

C_{2v}	A ₁	A ₂	B ₁	B ₂
A ₁	A ₁	A ₂	B ₁	B ₂
A ₂		A ₁	B ₂	B ₁
B ₁			A ₁	A ₂
B ₂				A ₁

C_{6v}, D₆, D_{6h}	A ₁	A ₂	B ₁	B ₂	E ₁	E ₂
A ₁	A ₁	A ₂	B ₁	B ₂	E ₁	E ₂
A ₂		A ₁	B ₂	B ₁	E ₁	E ₂
B ₁			A ₁	A ₂	E ₂	E ₁
B ₂				A ₁	E ₂	E ₁
E ₁					A ₁ +A ₂ +E ₂	B ₁ +B ₂ +E ₁
E ₂						A ₁ +A ₂ +E ₂

C_{5v}, D₅					C₆
A ₁	A ₁	A ₂	E ₁	E ₂	C_{6h} A B
A ₂		A ₁	E ₁	E ₂	A A B B
E ₁			A ₁ +A ₂ +E ₂	E ₁ +E ₂	A E ₁
E ₂				A ₁ +A ₂ +E ₁	

C_{4v}, D₄, D_{2d}	A ₁	A ₂	B ₁	B ₂	E
A ₁	A ₁	A ₂	B ₁	B ₂	E
A ₂		A ₁	B ₂	B ₁	E
B ₁			A ₁	A ₂	E
B ₂				A ₁	E
E					A ₁ +A ₂ +B ₁ +B ₂

O_h, T_d	A ₁	A ₂	E	T ₁	T ₂
A ₁	A ₁	A ₂	E	T ₁	T ₂
A ₂		A ₁	E	T ₂	T ₁
E			A ₁ +A ₂ +E	T ₁ +T ₂	T ₁ +T ₂
T ₁				A ₁ +E+[T ₁]+T ₂	A ₂ +E+T ₁ +T ₂
T ₂					A ₁ +E+[T ₁]+T ₂

D₂, D_{2h}	A	B ₁	B ₂	B ₃
A	A	B ₁	B ₂	B ₃
B ₁		A	B ₃	B ₂
B ₂			A	B ₁
B ₃				A

C₄, C_{4h}, S₄			
A	A	B	E
B		A	E
E			[A]+A+E

C_{3v}, D₃			
A ₁	A ₁	A ₂	E
A ₂		A ₁	E
E			A ₁ +A ₂ +E

Appendix D: Standard Valence Orbital H_{ii} values (eV).

Atom	ns	np	(n-1)d	n
H	-13.6			1
B	-15.2	-8.5		2
C	-21.4	-11.4		
N	-26.0	-13.4		
O	-32.3	-14.8		
F	-40.0	-18.1		
Si	-17.3	-9.2		3
P	-18.7	-14.0		
S	-20.0	-13.3		
Cl	-26.3	-14.2		
Sc	-8.9	-2.8	-8.5	4
Ti	-9.0	-5.4	-10.8	
V	-8.8	-5.5	-11.0	
Cr	-8.7	-5.2	-11.2	
Mn	-9.8	-5.9	-11.7	
Fe	-9.1	-5.3	-12.6	
Co	-9.2	-5.3	-13.2	
Ni	-9.2	-5.2	-13.5	
Cu	-11.4	-6.1	-14.0	
Zn	-12.4	-6.5		
Ga	-14.6	-6.8		
Ge	-16.0	-9.0		
As	-16.2	-12.2		
Se	-20.5	-13.2		
Br	-22.7	-13.1		
Mo	-8.3	-5.2	-10.5	5
Ru	-10.4	-6.9	-14.9	
Rh	-3.09	-4.6	-12.5	
Pd	-7.3	-3.8	-12.0	
Sb	-18.8	-11.7		
I	-18.0	-12.7		6
Te	-20.8	-13.2		
W	-8.3	-5.2	-10.4	
Re	-9.36	-6.0	-12.7	
Os	-8.5	3.5	-11.0	
Pt	-9.1	-5.5	-12.6	
Au	-10.9	-5.6	-15.1	

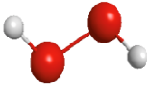
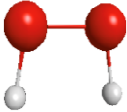

Point Groups and their Detailed List of Symmetry Elements are Included in the Below Table.

Point group	Order of group, h	Type of symmetry elements
C_1	1	E (=C ₁)
C_1	2	E, I (=S ₂)
C_1	2	E, σ
<i>C_n – groups: (h = n)</i>		
C_2	2	E, C ₂
C_3	3	E, C ₃ ¹ , C ₃ ²
C_4	4	E, C ₄ ¹ , C ₄ ² (=C ₂), C ₄ ³
C_5	5	E, C ₄ ¹ , C ₄ ² , C ₄ ³ , C ₄ ⁴
<i>C_{nv} – groups: (h = 2n)</i>		
C_{2v}	4	E, C ₂ , σ
C_{3v}	6	E, C ₃ ¹ , C ₃ ² , 3 σ_v
C_{4v}	8	E, C ₄ ¹ , C ₄ ² (=C ₂), C ₄ ³ , 2 σ_v , 2 σ_v^*
<i>C_{nh} – groups: (h = 2n)</i>		
C_{2h}	4	E, C ₂ , i=(S ₂), σ_h
C_{3h}	6	E, C ₃ ¹ , C ₃ ² , S ₃ ¹ , S ₃ ⁵ , σ_h
C_{4h}	8	E, C ₄ ¹ , C ₄ ² (=C ₂), C ₄ ³ , S ₄ ¹ , S ₄ ³ , σ_h , i=(S ₂)
<i>D_n – groups: (h = 2n)</i>		
D_2	4	E, C ₂ , C ₂ '
D_3	6	E, C ₃ ¹ , C ₃ ² , 3C ₂
D_4	8	E, 2C ₄ , C ₂ , 4C ₂

Point group	Order of group, h	Type of symmetry elements
<i>D_{nh}</i> – groups: (h = 4n)		
D _{2h}	8	E, C ₂ , 2C ₂ ['] , i=(S ₂), σ _h , 2σ _v
D _{3h}	12	E, 2C ₃ , 3C ₂ , σ _h , 3σ _v , 2S ₃ , (S ₃ ¹ , S ₃ ⁵)
D _{4h}	16	E, 2C ₄ , (C ₄ ¹ , C ₄ ²), C ₂ =(C ₄ ²), 2C ₂ ['] , 2C ₂ ^{''} , σ _h , 2σ _v , 3σ _d , i , 2S ₄ (S ₄ ¹ , S ₄ ³)
<i>D_{nd}</i> – groups: (h = 4n)		
D _{2d}	8	E, C ₂ , 2C ₂ ['] , 2σ _d , 2S ₄
D _{3d}	12	E, 2C ₃ (C ₃ ¹ , C ₃ ²), 3C ₂ , i, 3σ _d , 2S ₆ (S ₆ ¹ , S ₆ ³)
D _{4d}	16	E, 2C ₄ (C ₄ ¹ , C ₄ ³), C ₂ =(C ₄ ²), 4C ₂ ['] , 4σ _d , 4S ₈ (S ₈ ¹ , S ₈ ³ , S ₈ ⁵ , S ₈ ⁷)
<i>S_n</i> (n=even) – groups: (h = n)		
S ₄	4	E, S ₄ ¹ , S ₄ ³ , C ₂
S ₆	6	E, S ₆ ¹ , S ₄ ⁵ , C ₃ ¹ , C ₃ ² , i
S ₈	8	E, S ₈ , S ₈ ¹ , S ₈ ³ , S ₈ ⁵ , S ₈ ⁷ , C ₄ ¹ , C ₄ ³ , C ₂ =(C ₄ ²)
<i>Infinite- point group</i> (h=∞)		
C _{∞v}	∞	E, ∞, C _∞ , ∞σ _v
D _{∞v}	∞	E, ∞, C _∞ , ∞σ _v , σ _h , i

2.7 Terminal Questions

1. Draw *trans*, *cis* and *gauche* structures of H₂O₂, and assign these structures to point groups.

trans	cis	gauche
		
C_{2h}	C_{2v}	C₂

2. Write point group for Ethane staggered and Ethane eclipsed, Acetylene, Cis and Trans dichloroethylene.

Ethane in the staggered conformation has 2 C₃ axes, 3 perpendicular C₂ axes bisecting the C–C line, in the plane of the two C's and the H's on opposite sides of the two C's. 3σ_d, i, S₆. There is no σ_h, So Point group is **D_{3d}**.

Ethane in eclipsed conformation has two C₃ axes (the C–C line), three perpendicular C₂ axes bisecting the C–C line, in the plane of the two C's and the H's on the same side of the two C's. Mirror planes include σ_h and 3σ_d. So point group is **D_{3h}**.

Acetylene has a C_∞ axis through all four atoms, an infinite number of perpendicular C₂ axes, a σ_h plane, and an infinite number of σ_d planes through all four atoms. D_{∞h}.

cis-1,2-dichloroethylene has a C₂ axis perpendicular to the C–C bond, two mirror planes (one the plane of the molecule and one perpendicular to the plane of the molecule and perpendicular to

the C–C bond). C_{2v} .

trans-1,2-dichloroethylene has a C_2 axis perpendicular to the C–C bond and perpendicular to the plane of the molecule σ_h a mirror plane in the plane of the molecule, and an inversion center. C_{2h} .

3. Distinguish between the concepts of (a) symmetry operation, (b) symmetry element, and (c) symmetry point group.

(a) Symmetry operation: an action, such as a rotation through a certain angle, that leaves the molecule apparently unchanged.

(b) Symmetry element: a point, a line, or a plane with respect to which the symmetry operation is performed.

(c) Symmetry point group: All of the operations that leave at least one point of the molecule unmoved.

4. For the following molecules and ions, provide the Lewis structure, VSEPR geometry, and assign the point group (the central atom(s) are underlined) and write symmetry operations.

SF_5Cl (C_{4v}), BeH_2 ($D_{\infty h}$), OCS ($C_{\infty v}$), SO_4^{2-} (T_d), CH_3F (C_{3v}), NO_2F (C_{2v}), IF_4^- (D_{4h}), $[TeCl_6]^{2-}$ (O_h), NNO ($C_{\infty v}$), NHF_2 (C_s), XeF_6 (O_h), XeO_3 (C_{3v}), TeF_6 (O_h), $[BrBr_2]^-$ ($D_{\infty h}$), $BrNO$ (C_s), IF_5 (C_{4v}), OSO (C_{2v}), ICl_3 (C_{2v}), $SeCl_4$ (C_{2v}), KrF_2 ($D_{\infty h}$), SCN^- ($C_{\infty v}$)

6. Write point groups and symmetry operations of followings

a) NH_3 ; b) SF_6 ; c) CCl_4 ; d) $H_2C=CH_2$; e) $H_2C=CF_2$.

UNIT 3: X-RAY DIFFRACTION METHODS

CONTENTS:

3.1 Objectives

3.2 Introduction

3.3 Bragg condition

3.3.1 Derivation of Bragg's equation

3.4 Miller indices

3.5 Method of X-ray structural analysis of crystals

3.5.1 Laue's method

3.5.2 Bragg's method

3.5.3 Debye- Scherrer method

3.6 Procedure for an X-ray structure analysis

3.7 Procedure for absolute configuration of molecules

3.8 Ramchandran diagram

3.9 Electron Diffraction: Measurement technique

3.9.1 Scattering intensity vs scattering angle

3.10 Wierl equation

3.11 Elucidation of structure of simple gas phase molecules

3.12 Summary

3.13 Self-Assessment Questions (SAQ)

3.14 Glossary

3.15 Answers to SAQ's

3.16 References

3.17 Terminal Questions

3.1 OBJECTIVES

After reading this unit, we will be able to:

- Explain the concept of X-ray diffraction.
- Define Bragg's law.
- Describe Bragg's method of determining the structure of crystals.
- Describe Laue's method of determining the structure of crystals.
- Explain Debye-Scherrer method of determining the structure of crystals.
- Give the concept of Ramchandran diagram or plots.
- Write Wierl equation.
- Know about scattering intensity and scattering angle.
- Give the application of electron diffraction.

3.2 INTRODUCTION

In lower classes, we have discussed about the electromagnetic radiations. In these radiation, electric field and magnetic field are present and possess speed as the speed of light (3×10^8 m/s). X-ray is one of the electromagnetic radiation that possess energy of 9.0×10^3 kJ/mol. X-Rays have **wavelengths** in the **range** of 0.01 to 10 nanometers, corresponding to **frequencies** in the **range** 30 petahertz to 30 exahertz (3×10^{16} Hz to 3×10^{19} Hz)

Now moving to diffraction, it is defined as the bending (slight) of light when it passes around the edge of an object. The phenomenon of diffraction is shown by electromagnetic radiations like X-rays, gamma rays and microscopic particles like electron, neutron.

In this unit we will discuss about X-ray and its various method of determining the structure of crystals by using the concept of X-ray diffraction that involves Laue's method, Bragg's method,

Debye- Scherrer method. In addition to these methods, we will also discuss about the procedure for an X-ray structure analysis. In order to understand the three dimensional structure of protein in a clear way, Ramchandran diagram or plots were taken into consideration. General Introduction of Electron Diffraction that includes scattering intensity, scattering angle, Wierl equation, and measurement technique molecules are also discussed in this unit.

The entire unit will be clearer after going through each and every topic in detail given below. After reading the unit we will come to know that how on passing X-ray on the surface of crystals we can determine the structure of the crystal. Let us now discuss these topics one by one.

3.3 X- Ray and BRAGG'S CONDITION

Before discussing the Bragg's condition, let us have some general idea of X-rays. X-ray as we know are the electromagnetic radiations having wavelength of 0.1 to 150 \AA . X rays are produced whenever high-speed electrons collide with a metal target in a vacuum tube. A hot tungsten filament is source of electrons. A current is passed through the tungsten filament and heats it up. As it is heated up the increased energy enables electrons to be released from the filament (**thermionic emitter**). High voltage (in kilo volts) is applied between cathode and anode. A small increase in voltage causes large increase in tube current which accelerates high speed electrons from negative cathode (tungsten filament cathode) to positive anode (target tungsten metal anode). The electrons are attracted towards the positively charged anode, the tungsten target with a maximum energy (Energy can be controlled or determined by the tube potential).

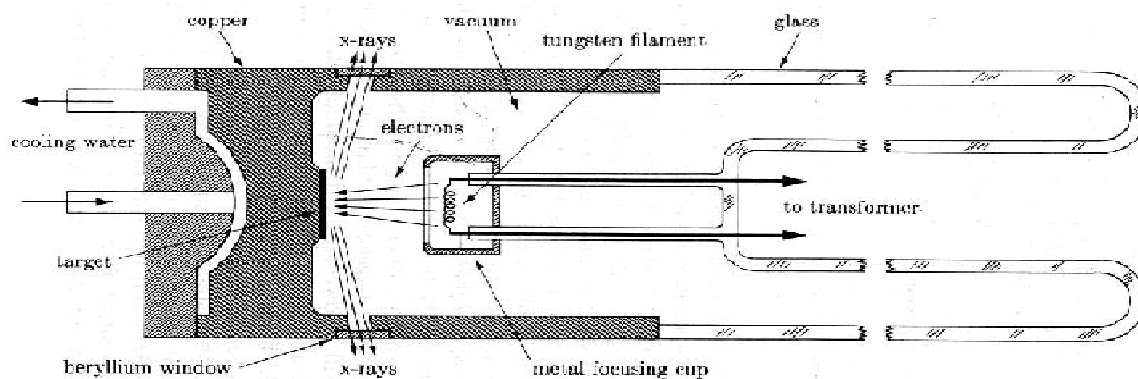


Figure 1: X-ray Tube Crosssection

Ref: Univ of Queensland Physics Museum

When an high energy electron collides with electron present in the inner shell of tungsten then both are ejected out from tungsten atom leaving a hole in inner shell which is filled by electron in outer shell accompanying loss of energy emitted out as an X-ray photon. More than 99% of the electron energy is converted into heat and less than 1% of energy is converted into x-rays.

When high-energy electron interacts with K-shell electron leads to the production of X-ray photons, **K characteristic radiation**..

When high-energy electron interacts with nucleus i.e. it passes near nucleus it is slowed and its path is deflected. Energy lost is emitted as bremsstrahlung X ray photon leads to **bremsstrahlung radiations**.

When high-energy electron interacts with outer shell electrons causes line spectrum

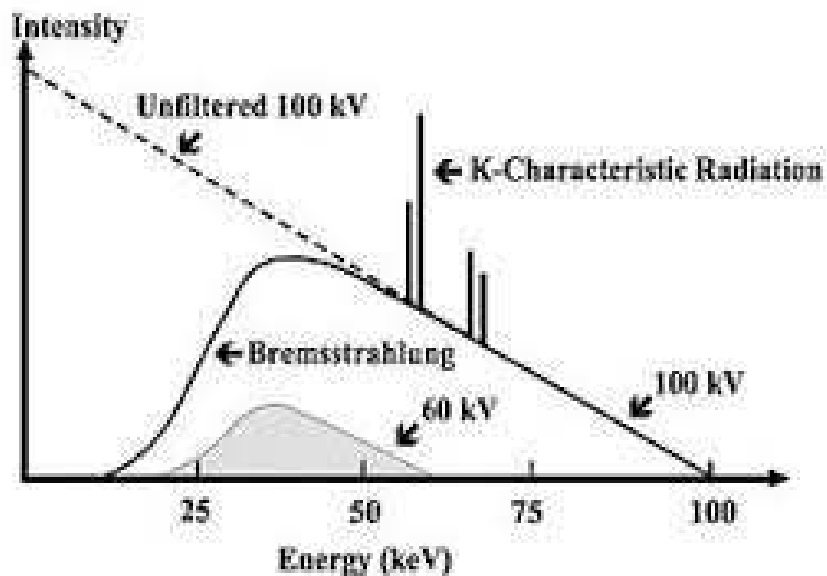


Figure 2 : X-ray Spectrum

Ref : Wikibooks, Basic Physics of Digital Radiography

X-ray diffraction method is used for determining the structure of a molecule. In X-ray diffraction method, X-rays are allowed to pass through the crystal there is interaction between X-rays and atoms of the crystal. The atomic planes of crystal are responsible for interference of incident beam of X-rays with one another when they leave the crystal. So there is scattering of X-rays (electromagnetic radiations). These emitted radiations either overlap with each other to form strong peaks in which bright spots (constructive interference) are developed on the photographic

plate or get subtracted from each other resulting in dark spot (destructive interference) on the photographic plate (interference). This results in diffraction pattern. This is X-ray diffraction or Bragg's diffraction.

From this diffraction pattern we can determine the distribution of electrons (position of particles), nature of interactions within the molecule. Study of X ray diffraction pattern helps to measure the average spacing between layers or rows of atoms. We can measure the size, shape and internal stress of small crystalline regions. In addition to these, we can also have knowledge regarding the atomic vibrations, which are thermal in nature. This method is used generally to reveal atomic structure of crystals and is easy in comparison to other methods. By using X-rays, the structure of crystal is studied which comes under the heading X-rays crystallography.

W. H. Bragg and W. L. Bragg (father and son) explained the X-ray diffraction of crystals. According to them, when X-ray is allowed to fall on the crystal (consisting of the layers of atoms which are equidistant from each other with spacing 'd') surface making an angle (θ) with the crystal, a small portion of X-ray is scattered by the layer of atoms, which are present on the surface. The remaining portion of the X-ray moved to another layer (second layer) of the atom from which some part is scattered and remaining part move to another layer. This scattering refers to diffraction of X-ray. One point to be noted here is that the X-ray diffracted from the layers of atoms obeys the laws of reflection. In short we can say that at a particular wavelength of X-ray and at a particular angle on the crystal surface, there is reflection of X-ray. Bragg's diffraction is also known by name Bragg formulation of X-ray diffraction. On the basis of diffraction of X-ray, they put forward an equation known as Bragg's equation. Diffraction occurs when Bragg's law is satisfied condition for constructive interference from planes with spacing 'd'. Different planes have different spacing and to satisfy Bragg's law θ must change as d changes. Let us derive the equation:

3.3.1 Derivation of Bragg's equation:

In order to derive Bragg's equation, let us consider a crystal that consist of layer (planes) of atoms which are arrange parallel to each other. The two successive layers or plane of the atom are separated by distance (d). X-ray of a particular wavelength falls on the first layer of the

crystal at a point A making an angle θ . It has been seen that some X-ray is reflected at the same angle θ while some penetrate the second layer of the crystal at a point B by travelling the extra distance BC which is shown in Figure 3.

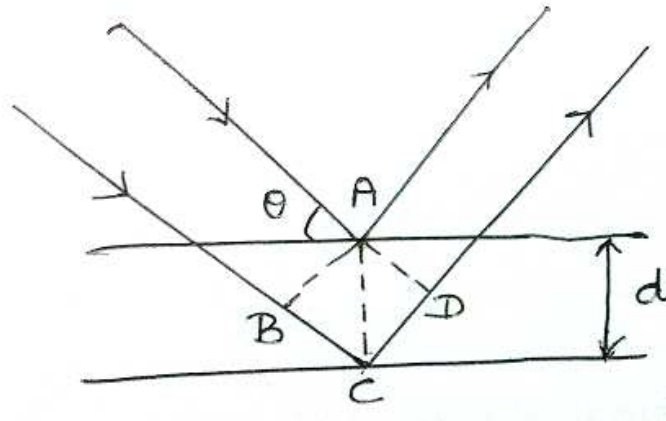


Figure 3. Diagram explaining Bragg's equation

The reflected ray also travels an extra distance CD. Therefore the extra distance travelled by X-ray (BC + CD) is given by:

$$n\lambda = BC + CD \quad \text{here, } n \text{ is any integer}$$

$$BC + CD = 2 d \sin \theta$$

$$\text{As } BC = CD$$

$$n\lambda = 2 BC$$

$$BC = d \sin \theta \quad \sin \theta = BC / d$$

$$2 BC = 2 d \sin \theta$$

Here θ is known as glancing angle. When the reflected rays are in the same phase then extra distance travelled by X-ray is equal to integral multiple (n) of wavelength. Therefore:

$$n \lambda = 2 d \sin \theta \dots\dots\dots (1)$$

Equation 1 is known as Bragg's equation. The integer n has value 1, 2, 3 and so on. For n equal to one, diffraction (reflection) is first order. For n equal to two, the diffraction is said to be

second order. Similarly third order and so on. In order to obey Bragg's equation, the wavelength of X-ray should be less than or equal to $2d$, where d is distance between two successive layers or plane of the atom. It is clear from equation 1 that if glancing angle and wavelength is known, we can find the value of d . Therefore by using Bragg's equation, we can determine the distance between two successive layers or plane of the atom (d) in order to explain its crystal structure. The concept of Bragg diffraction applies equally to neutron diffraction as well as to electron diffraction.

3.4 MILLER INDICES

A crystal consists of a periodic arrangement of the unit cell into a lattice. The unit cell can contain a single atom or atoms in a fixed arrangement. Crystals consist of planes of atoms that are spaced a distance d apart, but can be resolved into many atomic planes, each with a different d -spacing. a, b and c (length) and α, β and γ angles between a, b and c are lattice constants or parameters which can be determined by XRD.

In order to define a particular plane in a crystal, there are some set of integers known as miller indices. These integers are represented by h, k, l . The number of indices is equal to the dimension of the crystal. Like for one dimension crystals, the number of miller indices is one. Similarly for two dimension crystals, the number of miller indices is two. Miller indices of a crystal are always written within the bracket and with negative value of Miller indices, a bar is placed. Miller indices are used to specify directions and planes. These directions and planes could be in lattices or in crystals.

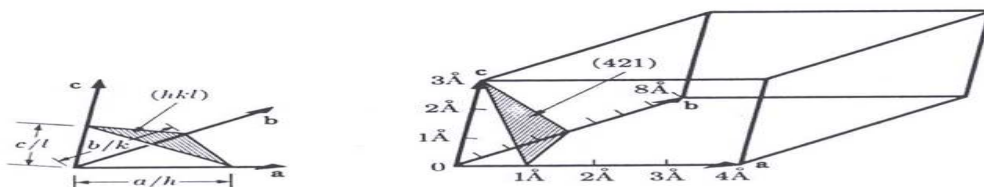


Figure 4: Miller indices are the reciprocals of the fractional intercepts, which the plane makes with crystallographic axes

Indices of a plane of a crystal are obtained by taking the reciprocal of the intercept. The intercept is known as Weiss indices. After taking the intercept, its reciprocal is taken and finally it is multiplied by the smallest number in order to convert reciprocal into integer. For example, there is a crystal plane which cut through the crystal X -axis at point a, Y- axis at point 2b and Z-axis at point 3c. Therefore the crystal plane is given by a, 3b, 2c. In order to find Miller indices we have to follow the procedure discussed above.

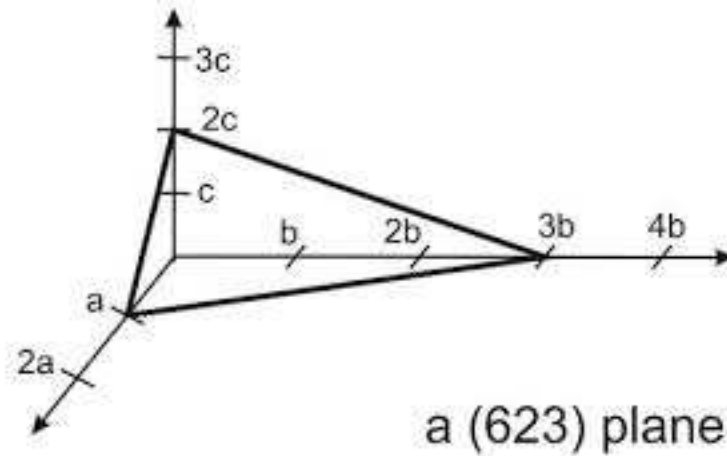


Figure 5: Crystal plane

Crystal plane: a, 3b, 2c

Intercept or Weiss indices: 1, 3, 2

Reciprocal of intercept: $1/1$, $1/3$, $1/2$

Multiplying by the smallest number to get an integer: $6 \times 1/1$, $6 \times 1/3$, $6 \times 1/2$

Miller indices: (6, 2, 3)

The concept of Miller indices is very important and always a question is asked in exam. Another example, for crystal plane: $2a$, $-3b$, $-6c$, Miller indices is obtained by:

Crystal plane: $2a$, $-3b$, $-6c$

Intercept or Weiss indices: 2, -3, -6 so, reciprocal of intercept: $1/2$, $1/-3$, $1/-6$

Multiplying by the smallest number to get an integer: $6 \times 1/2$, $6 \times 1/-3$, $6 \times 1/-6$

Miller indices: $(3, \bar{2}, \bar{1})$

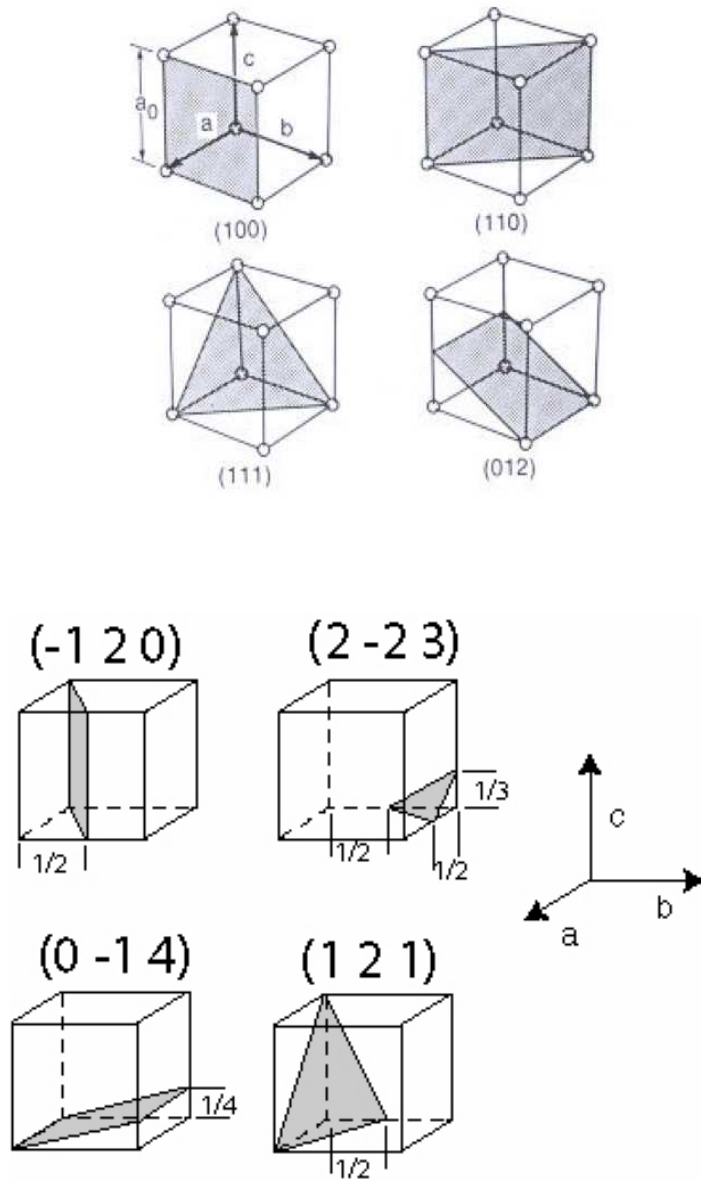


Figure 6: Several atomic planes in cube

3.5 METHODS OF X-RAY STRUCTURAL ANALYSIS OF CRYSTALS

There are following methods of X-ray structural analysis of crystals that includes Laue's method, Bragg's method and Debye- Scherrer method. Let us discuss these methods one by one:

3.5.1 Laue's method:

This is one of the diffraction method which is used for determining the structure of single crystal. According to Laue in 1913, there should be ordered arrangement of atoms in the crystal. In this method, a homogeneous beam of X-ray (narrow) which is produced from tungsten is allowed to fall on a crystal which is small in size. This result in diffraction of X-ray from the crystal. The diffracted beam is then allowed to fall on a photographic plate. In a photographic plate, series of spots appears which are arranged in a proper manner. It was observed that if the diffracted rays are in the same phase then series of bright spots are seen on the photographic plate whereas when diffracted rays are not in the same phase then a series of dark spots are observed on the photographic plate. This array of spots is called a diffraction pattern and the arrangement of spots follows certain curves. Based on how the film is positioned there are two types in Laue method, as seen in figure 4

1. Transmission Laue method - The film is placed in the forward direction (behind the crystal) so that forward scattered radiation is detected. The film is flat and perpendicular to the incident beam and the diffracted beams are partially transmitted through the sample before striking the film and hence the name.

2. Back reflection Laue method - The film is placed between the incident beam and the crystal so that the back reflected rays are used to form the image. The incident beam passes through a hole in the film and falls on the crystal

Laue reflections from a single crystal lie on the surface of a cone whose axis is the zone axis. Zone axis represents the common direction for a set of planes, the direction is said to lie in all the planes belonging to that zone axis. When this cone of reflections intersects with the film placed perpendicular to the primary beam, we get the spot pattern which is clear from Figure 2. Laue

method can be used for checking crystal orientation and quality but cannot be used for any quantitative analysis since the wavelength of the X-ray causing a particular spot is not known.

As in this method photographic plate is used hence this method is also known as Laue's photographic method.

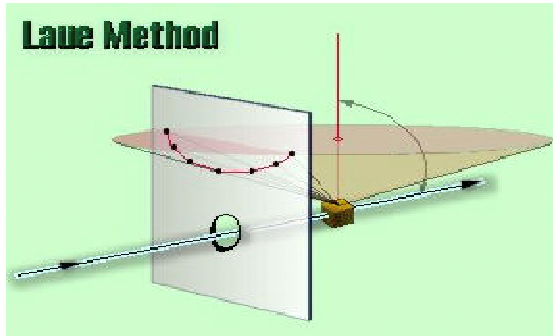


Figure 7: Transmission Laue method

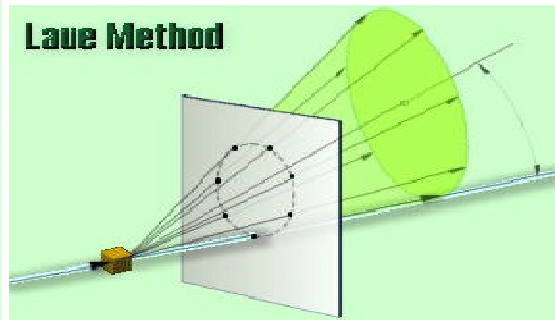


Figure 8: Back Reflection Laue Method

Ref: www.inel.us/index.php/xrd-products/laue-diffraction

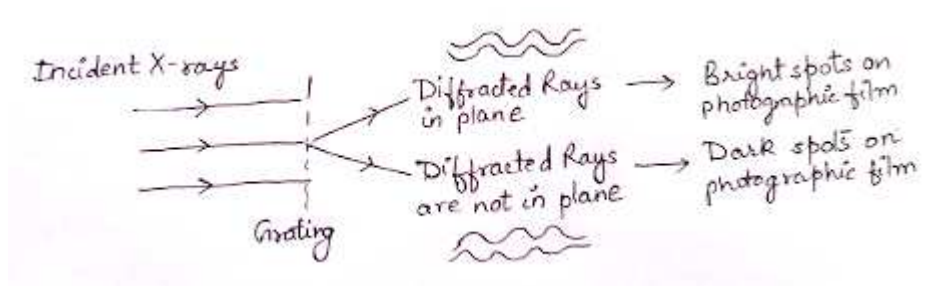


Figure 9. Laue's photographic method

Laue's method is too complicated and difficult to construct. Therefore a simple method of determining the structure of a crystal was given by Bragg.

3.5.2 Bragg's method:

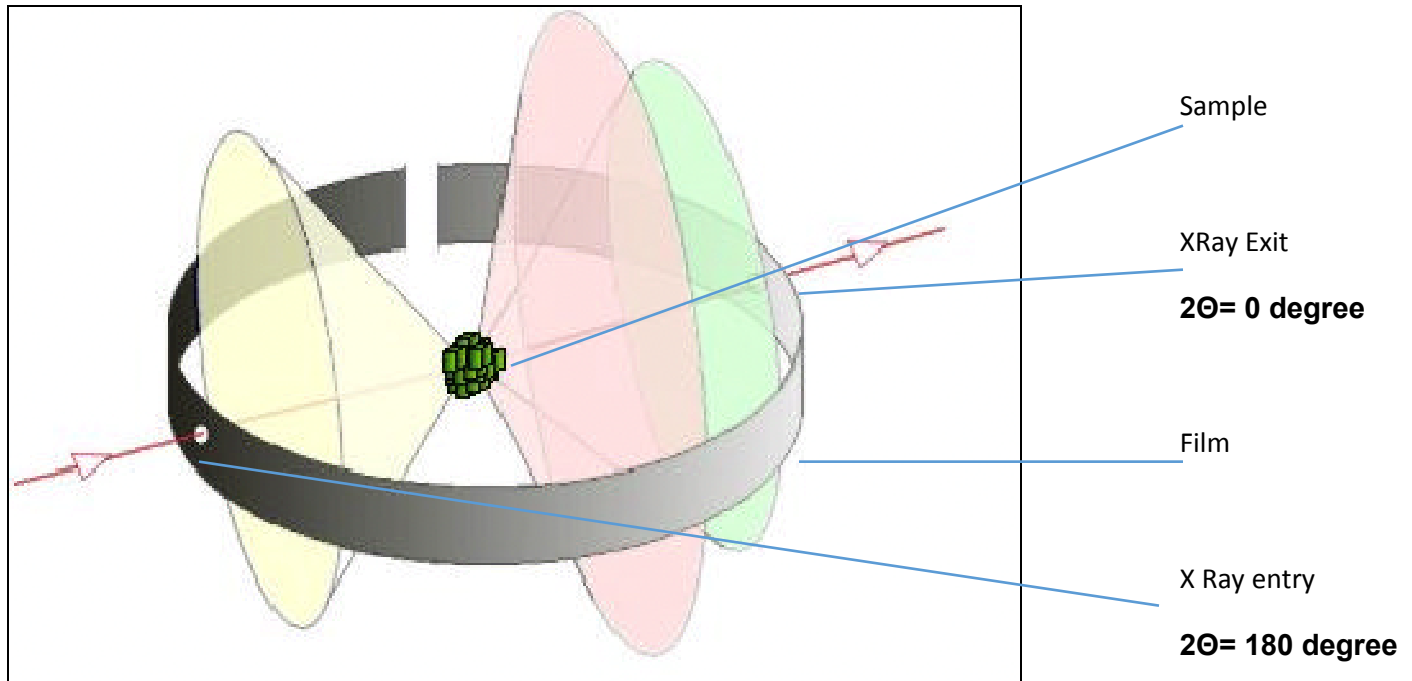
As we have discussed earlier that according to Bragg, X-ray are allowed to fall on a crystal and diffraction from successive plane of the crystal is observed and glancing angle (θ) for every diffraction from the plane is observed by using Bragg's X-ray spectrometer. Then the interplanar distance or distance between two successive plane or layer of an atom of a crystal is determined by using Bragg's equation which is $n \lambda = 2 d \sin \theta$. In this way, 1 for different arrangement of atom in a layer is determined and their ratio is obtained. It has been found that this ratio is

different for crystal having different arrangement. Thus we can predict the structure of a crystal by comparing the experimentally observed ratio with the ratio obtained after calculation by using Bragg's law. For example, for simple cubic crystal:

$$d_{100} : d_{110} : d_{111} = 1 : 1(2)^{1/2} : 1(3)^{1/2}$$

3.5.3 Debye- Scherrer method:

This method as the name indicates was given by Debye and Scherrer in 1916. According to them, crystal is taken in a form of a fine powder into a capillary tube which is placed in a camera. Here photographic plate is taken in a form of circular arc. In this method, on applying X-ray, there is no rotation as in powder form crystal have arrangement in all possible orientations resulting in the maximum diffraction of X-ray from the crystal. A narrow beam of X-ray passes through collimator and strikes the sample crystal present in a capillary tube in the form of powder. A film strip placed along the cylindrical camera wall will be exposed by the scattered radiation and display the diffraction pattern after some time. The diffracted (reflected) X-rays are allowed to fall on the photographic plate of film which is in a form of circular arc as given in Figure 10 and 11. The unscattered radiation leaves the camera via the exit port.



Ref :Physics.stackexchange.com

Figure 10: Debye Scherrer Method A sample of some hundreds of crystals show that the diffracted beams form continuous cones. A circle of film is used to record the diffraction pattern as shown. Each cone intersects the film giving diffraction lines. The lines are seen as arcs on the film

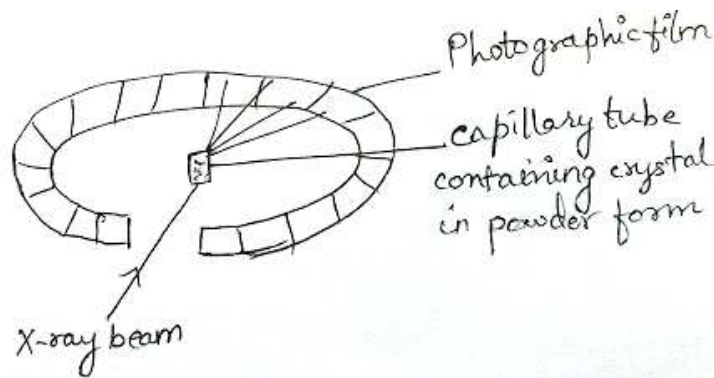


Figure 11. Diffraction for powder method

On the photographic plate bright areas are observed which are at different distance from each other. Then by using Bragg's equation, we can convert distance (d) into glancing angle (θ) where camera is used for measurement of accurate lattice parameters.

3.6 PROCEDURE FOR AN X-RAY STRUCTURE ANALYSIS

From above discussion it is clear that from diffraction pattern of X-ray from a crystal is used for structure analysis of crystal. How this structure elucidation takes place or procedure is discussed here. The procedure involves Bragg's law, scattering factors, factors related to structure and finally its relation to electron density. The concept of Bragg's law has been discussed earlier through which we can determine interplanar distance (d) and glancing angle. Now we will move to scattering factor. The diffraction of X-ray is due to the interaction between X-ray and electrons of an atom. Thus the diffraction depends on the number of electrons which is expressed in terms of scattering factor. It is represented by f . It has been observed that larger the value of f , more will be the diffraction of X-ray from the crystal. Thus the scattering factor (f) depends on the distribution of electron density (ρ) of an atom.

Let us now move to another step i.e. factor related to structure or simply structure factor. Each structure factor refers to amplitude of diffraction from a particular plane having set of integer h along X-axis, k along Y-axis while l along Z- axis. Structure factor is represented by F_{hkl} . If f_i represents the scattering factor where i indicates number of atoms, then the diffraction from plane having set of integer (hkl), structure factor is given by:

$$F_{hkl} = \sum f_i e^{i\phi_{hkl}^{(i)}}$$

Where $\phi_{hkl}^{(i)} = 2\pi(hx + ky + lz)$ and represents phase difference. It has been observed that the intensity of diffracted X-ray is proportional to the square of structure factor. Therefore we can calculate structure factor by taking the square root of intensities of diffracted ray. Now by knowing the structure factor, we can determine the electron density (ρ) by the relation given below:

$$\rho = 1/V \sum F_{hkl} e^{-2\pi i(hx + ky + lz)}$$

where V represents the volume of the unit cell. Electron density refers to probability of finding the electrons around the nucleus of an atom. In this way starting from Bragg's law to electron density, we can predict the structure of a crystal.

3.7 PROCEDURE FOR ABSOLUTE CONFIGURATION OF MOLECULE

In order to predict an absolute configuration of a molecule or a crystal i.e. whether it is simple cubic lattice or body-centered cubic lattice or face-centered cubic lattice, the following data is taken into consideration for each type of crystal lattice. We know that interplanar distance (d_{hkl}) in case of cubic lattice is given by expression:

$$d_{hkl} = (h^2/a^2 + k^2/b^2 + l^2/c^2)^{-1/2}$$

For cubic lattice: $a = b = c$

$$\text{Therefore: } d_{hkl} = (h^2/a^2 + k^2/a^2 + l^2/a^2)^{-1/2}$$

$$d_{hkl} = a / (h^2 + k^2 + l^2)^{1/2} \dots\dots\dots (1)$$

From Bragg's equation: $n \lambda = 2 d_{hkl} \sin \theta$

$$\text{For } n = 1; \lambda = 2 a \sin \theta / (h^2 + k^2 + l^2)^{1/2}$$

$$\text{On squaring: } \lambda^2 = 4 a^2 \sin^2 \theta / h^2 + k^2 + l^2$$

$$\sin^2 \theta = (\lambda^2 / 4 a^2) (h^2 + k^2 + l^2) = K (h^2 + k^2 + l^2) \dots\dots\dots (2)$$

Where K is constant for a particular cubic lattice and possess value equal to $(\lambda^2 / 4 a^2)$. Let us now discuss different cases:

Case 1: For simple cubic lattice:

For different set of integers (hkl), the interplanar distance (d_{hkl}) is calculated from equation 1 and $\sin^2 \theta$ value from equation 2. For example for hkl value 100:

$$d_{hkl} = a / (h^2 + k^2 + l^2)^{1/2}$$

$$d_{hkl} = a / (1^2 + 0^2 + 0^2)^{1/2}$$

$$d_{hkl} = a$$

$$\sin^2 \theta = K (1^2 + 0^2 + 0^2) = K$$

In this way, the value of d_{hkl} is found to be a while the value of $\sin^2 \theta$ is K . Similarly for other values of hkl , the value of d_{hkl} and $\sin^2 \theta$ is calculated for set 110, 111, 200, 210, 211, 220 etc. From the calculated values it is clear that diffraction pattern consist of series of six equally spaced six lines for every all values of hkl . Here spacing between series of lines is K .

Case 2: For face- centered cubic lattice:

In this case also like simple cubic lattice, the value of d_{hkl} and $\sin^2 \theta$ is calculated for different set of integer starting from 100, 110, 111, 200, 210, 211 and so on. It has been found that in this case, diffraction lines are obtained only from those planes having the value of all integer (h,k,l) either odd or even. Like for set 111, here all integer possess odd value therefore lattice is face –centered cubic. Similarly for all even value like 200, 222, again face –centered cubic lattice. The value of d_{hkl} and $\sin^2 \theta$ for all type of cubic crystal is same for same set of integers.

Case 3: For body- centered cubic lattice:

In case of body- centered cubic lattice, where in addition to atoms at the corner, the body of the crystal contains an atom. Here also the value of d_{hkl} and $\sin^2 \theta$ is same like other cubic system but has been observed that diffraction lines are obtained only from those planes having the sum of integer ($h+k+l$) equal to even number. Like for 100, ($h+k+l$) value is 1, it is odd, hence diffraction line will not observed. Next to 100, there is another set of integer 110, here the sum of ($h+k+l$) is two, hence from this plane, the diffraction will takes place as it is even. Hence it is concluded that in body-centered cubic lattice, the spacing between the series of lines is equal to $2K$.

Finally it is summarized that by observing the diffraction pattern or lines or separation between the series of lines, we can predict the structure of crystal/ lattice.

3.8 RAMCHANDRAN DIAGRAM

G. N. Ramchandran was an Indian biophysicist who proposed a Ramchandran diagram or Ramchandran plot or $[\Phi, \Psi]$ plot which clearly defines the three dimensional structure of protein. This diagram Ramchandran diagram visualize backbone diheadral angle Ψ against Φ of amino acids residues in protein structures. Proteins are naturally occurring polypeptide that composed of alpha amino acid linked together by $-\text{NH}_2\text{CH}$ bond. Protein as we known contains amino acids

as a repeating unit and consist of polypeptide chain. Amino acid contains two functional group: amino group (-NH₂) as well as carboxylic acid group (-COOH), hence the name amino acid. In protein, the bond of the main chain N-C^α and C^α-C bonds are free to move and can undergoes rotation by making an angle ϕ and ψ known as torsional angle as shown in Figure 4. Ramchandran et al. describe the rotations of polypeptide backbone around the bonds between N-C^α called phi Ψ and C^α-C bonds called psi ψ . Plot provides an easy way to view the distribution of torsion angles in protein structure. it also provides an overview of excluded regions that show which rotations of the polypeptide are not allowed due to steric hindrance. Torsional angles are important parameters which control protein foldings. or each conformation, the structure was examined for close contacts between atoms. Here atoms were treated as hard spheres with dimensions corresponding to their van der Waals radii. According to Ramchandran with the help of computer models the variation of torsional angle (ϕ and ψ) is studied with the aim of finding stable conformations of small polypeptide.

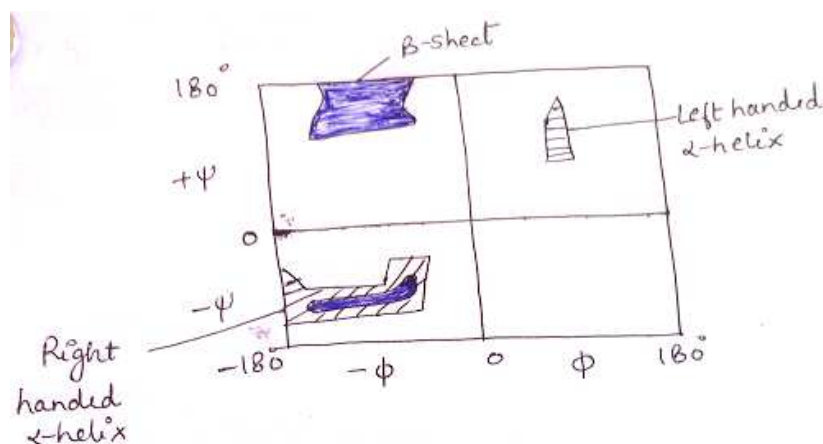


Figure 4. Ramchandran plot

Torsional angle refers to conformation of the polypeptide backbone, which are sterically not allowed.

Form the Figure 4, it is clear that the white part represents conformations where the polypeptide chain contains atoms which are close to each other. This white region refers to all amino acid excluding glycine as the region is sterically not allowed. Blue region refers to region having α -helical and β -sheet conformations as there is no static clashes. The remaining portion other than blue and white represents allowed regions as atoms are allowed to come slight closer to each

other. Finally it was concluded that steric hindrance between the side chains corresponds to disallowed region. Side chain involves the interaction between main chain atoms and side chain C-beta methylene group. Among amino acid, glycine is one which do not possess side chain having large area. In this way Ramchandran explains three dimensional structure of protein.

3.9 ELECTRON DIFFRACTION: MEASUREMENT TECHNIQUE

In X-ray diffraction, the diffraction pattern is observed by using X-ray. Similarly in electron diffraction, the diffraction pattern is observed by using electrons. According to de Broglie concept of dual nature i.e. electron possesses both particle nature having a particular mass and wave nature having a particular wavelength. The electron diffraction phenomenon was shown by C. J. Davison and I. H. Germer by using nickel crystal. On applying a known potential difference, electrons can be accelerated to controlled energy. It has been observed that electrons accelerated to an energy in the order 40keV can show the phenomenon of electron diffraction. The crystal sample taken for electron diffraction is very small in amount and this amount of sample is several million times smaller than the sample used in X- ray diffraction. On applying electron diffraction to gases, it is useful in measuring the bond distance and bond angle.

By using the concept of electron diffraction, we can determine shape, bond angle and bond distance. In order to measure electron diffraction, following steps are considered. Firstly by using a hot filament, electrons are produced then these electrons are accelerated by applying a constant electric field of 40 kV/cm and finally these accelerated electrons are passed through the sample which is in a form of solid crystal, gas or liquid in a form of vapour where the sample is maintained at low pressure as given in Figure 5. Now on the photographic plate, after several hours the diffraction pattern will be observed that consists of concentric rings which is similar to the pattern that was observed when Debye-Scherrer powder method of X-ray diffraction was considered. The diffraction pattern on the photographic plate consists of central region which is dark around which concentric rings of varying intensities are observed. Dark region represents un-diffracted beam. By using idea of intensities and concentric rings, we can predict whether a molecule or a sample that undergoes diffraction possess linear or non- linear, planar or non-planar structure along with its bond angle and bond length.

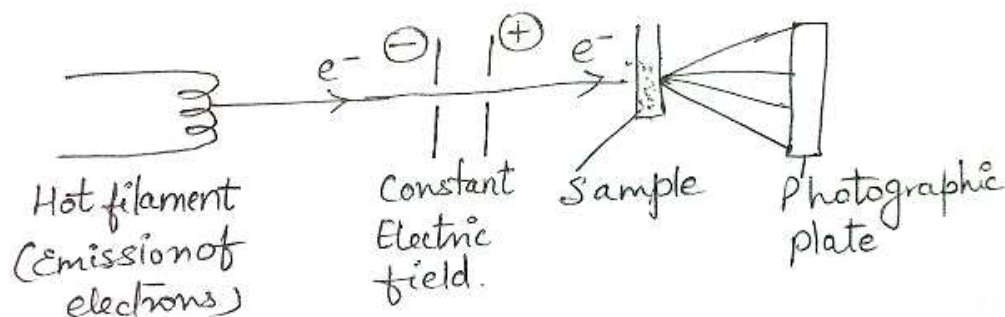


Figure 5. Apparatus for measurement of electron diffraction

The intensities of the diffracted lines depends on the structure of molecules, atomic radii and on the geometry of the apparatus used for the measurement of electron diffraction. As every molecule possesses a particular diffraction pattern hence it is concluded that by using electron diffraction, we can obtain data that predict the structure of molecule.

3.9.1 Scattering intensity vs scattering angle:

Scattering intensity refers to intensities of scattered or diffracted beam which depends on the structure of molecules, atomic radii and the geometry of the apparatus used for measuring electron diffraction. The electron is a low-mass, negatively charged particle. So passing close to other electrons or the positive nucleus of an atom can easily deflect it. These Coulomb (electrostatic) interactions cause electron scattering, which is the process that makes Electron microscopy feasible. Electron scattering is fundamental to all electron microscopy. We know well that your eyes cannot see any object unless it interacts with visible light in some way, for example through reflection or refraction, which are two forms of scattering (e.g., we can't see a light beam unless it is scattered by dust within it or it interacts an object surface). X-rays are scattered at the electrons of the atomic shell. During the scattering process the electron is starts oscilating. It becomes a dipole and a spherical wave is sent out. If the wavelength and energy of the scattered wave does not change then it is elastic scattering. Scattering angle refers to angle by which diffraction occurs. In order to understand the concept of scattering intensity and scattering angle, let us consider a diatomic molecule CD as a sample that undergoes the phenomenon of electron diffraction. Assuming that atom C is present at the origin and atom D is present at distance r from C atom. Φ and α represents the angle through which molecule CD rotate or

orient. Now a beam of electron enters the molecule which is parallel to Y-axis and a diffraction pattern is observed having an angle θ . This is shown in Figure 6.

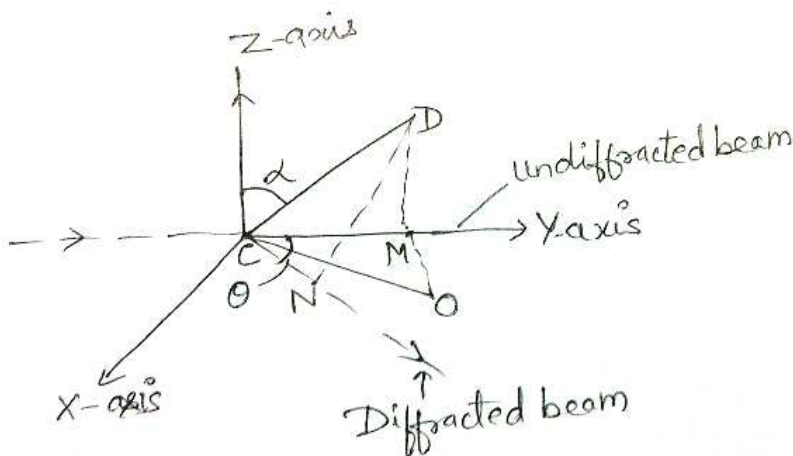


Figure 6. Diffraction through a diatomic molecule (CD)

It is clear from above figure that DN is perpendicular to diffracted beam while DM is perpendicular to undiffracted beam. The average scattering intensity of a randomly oriented group of molecules is given by

$$I_{av}. I_{av} = 2A^2 (1 + \sin sr / sr)$$

as $s = 4\pi/\lambda (\sin \theta/2)$ Where A represents the amplitude of the scattered beam and r is the distance between diatomic molecule C and D.

3.10 WIERL EQUATION

Mark and Wierl in 1930 observed that when accelerated electrons strikes a molecule, there is elastic collision that leads to coherent scattering which is represented by the concentric rings on the photographic plate. For a very complex molecule, the average scattered intensity is obtained by using Wierl equation which is given below:

$$I(\theta) = \sum_{i=1}^n \sum_{j=1}^n f_i f_j \frac{\sin s r_{ij}}{s r_{ij}}$$

Where i and j represents atoms in a molecule, f_i and f_j represents its form factor, r_{ij} represents the distance between the atom in a molecule while n represents the number of atoms and $I(\theta)$ represents the intensity of scattered beam having a particular scattering angle (θ).
 formula for the intensity of an electron beam scattered through a specified angle by diffraction from the molecules in a gas. Wierl equation calculate only intensities but not the intermolecular distance between atoms in the molecule.

3.11 ELUCIDATION OF STRUCTURE OF SIMPLE GAS PHASE MOLECULES

As we have discussed earlier that by using the concept of electron diffraction, we can obtain bond angle and bond length in simple gaseous molecules. For simple molecules, the result obtained from X-ray diffraction pattern and electron diffraction pattern are compared. It has been observed that when number of atom in a molecule increases, it is not possible to obtain all structural parameters that are necessary. X-ray diffraction pattern is more useful than electron diffraction pattern which indicates that X-ray diffraction method provide more information regarding the structure elucidation of simple gaseous molecules. Though there are several molecules like CCl_4 , TiCl_4 , SF_6 , C_2H_4 , CH_3Cl which shows the phenomenon of electron diffraction to obtain parameters related to structure elucidation (bond length, geometry etc). Table 1 shows bond length of simple gas molecules.

S. No.	Gas molecule	Bond	Bond Length (pm)
1.	CCl_4	Cl-Cl	285
2.	TiCl_4	Ti-Cl	218
3.	SF_6	S-F	158
4.	C_2H_4	C-H	133
5.	CH_3Cl	C-H	111

3.12 SUMMARY

In the present unit we have discussed about X-ray diffraction that includes Bragg's law along with the Laue's method and Debye-Scherrer method for determining the structure of crystal. To understand the three dimensional structure of protein in more simple way. Ramchandran diagrams are taken into consideration. In addition to these, general idea of electron diffraction which includes measurement techniques has been discussed in the unit. In order to make the unit interesting, the application of X-ray diffraction along with electron diffraction are taken in account in order to study the structure of crystal.

3.13 SELF-ASSESSMENT QUESTIONS (SAQ)

- Fill in the blanks

1. Bragg's equation is
2. In Bragg's equation, θ is known as angle
3. For n equal to one in Bragg's equation, the diffraction is known as order diffraction.
4. Bragg's equation is applied to both and diffraction.
5. Miller indices are always written within the
6. Miller indices having negative value are written by placing over the indices.
7. Weiss indices is equal to the of the crystal plane.
8. In body- centered cubic crystal, diffraction takes place only from those plane of crystal having sum of integer (h+k+l) is equal to number.
9. The separation between the series of lines in simple or primitive cubic lattice is
10. Ramchandran explains the dimensional structure of protein.
11. The repeating unit of protein is
12. In method, crystal in a powder form is taken in a capillary tube having camera.

13. According to method, when diffracted rays are present in phase, we observe bright spots on the photographic plate.

14. The glancing angle is measured by Bragg's X-ray

15. In electron diffraction, the sample used is than X-ray diffraction.

16. Wierl equation is used for determination of intensity of molecules.

17. X-ray diffraction method is Than electron diffraction method in determining the structure of a molecule.

18. Scattering intensity is proportional to the of the amplitude.

3.14 ANSWERS TO SAQ's

- Fill in the blanks

1. $n\lambda = 2l \sin \theta$; 2. Glancing; 3. First; 4. Electron and neutron; 5. Brackett; 6. Bar; 7. Intercept; 8. Even; 9. K; 10. Three; 11. Amino acid; 12. Debye- Scherrer; 13. Laue's; 14. Spectrometer; 15. Smaller; 16. Complex; 17. Better; 18. Square.

3.15 REFERENCES

Kaur, H. (2016), Instrumental Methods of Chemical Analysis, Pragati Prakashan, Meerut.

Raj, G. (2017), Advanced Physical Chemistry, Goel Publishing House, Meerut.

Puri, B. R., Sharma, L. R. and Pathania, M. S. (2012), Principles of Physical Chemistry, Vishal Publishing Co., Delhi.

Bahl, A., Bahl, B. S. and Tuli, G. D (2014), Essentials of Physical Chemistry, S. Chand & Company Pvt. Ltd., New Delhi.

Sharma, B. K. (2006), Spectroscopy, Goel Publishing House, Meerut.

3.16 TERMINAL QUESTIONS

1. Explain the concept of X-ray diffraction along with its application in determining the structure of a crystal that includes Bragg's method, Laue's method and Debye- Scherrer method.
2. Describe in detail the electron diffraction along with its measurement technique.
3. What are X-Rays? Why do we use x-rays to examine crystal structures? How are the X-rays produced in experiment?
4. X rays of wavelength 0.154 nm are diffracted from a crystal at an angle of 14.17° . Assuming that $n = 1$, what is the distance (in pm) between layers in the crystal?

UNIT 4: CHROMATOGRAPHIC METHODS

CONTENTS:

- 4.1 Objectives
- 4.2 Introduction
 - 4.2.1 Principle
 - 4.2.2 Classification of Chromatography Methods
- 4.3 Gas Liquid Chromatography
 - 4.3.1 Principle of GLC
 - 4.3.2 Instruments / Apparatus / Gas Chromatography
 - 4.3.3 Application of GLC
- 4.4 Principle and application of TLC
 - 4.4.1 Experimental procedure:
 - 4.4.2 Applications of TLC
- 4.5 Paper Chromatography
 - 4.5.1 Principle
 - 4.5.2 Pure cellulose paper
 - 4.5.3 Mechanism of paper chromatography
 - 4.5.4 Types of paper Chromatography
 - 4.5.5 Ascending method
 - 4.5.6 Drying the Paper Chromatogram
 - 4.5.7 Applications
- 4.6 Column Chromatography
 - 4.6.1 Preparation of columns
 - 4.6.2 Preparation of the Material
 - 4.6.3 Pouring of the column
 - 4.6.4 Development of chromatogram
 - 4.6.5 Application of column chromatography
- 4.7 high Performance Liquid Chromatography
- 4.8 Ion Exchange chromatography
 - 4.8.1 Ion exchange matrix

4.8.2 Ion Exchangers are Resinous Organic Materials

4.8.3 Ionisable Groups

4.8.4 Applications of Ion exchange chromatography

4.8.5 Separation of Amino Acids

4.8.6 Separation of rare earth elements

4.9 Summary

4.10 Terminal Questions

4.11 References / Bibliography

4.1 OBJECTIVES

For separation and purification of mixtures a very good technique is available which is known as Chromatography. It has been derived from a Greek word: chromo means colour and Gram stands for Bands. Therefore, in chromatography formation of coloured bands are visible. Earlier this technique was used only for the separation of coloured substances which include plant pigments chlorophyll, xanthophyll and carotenoids. However, the technique is now equally well applied to colourless substance.

4.2 INTRODUCTION

In the year 1903 Dr. Michal Tswett used the technique to show the separation of plant pigments by percolating a petroleum ether extract through a glass column packed with powdered calcium carbonate. The various pigments migrated through the column at different rates giving rise to different coloured zones. Extrusion followed by sectioning of the calcium carbonate packing resulted into separation of the components.

The more complex modern chromatographic techniques are used for separations of colourless materials also, but the original term is retained.

Chromatography involves two mutually immiscible phases. One is called stationary & other mobile phase. The stationary phase can be solid or a liquid and the mobile phase can be liquid, gas or a supercritical fluid, which are brought into contact. A sample to be analysed is introduced into a mobile phase. It is then carried through a stationary phase which is packed in the form of column or thin-layer. Different constituents of the sample undergo repeated partitions between

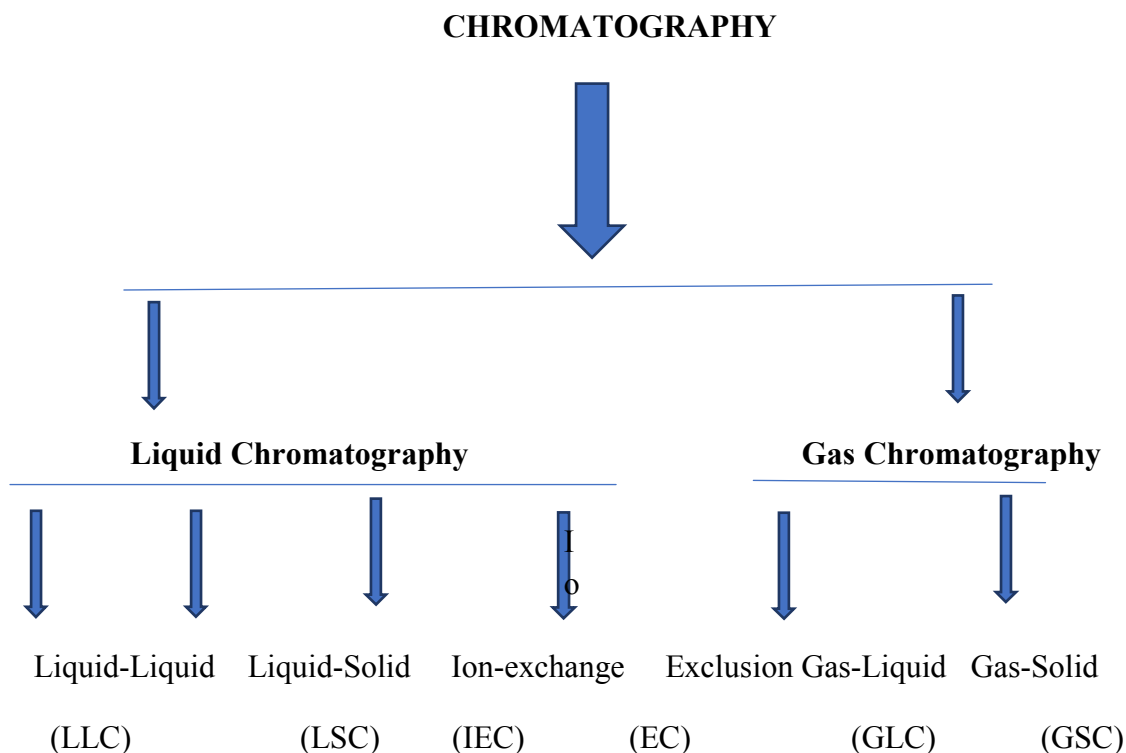
the mobile phase and the stationary phase. This separates different constituents of the sample into bands in the mobile phase. This method has advantage due to the wide choice of materials for the stationary and mobile phases. Therefore, it is possible to separate molecules which differ slightly in their chemical and physical properties.

4.2.1 Principle:

Chromatography is based on the principle of separation of compounds into different bands (colour graphs) and then identification of those bands. Variation in the rate at which different constitutions of mixture sample migrate through a stationary phase under the influence of a mobile phase (because of difference in distributions) is the basic principle of all the chromatographic techniques. After separation of the compounds, they are identified by suitable detection methods.

4.2.2 Classification of chromatography methods:

On the basis of types of mobile & stationary phase and the mechanism of separation the Chromatographic methods can be classified into two categories viz Liquid chromatography and Gas chromatography. They are further divided as below:



- LLC: Separation involves predominantly a simple partitioning between the immiscible liquid phase, one stationary & the other mobile.
- LSC (Adsorption): The stationary phase is solid and the mobile phase is liquid. The retentive ability of the stationary phase is due to the physical surface forces.
- IEC: Counter ions of the stationary phase are selectively exchanged with the ionic compounds of the sample.
- EC: Solute molecules are separated based on molecular geometry and size. Here stationary phase has porous, gel structure which is used in the form of exclusion packings.
- GLC: The mobile phase is a gas and stationary phase is a liquid.
- GSC: The mobile phase is a gas and stationary phase is a solid.

4.3 GAS-LIQUID CHROMATOGRAPHY (GLC):

Martin and Synge introduced in 1941 first introduced the idea of fractionating gases by passage over a solid or a thin layer of non-volatile liquid held on a solid support. The technique of separating the components of vaporised sample has been given a name Gas chromatography. The GLC has become an important technique of Gas-chromatography since the year 1950.

4.3.1 Principle of GLC:

The mixture to be separated is vaporized and sent into the column by a flowing inert gas, called the carrier gas, such as hydrogen or helium. The mixture of the vaporized sample and the carrier gases create mobile phase. The column is packed with a solid on the surface of which is adsorbed a very low volatile liquid which acts as the stationary phase.

The components of the mixture partition between the two phases and get separated in the increasing order of their partition coefficients. The principle of GLC is almost similar to that of liquid-liquid partition chromatography except that the mobile phase in GLC is a gas.

4.3.2 Instruments / Apparatus used in Gas Chromatography:

Gas chromatography is used to separate and analyse mixture of volatile organic liquids. The instrument used in this technique is given below:

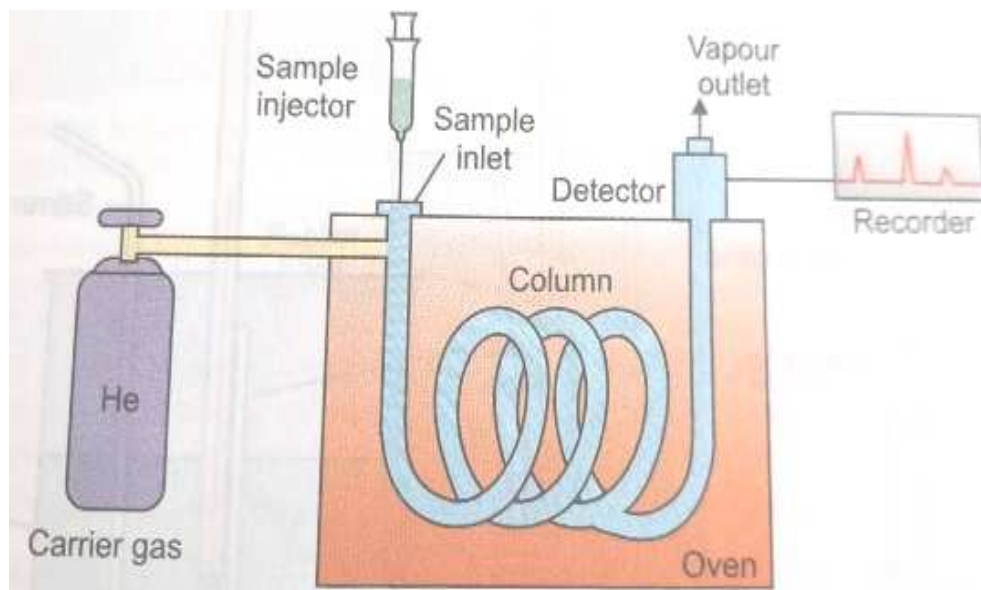


Fig.1. Schematic diagram of a gas chromatograph

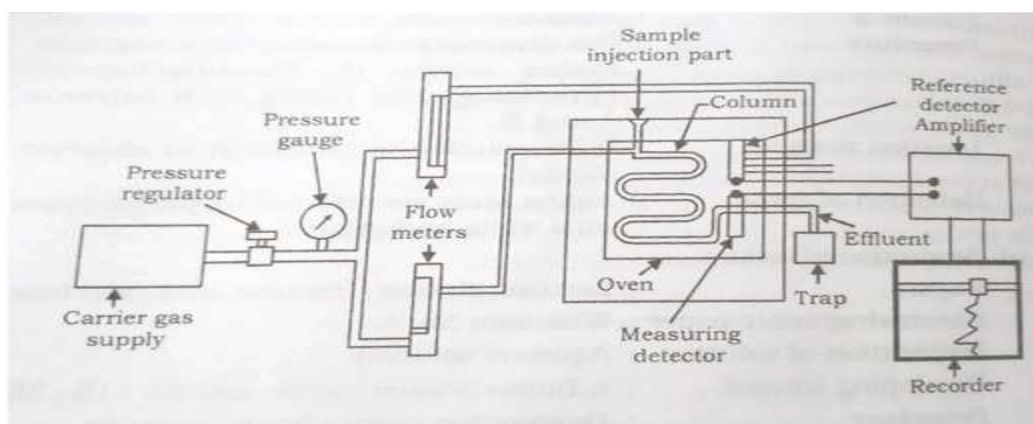


Fig.2. Gas chromatograph

The column packing is made up of porous, inert powder (crushed fire brick) soaked with an oil (silicone or petroleum jelly). The inert gas Helium is passed through the heated column, and sample of the liquid mixture, to be separated, is injected into the helium stream. The mixture vaporises and is carried through the column by the helium flow.

The component of the mixture pass through the column at different rates, some quickly, some slowly. As soon as the component reaches the detector, a peak is provided by the recorder. Therefore, the recorder gives different peaks dedicated to different components as they reach to detector. A gas chromatogram of a mixture of hydrocarbons is shown in Fig. 2.

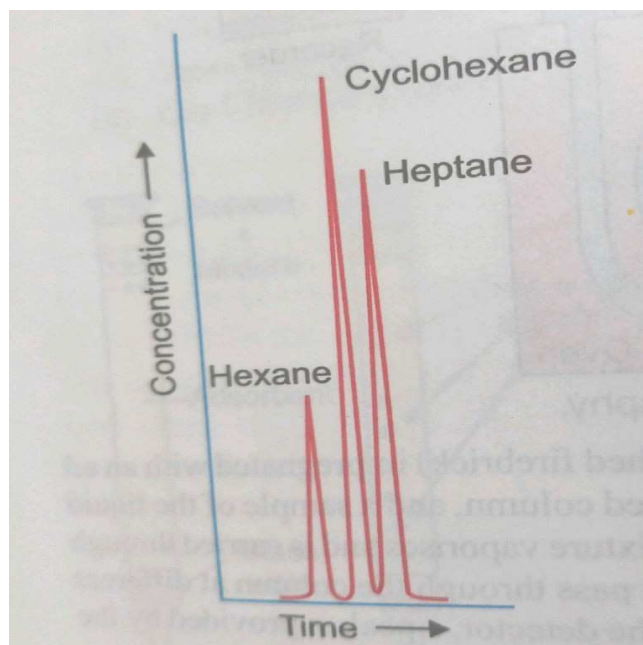


Fig. 3 A gas chromatogram of a mixture of hydrocarbons

The area under the peak is proportional to the concentration of that component in the sample. A component is identified by its retention time which is, the length of time it remains in the column by chilling the exit stream of helium (from vapour outlet) at appropriate time, each individual component can be condensed and collected.

4.3.3 Application of GLC:

For the qualitative and quantitative analysis of the compounds in a mixture GLC is very useful. It helps in the analysis of atmospheric pollutants. In the separation of aromatic compounds GLC has been found to be very effective. Using GLC, complex compounds like vitamins and hormones have been separated and purified. GLC also has been used for ascertaining the homogeneity of a compound (as a pure compound will give only a single bond. Thus, GLC can be used for purification of a compound as an alternative to distillation.

4.4 PRINCIPLE AND APPLICATIONS OF TLC

TLC is a special technique of adsorption chromatography. It is an indispensable tool in the hands of organic chemist for separation of organic mixture fast and convenient. The basic principle of thin layer chromatography is the same as that of column-adsorption chromatography.

A thin layer of adsorbents, such as silica and alumina are spread on glass plate. However, TLC has added advantage over bath column and paper chromatography, these are:

- A variety of adsorbents e.g. alumina, silica gel, cellulose, acetylated cellulose, ion-exchange cellulose powders, etc. and even mixture can be used for preparing thin layers.
- The thickness of the coated adsorbent layer can be varied at will.
- Separation can be by adsorption, Ion-exchange partition or gel filtration depending upon the medium used.
- The spots are very compact, so it is possible to detect compounds at lower concentration than on paper.
- The separated compound can be recovered with greater ease scrapping off the plate and extracting with a suitable solvent.

4.4.1 Experimental procedure:

Following steps are involves TLC technique:

- 1. *Adsorbents:*** Adsorbents possess two important properties. One is its particle size and other is its homogeneity. Most commonly used adsorbents are silica gel, alumina and Kieseluhr. In most cases a binder such as plaster of Paris is used to give plate more strength.
- 2. *Preparation of TLC plates:*** Slurry of adsorbent in organic solvent, preferable petroleum ether is prepared. An organic solvent mixture of low polarity may be used; use of water in preparation of slurry should be avoided. The slurry thus obtained is coated on the glass plate of the size 5x20sqcm.
- 3. *Application:*** A starting line is drawn with a capillary as pipette at a distance of 2 cms from one side of the glass plate. A spot of the solution of the mixture to be separated is kept on this line with the help of capillary. The size of spot should be as small as possible. The solvent is then evaporated.
- 4. *Development of chromatogram:*** The glass plate is then placed vertically in a developing chamber containing solvent at the bottom (about 1 cm deep) so that the solvent is in contact with the adsorbent layer at the end near the sample spot. The chamber is then closed and the solvent is allowed to ascend the layer up to the height of 10-15 cms. If

properly selected, the solvent will resolve the original spot of mixture into a series of spots; each corresponding to a single component the plate is taken of and dried.

5. **Location:** The colourless spots are located (visualised) by keeping the dry plate in an Iodine chamber. The spots begin to shine against brown back ground. Thus, the number of the compounds present in the mixtures is characterised by dividing the distance moved by compound by the solvent point which is also known as retention factor (R_f).

$$R_f = \frac{\text{Distance moved by the compound}}{\text{Distance moved by the solvent}}$$

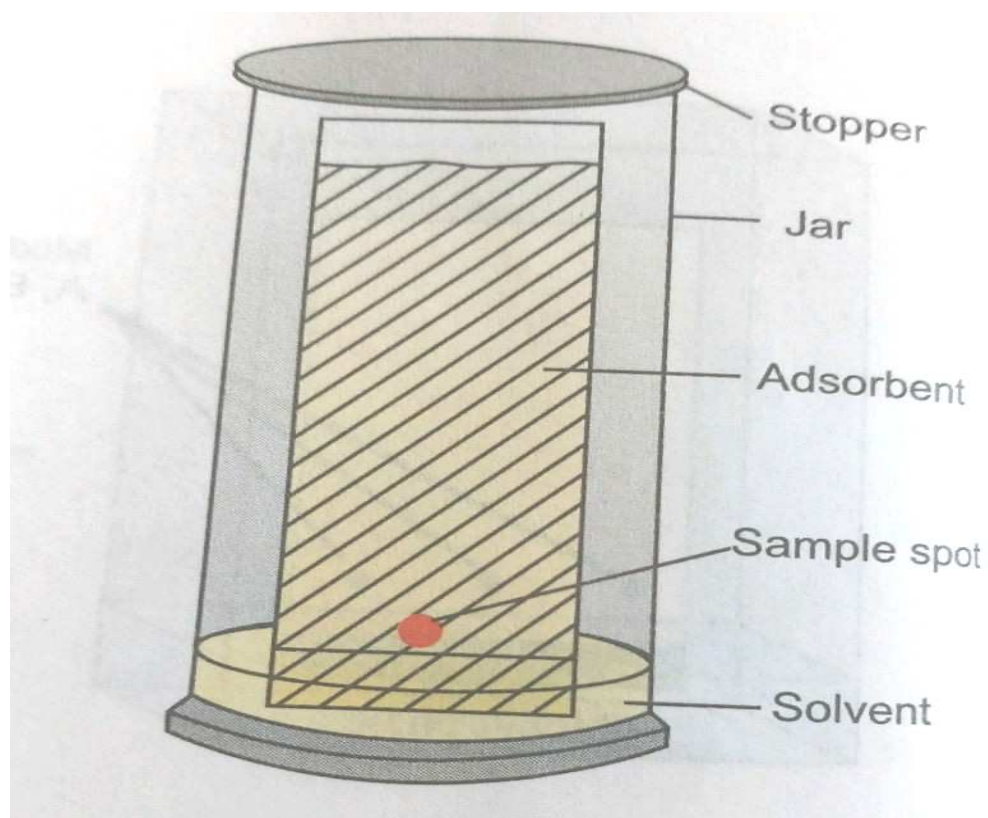


Fig.4 TLC

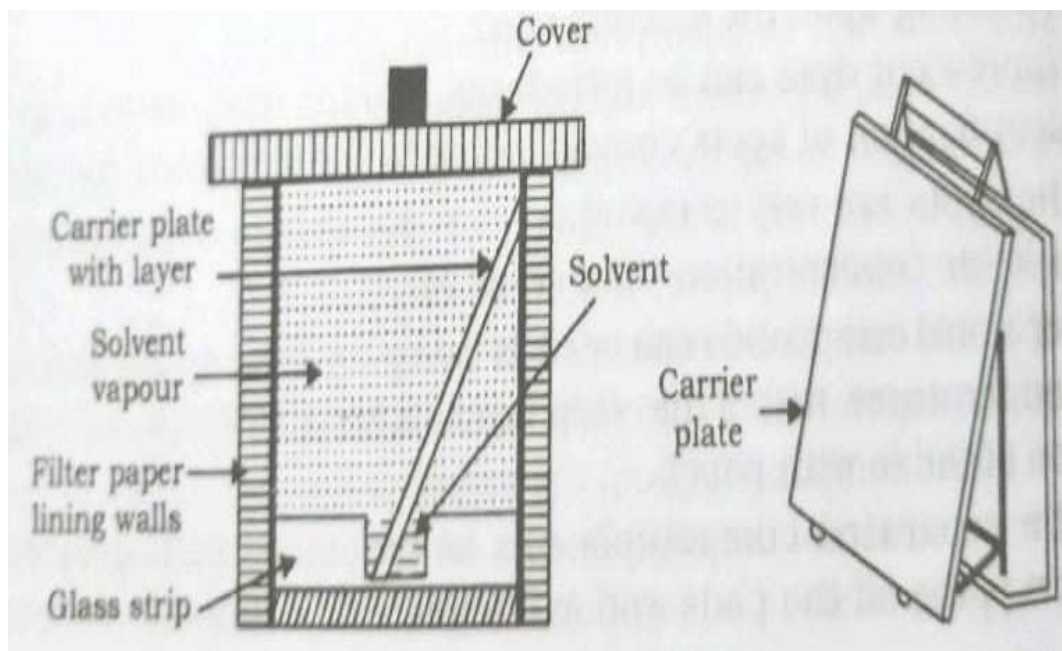


Fig.5. Developing Chamber for Chromatoplates

Mixture no. 1 & 2 contain 3 components each which are different while mixture no.3 contains only one component.

4.4.2 Applications of TLC:

This technique separates and Isolates molecular species. It can be used in place of GLC of natural products.

a. Separation of dyes:

Slurry of silica gel (containing gypsum) in distilled water is prepared and spread over uniformly on the glass plate.

B. Separation of Inorganic ions (copper groups):

Silica gel with starch as the binder in distilled water. Prepared in n-butanol and 1.5N hydrochloric acid. The chromate plate is inserted in a tank filled with hrs gas containing a mixture of n-butanol and 1.5 N HCl appropriate coloured spots.

4.5 PAPER CHROMATOGRAPHY

It is one of the most commonly used methods in the separation and identification of complex organic compounds such as amino acid and proteins.

4.5.1 Principle:

In this technique sheets are used for separation of substances. According to Ingalls (1953) the greater advantage of paper chromatography lies in the simplicity of the operations and many cases relatively small expense of equipment. Even an amateur can assemble a homo-made piece of equipment to demonstrate the separation of organic compound. Paper chromatography is the separation of substances from a mixture by the passage of solvent in a definite direction and selective fixation (Weil, 1953). The cellulose in the form of paper sheets makes an ideal support medium where water is absorbed between the cellulose fibres and forms the stationary hydrophilic phase. The mobile phase (flowing solvent) which is immiscible or slightly miscible with water percolates within the porous structure of the paper.

The technique is based on the different partition coefficient of the components between the two phases. Although separation is affected mainly by partition, but adsorption effected also occur to certain extent (when the stationary phase is not completely inert). On a strip of filter paper (about 30X 5sqcm) a pencil line is drawn at a distance of 5 cm from one end. A minute drop of the test solution is placed at the centre of this line which is called the original line. The solvent is then evaporated and the strip is suspended from the other end in a tall vessel.

Fig. 3 containing the developing solvent (i.e. the mobile phase a solvent mixed with water to prevent dehydration of the paper during the flow of solvent) so that the pencil line end dips in the solvent to a depth 1-2 cm. the top of the vessel is closed with card board so that the suspended paper strip is in as atmosphere of the vapours of both solvents.

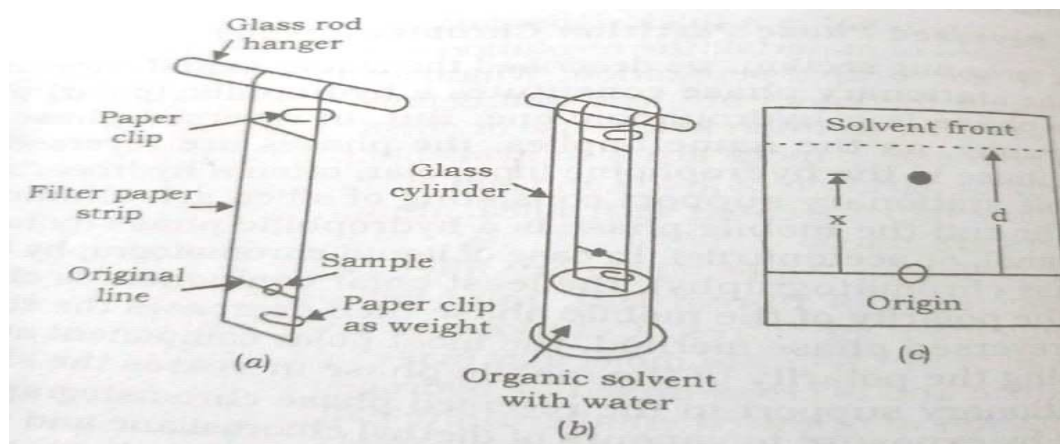


Fig.6. Strip prepared for chromatography analysis. (b) Paper Strip suspended in a container and dipping into a solvent and (c) Determination of the R_f value.

Upon standing the solvent rises over the strip by capillary action. Carrying the components of the sample at different rates according to their partition coefficients. When the solvent reaches almost the top edge of the paper strip, the is removed and the solvent front is marked. The strip is then dried and the positions of the components are visualised by a suitable staining reagent. The dry paper strip is called chromatogram.

The distance of the solvent front and the separated spots (corresponding to different sample components) marked on the strip are measured from the original line. The ratio of the distance moved by a component to that moved by the solvent is known as the R_f value, i.e. $R_f = X / d$, and is more or less constant for particular compound, solvent system, and paper under carefully controlled conditions of solute concentration, temperature the direction of flow of the solvent (i.e. ascending descending, or radical) and PH.

Paper chromatography is useful for the quantitative as well as qualitative separation of the components of the mixture. This technique does not provide a means of separating the sample in pure form but it helps to know how pure is the sample.

1). Choice of Paper: The quality of result depends upon the nature of the paper used. Since a paper strip serves as a stationary support to hold a polar solvent stationary and the mobile phase percolates within the porous structure of the paper. Some of the types of paper use in this technique are described below:

Whatman No. 1 is the paper most frequently used for analytical purposes. Whatman No. 3 MM is a thick paper and is best used for separating large quantities of materials. The resolution is however, inferior to Whatman No. 1. For rapid separation, Whatman No. 4 and 5 are convenient, although the spots are less well defined. The paper may be treated with a buffer solution before use or chemically modified by acetylation. Ion-exchange papers are also available commercially. The separation of liquids and similar hydrophobic substances, pure cellulose paper treated with silica, alumina etc. and Kieselguhr filter paper are used.

4.5.2 Pure cellulose paper:

Ordinary filter paper is a somewhat random arrangement of fibres. But pure cellulose paper is a systematic arrangement of fibres. Cellulose is a fibre network of polymeric carbohydrate chain of molecular mass up to 500,000. It is hydrophilic in nature and cross linked by hydrogen bonding system. These fibres have regions of a high degree of order called crystallites and regions of a low degree of order called amorphous regions. The stationary liquid phase is not uniformly distributed over these fibres but appears to be concentrated in the amorphous regions. The surface and the amorphous regions that hold the adsorbed water are mutually responsible for retention.

4.5.3 Mechanism of Paper Chromatography:

The paper chromatographic separation is affected purely by partition and partly by adsorption. A solution of the mixture of compounds to be separated is prepared by dissolving in water (if soluble) or in an aqueous solution of a suitable organic solvent (if not soluble in water) and a minute drop is placed on a strip of chromatographic paper. The developing solvent (mobile phase) taken in a container is allowed to travel along the strip suspended in the container.

The compounds of the mixture partition between the two phases. Because of the difference in the partition coefficient of the compounds, the rate of travel of each compound is different. As a result, the separation of the compound of the mixture occurs on the strip of paper.

4.5.4 Types of Paper Chromatography:

Essentially the paper chromatography apparatus consists of a sheet of paper supported on a frame so that one edge is dipped into the solvent taken in a container. Depending upon the separation of

the components being carried out in one direction or two directions, there are one dimensional or two-dimensional paper chromatography. One-dimensional paper chromatography: When a sheet of paper is suspended in a container containing the solvent (i.e. mobile phase), the latter travels along the sheet of paper. Depending upon the direction of flow of the mobile phase, one dimensional paper chromatography is known. These are:

4.5.5 Ascending method:

The procedure outlined in the beginning of the section is known as ascending paper chromatography. As stated on a strip of filter paper (about 30x5 cm). a pencil line (called the base line) about 5 cm from one end is drawn along the width of the paper strip. If the mixture to be separated contains four compounds (A, B, C and D) then mark five points on the base line on the filter paper. A minute drop of the solution of the mixture is put on one point and similar drops of solution of each of known compounds (A, B, C and D) are placed in the remaining four points. After allowing the solvent to dry, the strip is then suspended by means of a glass rod tied with paper clip in an air tight glass tank above the developing solvent called eluting solvent (the mobile phase) so that the pencil line end dips in the solvents because of the capillary action the solvent rises. The solvent dissolves the compound and flows up until the force due to capillarity is counter balanced by the downward force due to gravity. When the solvent reaches a suitable height or the top the paper is taken out of the tank the solvent front is marked with a pencil and the paper is allowed to dry. Different compounds in the mixture travel through different heights on the filter paper and correspond to the heights travelled by known compounds (Fig.6). Thus, different compounds in a mixture are separated as well as identified.

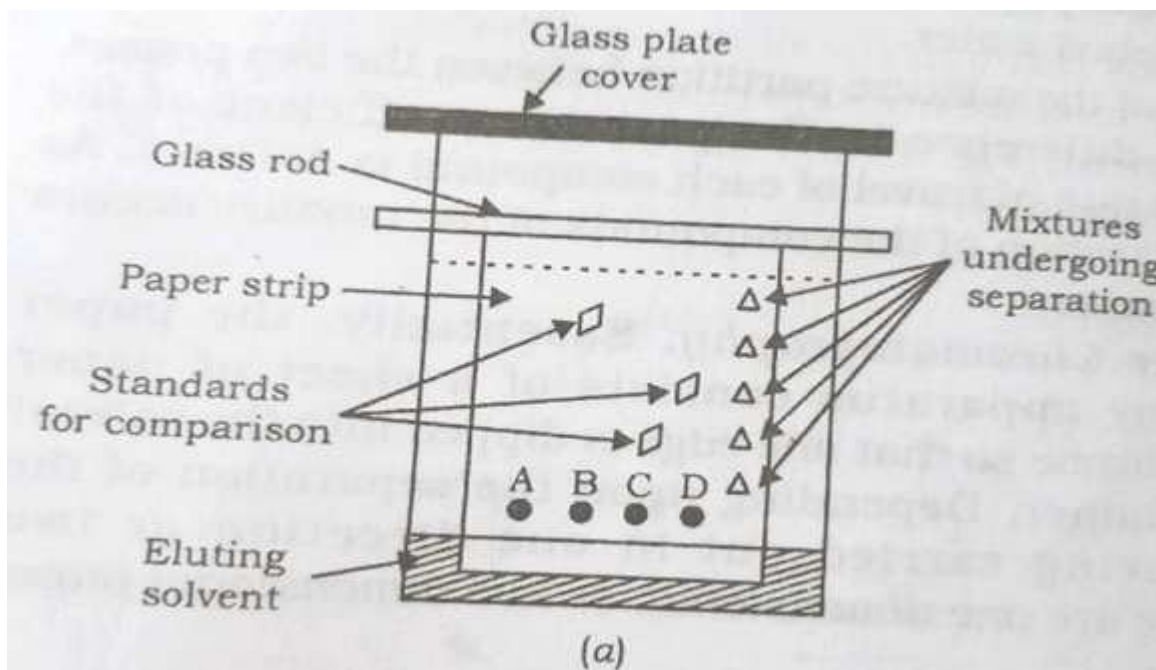


Fig.7. Developing Chamber for Chromatoplates

Two - dimensional paper chromatography: This is convenient for the separation of the compounds which have similar R_f values. Two - dimensional paper chromatography consists of re-chromatography of the separated mixture at right angle to the first direction of development using a different solvent for the second development. The mixture is separated by ascending chromatography in one direction with a solvent which should be volatile than after drying, the paper is turned through a right angle and separation is carried out in the second solvent. After locating the spots, a map is obtained and compounds can be identified by comparing their positions with a map of know compound developed under the same conditions.

4.5.6 Drying the Paper Chromatogram:

After the development of the chromatogram, the paper is removed from the glass tank and the solvent front is marked with a pencil. If the solvent (i.e. the mobile phase) is sufficiently volatile the paper can be dried in a fume-cup board. In case of less volatile solvents the paper can be dried by a hot-air blower or by a drying oven.

4.5.7 Applications:

It is used the separation of organic acids, like oxalic, maloric, succinic tartaric, adipic malic, fumaric and citric acid etc. by used Whatman paper No. 1. It is also used for the separation of sugars.

4.6 COLUMN CHROMATOGRAPHY

This is also known as adsorption chromatography which is the oldest form of chromatography first used in 1903 by Russian Botanist Tswett for the separation of plant pigments. It was further developed by Kuhn and Lederer for the separation of carotene and xanthophylls on a preparative scale.

The column chromatography is based on the fact that when a solution of a mixture is allowed to come in contact with some adsorbent, different constituents of the mixture get adsorbed to the different extent due to difference in the power of adsorption.

We know that in adsorption chromatography, the mobile phase containing the components (of a mixture) to be separated is either a liquid or a gas and the stationary phase is solid. This chromatography may be termed as liquid-solid, or gas-solid, adsorption chromatography.

This liquid-solid adsorption chromatography, the stationary phase is column of solid adsorbent and the mobile phase a solvent, carrying with it the components to be separated is percolated through the column. The separation of the components is based on the selective adsorption of the components from the mobile phase on the stationary phase.

Experimental Procedure: The separation of component in a mixture by adsorption from the mobile phase (liquid) on to the surface of solid adsorbent packed in a column is termed as column adsorption chromatography. Since the separation depends upon the continuous adsorption and desorption of the components on the column by continuous percolating of the solvent down the column. The success of column chromatographic separation depends upon:

- The geometry of the column
- Nature of the eluent
- The rate of flow of eluent and temperature we described the complete procedure which involves the following steps:

4.6.1 Preparation of columns:

Chromatographic columns are usually made of glass. However, long columns give good resolution of components but wide columns are better for dealing with large quantities of

material. A glass tube (burette type) of 1-4 cm diameter and 20-100 cm long with a stop cock at the bottom is used. In this tube the column packing material is taken as the stationary phase. In the lower end of the tube, a sintered glass disc is sealed to keep the stationary phase in position. Alternatively, a cotton plug or a glass-wool plug covered with a filter paper disc can be used as a support for the stationary phase. Usually a ground glass socket is introduced at the top in order to insert a separating funnel which serves as solvent reservoir.

4.6.2 Preparation of the material: They should have the following characteristics:

1. It should be chemically inert.
2. It should have particles of uniform size.
3. It should be finely divided to give high surface area per unit mass.
4. It should have selective adsorption power.

Adsorbents need to be activated by heating or by acid treatment, some gel filtration materials need to be swollen, and ion-exchange resins have to be obtained in the required ionized form by washing. After equilibration thin slurry of the material is prepared.

4.6.3 Pouring of the column: The tube is cleaned with chromic acid, washed with water and dried. The tube is packed with adsorbent by filling it about one-third full with solvent and slowly adding slurry of the adsorbent in the solvent. The stop air bubbles being trapped in the column.

The suspension is allowed to settle and excess solvent runs off to avoid layering effects. The surface of the already packed adsorbent should be stirred with a glass rod before pouring next portion of the slurry. A small amount of the solvent is poured into the column so that it is covered with a layer of solvent. In no case the column, be allowed to dry.

4.6.4 Development of chromatogram: The components from the solution get adsorbed at the top of the column. As the solvent (eluent) passes through the column it re-dissolves the components and carries with it and passes down the column. Because of the selective adsorption capacity of the adsorbent, the various components in the mobile phase will move down the column at different rates and separate out into different bands on the column. The development of the column and the whole column with bands of separated components is termed as chromatogram (shown in Fig.7)

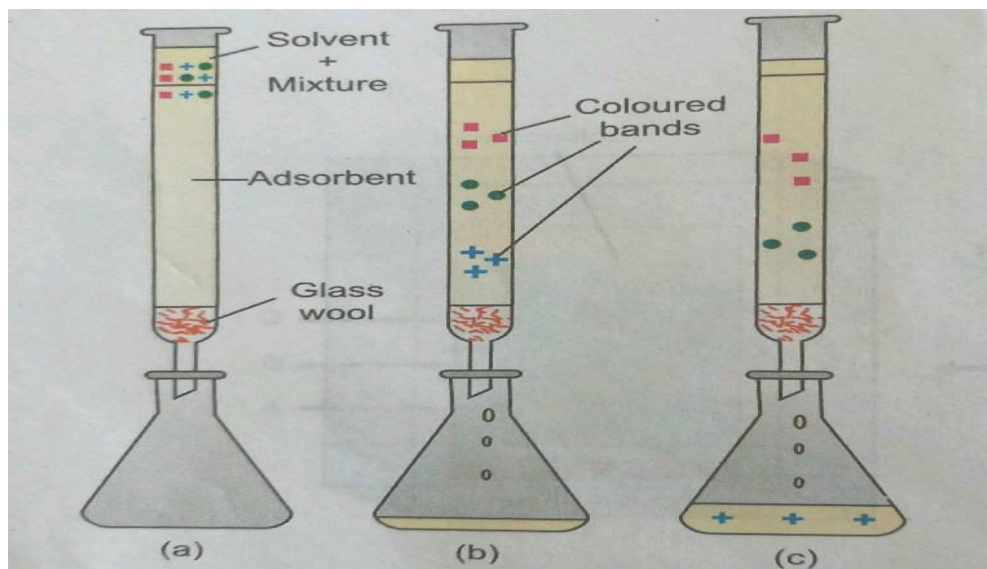


Fig.8. Apparatus for column chromatography: (a) Mixture added to top of the column; (b) appearance of bands; (c) collection of fractions.

4.6.6 Application of column chromatography:

- The solvents for the spectroscopy are purified by these techniques. Since silica gel selectively removes aromatics by using silica gel column. Alumina can be used for the separation of aromatics.
- For the separation of components different adsorbents are used. For example, magnesium silicate is used to separate acetylated sugars, Steroids and essential oils. Alumina or silicic acid is used to separate the compounds containing the oxygenated groups.
- The technique can be used to separate optical geometrical and diastereomers. The homogeneity of a coloured compound can be tested easily.

4.7 HIGH PERFORMANCE LIQUID CHROMATOGRAPHY (HPLC)

High performance liquid chromatography (HPLC) is developed from column chromatography. Much higher efficiencies and better resolutions are obtained in HPLC through the use of smaller particles of stationary phase and pumping of the mobile phase through the column under pressure for achieving higher flow rate. The mobile phase is typically pumped at pressure up to about 3000 psi (2000bar) and flow rates of $1-5 \text{ cm}^3 \text{ min}^{-1}$ can be achieved through

columns (10-25 cm long). The columns are usually packed with smaller sized particles (diameter of the order of 3 μ m).

a. Principle:

Microgram to gram quantities of mixtures can be separated by passage of sample by means of a pressurized flow of the liquid mobile phase through a column containing a stationary solid phase; components of the mixture migrate through the column at different rates due to different relative affinities for the mobile and stationary phase based on charge size or adsorption.

b. Apparatus and instrumentation:

The equipment required for the HPLC is quite different for various components. The equipments used for HPLC are described below:

c. Pumps: To have reasonable flow rate of about 3 ml/min, high pressure of the order of 1000-6000 psi are required. Pressure pumps having outputs of this range are employed for HPLC. Two types of mechanical pumps are used. One is a screw driver syringe and other is a reciprocating pump. The latter pumps are widely used.

d. Solvent reservoir and degasser: An appropriate solvent (mobile phase) is taken in the reservoir for the chromatographic separation. Since the solvent contains dissolved gases usually oxygen and Nitrogen that interfere by forming bubbles in the column (hence band spreading) and the detector system, it is necessary to remove. This is done either by filtering the solvent through a multi pore filter under vacuum before introduction into the reservoir or by means of degassers filled with the reservoirs. Degassers may consist of a vacuum pumping system, a distillation system or an arrangement for heating and stirring the solvent.

e. Mixer: In an isocratic elution (chromatographic separation making use of single solvent, either pure or mixed solvent of fixed composition) mixer is however, not required. But in case of gradient elution, in which two or more solvents of different polarities are used (i.e. mixed solvents of varying composition) a mixer is provided with the modern HPLC equipment.

f. Columns: These are made up of heavy-walled glass tubing or precision bore stainless steel of 15 to 150 cm in length with an inside diameter of about 2 to 3 mm. silica gel is used for packing the columns. Alumina and celite are also used. Particles of uniform size in the range of 3 to

10 μ m are preferable. Column for HPLC employed for separation are often called analytical columns.

Sample injection system: the sample can be introduced into the column by means of a syringe and a self-sealing septum of silicone.

The flow of the mobile phase through the column can be stopped momentarily a cap on the column is removed and the sample introduced directly on to the head of the column by means of a syringe. In modern HPLC, rotary sampling valves are provided for sample injection in to the column.

g. Temperature control: The chromatographic separation by liquid chromatography is carried out at room temperature without their mostating, when temperature control is required, water jacketed column available commonly may be employed for maintaining the desired temperature.

h. Detector and recorder: The detector system employed in HPCL depends on the nature of the sample under investigation and various detectors based on the characteristic properties. Such as absorption of UV, visible IR. Conductance, refractive index, mass spectrometry etc of the sample are used for detecting and recording purposes. The diagram includes a flow splitter which produces different output streams. High-pressure pumps employed to force the mobile phase into the column.

Application of HPLC: Complementary to gas chromatography; it is used largely for the separation of non-volatile substances including ionic and polymeric samples. It is used in pharmaceutical bio-chemical, food product, forensic chemistry and pesticide industries. It is also extremely useful in clinical and environmental studies.

Disadvantages of HPLC:

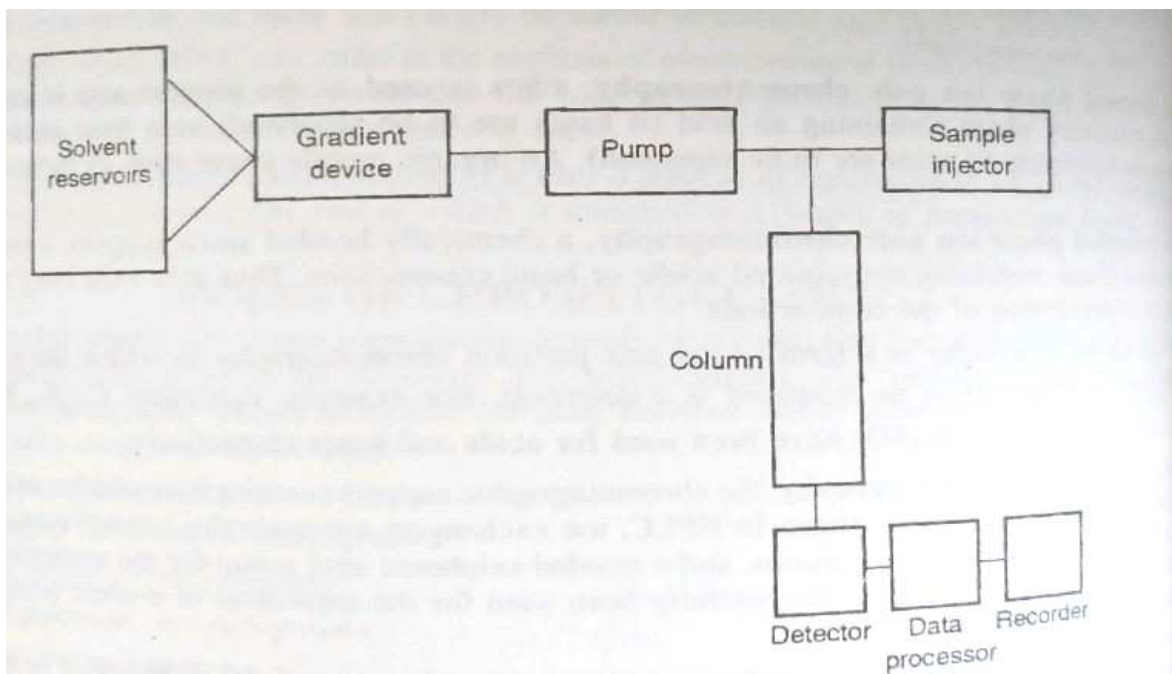


Fig.9. Diagram of a modern HPLC

- Universal detection system not available
- Column performance very sensitive to setting of the packed bed or accumulation of particulate matter or strong adsorbed materials at the top.

i. Comparison between HPLC and GC

HPLC is comparable to GC in speed, efficiency and resolution.

- HPLC is more versatile and thermally stable samples but HPLC is not limited.
- In HPLC choice of stationary phase includes solid adsorbents, chemically modified adsorbents, ion-exchange and exclusion materials thus allowing all four sorption mechanisms to be exploited.
- HPLC has much wider choice of mobile phase than in GC. It facilitates a very considerable variation in the selectivity of the separation process.

4.8 ION EXCHANGE CHROMATOGRAPHY

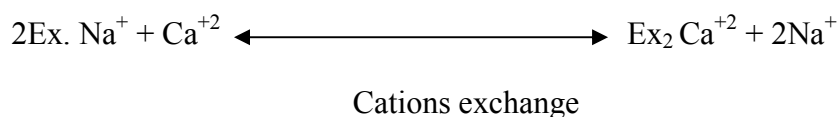
The technique of separating the compounds in a mixture through a reversible process is known as ion exchange chromatography. In this process ions of like signs are exchanged between liquid and solid. Solid is a highly insoluble material which is also termed as ion

exchange. The solids capable of exchanging the ions with other ions in the surrounding solution without any apparent physical change taking place in the matrix is termed as ion exchange chromatography. Alternatively, the exchange of ions of like sign between the solution (mobile phase) and the adsorbent (stationary phase) in contact with the solution takes place in this technique. In column adsorption chromatography the stationary phase consists of an ion-exchange column.

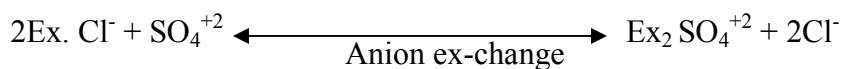
In this technique, the compounds to be separated are previously adsorbed by the adsorbent column and then displaced from the column by percolating a solution containing the component which is more strongly adsorbed by the adsorbent than the previously adsorbed components. This is also called the displacement chromatography.

4.8.1 Ion-exchange matrix:

We learnt that ion-exchange matrix is an insoluble material containing labile ions which can exchange these ions with other ions in the surrounding solution. The substance in which labile ions are cations and exchange cations are known as cation exchanger, eg.



While the substance which exchange anion is known as anion exchanger eg.



Where exchanger represents the solid insoluble matrix containing the monovalent functional (i.e. ionizable) groups responsible for the exchange.

4.8.2 Ion-exchangers are Organic materials:

A wide variety of materials shows ion exchange properties. These properties in clays and zeolites have long been observed. But several synthetic materials which act as ion exchangers have been prepared. Adams and Holmes produced the first synthetic ion-exchange material in 1935 by condensing phenol, sulphonic acid and formaldehyde to form an insoluble resin. Other resins have been prepared largely from aromatic compounds by introducing ionizable group into the aromatic rings by suitable chemical reactions. Thus, an ion exchange consists of an insoluble resinous organic material containing numerous charged group (i.e. ionizable groups).

But the resin as a whole must, of course be electrically neutral, which means that every ionizable group must have an equivalent oppositely charged group associated with it. The identity of the latter is determined solely by chemical equilibrium with the surrounding aqueous solution.

4.8.3 Ionizable groups:

Synthetic ion exchanger resins are high molecular weight polymeric materials containing numerous ionizable groups, ionizable in water and capable of exchanging mobile ion with an aqueous solution. Depending on the nature of the mobile ion (i.e. exchange ions) but not on the nature of the ionizable group (i.e. exchange ions), but not on the nature of the ionizable group (i.e. the charge carried on the matrix) ion exchange resins can be divided into two types: cation and anion exchanger.

The ionizable groups, as the cation exchanger contains numerous negatively charged group, it is called as acid cation exchanger or acid resin and similarly anion exchanger containing positively charged group is called as base anion exchanger or base resin. Anion and cation exchange resins are shown as Fig. 9.

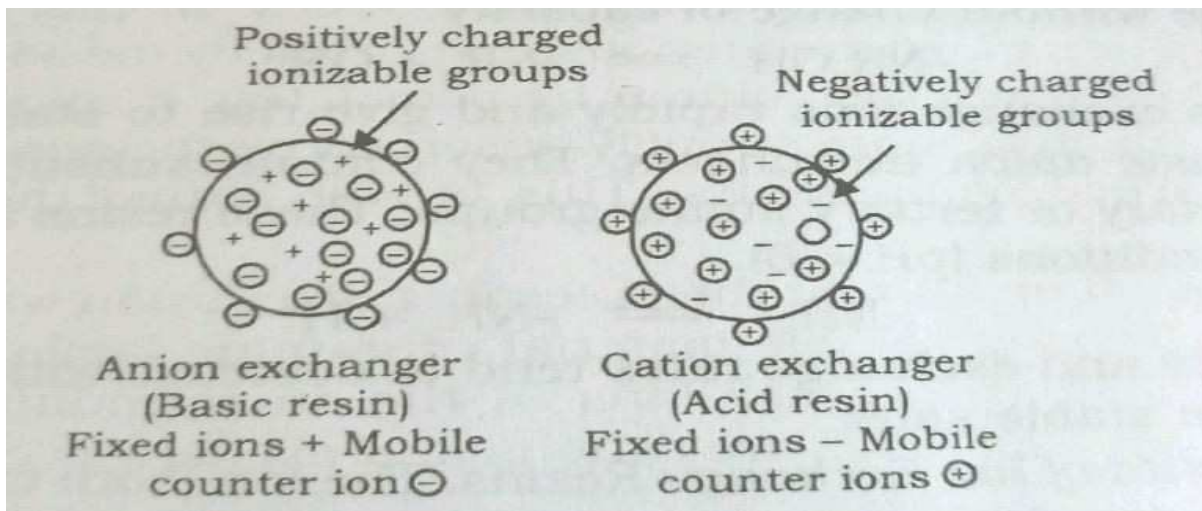
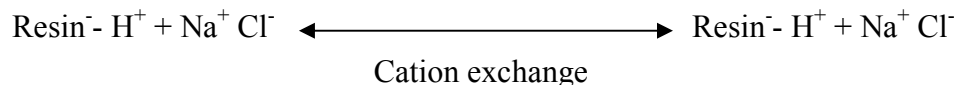


Fig.10. Anion and cation exchange Resins

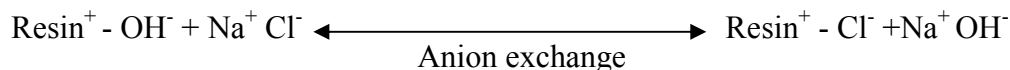
For example, cation exchange resins contain fixed anion sites in such forms as $R_n-SO_3^-H$ (sulphonic acids), $R-COOH$ (carboxylic acid), etc. where R_n represents the polymerized hydrocarbon in acid group, hydrogen is the exchangeable cation and SO_3^-H is the ionizable group. In fact, the anion exchangers usually contain high molecular weight insoluble base with

the fixed ion in form of a cation, such as amino group: R_n-NH_2 , R_n-NH-R , and R_2-NR_2 : and quaternary ammonium groups $R_n.NR_3^+$.

These resins exchange hydroxyl groups (OH^-). The cation exchange reaction of the resin can be written as:



The reaction on anion exchange resin can be shown as:



A cation exchanger replaces all cations with H^+ ion, while the anion exchanger removes the anion, replacing with OH^- ions.

4.8.4 Application of ion-exchange chromatography:

Application of bonded phase-chromatography to higher polar, multi charged and strongly basic samples has been extended by following ion-exchange chromatography. Because of higher efficiency, speed and reproducibility Ion-exchange chromatography has found many applications to separation of compounds.

4.8.5 Separation of amino acids:

For the separation of amino acid preparation of the column by a strongly acidic resin swollen in 4 mol/dm^3 HCl, decanted and washed with 0.1 mol/dm^3 HCl is used for preparing the column of separation of amino acid.

4.8.6 Separation of rare earth elements:

A cation exchange resin is treated with HCl and used for column. A solution of the mixed rare earth chlorides in water is prepared and a citrate buffer is added to it and the mixture is passed through the cation-exchange column. The column is washed continuously with more of the buffer solution and the effluent collected is analysed.

4.9 SUMMARY

- This chapter provides a systematic and comprehensive description of a series of chromatographic separation techniques.
- This chapter covers different techniques, such as partition (or liquid-liquid and gas-solid) adsorption (or liquid-solid and gas-solid), paper thin-layer, and ion-exchange chromatography, moving boundary zone.
- An exciting chromatographic technique such as high-performance liquid chromatography (HPLC) which is relatively new is also presented in this chapter. Applications of various techniques for the separation of organic compounds, Inorganic mixtures and leaf pigments are documented under the appropriate chromatographic technique.

4.10 TERMINAL QUESTIONS

1. What is the basic principle of gas liquid chromatography and its applications?
2. Discuss in brief, TLC and its application.
3. Explain paper chromatography and its mechanism.
4. Write explanatory note on column chromatography.
5. Write short notes on:
 - 5.1. HPLC
 - 5.2. Ion-exchange chromatography.

4.11 TERMINAL ANSWERS:

1. The mixture to be separated is vaporized and sent into the column by a flowing inert gas, called the carrier gas, such as hydrogen or helium. The mixture of the vaporized sample and the carrier gases create mobile phase. The column is packed with a solid on the surface of which is adsorbed a very low volatile liquid which acts as the stationary phase. The components of the mixture partition between the two phases and get separated in the increasing order of their partition coefficients. The principle of GLC is almost similar to that of liquid-liquid partition chromatography except that the mobile phase in GLC is a gas.

Application of GLC:

- For the qualitative and quantitative analysis of the compounds in a mixture GLC is very useful.

- It helps in the analysis of atmospheric pollutants. In the separation of aromatic compounds GLC has been found to be very effective.
- Using GLC, complex compounds like vitamins and hormones have been separated and purified.
- GLC also has been used for ascertaining the homogeneity of a compound (as a pure compound will give only a single bond. Thus, GLC can be used for purification of a compound as an alternative to distillation.

2. TLC is a special technique of adsorption chromatography. It is an indispensable tool in the hands of organic chemist for separation of organic mixture fast and convenient. The basic principle of thin layer chromatography is the same as that of column-adsorption chromatography. A thin layer of adsorbents, such as silica and alumina are spread on glass plate. However, TLC has added advantage over both column and paper chromatography, these are:

- A variety of adsorbents e.g. alumina, silica gel, cellulose, acetylated cellulose, ion-exchange cellulose powders, etc. and even mixture can be used for preparing thin layers.
- The thickness of the coated adsorbent layer can be varied at will.
- Separation can be by adsorption, Ion-exchange partition or gel filtration depending upon the medium used.
- The spots are very compact, so it is possible to detect compounds at lower concentration than on paper.
- The separated compound can be recovered with greater ease scrapping off the plate and extracting with a suitable solvent.

PROCEDURE:

1. Adsorbents: Adsorbents possess two important properties. One is its particle size and other is its homogeneity. Most commonly used adsorbents are silica gel, alumina and Kieseluhr. In most cases a binder such as plaster of Paris is used to give plate more strength.

2. Preparation of TLC plates: Slurry of adsorbent in organic solvent, preferable petroleum ether is prepared. An organic solvent mixture of low polarity may be used; use of water in preparation

of slurry should be avoided. The slurry thus obtained is coated on the glass plate of the size 5×20 sqcm.

3. Development of chromatogram: The glass plate is then placed vertically in a developing chamber containing solvent at the bottom (about 1 cm deep) so that the solvent is in contact with the adsorbent layer at the end near the sample spot. The chamber is then closed and the solvent is allowed to ascend the layer up to the height of 10-15 cms. If properly selected, the solvent will resolve the original spot of mixture into a series of spots; each corresponding to a single component the plate is taken of and dried.

4. Location: The colourless spots are located (visualised) by keeping the dry plate in an Iodine chamber. The spots begin to shine against brown back ground. Thus, the number of the compounds present in the mixtures is characterised by dividing the distance moved by compound by the solvent point which is also known as retention factor (Rf).

$$R_f = \frac{\text{Distance moved by the compound}}{\text{Distance moved by the solvent}}$$

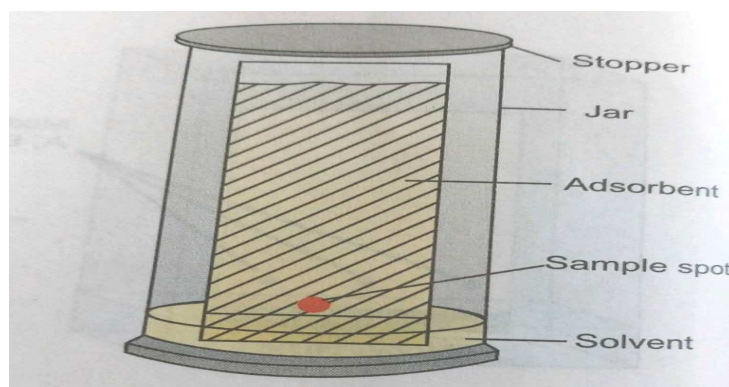


Fig.11 TLC

APPLICATIONS:

This technique separates and Isolates molecular species. It can be used in place of GLC of natural products.

(A). Separation of Dyes:

A slurry of silica gel (containing gypsum) in distilled water is prepared and spread over uniformly on the glass plate.

(B). Separation of Inorganic Ions:

Silica gel with starch as the binder in distilled water. Prepared in n-butanol and 1.5N hydrochloric acid. The chromate plate is inserted in a tank filled with hrs gas containing a mixture of n-butanol and 1.5 N HCl appropriate coloured spots.

3. It is one of the most commonly used methods in the separation and identification of complex organic compounds such as amino acid and proteins. In this technique sheets are used for separation of substances. According to Ingalls (1953) the greater advantage of paper chromatography lies in the simplicity of the operations and many cases relatively small expense of equipment. Even an amateur can assemble a homo-made piece of equipment to demonstrate the separation of organic compound. Paper chromatography is the separation of substances from a mixture by the passage of solvent in a definite direction and selective fixation (Weil, 1953). The cellulose in the form of paper sheets makes an ideal support medium where water is adsorbed between the cellulose fibres and forms the stationary hydrophilic phase. The mobile phase (flowing solvent) which is immiscible or slightly miscible with water percolates within the porous structure of the paper. The technique is based on the different partition coefficient of the components between the two phases. Although separation is affected mainly by partition, but adsorption effected also occur to certain extent (when the stationary phase is not completely inert). On a strip of filter paper (about 30X 5 sqcm) a pencil line is drawn at a distance of 5 cm from one end. A minute drop of the test solution is placed at the centre of this line which is called the original line. The solvent is then evaporated and the strip is suspended from the other end in a tall vessel.

Mechanism: The paper chromatographic separation is affected purely by partition and partly by adsorption. A solution of the mixture of compounds to be separated is prepared by dissolving in water (if soluble) or in an aqueous solution of a suitable organic solvent (if not soluble in water) and a minute drop is placed on a strip of chromatographic paper. The developing solvent (mobile phase) taken in a container is allowed to travel along the strip suspended in the container. The compounds of the mixture partition between the two phases. Because of the difference in the partition coefficient of the compounds, the rate of travel of each compound is different. As a result, the separation of the compound of the mixture occurs on the strip of paper.

4. This is also known as adsorption chromatography which is the oldest form of chromatography first used in 1903 by Russian Botanist Tswett for the separation of plant pigments. It was further developed by Kuhn and Lederer for the separation of carotene and xanthophylls on a preparative scale.

The column chromatography is based on the fact that when a solution of a mixture is allowed to come in contact with some adsorbent, different constituents of the mixture get adsorbed to the different extent due to difference in the power of adsorption.

We know that in adsorption chromatography, the mobile phase containing the components (of a mixture) to be separated is either a liquid or a gas and the stationary phase is solid. This chromatography may be termed as liquid-solid, or gas-solid, adsorption chromatography.

This liquid-solid adsorption chromatography, the stationary phase is column of solid adsorbent and the mobile phase a solvent, carrying with it the components to be separated is percolated through the column. The separation of the components is based on the selective adsorption of the components from the mobile phase on the stationary phase.

Experimental Procedure: The separation of component in a mixture by adsorption from the mobile phase (liquid) on to the surface of solid adsorbent packed in a column is termed as column adsorption chromatography. Since the separation depends upon the continuous adsorption and desorption of the components on the column by continuous percolating of the solvent down the column. The success of column chromatographic separation depends upon:

- The geometry of the column
- Nature of the eluent
- The rate of flow of eluent and temperature we described the complete procedure which involves the following steps:

Chromatographic columns are usually made of glass. However, long columns give good resolution of components but wide columns are better for dealing with large quantities of material. A glass tube (burette type) of 1-4 cm diameter and 20-100 cm long with a stop cock at the bottom is used. In this tube the column packing material is taken as the stationary phase. In the lower end of the tube, a sintered glass disc is sealed to keep the stationary phase in position.

Alternatively, a cotton plug or a glass-wool plug covered with a filter paper disc can be used as a support for the stationary phase.

5. (a) High performance liquid chromatography (HPLC) is developed from column chromatography. Much higher efficiencies and better resolutions are obtained in HPLC through the use of smaller particles of stationary phase and pumping of the mobile phase through the column under pressure for achieving higher flow rate. The mobile phase is typically pumped at pressure up to about 3000 psi (2000bar) and flow rates of $1-5 \text{ cm}^3 \text{ min}^{-1}$ can be achieved through columns (10-25 cm log). The columns are usually packed with smaller sized particles (diameter of the order of $3\mu\text{m}$). Microgram to gram quantities of mixtures can be separated by passage of sample by means of a pressurized flow of the liquid mobile phase through a column containing a stationary solid phase; components of the mixture migrate through the column of different rate due to different relative affinities for the mobile and stationary phase, bases on charge size or adsorption.

(b)The technique of separating the compounds in a mixture through a reversible process is known as ion exchange chromatography. In this process ions of like signs are exchanged between liquid and solid. Solid is a highly insoluble material which is also termed as ion exchanger. The solids capable of exchanging the ions with other ions in the surrounding solution without any apparent physical change taking place in the matrix is termed as ion exchange chromatography. Alternatively, the exchange of ions of like sign between the solution (mobile phase) and the adsorbent (stationary phase) in contact with the solution takes place in this technique. In column adsorption chromatography the stationary phase consists of an ion-exchange column.

In this technique, the compounds to be separated are previously adsorbed by the adsorbent column and then displaced from the column by percolating a solution containing the component which is more strongly adsorbed by the adsorbent than the previously adsorbed components. This is also called the displacement chromatography.

4.11 REFERENCES:

1. I. L. Finar, Organic chemistry, ELBS publication 1962.
2. U. N. Dash, Analytical Chemistry; sultan chand& sons publisher 2005.
3. Arun Bahl& B. S. Bahl, Organic Chemistry, S. Chand & Company-2006.
4. Harbans Singh & M. M. N. Tandon; Organic Chemistry; ShivalalAgrawala&Company ,Publishers 1985.
5. Shashi Chawla; Engineering Chemistry; Dhanpat Rai Co. Publishers – 2001.
6. Vogels Analytical Chemistry. Sultan Chand & Sons publishers 2005.

UNIT 5: GAS CHROMATOGRAPHY

CONTENTS:

- 5.1 Objectives
- 5.2 Introduction
 - 5.2.1 Gas Solid Chromatography (GSC)
 - 5.2.2 Gas Liquid Chromatography (GLC)
 - 5.2.3 Theory of Gas Chromatography
 - 5.2.4 Specific Retention
 - 5.2.5 Relative Retention
 - 5.2.6 Retardation Factor
- 5.3 Techniques of gas liquid chromatography
- 5.4 Parts of Gas Chromatography
- 5.5 Detectors
- 5.6 General Characteristic and applications
- 5.7 Thermal Conductivity Detectors (TCD)
- 5.8 General Characteristics of (TCD) and applications
- 5.9 Flame Ionisation Detectors (FID)
- 5.10 General Characteristics of (FID) and application
- 5.11 Electron Capture Detector (ECD)
- 5.12 General Characteristics and applications
- 5.13 Concept of HEPT: Height equivalent plate theoretical
 - 5.13.1 Particle size and surface area
 - 5.13.2 The rate of flow of carrier gas
- 5.14 Factors affecting separation or (HETP)
- 5.16 Type and amount of stationary phase
- 5.17 Van-Demeter equation
- 5.18 Summary
- 5.19 Terminal Questions
- 5.20 References / Bibliograph

5.1 OBJECTIVES

In this chapter we learn about general characteristics of Gas chromatography and also discuss about its applications.

5.2 INTRODUCTION

In the chromatographic technique a solute undergoes distribution between fixed (stationary) and the moving (mobile) phases. It is the mobile phase which translates the solute down the column unit it eventually emerges from the end of the column separated from other solutes which elute earlier or later.

However, Gas chromatograph (GC) is the technique of chromatography in which the moving phase is a gas. Gas Chromatography was originally suggested by Martin and Synge in the year 1941. It is basically a separation technique in which the compounds of a vaporised sample are separated and fractionated as a consequence of partition between a mobile gaseous phase and a stationary phase held in column. The partition takes place between a gas and liquid or gas or solid.

According to the nature of stationary phase, gas chromatography is divided into two classes. When the fixed phase consists of a solid material, such as granular silica, alumina or carbon, the chromatography is termed as gas solid chromatography (GSC) since the fixed or stationary phase is solid. In another, the fixed phase is a non-volatile liquid held as a thin layer on a solid support. The most common support is diatomaceous earth or Kieselquar. When the stationary phase is a liquid the technique is known as gas liquid chromatography (GLC).

In principle, gas liquid and liquid partition chromatography differ only in the sense that the mobile phase of the former is a gas rather than liquid.

5.2.1 Gas Solid Chromatography (GSC):

The gas solid chromatography is quite limited in applicability, largely because of the difficulty in reproducing surface areas and because of excessive retention of active gases on solid surfaces, which reduce the available area. Gas solid chromatography has also not been used widely because of tailing of elution peaks i.e. because of tailing caused by non-linear adsorption Isotherms.

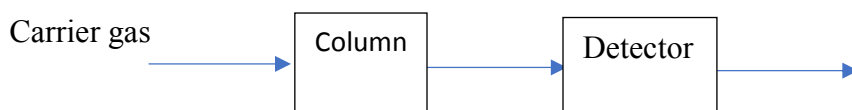


Fig. (A) Basic Chromatograph Arrangement

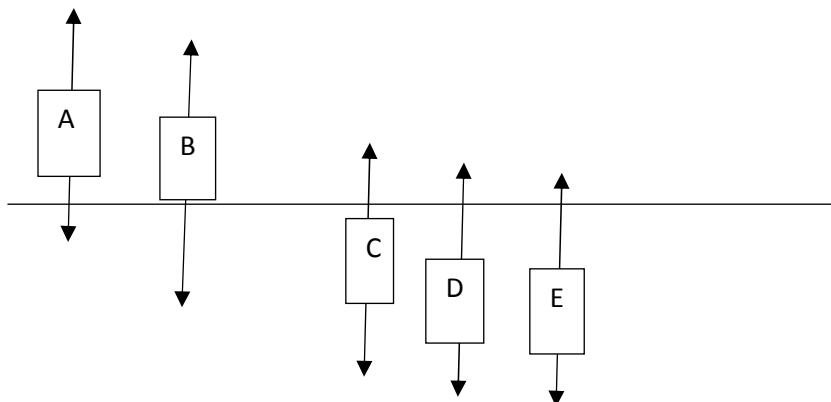


Fig.1. Schematic Representation of partition coefficient

5.2.2 Gas Liquid Chromatography (GLC):

G.L.C. is a widely used technique by all types of chemists especially in pharma industries.

1. In gas chromatography the components of a vaporised sample are fractionated as a consequence of a partition between a mobile gaseous phase and a stationary phase held in a column.
2. According to the nature of stationary phase, gas chromatography may be classified as gas solid and moving phase is gas, the chromatography is called gas-solid chromatography.

Similarly, when the stationary phase is liquid supported on an inert solid matrix and the moving phase is gas, the chromatography is known as gas-liquid chromatography.

3. The principle of GLC is almost similar to that of liquid-liquid partition chromatography, except that the mobile phase in GLC is a gas rather than a liquid.
 4. Gas chromatography provides a quick and easy way of determine the number of components in a mixture, the presence of impurities in a substance and in many cases prima facie evidence of the identify of a compound. The only requirement is same degree of stability at a temperature necessary to maintain the substance in the gas phase.
 5. A gas chromatography (GC) is an essential tool for the chemists who are interested in the synthesis and characterisation of covalent compounds of moderate molecular weight.
 6. In gas-solid chromatography, the stationary phase consists of solid material such as granular silica, alumina or carbon. The separation process involves adsorption on the solid surface. GSC is limited in its applications largely because of tailing caused by non-linear adsorption isotherms and partly because of the difficulty in reproducing surface conditions and partly because of excessive retention of reactive gases which reduces the available area. Surface catalysts also play a restricting role.
 7. Gas chromatography is a general term which is generally applied to separations of gas phase mixtures by interaction with a high boiling liquid or a solid.
 8. The gas separation method now has been developed to the point where it is applicable to almost every area involving the physical and chemical Sciences.
 9. The various important advantages of gas chromatography include:
 - a) Both qualitative and quantitative analysis ae possible.
 - b) The time of analysis is short.
 - c) The instrument is simple.
 - d) The sensitivity is high.
 - e) The method is applicable to about 60% of the organic compounds known to man.
 - f) Separations which are very difficult or virtually impossible by other techniques can be simple and straight forward with gas chromatography. The successful separation of cis and trans isomers, separation of oxygen isotopes or separation of pesticides is a matter of selecting the proper column and conditions.
-

5.2.3 Theory of Gas Chromatography:

Volatile sample which are thermally stable at the operating temperature are introduced into the gas stream via an injection port. Due to different distribution ratios between the mobile phases and stationary phase, different constituents migrate at different rates. A continuous flow of gas elutes the components from the column. In order of increasing distribution ratio, a detector connected to a recording system records the chromatogram.

Generally, a chemically inert gas is employed like He, N₂, Ar or H₂. The temperature of the separating column must be equal to or slightly above the average boiling point of the least volatile component of the sample. A good stationary phase is thermally stable & chemically inert. It has low volatility and good degree of compatibility with analysis.

Generally, thermal conductivity detector, flame ionisation detector or electron capture detectors are used as they have good precision and accuracy, high signal to noise ratio. They are non-destructive and exhibit good response over a wide range of temperature. Detectors are selected according to carries gas. For example: due to high thermal conductivity of He, thermal conductivity detectors are preferred for He. Ultimate accuracy of the chromatogram depends on the choice of the detector recorder which is fed by the output of the detector. The availability of versatile and specific detectors and the possibility of coupling the gas chromatography to an infra-red spectrometer or a mass spectrometer further enhance the usefulness of gas chromatography. The time required for the maximum for solute peak to reach the detector in a gas chromatography column is called retention time. The retention volume V_R , of a component is the volume of gas required to carry a component maximum through the column. We can use the partition ratio K , as the point of departure for calculating the retention volume V_R , or retention time t_R , which is required to describe the chromatographic behaviour of a solute. Conversely, if V_R^0 or t_R is expressed in sufficient reduced terms, such information can be used to identify the solute in unknown mixtures. Experimentally, the retention volume is calculated from the product of retention time t_R and gas flow rate at the column outlet adjusted to column temperature F_c . An additional correction factor, j is applied in order to correct the gas volume for compressibility in the column. Thus,

$$V_R = t_R \cdot F_c \dots\dots\dots(1)$$

Where tR = Retention time

Fc = Volume flow rate of gas corrected retention volume V_R^0 is the volume corrected to the average pressure in the column. According to James and Martin.

$$V_R^0 = j \cdot tR \cdot Fc \dots\dots\dots(2)$$

Where:

$$j = \frac{3 \left[\left(\frac{P_1}{P_2} \right)^2 - 1 \right]}{2 \left[\left(\frac{P_1}{P_2} \right)^3 - 1 \right]}$$

Where P_1 and P_2 are the gas pressure at the inlet and outlet of the column.

If one half the solute is in the column at the point corresponding to V_R^0 and one half the solute has been eluted in the volume, V_R^0 , the one-half portion of solute remaining in the column is distributed between the gas space in the column, V_M and the volume of liquid phase V_L . we also know that the solute in the column is distributed between V_M and V_L in the proportion required by the partition coefficient K. thus if C_L and C_M be the concentrations of the solute in the liquid and gas phase respectively than:

$$K = \frac{C_L \times V_L}{C_M \times V_M} = K \frac{V_L}{V_M} \dots\dots\dots(3)$$

Where $K = C_L / C_M$ is the partition coefficient and unlike K, independent of the amount of liquid phase in the column.

It depends only on the solute, solvent and temperature taking the ratio of the amount of solute that remains in the column to that which had eluted from the column. We get:

$$\frac{0.5}{0.5} = 1 = \frac{(C_M \times V_M) + (C_L \times V_L)}{(C_M \times V_R^0)}$$

$$= \frac{(V_M + KV_L)}{(V_R^0)}$$

$$\text{And } V_R^0 = V_M + KV_L \dots \dots \dots (4)$$

This is the fundamental retention equation required for all chromatographic procedures which have two immiscible phases corresponding to V_M and V_L a partition equilibrium factor corresponding to K .

5.2.4 Specific Retention:

It is very difficult to determine absolute retention either by equation (2) or equation (4), because of the fact that V_R^0 is specific for a single column having specific V_M and V_L values.

It is, therefore, desirable retention in some reduced specific form, called specific retention volume in gas chromatography so that it is easy to calculate the experimental retention volume or retention time by inserting the particular column parameters. The specific retention volume, V_G is defined as.

$$V_G = \frac{V_R^0 - V_M}{w_L} \cdot \frac{273}{T_c} = \frac{j F \rho (t_R - t_A)}{w_L} \cdot \frac{273}{T_c} \dots \dots \dots (5)$$

Where W_L = weight of liquid phases in the column.

T_C = absolute column temperature

t_A = retention time for non-retention solutes such as air.

V_G is equal to the volume of gas at 0^0_C , required to elute one half of a specific solute from a column which contains one gm of a specific liquid phase and contain no gas phase.

5.2.5 Relative Retention:

The calculation of absolute retention volume v^0_R or V_G requires the use of experimental quantities t_R , t_A , T_C , W_L and j . The use of the precise values of these parameters may be evaluated by expressing the retention behaviour in such a manner as to be useful and reproducible for an analyst. Such a simple treatment of retentions is called relative retention.

which is the retention of a specific solute relative to that of a specified standard solute on the same column at the same temperature. Thus:

$$\alpha = \frac{(t_{RI} - t_A)}{(t_{RII} - t_A)} = \frac{(V_{RI}^0 - V_M)}{(V_{RII}^0 - V_M)} = \frac{V_{GI}}{V_{GII}} \neq \frac{V_{RI}^0}{V_{RII}^0}$$

Here it should be noted that is not equal to the ratio of V_R^0 values because these contain the gas volume of the column (V_M). whereas other ratios are corrected for the gas volume. In expressing relative retention information of the solute liquid phase, temperature and standard solute should also be given. If the absolute retention for the standard solute is calculated from equation (2) and (5) the absolute retention of any solute referred to this standard at temperature and liquid phases.

$$V_R^0 (\text{solute}) = \alpha V_R^0 (\text{standard})$$

$$V_G (\text{Solute}) = \alpha V_G (\text{standard})$$

$$(t_R - t_A) \text{ Solute} = \alpha (t_R - t_A) \text{ standard}$$

5.2.6 Retardation Factor:

The retardation factor R_F is the ratio between the rate of movement of sample and carrier gas through the column. For a column having a length of unity

$$\text{The rate of movement for carrier gas} = \frac{1}{t_G}$$

$$\text{Then } R_F = \frac{1/t_R}{1/t_G} = \frac{t_G}{t_R} \dots \dots \dots (6)$$

$$\text{But we know that } V_R^0 = J. t_R. F_C \dots \dots \dots (7)$$

$$\text{And } V_G^0 = J. t_G. F_C \dots \dots \dots (8)$$

Where V_G^0 is the corrected retention volume of the component that does not partition appreciably with the liquid phase.

Combining equation (7) (8) and (9), we get

$$R_{12} = \frac{V_G^0}{V_R^0} \dots \dots \dots (9)$$

The plate theory, employed for liquid partition chromatography can be readily adopted to the gas liquid system.

5.3 TECHNIQUES OF GAS LIQUID CHROMATOGRAPHY

The principle of gas liquid chromatography is almost similar to that of liquid partition chromatography, except that the mobile phase in gas liquid chromatography is a rather than liquid.

The sample is introduced as a gas at the head of the column. As a result, components having finite solubility in the stationary (fixed) liquid phase distribute themselves between this phase and the mobile gas phase, in accordance with the (fixed) liquid phase distribute themselves between this phase and the mobile gas phase, in accordance with the equilibrium law. Elution is then carried out by forcing an inert gas (such as nitrogen or helium) through the column. The rate of movement of the various components along the column depends upon their tendency to dissolve in the stationary liquid phase. Components having a negligible solubility in the stationary phase move rapidly through the column, while those components whose distribution coefficient favours the solvent liquid phase, move with a low rate through the column. Ideally bell-shaped elution curves are obtained. Qualitative identification of the components is based upon the time required for the peak to appear at the end of the column. Quantitative data are obtained from the evaluation of peak areas.

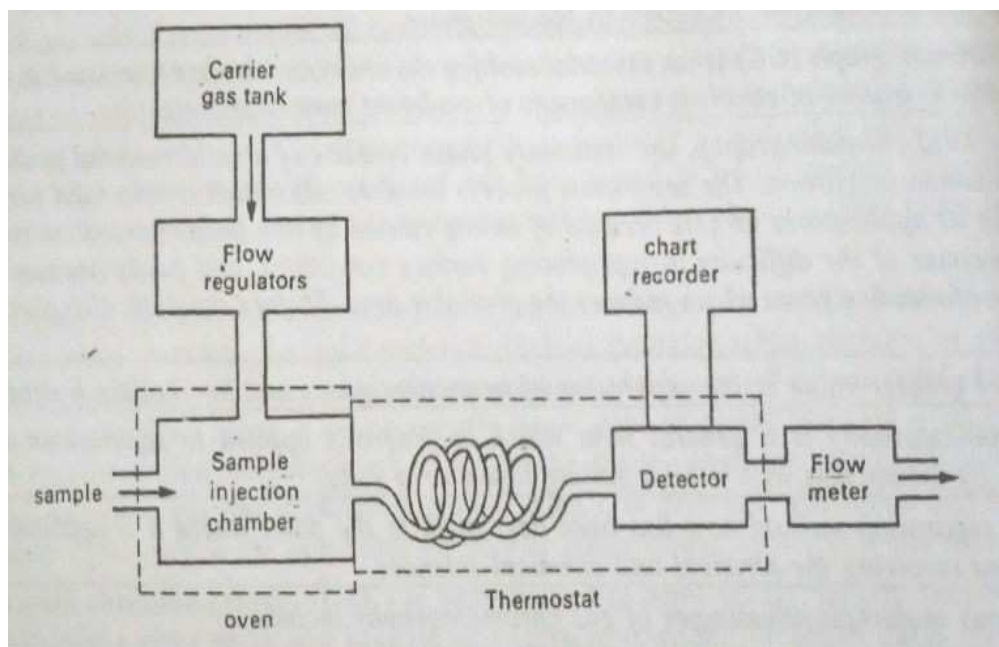


Fig. 2. Block diagram of a gas chromatographic apparatus

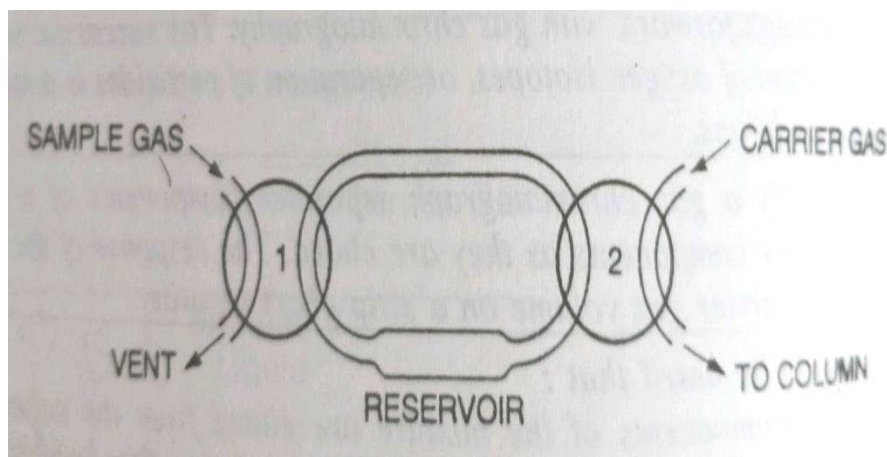


Fig. 3. Gas sampling valve

The mobile gas phase is continuously moving over stationary liquid phase, when a sample of solute vapour is introduced into the gas stream at the head of the column. It is swept into the column and undergoes distribution between the gas and liquid phase in a more or less stepwise manner indicated by the arrows. The greater the total amount of liquid phase in the column, the more solute will dissolve in it and the greater will be the partition ratio. Similarly, an increase in the amount of liquid phase results in increased residence time in that phase, and the partition ratio increases. It should be noted that the gas phase simply serves to move the solute down the

column between excursion into the liquid phase. Consequently, all solutes spend the same time in the gas phase in any particular column.

The partition ratio is simply the ratio of the amount of solute in the stationary phase to the amount of solute in the mobile phase.

$$K = \frac{\text{Amount of solute in the liquid phase}}{\text{Amount of solute in the gas phase}}$$
$$= \frac{\text{Time in the liquid phase}}{\text{Time in the gas phase}}$$

The partition depends on:

- (a). The nature of solute.
- (b). The nature of liquid phase.
- (c). The amount of liquid phase and
- (d). The temperature.

5.4 PARTS OF GAS CHROMATOGRAPY

The equipment consists of following components:

- CARRIER GAS
- COLUMN (COLUMN THERMOSTATING)
- DETECTOR

a. Carrier gas:

The mobile phase (carrier gas) in gas liquid chromatography is usually oxygen or nitrogen, although carbon dioxide and hydrogen from tank sources have also been tried. Hydrogen has the obvious disadvantage of explosion danger.

The most important requirement of a carrier gas is:

It should be inert

1. It should be available at low cost, because large quantities are used
2. It should allow the detector to respond in an adequate manner.

Helium or nitrogen gas cylinder (in which carrier gas is filled in compressed form) is used as a carrier gas reservoir. The cylinder is also attached with a pressure regulator to reduce and control the gas flow through the separation column. A soap bubble meter is also an accurate device for reproducing the flow rate of the carrier gas.

In a soap bubble meter, a soap film is formed in the path of the gas when a rubber bulb containing an aqueous solution of soap or detergent is squeezed. The time required for the soap film to move between two graduations on the burette is then measured and converted to flow rate.

b. Columns:

The column is the most essential component of chromatography. In a column, the different solutes in the vaporised samples are separated from each other by virtue of their different interactions with the column packing.

c. Column thermostating:

It is necessary to control the column temperature to a few tenths of a degree. The temperature has been controlled by using circulating air baths, electrically heated metal blocks and jackets fed with the vapours from a constant boiling liquid.

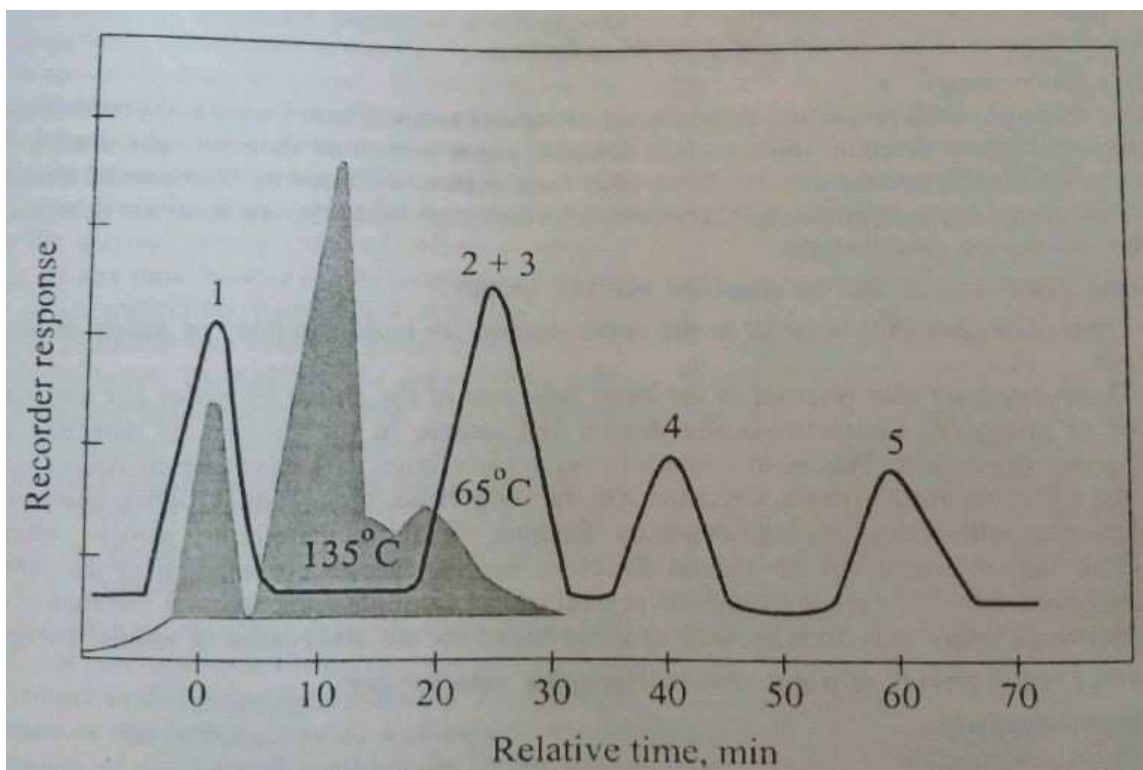


Fig. 4. Effect of temperature on the separation of hexane isomers

Boiling point of the sample and the degree of separation achieved are the two important factors upon which the optimum column temperature depends. Normally a temperature equal to or slightly greater than the average boiling point of sample gives a reasonable elution period (10-30 minutes).

If sample containing components with a broad boiling range is to be fractionated, it is desirable to increase the temperature of the column as the separation proceeds. It may be carried out either continuously or in steps'

In general, optimum resolution is associated with minimal temperature. The cost of lowered temperature. However, is an increase in elution time and hence the time required to complete an analysis.

5.5 DETECTORS

Detection devices for a gas-liquid Chromatography must respond rapidly and reproducibility to the low concentrations of the solute emitted from the column. The solute concentration in a

carrier gas at any instant is only a few paths in a thousand at most. Frequently detector is used to concentrations that are smaller by one or two orders of magnitude. In addition, the interval during which a peak passes the detector is usually a second or less. Hence, detector must be capable of exhibiting its full response during this brief period.

Any physical property which varies widely from one gas to another and which can easily be monitored can be made the basis of the detector.

The most important physical properties are:

- Thermal conductivity.
- The velocity of sound.
- Gas density measurements.
- Magnetic susceptibility.
- Ionisation of α and β rays from a radioactivity source.
- Combustion properties of gases.
- The absorption of ultraviolet and infrared radiations.
- Mass spectrometry.

Based on these physical properties, detectors are of various type. These include, thermal conductivity detector, electron capture detector, cross section detector, argon ionisation detector, gas density balance, microwave excited discharge detectors etc. some other main detectors offered by commercial manufactures are helium ionisation flame thermos-couple, piezoelectric absorption, argon triode detector, micro detector, photo ionization detectors etc.

5.6 GENERAL CHARACTERISTIC AND APPLICATIONS

The general requirements of a gas chromatographic detector are:

(a) High sensitivity (b) Physically suitable (c) Capable of operable up to a maximum column temperature of e.g. 350⁰C (d) Ease of operation (e) No response to undesired compounds (f) An output signal which is a linear function of the concentration of sample in the detector (g) Linear response extending to high concentration.

However, no single detector meets all these requirements. The most desirable properties of the detector include sensitivity, stability over extended periods, linear response characteristics over

range of sample concentrations and uniform response for a wide variety of compounds. Fortunately, thermal conductivity detectors, flame ionisation detectors and electron capture detectors have these properties and hence they are most widely used.

5.7 THERMAL CONDUCTIVITY DETECTORS (TCD)

Since the thermal conductivities of gases are different, a change in composition of the gas causes the thermal conductivity to change. In other words, thermal conductivity of the gas stream. The instrument used for this purpose is also called as katharometer. The sensing element of this device is an electrically heated wire (hot wire) made of a metal whose resistance change practically with temperature. Thus, the hot wire may be a fine platinum or tungsten wire or alternatively a semiconducting thermistor when a current is forced to flow through the heating element, thermal conductivity of the surrounding gas change with a change in the temperature of the wire. By measuring the resistance of the wire or thermistor at a constant current the thermal conductivity of the gas can be calculated and the change in the gas mixture in the effluent stream can be monitored.

In chromatographic applications, a double detector is always employed. The detector usually consists of a block of metal with two cylindrical cavities machined into it, each cavity is provided with a centrally positioned thin wire filament made of platinum or tungsten or thermistor, one element is placed in the gas stream of the sample injection chamber and other immediately beyond the column. The two hot wires are placed in a wheat stone bridge arrangement (as shown in the Fig...4)

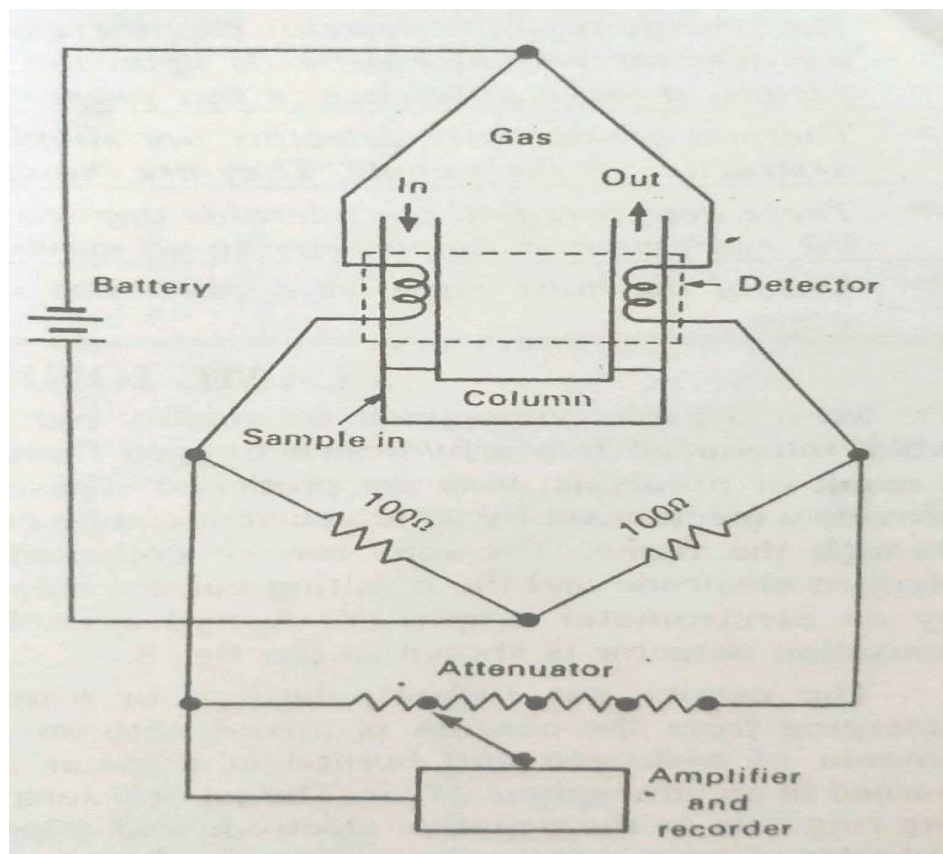


Fig. 5. Schematic diagram of a thermal conductivity detector (TCD)

Such that an increase in sensitivity can be obtained. In this way, the thermal conductivity of the carrier gas is nullified and the effects of variation of column temperature pressure and electrical power are minimised. Since the thermal conductivity of helium is about 6 to 10 times greater than most organic compound, the sensitivity of the detectors is also increased by using helium as a carrier gas, Helium by virtue of its molecular weight and size, has a very high thermal conductivity, whereas gases such as nitrogen and carbon dioxide have almost the same conductivity as a carrier gas, the presence of small amounts of thermal conductivity of the column effects. As a result of this decrease in conductivity, the detector undergoes a marked rise in temperature.

Since the detector response depends upon the difference in thermal conduction of the carrier gas and sample a large difference in this property is essential. An increase in temperature of the detector causes a change in the resistance of wire or thermistor and this resistance gives a measure of the thermal conductivity of the gas.

5.8 GENERAL CHARACTERISTICS OF (TCD) AND APPLICATIONS

- The thermal conductivity detector is based upon changes in thermal conductivity of the gas stream. In other words, this type of detector senses a change in surrounding by the amount of heat which can be dissipated.
- Because the thermal conductivity of different gases are different, a change in composition of the gas causes the thermal conductivity to change.
- The detector is a hot wire through which a current is allowed to flow. As the thermal conductivity of the gas environment change, the temperature of the wire also changes. In other words, the sensing element of this detector is an electrically heated source whose temperature at constant electrical power depends on the thermal conductivity of the surrounding gas. The heated element or hot wire is a fine platinum or tungsten wire whose resistance change drastically with temperature or alternatively semiconducting gas.
- The resistance of the wire or thermistor gives a measure of the thermal conductivity of the gas. Thus, by measuring the resistance of the wire at a constant current, the change in a gas mixture in the effluent stream can be monitored.
- In the gas chromatographic technique, hot wires are placed in a Wheatstone bridge arrangement in such a manner that an increase in sensitivity can be gained.
- The detector must be held at a temperature at least as high as that of the column to prevent condensation.
- Helium and hydrogen conduct heat considerably better than any other gases. Therefore, one of these are the best choice. Helium by virtue of its molecular weight and size, has a very high thermal conductivity, whereas gases such as N₂ have almost the same conduction of the carrier gas and sample, a large difference in this property is essential.
- The sensitivity of the detector can, therefore be increased by using helium as a carrier gas because the detector response depends upon the difference in thermal conduction of the carrier gas and sample a large difference in this property is essential.

- Thermal conductivity detectors are simple, rugged, inexpensive, non-selective, accurate and non-destructive to the sample. They are however, not as sensitive as some of the other detectors.
- These detectors find considerable applications in industry apart from gas chromatography, because the equipment is simple, having no moving parts and the precision is good.
- One of the most extensive applications of these detectors is in the determination of CO₂ in flue gases.

5.9 FLAME IONISATION DETECTORS (FID)

Most organic compounds are readily pyrolyzed when introduced into a hydrogen-oxygen flame. As a result of pyrolysis, ions are produced. These ions provide a mechanism by which current can be carried through the flame. The ions can be collected at a charged electrode and the resulting current measured by an electrometer amplifier. A hydrogen flame ionisation detector is shown in:

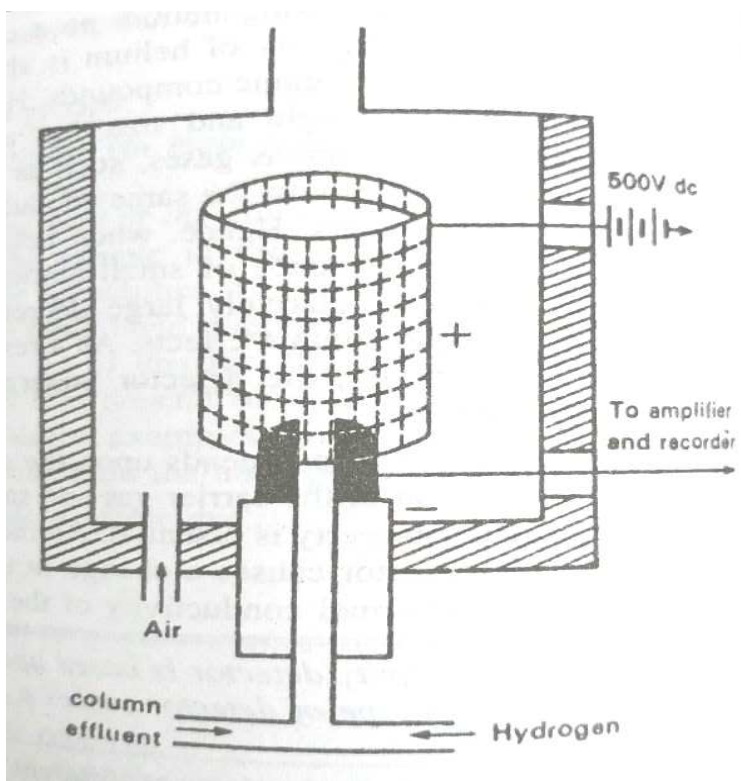


Fig.6. Hydrogen flame ionisation detector

The carrier gas (usually helium or nitrogen) emerging from the column is mixed with an equal amount of hydrogen and burned at a metal jet or a surrounding ring acts as the negative electrode and a loop of cylinder of inert metal surrounding the flame acts as the positive electrodes. As the composition of the gas in the flame change. Thus, the current flow will vary with the change in composition of the gas eluted from gas chromatographic column.

The ionisation of carbon compounds in a flame is roughly proportional to the number of reduced carbon atoms in the flame. Oxidised carbon atoms (e.g. carbonyl, amino or alcoholic) produce fewer ions or none out all.

The flame ionisation detector has now because one of the most popular detectors, superior for quantitative analysis, since there is no flow dependence. Moreover, it is highly sensitive and has a wide range of linear response.

Hartmann has showed that the flame ionisation detector may be almost solely responsive for organo-phosphorus compounds even in sub microgram amounts, by incorporating a block of CsBr or Rb₂SO₄ a part of the structure immediately surrounding the flame. Here, the flame is operated in an atmosphere of hydrogen rather than in an air atmosphere as described in case of ordinary flame ionisation detector.

5.10 GENERAL CHARACTERISTICS OF (FID) AND APPLICATION

- A large number of organic compounds are readily pyrolyzed when introduced into a hydrogen oxygen flame and produce ions in the process. The ions can be collected at a charged electrode and the resulting current measured by an electrometer amplifier. This happens in a flame ionisation detector, which operates on the principle of the change in conductivity of a flame when the composition of a gas mixture is changed.
- The N₂O or He acting as carrier gas emerging from the column is mixed with about an equal amount of hydrogen and burned at a metal jet in an atmosphere of air. In other words, the effluent gases are mixed with the hydrogen and burned in the presence of an oxidant such as air.

- As the composition of the gas in the flame changes, the number of ions and electrons will change. Thus, the current flow will vary with the change in composition of the gas eluted from the gas chromatographic column.
- The flame ionisation detector is destructive of the flame ionisation detector is destructive of the sample.
- It is more complicated and more expensive than the thermal conductivity detector and is more sensitive than thermal conductivity detector.
- The sensitivity to organic solutes varies roughly in proportion to the number of carbon atoms.

5.11 ELECTRON CAPTURE DETECTOR (ECD):

The electron capture or electron affinity detector operates on the basis of differential electron absorption by molecules or functional groups with the molecules. The detector is composed of a radioactive source which emits electrons (β emitter such as strontium-90 or tritium), a cathode which repels the electrons and an anode wire screen which collects electrons. A titanium foil containing adsorbed tritium make the most satisfactory radioactive source which emits electrons, when the effluent from the chromatographic column is allowed to flow through the chamber containing a β -emitter (radioactive source which emits electrons), the carrier gas (argon or nitrogen in this case) tends to be promoted to a meta stable state by β radiation. As a result, electrons are absorbed and a decrease in current is observed at the anode.

The excited argon (or nitrogen) atom cause ionisation of the sample molecules by collisions. The resulting ion current can be measured in essentially the same way as in the hydrogen flame detector.

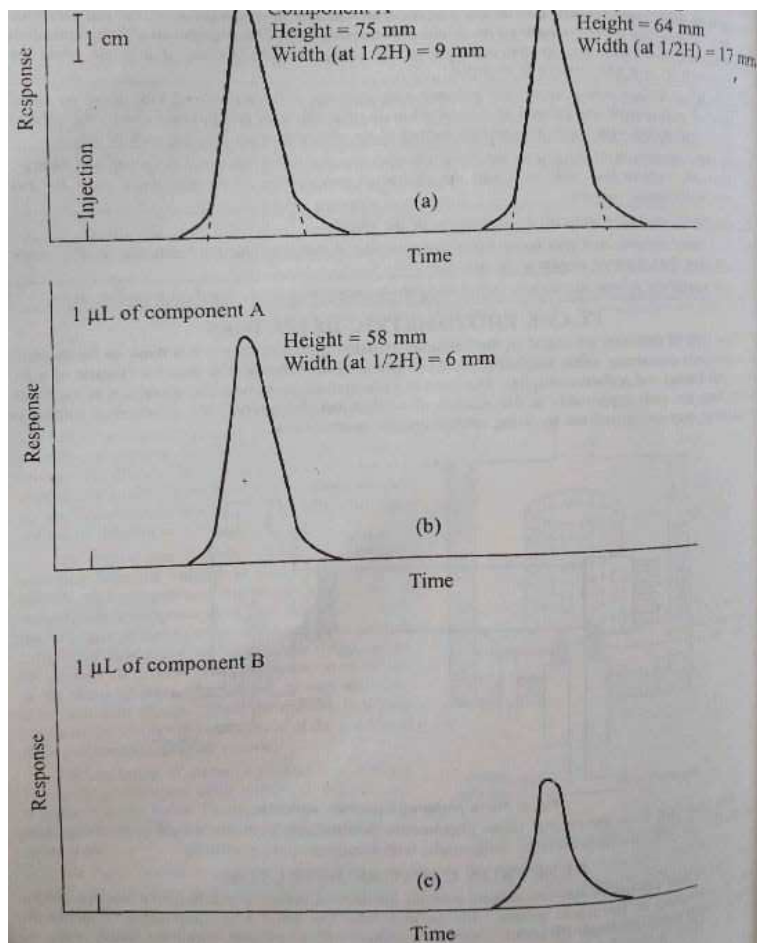


Fig.7. Determination of detector response

5.12 GENERAL CHARACTERISTICS AND APPLICATIONS

- The electron affinity or electron capture detector is composed of a radioactive source which emits electrons, a cathode which repels the electrons, an anode and wire screen which collects electrons.
- As a compound enters the chamber, electrons are absorbed and a decrease in current is observed at the anode.

- If the concentration of the sample increases a resultant decrease in current will also observed. Hence, the detector will respond to change in the amount of sample being eluted from the column.
- These detectors are also known as β -ray detectors because of radioactive source which emits electron (β -emitters are Sr-90, H-3 etc.)
- These detectors are highly sensitive and possess the advantage of not altering the sample significantly.
-

5.13 CONCEPT OF HEPT

HEIGHT EQUIVALENT PLATE THRETICAL:

The ability of column to separate mixture is measured by either number of theoretical plates or the height equivalent to a theoretical plate (HETP). For a given column length, the separation efficiency increases as the number of theoretical plates increases. The efficiency of a column, to resolve components of a mixture is dependent upon the following factors.

5.13.1 Particle size and surface area:

The number of theoretical plate increases with a decrease in particle size or increase in surface area. In general, a 60/80 mesh particle size is used in a 0.25-inch column.

5.13.2 The rate of flow of carrier gas:

The maximum efficiency can be obtained with an optimum flow rate i.e. the rate of flow should neither be too low nor too high. The elute peaks will tend to be broad the gas flow rate is too low and the peaks will not be resolved if the rate of flow of the gases too fast.

5.14 FACTORS AFFECTING SEPARATION OR (HETP)

(a) Type and amount of stationary phase:

This is a very important factor in determining the efficiency of the column. The best separation is obtained when the liquid stationary phase is structurally similar to the compounds being separated.

The amount of stationary phase also affects the efficiency of the column. As the concentration of the column. As the concentration of the liquid phase increases, the number of theoretical plates for the column also increases. Excessive liquid support can also cause tailing of peaks in the chromatogram. In general, light loading is preferred and most columns contain 1-15% liquid.

(b) Column length:

The number of theoretical plates increases with an increase in the length of the column. But there is a practical limit to the length of the column, because long columns may create problems relating to the gas flow. 1-10m long columns are most common.

(c) Column diameter:

The number of theoretical plates increases with a decrease in the diameter of the column. Columns having a diameter of 0.25 inch are most common.

(d) Column temperature:

The maximum operable temperature for the column is determined by the following factors.

- Vapour pressure of the liquid phase
- Vapour pressure of the sample
- Efficiency of separation

The temperature should not be as high as to vaporise the stationary phase, under this condition the uniformity of the column is destroyed. On the other hand, it is a must that the temperature should be high enough to maintain the sample in the vapour state.

5.17 VAN-DEEMTER EQUATION

Van-Deemter for a packed gas chromatography column that the broadening of a peak is the summation of somewhat inter dependent effects from several sources. The Van-Deemter equation expresses these in terms of the plate height H .

$$H = A + \frac{B}{\bar{u}} + C\bar{u}$$

Where A , B and C are constants for a given system and are related to the three major factors affecting H and \bar{u} is the average linear velocity of the carrier gas in cm/s. while the Van-Deemter

equation was developed for gas chromatography, it in principle holds for liquid chromatography as well. Although the diffusion term becomes less important while equilibration term become more critical. For liquid chromatography \bar{u} represent the liquid mobile phase velocity.

The value of \bar{u} is equal to the column length L , divided by the time for an unretained substance to elute t_M .

$$\bar{u} = \frac{L}{t_M}$$

The general flow term for chromatography is the mobile-phase velocity u . But in gas-chromatography, the linear velocity will be different at different positions along the column due to compressibility of gases. So, we use the average linear velocity \bar{u} in liquid chromatography, compressibility is negligible and $\bar{u} = u$. because of this we will generally use the term u .

The significance of the three terms A , B and C in packed-column gas-chromatography is illustrated, which is a plot of H determined as a function of carrier gas velocity. Here A represents eddy diffusion and is due to the variety of tortuous (variable-length) pathways available between the particles in the column and is independent of the gas or mobile-phase velocity. The heterogeneity in axial velocities (eddy diffusion) is related to particle size and geometry of packing by

$$A = 2 \lambda dp$$

Where λ is an empirical column constant (packing-factor), with typical values of 0.8 to 1.0 for a well packed column and dp is the average particle diameter. It is minimized by using small and uniform particles, and tighter packing (which creates back pressure though). But an extremely fine solid supports is difficult to pack uniformly, which affects eddy diffusion.

5.18 SUMMARY

- This chapter provides a systematic and comprehensive description of a series of chromatographic separation techniques.
- This chapter covers different techniques, such as partition (or liquid-liquid and gas-solid) adsorption (or liquid-solid and gas-solid), paper thin-layer, and ion-exchange chromatography, moving boundary zone.

- An exciting chromatographic technique such as high-performance liquid chromatography which is relatively, new is also presented in this chapter. Applications of various techniques for the separation of organic compound. Inorganic mixtures and leaf pigments are documented under the appropriate chromatographic technique.

5.19 TERMINAL QUESTIONS

1. What is the basic technique of GLC?
2. With reference to GLC explain (a) Specific retention (b) Retardation factor (c) Relative retention.
3. Write short notes on:
(a) Column (b) Detector
4. Explain the Van-Deemter equation.
5. Write the explanatory note on Electron Capture Detector (ECD).

TERMINAL ANSWERS:

1. The sample is introduced as a gas at the head of the column. As a result, components having finite solubility in the stationary (fixed) liquid phase distribute themselves between this phase and the mobile gas phase, in accordance with the (fixed) liquid phase distribute themselves between this phase and the mobile gas phase, in accordance with the equilibrium law. Elution is then carried out by forcing an inert gas (such as nitrogen or helium) through the column. The rate of movement of the various components along the column depends upon their tendency to dissolve in the stationary liquid phase. Components having a negligible solubility in the stationary phase move rapidly through the column, while those components whose distribution coefficient favours the solvent liquid phase, move with a low rate through the column. Ideally bell-shaped elution curves are obtained. Qualitative identification of the components is based upon the time required for the peak to appear at the end of the column. Quantitative data are obtained from the evaluation of peak areas.

The mobile gas phase is continuously moving over stationary liquid phase, when a sample of solute vapour is introduced into the gas stream at the head of the column. It is swept into the column and undergoes distribution between the gas and liquid phase in a more or less stepwise manner indicated by the arrows. The greater the total amount of liquid phase in the column, the

more solute will dissolve in it and the greater will be the partition ratio. Similarly, an increase in the amount of liquid phase results in increased residence time in that phase, and the partition ratio increases. It should be noted that the gas phase simply serves to move the solute down the column between excursion into the liquid phase. Consequently, all solutes spend the same time in the gas phase in any particular column.

2. (a) Specific retention:

It is very difficult to determine absolute retention either by equation (2) or equation (4), because of the fact that V_R^0 is specific for a single column having specific V_M and V_L values.

It is, therefore, desirable retention in some reduced specific form, called specific retention volume in gas chromatography so that it is easy to calculate the experimental retention volume or retention time by inserting the particular column parameters. The specific retention volume, V_G is defined as.

$$V_G = \frac{V_R^0 - V_M}{w_L} \cdot \frac{273}{T_C} = \frac{J F \rho (t_R - t_A)}{w_L} \cdot \frac{273}{T_C} \dots\dots\dots (5)$$

Where W_L = weight of liquid phases in the column.

T_C = absolute column temperature

t_A = retention time for non-retention solutes such as air.

V_G is equal to the volume of gas at 0^0_C , required to elute one half of a specific solute from a column which contains one gm of a specific liquid phase and contain no gas phase.

2. (b) Retardation Factor:

The retardation factor R_F is the ratio between the rate of movement of sample and carrier gas through the column. For a column having a length of unity

The rate of movement for carrier gas = $\frac{1}{t_G}$

Then $R_F = \frac{1/t_R}{1/t_G} = \frac{t_G}{t_R} \dots\dots\dots (6)$

But we know that $V_R^0 = J \cdot t_R \cdot F_C \dots\dots\dots (7)$

$$\text{And } V_G^0 = J. t_G. F_C \dots \dots \dots (8)$$

Where V_G^0 is the corrected retention volume of the component that does not partition appreciably with the liquid phase.

Combining equation (7) (8) and (9), we get

$$R_{12} = \frac{V_G^0}{V_R^0} \dots \dots \dots (9)$$

The plate theory, employed for liquid partition chromatography can be readily adopted to the gas liquid system.

2. (c) Relative retention:

The calculation of absolute retention volume v_R^0 or V_G requires the use of experimental quantities t_R , t_A , T_C , W_L and j . The use of the precise values of these parameters may be evaluated by expressing the retention behaviour in such a manner as to be useful and reproducible for an analyst. Such a simple treatment of retentions is called relative retention. which is the retention of a specific solute relative to that of a specified standard solute on the same column at the same temperature. Thus:

$$\alpha = \frac{(t_{RI} - t_A)}{(t_{RII} - t_A)} = \frac{(V_{RI}^0 - V_M)}{(V_{RII}^0 - V_M)} = \frac{V_{GI}}{V_{GII}} \neq \frac{V_{RI}^0}{V_{RII}^0}$$

Here it should be noted that is not equal to the ratio of V_R^0 values because these contain the gas volume of the column (V_M). whereas other ratios are corrected for the gas volume. In expressing relative retention information of the solute liquid phase, temperature and standard solute should also be given. If the absolute retention for the standard solute is calculated from equation (2) and (5) the absolute retention of any solute referred to this standard at temperature and liquid phases.

$$V_R^0 (\text{solute}) = \alpha V_R^0 (\text{standard})$$

$$V_G (\text{Solute}) = \alpha V_G (\text{standard})$$

$$(t_R - t_A) \text{ Solute} = \alpha (t_R - t_A) \text{ standard}$$

3.(a) COLUMN

The column is the most essential component of the chromatography. In column, the different solutes in the vaporised samples are separated from each other by virtue of their different interaction with the column packing.

It is necessary to control the column temperature to a few tenths of a degree. The temperature has been controlled by using circulating air baths, electrically heated metal blocks and jackets fed with the vapours from a constant boiling liquid.

Boiling point of the sample and the degree of separation achieved are the two important factors upon which the optimum column temperature depends. Normally a temperature equal to or slightly greater than the average boiling point of sample gives a reasonable elution period (10-30 minutes).

If sample containing components with a broad boiling range is to be fractionated, it is desirable to increase the temperature of the column as the separation proceeds. It may be carried out either continuously or in steps'

In general, optimum resolution is associated with minimal temperature. The cost of lowered temperature. However, is an increase in elution time and hence the time required to complete an analysis.

3. (b) DETECTOR

Detection devices for a gas-liquid chromatography must respond rapidly and reproducibility to the low concentrations of the solute emitted from the column. The solute concentration in a carrier gas at any instant is only a few parts in a thousand at most. Frequently detector is used to concentration that are smaller by one or two orders of magnitude. In addition, the interval during which a peak passes the detector is usually a second or less. Hence, detector must be capable of exhibiting its full response during this brief period.

Any physical property which varies widely from one gas to another and which can easily be monitored can be made the basis of the detector.

The most important physical properties are:

- Thermal conductivity.

- The velocity of sound.
- Gas density measurements.
- Magnetic susceptibility.
- Ionisation of α and β rays from a radioactivity source.
- Combustion properties of gases.
- The absorption of ultraviolet and infrared radiations.
- Mass spectrometry.

Based on these physical properties, detectors are of various type. These include, thermal conductivity detector, electron capture detector, cross section detector, argon ionisation detector, gas density balance, microwave excited discharge detectors etc. some other main detectors offered by commercial manufactures are helium ionisation flame thermos-couple, piezoelectric absorption, argon triode detector, micro detector, photo ionization detectors etc.

4. Van-Deemter Equation

Van-Deemter for a packed gas chromatography column that the broadening of a peak is the summation of somewhat inter dependent effects from several sources. The Van-Deemter equation expresses these in terms of the plate height H .

$$H = A + \frac{B}{\bar{u}} + C\bar{u}$$

Where A , B and C are constants for a given system and are related to the three major factors affecting H and \bar{u} is the average linear velocity of the carrier gas in cm/s. while the Van-Deemter equation was developed for gas chromatography, it in principle holds for liquid chromatography as well. Although the diffusion term becomes less important while equilibration term become more critical. For liquid chromatography \bar{u} represent the liquid mobile phase velocity.

The value of \bar{u} is equal to the column length L , divided by the time for an unretained substance to elute t_M .

$$\bar{u} = \frac{L}{t_M}$$

The general flow term for chromatography is the mobile-phase velocity u . But in gas-chromatography, the linear velocity will be different at different positions along the column due

to compressibility of gases. So, we use the average linear velocity \bar{u} in liquid chromatography, compressibility is negligible and $\bar{u} = u$. because of this we will generally use the term u .

The significance of the three terms A, B and C in packed-column gas-chromatography is illustrated, which is a plot of H determined as a function of carrier gas velocity. Here A represents eddy diffusion and is due to the variety of tortuous (variable-length) pathways available between the particles in the column and is independent of the gas or mobile-phase velocity. The heterogeneity in axial velocities (eddy diffusion) is related to particle size and geometry of packing by

$$A = 2 \lambda dp$$

Where λ is an empirical column constant (packing-factor), with typical values of 0.8 to 1.0 for a well packed column and dp is the average particle diameter. It is minimized by using small and uniform particles, and tighter packing (which creates back pressure though). But an extremely fine solid supports is difficult to pack uniformly, which affects eddy diffusion.

5.ELECTRON CAPTURE DETECTOR (ECD):

The electron capture or electron affinity detector operates on the basis of differential electron absorption by molecules or functional groups with the molecules. The detector is composed of a radioactive source which emits electrons (β emitter such as strontium-90 or tritium), a cathode which repels the electrons and an anode wire screen which collects electrons. A titanium foil containing adsorbed tritium make the most satisfactory radioactive source which emits electrons, when the effluent from the chromatographic column is allowed to flow through the chamber containing a β -emitter (radioactive source which emits electrons), the carrier gas (argon or nitrogen in this case) tends to be promoted to a meta stable state by β radiation. A result, electrons are absorbed and a decrease in current is observed at the anode.

The excited argon (or nitrogen) atom cause ionisation of the sample molecules by collisions. The resulting ion current can be measured in essentially the same way as in the hydrogen flame detector.

5.20 REFERENCES

1. B.K. Sharma, Instrumental methods of chemical analysis; Krishna Prakashan India 1972.
2. I. L. Finar, organic chemistry, ELBS publication 1962.
3. U. N. Dash, Analytical Chemistry; sultan chand& sons publisher 2005.
4. Arun Bahl& B. S. Bahl Organic Chemistry S. Chand & Company-2006.
5. Harbans Singh & M. M. N. Tandon; Organic Chemistry; ShivalalAgrawala& Company Publishers 1985.
6. Shashi Chawla; Engineering Chemistry; Dhanpat Rai Co. Publishers – 2001.

UNIT- 6 RADIOANALYTICAL METHODS

CONTENTS:

- 6.1 Objectives
- 6.2 Introduction
- 6.3 Detection and Measurement of Radioactivity
- 6.4 Isotopic dilution analysis
- 6.5 Advantages of isotopic dilution method
- 6.6 Radiometric Titrations
- 6.7 Radiochromatography
- 6.8 Summary
- 6.9 Terminal Questions
- 6.10 Reference

6.1 OBJECTIVES

After going through this you will be able to:

- Describe radiometric analysis
- Describe various techniques of the measurement of radioactivity
- Define isotopic dilution analysis
- Define Radiometric titrations and radio chromatography

6.2 INTRODUCTION

The discovery and production of both natural and artificial radioactive isotopes and their use as tracer atoms have led to the development of analytical methods, called radiochemical methods. In general, nuclear analytical methods based on the measurement of radioactivity is studied in two ways broadly

- (1) Radioactivity can be measured directly in the system
- (2) Sample can be made radioactive by the bombardment of the sample in a nuclear reactor or accelerator facility or by adding radioactivity in the system

Radiochemical methods are based upon natural and artificial radioactivity and are sensitive and specific. These methods can be classified into three categories.

- (1) In activation analysis, the resulting radioactivity is measured when one or more elements of a sample are made active by irradiating them with thermal neutrons from a thermal reactor.
 - (2) In isotopic dilution, a pure but radioactive form of the substance to be determined is mixed with a known amount of the sample. After equilibrium has been attained, the activity of an isolated fraction of interest is measured; and
 - (3) In radiometric analysis, a radioactive reagent is employed for separating the component of interest completely from the bulk of the sample and then its activity is measured.
- Radioactive determinations of more than thirty elements have been made.

Radiochemical methods are more specific and sensitive and can more easily be applied than any other analytical methods.

The importance of radioanalytical chemistry is dispersed in many fields as Chemistry, Physics, Medicine, Human and Plant physiology, Forensic science, Geology, Earth atmosphere etc. Radioanalytical Chemistry is applied in preparation and characterizing new elements, determination of course of physical processes and chemical reactions, determination of age of materials i.e. radiocarbon dating, and creation of various radioactive cations, anions, or compounds which are used as tracer in human and plant tissues and organs, leak detection investigations in industry, ground water studies, oil well tracing etc.

6.3 DETECTION AND MEASUREMENT OF RADIOACTIVITY

The measurement of radioactivity of any radioactive substance simply means the measurement of the rates of emission of radioactive rays (α , β and γ -particles). It is well known that the gases when exposed to radioactive radiations become good conductor since they are

ionized. It is on this basis that the detection of radioactive particles is made and rates of their emission measured.

The detection and measurement of the radiations may be done in several ways-(1) the observation of visible light produced by scintillation, (2) the photographic detection, (3) detection based on ionization of gas (electrical measurement), and (4) displacement of electrons in semiconductor crystals.

The direct effects of nuclear radiations on photographic plates or film is especially useful for mapping the distribution of a radioactive substance in the surface layers of a solid material, such as minerals or biological specimens. This process is referred to as radio autography or autoradiography. The photographic detection is also very useful in film badges which the nuclear workers wear to detect the excessive exposure to radiation.

(1) Scintillation Counters

The counters based on scintillation have long been used for quantitative measurements but the technique has been modernized to become highly accurate and convenient only recently. In olden days, the technique involved the manual counting of the flashes that resulted when individual α -particles struck the zinc sulphide screen. The process of counting was tiring and cumbersome and was, therefore given up and gas ionizing devices were adopted. The use of scintillation counters has again been in use with the discovery of photomultiplier tubes, which convert the radiant energy to an electric signal that is measured by an amplifier. Anthracene, Stilbene, paraterphenyl and Sodium Iodide are the useful scintillation crystals. Sodium iodide is made more sensitive by introducing one percent thallium. The crystal is often shaped as a cylinder whose one of the plane surfaces faces the cathode of a photomultiplier tube. The incoming radiation during its passage through crystal loses its energy to the scintillator which is subsequently released as photons of radiation energy. The decay time is approximately 0.25 μ sec. the flashes of light produced in the scintillator crystal are reflected to the photocathode of the photomultiplier tube where they are converted into electrical pulses. These are amplified and counted. The number of photon produced in each flash is approximately proportional to the energy of impinging radiation. A pulse –height analyzer is incorporated which monitors the output of a scintillation counter and forms the basis of a beta or a gamma ray spectrometer.

Scintillator crystals other than sodium iodide have still shorter decay times (0.01 to 0.1 μsec). Organic liquid scintillators count weak beta emitters as ^{14}C , ^{35}S and especially ^3H . Their use is advantageous because they exhibit less self-absorption of radiation than do the solids. A number of organic compounds, when dissolved in suitable solvents, act as liquid scintillators. These are anthracene, p-terphenyl, 2,5-diphenyloxazole, α -naphthyl phenyloxazole (NPO) and phenylbiphenyloxadiazone (PBD). Of these the PPO (2,5-diphenyloxazole) is the most effective, but the radiation that it emits is ultraviolet. It is, therefore usually mixed with a secondary scintillator such as 4-bis, 2- (5-phenyl oxazolyl) benzene (POPOP), which translates the ultraviolet scintillations into the visible range. A recommended solution contains 5 gram per litre POP and 0.3 gram per litre dimethyl POPOP in toluene.

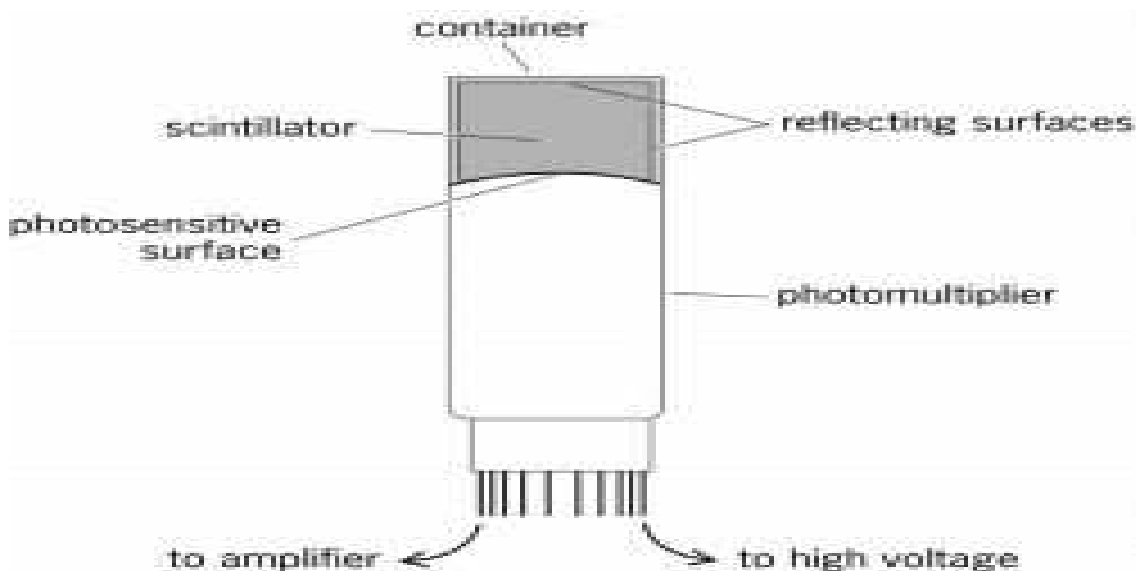


Figure1. Well scintillation counter

A considerable high efficiency is achieved with a well counter (figure1). This consists of a block of crystalline or plastic scintillator with a hole drilled into it to fit a test tube in it. A fixed volume of the solution has to be used in well counting for reproducible results.

While working with low-level activity, a distinction has to be made between pulses from scintillations with those spurious pulses arising in the photomultiplier, caused either by stray or cosmic radiation, or by short-effect noise. For this, two identical photomultipliers looking at the same scintillator connected in a coincidence circuit (figure 2) are used.

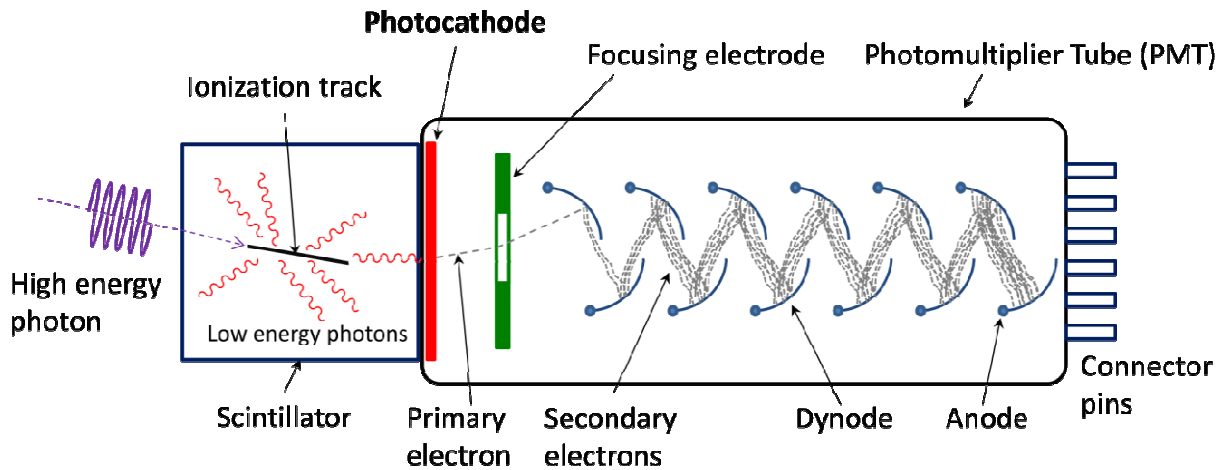


Figure 2. Liquid scintillation counting assembly

This consists of an electronic unit called an AND-gate, which transmits a signal to the counting equipment only when it receives simultaneous pulses from both multipliers.

(2) Photographic detection

The history of photographic detection of radiation lies in the exposure of ordinary photographic emulsions. Photographic detection is still very useful for the monitoring of total dosage in the detection of flaws in metals and in the study of cosmic radiation in spite of the fact that it has been supplanted by other methods of radio assay work.

(3) Gas ionization detectors

When ionizing radiation is passed through a gas it creates positive ions and electrons. If an electric field is applied, the positive ions start to migrate to the electrode of the opposite charge, where they are collected. The ionizing power varies from particle to particle; alpha particle being most effective and X- and γ -rays being the least. The living cells are damaged from radiation on account of this ionization. By the detection of this electric charge using a suitable electric circuit, an indication of the presence of ionizing radiation is given. Gas ionization detectors can be categorized under three large groups, namely ion chambers, Geiger-Müller counters and proportional counters.

The ionization of gas gives rise to electrical conductivity in it and the measurement of radioactivity, in fact, is the measurement of this electrical conductivity. Several detectors –

ionization chambers, Geiger-Muller counter and proportional counters have been used. All of these are gas ionization chambers. A typical gas-ionization chamber has been shown in figure 3.

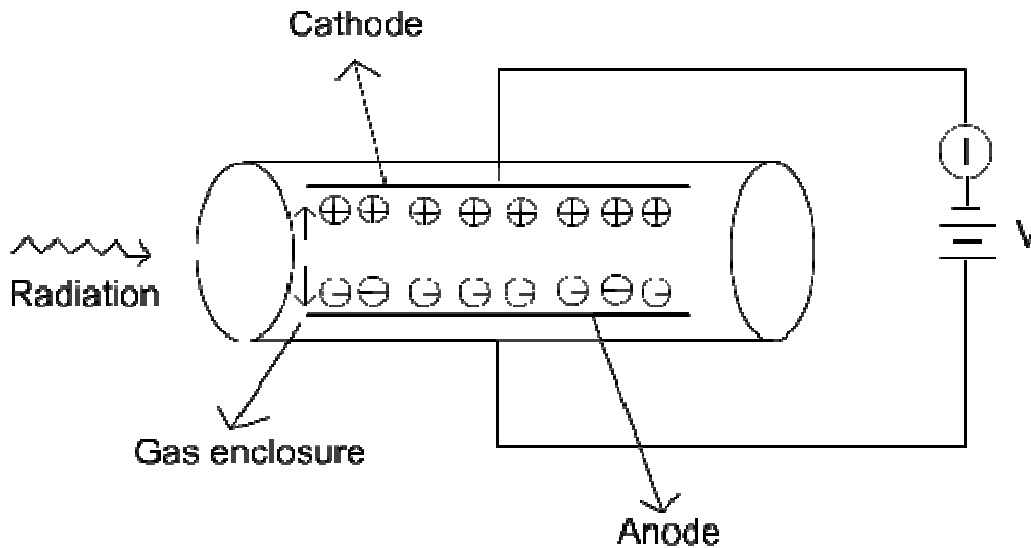


Figure 3. A typical gas-ionization chamber

It consists of a window made of mica, aluminium, lead, or other transparent material through which the radiation enters the chamber, which usually contains argon.

1. Ionization chamber

We have seen that radiation from radioactive substances produce ionization of gas. The degree of ionization produced in this manner is thus a measure of rate of emission of alpha, beta and gamma rays by the radioactive substance. The radioactivity of a substance can, therefore be detected and even measured by the ionization of gases caused by alpha and beta particles and also indirectly by gamma rays.

In an ionization chamber, the radioactive substance produces ions by collisions with the gas molecules. These ions are collected on electrodes at a relatively low potential in order to produce a current that is amplified by an electrometer to be read on a galvanometer. The current produced is a measure of the average rate of ion production and, therefore, the radioactivity the gas. Gamma and X-ray radiations are usually measured by ionization chamber. The ionization chamber has the advantage of requiring

only a low voltage (100-200V), but the current it passes is very small (perhaps, 10^{-8} A) and therefore, requires either a high gain amplifier or a highly sensitive electrometer. In practice, it is most useful for pulse work with the highly ionizing alpha particle.

2. Geiger Muller Counter

In Geiger Muller counter, the radiations emitted by a solid sample cause ionization of the gas with which the counter is filled.

It essentially consists of a cylinder of copper closed with a thin mica window at one end. The cylinder is filled with a gas, which is ionizable, e.g. argon and alcohol vapor, at a pressure of about 10 mm of Hg. In the middle of copper cylinder, but insulated from it, is a wire of tungsten which is usually kept at a potential of about 1500 volts. The radioactive substance is placed in a small tube and the tube is placed very close to the mica window. Radiation i.e. α or (or γ ray) emitted by the radioactive substance when enters the tube causes ionization of the gas present in the cylinder. A momentary flow of current takes place which can easily be amplified. This current causes a flash of light in a neon tube and a unit movement over a mechanical register. These types of counters measure each β or γ ray entering the counter.

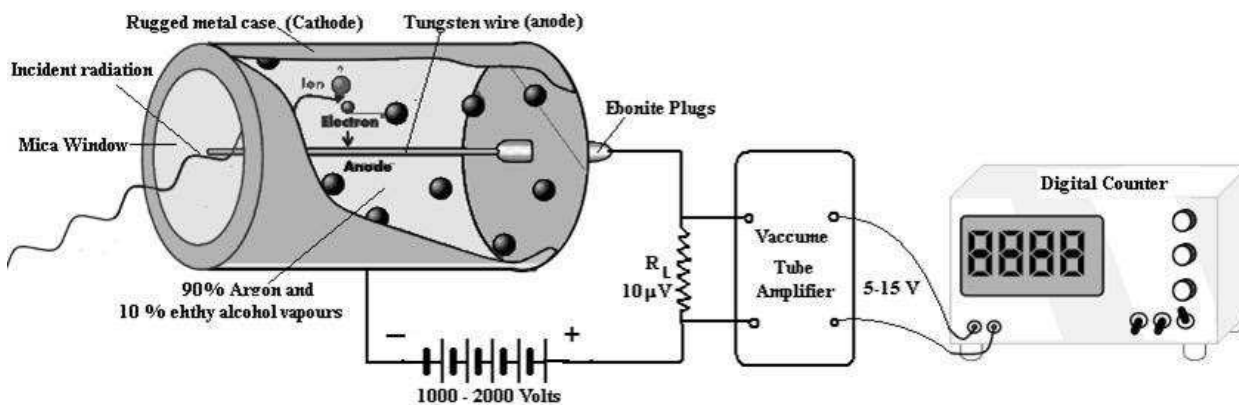


Figure 4. Geiger Muller counter

3. Proportional counter

The **proportional counter** is a type of gaseous ionization detector device used to measure particles of ionizing radiation. The size of output pulse of the counter is proportional to the number of electron pairs produced by the ionizing event and thus to the energy of radiation. So we can differentiate between different forms of ionizing radiation. In proportional counter too,

the radiations emitted by the solid sample cause ionization of the gas inside the counter. Gas of the chamber is an inert gas which is ionised by incident radiation, and a quench gas to ensure each pulse discharge terminates; a common mixture is 90% argon, 10% methane, known as P-10

A very high voltage is applied across the electrodes so that the ions produced are multiplied several thousand times and when the potential is discharged, the discharge is registered over a mechanical register. The magnitude of this discharge is proportional to the energy of the entering radiation..

(4) Semiconductor Counters

Silicon or germanium crystals have also been used as detectors. These semiconductor detectors are employed in a manner similar to the gas ionization chamber. Radiation which is absorbed by semiconductors produces pairs of electrons and holes which are accelerated towards the respective electrodes. The resulting current is proportional to the energy of the primary radiation. A sensitive and noise-free electrometer amplifier is warranted for the currents are of the magnitude of few microamperes.

Generally speaking, these semiconductors have been used for counting alpha and beta particles; γ ray spectrometers make use of lithium drifted germanium crystal, which has a greater resolving power than one using sodium iodide crystal. The disadvantage with the use of germanium-lithium detector is that it has to be operated at liquid temperature (77^0K) which is not convenient.

(5) Background radiation

A significant amount of radiation is always present in the atmosphere because of which the detector shows positive response even if the sample is absent. Natural radioactivity of the surrounding and cosmic rays are both responsible in part for this unwanted radiation. If the counter is shielded by 2-3 inch lead, the radiation reduces to the extent of 15-20 cpm. If there is a sudden increase in the count or else background count is higher, an accidental contamination of the immediate surrounding is the reason for it. It may also be due to the sudden failure of the counter. A correction for background counter is essential. Thus,

$$R_C = R_S - R_B$$

Where R_C is corrected counting rate and R_S and R_B are the rates for the sample and background respectively.

(6) Auxillary instrumentation

It is apparent that scintillation, proportional and Geiger counters all deliver information in form of pulses (electrical) and, therefore, an arrangement for counting these pulses has to be made. Low counting, 500-1000 cpm, is done with electromechanical device. For faster counting a glow-transfer tube has to be used. This is gas filled glow discharge tube in which 10 successive pulses transfer the glow successively to 10 positions in the tube. The fastest counting is, however, done by electronic scalar which is an electronic amplifier circuit so arranged that its output terminals transmit only every other input signal. Several such scalars are arranged in series so the pulses finally counted represent only 1/2, 1/4, 1/8 etc, of the number produced in the detector. Decade counters are the latest additions whose response to pulse is in factors of ten.

(7) Gamma Ray Spectrometers

This spectrometer makes use of either sodium iodide or germanium-lithium semiconductor. The electrical output pulses are proportional in magnitude to the energy of primary radiation. These pulses are amplified by linear amplifier and appear as voltage signals which are fed to a linear pulse height analyzer, the first stage of which consists of two discriminator circuits, each one is capable of rejecting any signal below a certain voltage. Thus, only one signal reaches the counter which lies in the voltage range V , which is referred to as window of a pulse height analyzer. A range of single channel analyzer has a voltage range of 100 volts or more and a window of 5-10 V. The window can be either manually or automatically adjusted to scan the entire voltage range. A multichannel analyzer has 2-100 separate channels, each acting as a single channel set for a different voltage span or width. This arrangement is useful for simultaneous counting and recording of the entire spectrum.

Examples of radionuclides used to study various reactions are:

- b. ^{54}Mn to study uptake of manganese by plants

c. ^{131}I For the determination of solubility product of PbI_2 using radiotracer method.

d. ^{128}I For the treatment of thyroid cancer using a radiopharmaceutical.

e. ^{126}I For assaying of thyroid hormones by RIA.

. Radioisotope of manganese ^{54}Mn , $t_{1/2} = 312$ d is used as radiotracer for the determination of manganese in a steel sample in a chemical laboratory

6.4 ISOTOPIC DILUTION ANALYSIS

This method was originally used by G. Hevesy and R. Hobie (1932) in analyzing rocks for lead, and later by A. J. Freedman and D.N. Hume (1950) in determining cerium.

The isotopic dilution analysis is very useful in analytical chemistry. This is a technique which is suitable where a compound can be isolated in a pure state but only with a poor yield. The technique consists in adding a known amount of the pure compound containing an active isotope to the unknown and thoroughly mixed with it. Now a sample of pure compound is isolated from the mixture and its activity is measured with the help of an electroscope. The amount of substance in the original material is then calculated as follows.

Let us consider a solution containing **M grams** of a compound which is to be determined. A portion of the same compound which is tagged with a radioactive isotope is now added to the solution. Suppose the weight of added portion be **m grams** and its activity is **A cpm** and a specific activity be $S_0 = A/M$. Let after complete and thorough mixing **w grams** of this compound is isolated in a pure state and its activity and specific activity are **B cpm** and $S = (B/w)$ respectively. Now assuming loss by decay to be negligible, total amount of activity after mixing must be the same as before.

$$S_0 m = (M + m) S \quad \dots\dots(1)$$

Thus from which it follows that,

$$M = w \cdot A/B - m \quad \dots\dots(2)$$

Now if the added material is highly active, then the amount added m can be very small compared to M . hence equation (2) becomes

$$M = wA/B \quad \dots\dots(3)$$

But $M = A.w/B - m = A 1/S - m \quad \dots\dots(4)$

Since $w/B = 1/S$

Equation (4) is evidently an equation of a straight line having a slope of A and intercept m . the calculation may be made clear by the following example.

(1) Suppose we have to determine the amount of glycine in a mixture with other amino acids. Isotopic dilution analysis is the most appropriate technique in this case, because glycine can be isolated chemically but only with a poor low yield. A sample of glycine which contains an atom of C^{14} in perhaps one in every million of its molecules is first synthesized commercially. A 0.500 gram portion of this active preparation, whose specific activity corrected for background, is 25000 cpm (Counts per minutes) per gram is mixed with unknown substance. Suppose 0.200 gram of pure glycine having an activity of 1250 counts per 10 min is obtained from the mixture. The specific activity corrected for background of the pure glycine is 100 counts per 5 min. thus the data may be summarized as follows.

$$M = 0.0500g$$

$$A = m S_0 = 12500 \text{ cpm}$$

$$W = 0.200g$$

$$S_0 = 25000 \text{ cpm per g}$$

Thus substituting in equation (3)

$$\begin{aligned} M &= w.A/B - m = 0.200 \times 12500 / 105 - 0.500 \\ &= 23.3 \text{ g.} \end{aligned}$$

23.3 g is the required weight of glycine in the sample.

Isotopic dilution analysis is becoming increasingly important in analytical biochemistry and other fields also. Some other examples of the applications of isotopic dilution analysis are given below.

- (2) It is of great importance in the determination of blood volume in an animal or human being. The method involves the injecting of a tracer in a donor subject and allowing sufficient time, i.e., about two weeks, for incorporation of the tracer into the circulating red cells. The tracer which is the most effective in this case is Fe^{59} . A sample is taken out from the donor and its activity per milliliter of blood is determined and finally an aliquot portion of the sample is introduced into the recipient subject. After certain or rather sufficient time the sample from the recipient is examined for radioactivity. From the dilution of activity it is possible to calculate the volume of blood in the recipient.
- (3) Another example is the determination of amino acids in the products of protein hydrolysis. The method is same, as above the method requires the complete isolation of each amino acid in pure form. Suppose, however, a known amount of amino acid labeled with C^{14} or deuterium is added to the hydrolysate. After complete and thorough mixing a small sample of the given amino acid is isolated and its activity is measured. The total concentration of the acid in the hydrolysate can now be determined easily by the decrease in activity.
- (4) It is well known fact that the solubility of water in benzene or other hydrocarbons is so small that it can not be measured with the help of physical or chemical methods. Hence again isotopic dilution analysis is the appropriate technique. If a known and definite amount of radioactive tritium, H^3 in the form of tritium oxide is added to given amount of water, then the resulting solution has an activity that can be measured even in minute amounts which is dissolved in organic solvents like benzene and other hydrocarbons.
- (5) D. Salyer and T.R.Sweet (1957) have used isotopic dilution analysis in the electro gravimetric determination of cobalt in steel and the cobalt other alloys. The reason of using this technique was that cobalt, deposited anodically as Co_2O_3 is apt to form poorly adherent layer; this precludes a conventional gravimetric determination, but loss of some particles of oxides during washing and drying is not objectionable in isotopic dilution. A standard curve, as shown in figure 1, was prepared by adding equal aliquots of Co^{60} to samples containing various amounts of pure cobalt and then electrodepositing Co_2O_3 in a

standardized manner. The unknown was fortified by an aliquot of Co^{60} immediately upon the dissolution of sample. The interfering elements were removed by chemical treatment, the cobalt was then deposited, the deposit is weighed and counted, and the cobalt in the original sample ascertained by reference to the standard curve.

- (6) Isotopic dilution method is employed in elemental analysis and for the determination of individual compounds in mixtures. In either case it involves the addition of measured amount of isotopes to mixtures being analyzed and the subsequent determination of the extent of dilution of this isotope in the mixture, or with compounds, the extent of dilution of the enriched compound.

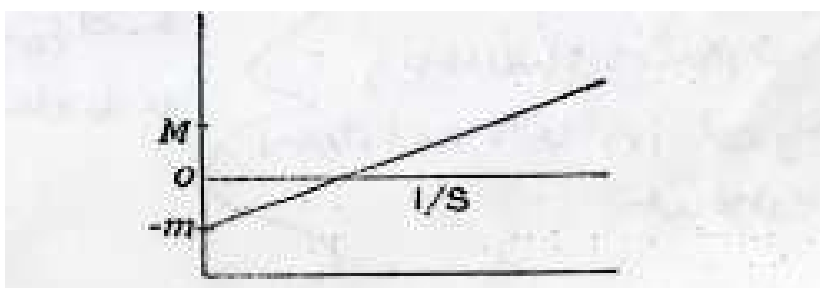


Figure 5

- (7) Biologists have used this method for a number of years to determine the elements and compounds of hydrogen, carbon and nitrogen. More recently the atomic energy commission has made available separately the isotope of many elements.

It is very interesting to note that it is possible to vary isotopic dilution method. Reverse isotopic dilution method can be used to determine the amount of radioactive constituent by adding a known amount of the inactive compound and then proceeding in the usual manner. Similarly, this method may be varied into isotope derivative method which employs a radioactive reagent of known specific activity to form an active derivative with the inactive substance which is to be determined.

6.5 ADVANTAGES OF ISOTOPIC DILUTION METHOD

There are several advantages of this method. Some of them are given below:

- (1) The method is absolute one and no corrections are therefore needed to obtain the final result.

- (2) The method is very sensitive.
- (3) Accuracy is of the order of one percent or higher.
- (4) In this method it is not necessary to separate the constituents. If during processing part of one isotope is lost, at the same time an equal proportion of the other isotope would be lost, as they can not be chemically separated.
- (5) Simple mass spectrometers can be used, as the method does not depend upon a particular source.

Limitations of IDA method

- (i) IDA is a destructive method. Since the sample is not preserved, it has a limited applicability.
- (ii) This method can be used for the determination of a single element at a time whereas for several samples multielemental methods are required.
- (iii) IDA has limited applicability for the elements for which suitable radiotracers are available.

6.6 RADIOMETRIC TITRATIONS

Radiometric titration is a volumetric titration method with a special opportunity for the indication of the end-point. Here the titrant gives an insoluble precipitate or a compound can be extracted easily, and when one of the reaction partners can be labeled.

It has the unique advantage of requiring no indicator. Radiotracer may be added to the titrant or the unknown analyte sample. At the end of reaction, there is sudden change in activity level wherefrom end point may be determined. Here titrant is radioactive. When we start titration radioactive ion form a compound and thus activity of titrant is reduced. We can determine concentration of radioactive ion at end point. It is of special importance where suitable indicators are not available in conventional titrimetry (neutralization, redox, precipitation and complexometric titrations). However, it is essential to have a separable phase and that puts great limitation on its use though in principle it can be used for all types of titrations. Therefore, most successful radiometric titration procedures have been developed for the precipitation reactions. A

typical example of radiometric titration is for the determination of halides using ^{110}Ag ($t_{1/2} = 252 \text{ d}$) as tracer where corresponding silver halide is precipitated.

6.7 RADIOCHROMATOGRAPHY

Radiochromatography is basically used on paper or thin layer chromatography though it could also be used in ion-exchange column or gas chromatography as well as in electrophoresis. It uses a radiotracer as cation or anion or radiolabelled organic compound and the separated cation, anion or a compound is identified on the basis of characteristic radiation emitted by the radionuclide. Thus, any chromatographic separation can be carried out using radiotracers. However, no developer/chromogenic reagent or detector is used to study the movement of initial spot except nuclear counting equipment or autoradiography. In this way distribution of activity is measured which increases sensitivity of the measurement by several fold. Radiochromatography has played an important role in the discovery of some trans- plutonium elements by Glen T. Seaborg. The identity of each element of $5f$ series was established beyond any doubt by the sequence of their appearance analogous to the sequence of the corresponding $4f$ elements. There are many instances of the use of radiochromatography for analytical purposes. Example of complete separation of the alkali metal ions Na^+ , K^+ , Rb^+ , and Cs^+ by using respective tracer solutions and scintillation gamma ray spectrometer.

6.8 SUMMARY

Radiochemical methods tend to be labor intensive and generate liquid waste due to the chemical separations that are involved. Furthermore, detection limits tend to be high for long-lived isotopes using radiochemical methods. Radiochemical methods are characterized by good accuracy and their ability to be adapted to a wide number of applications. Another advantage to this method is that they minimize or even eliminate the need for separations that are required in other analytical methods. Radiation from radioactive sources can be detected and measured in essentially the same way as X-radiation. Gas-filled transducers, scintillation counters and semiconductor detectors are all sensitive to alpha and beta particles and gamma rays because absorption of these particles induces ionization or photoelectrons, which can in turn, produce

thousands of ion pairs. Isotope dilution method is the most important class of radiochemical quantitative method. In this method, a weighed quantity of radioactively tagged analyte having a known activity is added to a measured amount of sample. The mixture is then mixed to homogeneity and then a fraction of the compound of interest is isolated and purified. The analysis is based upon the activity of the isolated fraction. Isotope dilution methods predate activation procedure and are still being extensively applied to problems in all branches of chemistry. Stable and radioactive isotopes are employed in the isotope dilution technique. The isotope dilution method is more convenient because of the ease with which the concentration of the isotope can be determined. Radiometric titrations and radio chromatography are important radio analytical techniques.

6.9 TERMINAL QUESTIONS

1. Explain the different techniques of the measurement of radioactivity.
2. Discuss isotopic dilution analysis.
3. Give the advantages of the isotopic dilution analysis.
4. Discuss Geiger Muller counter in brief.
5. Write a short note on gas ionization detectors.
6. Discuss semiconductor counter.
7. Write note on radiometric titrations and radio chromatography.
8. Discuss applications of isotope dilution analysis.

6.10 REFERENCE

1. Vogel's, Textbook of Quantitative Chemical Analysis, J. Mendham, R.C.Denny, J.D.Barnes and M.J.K. Thomas, The School of Chemical and Life Sciences, University of Greenwich, London. Pearson Education.
2. Instrumental Approach to Chemical Analysis, A.K. Srivastav and P.C. Jain, S. Chand Publication, 2000.
3. <http://egyankosh.ac.in/> Unit 13 Radioanalytical Methods.
4. <http://media.iupac.org/> Radioanalytical methods
Handbook of Analytical Techniques by Helmut Gunzler and Alex Williams WILEY-VCH Publishe

UNIT 7: DATA ANALYSIS

CONTANTS:

- 7.1 Learning Objectives
- 7.2 Introduction
- 7.3 Data Analysis
 - 7.3.1 The process of data analysis
 - 7.3.2 Techniques for analyzing data
 - 7.3.3 Initial data analysis
 - 7.3.4 Main data analysis
- 7.4 Error Analysis
 - 7.4.1 Types of errors
 - 7.4.2 Propagation of errors
- 7.5 Accuracy and Precision
- 7.6 Least Square Analysis
- 7.7 Average vs. deviation
- 7.8 Standard deviation
- 7.9 Let Us Sum Up
- 7.10 Answer to Check Your Progress
- 7.11 Further Readings
- 7.12 Model Questions

7.1 LEARNING OBJECTIVE

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

- Understand basic concepts of data analysis
- Define error analysis, data accuracy, and precision

- Students understand basic concepts of statistics and probability, comprehend methods needed to analyze and critically evaluate statistical arguments, and recognize the importance of statistical ideas.

7.2 INTRODUCTION

Data analysis, also known as **analysis of data** or **data analytics**, is a process of inspecting, cleansing, transforming, and modeling data with the goal of discovering useful information, suggesting conclusions, and supporting decision-making.

7.3 DATA ANALYSIS

Data analysis has multiple facets and approaches, encompassing diverse techniques under a variety of names, in a different business, science, and social science domains.

Data mining is a particular data analysis technique that focuses on modeling and knowledge discovery for predictive rather than purely descriptive purposes, while business intelligence covers data analysis that relies heavily on aggregation, focusing on business information. In statistical applications data analysis can be divided into descriptive statistics, exploratory data analysis (EDA), and confirmatory data analysis (CDA). EDA focuses on discovering new features in the data and CDA on confirming or falsifying existing hypotheses. Predictive analytics focuses on the application of statistical models for predictive forecasting or classification, while text analytics applies statistical, linguistic, and structural techniques to extract and classify information from textual sources, a species of unstructured data. All are varieties of the data analysis.

Data integration is a precursor to data analysis, and data analysis is closely linked to data visualization and data dissemination. The term *data analysis* is sometimes used as a synonym for data modeling.

7.3.1 THE PROCESS OF DATA ANALYSIS

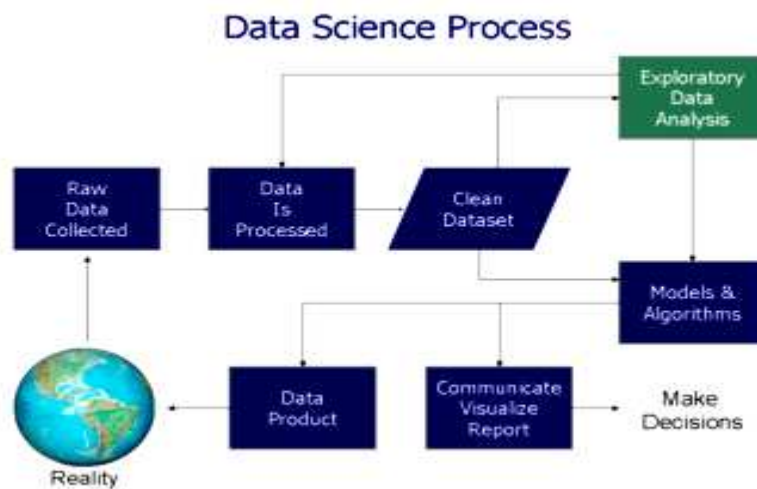


Figure: 1.1 Process of data analysis

Data science process flowchart from "Doing Data Science", Cathy O'Neil and Rachel Schutt, 2013

Analysis refers to breaking a whole into its separate components for individual examination. Data analysis is a process for obtaining raw data and converting it into information useful for decision-making by users. Data is collected and analyzed to answer questions, test hypotheses or disprove theories.

Statistician John Tukey defined data analysis in 1961 as: "Procedures for analyzing data, techniques for interpreting the results of such procedures, ways of planning the gathering of data to make its analysis easier, more precise or more accurate, and all the machinery and results of (mathematical) statistics which apply to analyzing data."

There are several phases that can be distinguished, described below. The phases are interactive, in that feedback from later phases may result in additional work in earlier phases.

Data requirements

The data is necessary as inputs to the analysis are specified based upon the requirements of those directing the analysis or customers who will use the finished product of the analysis. The general type of entity upon which the data will be collected is referred to as an experimental unit (e.g., a person or population of people). Specific variables regarding a population (e.g., age and income)

may be specified and obtained. Data may be numerical or categorical (i.e., a text label for numbers).

Data collection

Data is collected from a variety of sources. The requirements may be communicated by analysts to custodians of the data, such as information technology personnel within an organization. The data may also be collected from sensors in the environment, such as traffic cameras, satellites, recording devices, etc. It may also be obtained through interviews, downloads from online sources, or reading documentation.

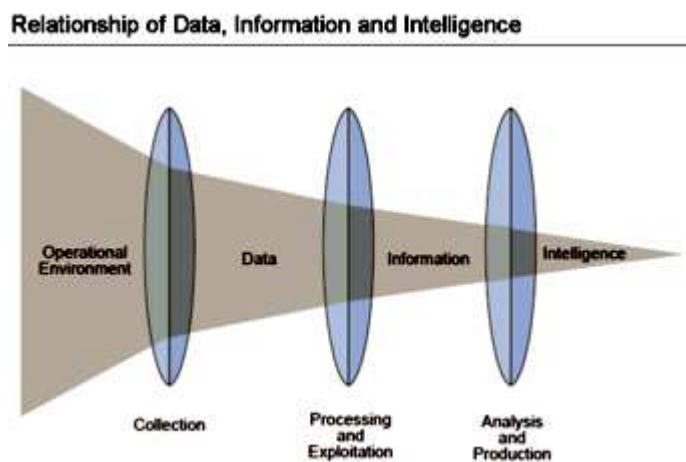


Figure 1.2: The phases of the intelligence cycle used to convert raw information into actionable intelligence or knowledge are conceptually similar to the phases of data analysis.

Data initially obtained must be processed or organized for analysis. For instance, these may involve placing data into rows and columns in a table format (i.e., structured data) for further analysis, such as within a spreadsheet or statistical software.

Data cleaning

Once processed and organized, the data may be incomplete, contain duplicates, or contain errors. The need for data cleaning will arise from problems in the way that data is entered and stored. Data cleaning is the process of preventing and correcting these errors. Common tasks include record matching, identifying inaccuracy of data, the overall quality of existing data, deduplication, and column segmentation. Such data problems can also be identified through a variety of analytical techniques. For example, with financial information, the totals for particular

variables may be compared against separately published numbers believed to be reliable. Unusual amounts above or below pre-determined thresholds may also be reviewed. There are several types of data cleaning that depend on the type of data such as phone numbers, email addresses, employers etc. Quantitative data methods for outlier detection can be used to get rid of likely incorrectly entered data. Textual data spell checkers can be used to lessen the amount of mistyped words, but it is harder to tell if the words themselves are correct.

Exploratory data analysis

Once the data is cleaned, it can be analyzed. Analysts may apply a variety of techniques referred to as exploratory data analysis to begin understanding the messages contained in the data. The process of exploration may result in additional data cleaning or additional requests for data, so these activities may be iterative in nature. Descriptive statistics such as the average or median may be generated to help understand the data. Data visualization may also be used to examine the data in graphical format, to obtain additional insight regarding the messages within the data.

Modeling and algorithms

Mathematical formulas or models called algorithms may be applied to the data to identify relationships among the variables, such as correlation or causation. In general terms, models may be developed to evaluate a particular variable in the data based on another variable (s) in the data, with some residual error depending on model accuracy (i.e., $\text{Data} = \text{Model} + \text{Error}$).

Inferential statistics includes techniques to measure relationships between particular variables. For example, regression analysis may be used to model whether a change in advertising (independent variable X) explains the variation in sales (dependent variable Y). In mathematical terms, Y (sales) is a function of X (advertising). It may be described as $Y = aX + b + \text{error}$, where the model is designed such that a and b minimize the error when the model predicts Y for a given range of values of X. Analysts may attempt to build models that are descriptive of the data to simplify analysis and communicate results.

Data product

A data product is a computer application that takes data inputs and generates outputs, feeding them back into the environment. It may be based on a model or algorithm. An example is an

application that analyzes data about customer purchasing history and recommends other purchases the customer might enjoy.

Communication



Figure 1.3: Data visualization to understand the results of a data analysis.

Main article: Data visualization

Once the data is analyzed, it may be reported in many formats to the users of the analysis to support their requirements. The users may have feedback, which results in additional analysis. As such, much of the analytical cycle is iterative.

When determining how to communicate the results, the analyst may consider data visualization techniques to help clear and efficiently communicate the message to the audience. Data visualization uses information displays such as tables and charts to help communicate key messages contained in the data. Tables are helpful to a user who might lookup specific numbers, while charts (e.g., bar charts or line charts) may help explain the quantitative messages contained in the data.

7.3.2 TECHNIQUES FOR ANALYZING DATA

Author Jonathan Koomey has recommended a series of best practices for understanding quantitative data. These include:

1. Check raw data for anomalies prior to performing your analysis;
2. Re-perform important calculations, such as verifying columns of data that are formula driven;
3. Confirm main totals are the sum of subtotals;

4. Check relationships between numbers that should be related in a predictable way, such as ratios over time;
5. Normalize numbers to make comparisons easier, such as analyzing amounts per person or relative to GDP or as an index value relative to a base year;
6. Break problems into component parts by analyzing factors that led to the results, such as DuPont analysis of return on equity.

For the variables under examination, analysts typically obtain descriptive statistics for them, such as the mean (average), median, and standard deviation. They may also analyze the distribution of the key variables to see how the individual values cluster around the mean.

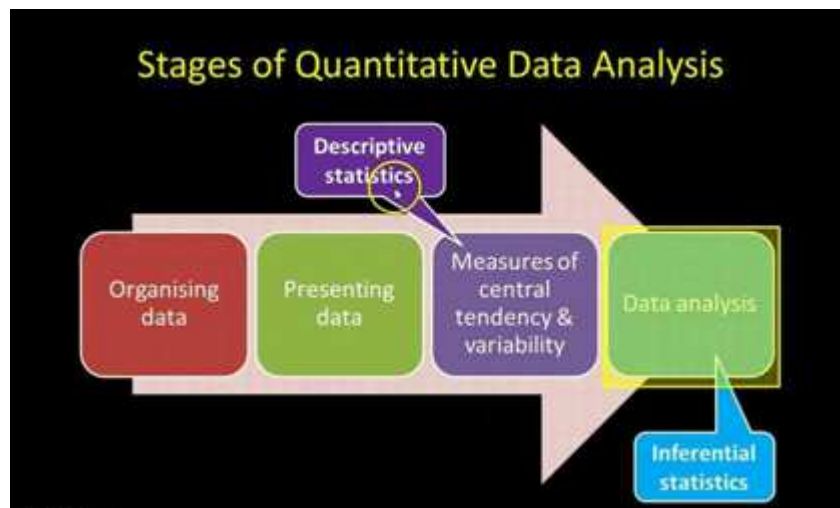


Figure 1.4: Stages of quantitative data analysis

An illustration of the MECE principle used for data analysis:

The consultants at McKinsey and Company named a technique for breaking a quantitative problem down into its component parts called the MECE principle. Each layer can be broken down into its components; each of the sub-components must be mutually exclusive of each other and collectively add up to the layer above them. The relationship is referred to as "Mutually Exclusive and Collectively Exhaustive" or MECE. For example, profit by definition can be broken down into total revenue and total cost. In turn, total revenue can be analyzed by its

components, such as revenue of divisions A, B, and C (which are mutually exclusive of each other) and should add to the total revenue (collectively exhaustive).

Analysts may use robust statistical measurements to solve certain analytical problems. Hypothesis testing is used when a particular hypothesis about the true state of affairs is made by the analyst and data is gathered to determine whether that state of affairs is true or false. For example, the hypothesis might be that "Unemployment has no effect on inflation", which relates to an economics concept called the Phillips Curve. Hypothesis testing involves considering the likelihood of Type I and type II errors, which relate to whether the data supports accepting or rejecting the hypothesis.

Regression analysis may be used when the analyst is trying to determine the extent to which independent variable X affects dependent variable Y (e.g., "To what extent do changes in the unemployment rate (X) affect the inflation rate (Y)?"). This is an attempt to model or fit an equation line or curve to the data, such that Y is a function of X.

Necessary condition analysis (NCA) may be used when the analyst is trying to determine the extent to which independent variable X allows variable Y (e.g., "To what extent is a certain unemployment rate (X) necessary for a certain inflation rate (Y)?"). Whereas (multiple) regression analysis uses additive logic where each X-variable can produce the outcome and the X's can compensate for each other (they are sufficient but not necessary), necessary condition analysis (NCA) uses necessity logic, where one or more X-variables allow the outcome to exist, but may not produce it (they are necessary but not sufficient). Every single necessary condition must be present and compensation is not possible.

7.3.3 INITIAL DATA ANALYSIS

The most important distinction between the initial data analysis phase and the main analysis phase is that during initial data analysis one refrains from any analysis that is aimed at answering the original research question. The initial data analysis phase is guided by the following four questions:

Quality of data

The quality of the data should be checked as early as possible. Data quality can be assessed in several ways, using different types of analysis: frequency counts, descriptive statistics (mean,

standard deviation, median), normality (skewness, kurtosis, frequency histograms, n: variables are compared with coding schemes of variables external to the data set, and possibly corrected if coding schemes are not comparable.

- Test for common method variance.

The choice of analyses to assess the data quality during the initial data analysis phase depends on the analyses that will be conducted in the main analysis phase.

Quality of measurements

The quality of the measurement instruments should only be checked during the initial data analysis phase when this is not the focus or research question of the study. One should check whether the structure of measurement instruments corresponds to structure reported in the literature.

There are two ways to assess measurement:

- Analysis of homogeneity (internal consistency), which gives an indication of the reliability of a measurement instrument. During this analysis, one inspects the variances of the items and the scales, the Cronbach's α of the scales, and the change in the Cronbach's alpha when an item would be deleted from a scale.

Initial transformations

After assessing the quality of the data and of the measurements, one might decide to impute missing data, or to perform initial transformations of one or more variables, although this can also be done during the main analysis phase.

Possible transformations of variables are:

- Square root transformation (if the distribution differs moderately from normal)
- Log-transformation (if the distribution differs substantially from normal)
- Inverse transformation (if the distribution differs severely from normal)
- Make categorical (ordinal/dichotomous) (if the distribution differs severely from normal, and no transformations help)

Did the implementation of the study fulfill the intentions of the research design?

One should check the success of the randomization procedure, for instance by checking whether background and substantive variables are equally distributed within and across groups. If the study did not need or use a randomization procedure, one should check the success of the non-random sampling, for instance by checking whether all subgroups of the population of interest are represented in the sample.

Other possible data distortions that should be checked are

- dropout (this should be identified during the initial data analysis phase)
- Item nonresponse (whether this is random or not should be assessed during the initial data analysis phase)
- Treatment quality (using manipulation checks).

Characteristics of the data sample

In any report or article, the structure of the sample must be accurately described. It is especially important to exactly determine the structure of the sample (and specifically the size of the subgroups) when subgroup analyses will be performed during the main analysis phase. The characteristics of the data sample can be assessed by looking at:

- Basic statistics of important variables
- Scatter Plots
- Correlations and associations
- Cross-tabulations

The final stage of the initial data analysis

During the final stage, the findings of the initial data analysis are documented, and necessary, preferable, and possible corrective actions are taken.

Also, the original plan for the main data analyses can and should be specified in more detail or rewritten.

In order to do this, several decisions about the main data analyses can and should be made:

- In the case of non-normals: should one transform variables; make variables categorical (ordinal/dichotomous); adapt the analysis method?
- In the case of missing data: should one neglect or impute the missing data; which imputation technique should be used?
- In the case of outliers: should one use robust analysis techniques?
- In case of items do not fit the scale: should one adapt the measurement instrument by omitting items, or rather ensure comparability with other (uses of the) measurement instrument(s)?
- In the case of (too) small subgroups: should one drop the hypothesis about inter-group differences, or use small sample techniques, like exact tests or bootstrapping?
- In case the randomization procedure seems to be defective: can and should one calculate propensity scores and include them as covariates in the main analyses?

Analysis

Several analyses can be used during the initial data analysis phase:

- Univariate statistics (single variable)
- Bivariate associations (correlations)
- Graphical techniques (scatter plots)

It is important to take the measurement levels of the variables into account for the analyses, as special statistical techniques are available for each level:

- Nominal and ordinal variables
 - Frequency counts (numbers and percentages)
 - Associations
 - circumambulations (crosstabulations)
 - hierarchical log-linear analysis (restricted to a maximum of 8 variables)
 - log-linear analysis (to identify relevant/important variables and possible confounders)

- Exact tests or bootstrapping (in case subgroups are small)
- Computation of new variables
- Continuous variables
 - Distribution
 - Statistics (M, SD, variance, skewness, kurtosis)
 - Stem-and-leaf displays
 - Box plots

Nonlinear analysis

The nonlinear analysis will be necessary when the data is recorded from a nonlinear system. Nonlinear systems can exhibit complex dynamic effects including bifurcations, chaos, harmonics, and subharmonics that cannot be analyzed using simple linear methods. Nonlinear data analysis is closely related to nonlinear system identification.

7.3.4 MAIN DATA ANALYSIS

In the main analysis phase analyses aimed at answering the research question are performed as well as any other relevant analysis needed to write the first draft of the research report.

Exploratory and confirmatory approaches

In the main analysis phase, either an exploratory or confirmatory approach can be adopted. Usually, the approach is decided before data is collected. In an exploratory analysis, no clear hypothesis is stated before analyzing the data, and the data is searched for models that describe the data well. In a confirmatory analysis, clear hypotheses about the data are tested.

Exploratory data analysis should be interpreted carefully. When testing multiple models at once there is a high chance of finding at least one of them to be significant, but this can be due to a type 1 error. It is important to always adjust the significance level when testing multiple models with, for example, a Bonferroni correction. Also, one should not follow up an exploratory analysis with a confirmatory analysis in the same dataset. An exploratory analysis is used to find ideas for a theory, but not to test that theory as well. When a model is found exploratory in a dataset, then following up that analysis with a confirmatory analysis in the same dataset could

simply mean that the results of the confirmatory analysis are due to the same type 1 error that resulted in the exploratory model in the first place. The confirmatory analysis, therefore, will not be more informative than the original exploratory analysis.

Stability of results

It is important to obtain some indication of how generalizable the results are. While this is hard to check, one can look at the stability of the results. Are the results reliable and reproducible?

There are two main ways of doing this:

- Cross-validation: By splitting the data into multiple parts we can check if an analysis (like a fitted model) based on one part of the data generalizes to another part of the data as well.
- Sensitivity analysis: A procedure to study the behavior of a system or model when global parameters are (systematically) varied. One way to do this is with bootstrapping.

Statistical methods

Many statistical methods have been used for statistical analyses. A very brief list of four of the more popular methods is:

- General linear model: A widely used model on which various methods are based (e.g. t-test, ANOVA, ANCOVA, MANOVA). Usable for assessing the effect of several predictors on one or more continuous dependent variables.
- Generalized linear model: An extension of the general linear model for discrete dependent variables.
- Structural equation modeling: Usable for assessing latent structures from measured manifest variables.
- Item response theory: Models for (mostly) assessing one latent variable from several binary measured variables (e.g. an exam).

7.4 *ERROR ANALYSIS*

Performing the experiment and collecting data is only the beginning of the process of completing an experiment in science. Understanding the results of any given experiment is always the central goal of the experiment. Presenting those results in a clear concise manner completes the experiment. This overview of the complete process is as valid in an instructional laboratory course as in a research environment. You will not have learned any physics if you did not understand the experiment. Presenting the results of your experimental work is fundamentally saying, **“This is what I did and this is what I learned.”** Putting together your presentation of the results should help you clarify the results for yourself. (If your instructor can clearly see what you did and what you learned, you might get a better grade.) **Data analysis should NOT be delayed** until all of the data is recorded. Take a low point, a high point and maybe a middle point, and do a quick analysis and plot. This will help one avoid the problem of spending an entire class collecting bad data because of a mistake in experimental procedure or an equipment failure.

First and foremost, data analysis means understanding what your results mean. When analyzing your data, try to think through the physical processes which are occurring. Write your train of thought down. Ultimately, the goal is for you to understand physics and the world a bit better. Your understanding of your results probably occurs in stages, with each stage being a refinement and possibly more mathematical than the previous stage. For example, one might first note that as time increases so does distance. Next, a quick graph of distance vs time might verify this understanding but the relationship is NOT linear, i.e. the data does not form a straight line. By further work, one might discover that distance increase linearly with the square of the time. Or sometimes the mathematical relationship may remain hidden.

Relate each successive stage of your understanding and interpretation of your results to the physical principles that are involved. In the above example, one might note that the change in position with time is caused by the velocity that is in turn caused by an acceleration of the gravitational force.

Finally, develop the related mathematics. Equations are nearly meaningly unless they are related to the physical laws. (Remember to identify all the variables and constants in your equations.)

Sometimes, your results will not support and may even contradict the physical explanation suggested by the manual or your instructor. Say so! But of course, then a few suggestions as to the reason for this apparent failure of the physical laws, would be in order. Do NOT just say "The equipment was a piece of sh_t!" Try to explain what went wrong or what competing effects have come into play.

One of the reasons that you are encouraged to record everything that is going on as it is going on, is that this information may help explain bad results. For example, partly for fun, you note each time your lab partner sneezes. Later while looking at the data, you discover that each data point that was being collected during a sneeze deviates from the pattern of the rest of the data. This may give you a good reason for dropping "bad" data. **The quality of the data, determines to a great extent, what conclusions can be reached from them.** If you are looking for a small effect, say a total change of 1 mm, and the uncertainties in your data are 2 mm then you really can not make any solid conclusions. (See the section on error analysis below.)

When one considers the **quality of a measurement** there are two aspects to consider. The first is if one were to repeat the measurement, how close would new results be to the old, i.e., how reproducible is the measurement? Scientists refer to this as the *precision* of the measurement. Secondly, a measurement is considered "good" if it agrees with the true value. This is known as the *accuracy* of the measurement. But there is a potential problem in that one needs to know the "true value" to determine the accuracy.

A good measurement must be close to the "true value" and be reproducible. In this experiment, if someone made one measurement of g_2 and got 9.79 m/s², it would be an accurate measurement. But if next time they tried they got 4.12 m/s², no one would believe that they were anything but lucky in the first measurement. Similarly, if one group 2 got values of 7.31, 7.30, 7.33, and 7.29 m/s² their results are reproducible but not really very good.

7.4.1 TYPES OF ERRORS

The words “error” and “uncertainty” are used to describe the same concept in measurement.

It is unfortunate that the term, “error” is the standard scientific word because usually there is no mistake or error in making a measurement. Frequently, the uncertainties are dominated by natural irregularities or differences in what is being measured.

All measurements have errors. Errors may arise from three sources:

a) **Careless errors:** These are due to mistakes in reading scales or careless setting of markers, etc. They can be eliminated by repetition of readings by one or two observers.

b) **Systematic errors:** These are due to built-in errors in the instrument either in design or calibration. Repetition of observation with the same instrument will not show a spread in the measurements. They are the hardest source of errors to detect.

c) **Random errors:** These always lead to a spread or distribution of results on repetition of the particular measurement. They may arise from fluctuations in either the physical parameters due to the statistical nature of the particular phenomenon or the judgment of the experimenter, such as variation in response time or estimation in scale reading.

Taking multiple measurements helps reduce uncertainties.

DETERMINING ERRORS:

Although it is interesting and reassuring to compare your results against “accepted values,” it is suggested that error analysis is done, before this comparison. One reason is that the accuracy of a set of measurements depends on how well the experiment was done, not how close the measurement was to the accepted value. One could get close to the accepted value, by sloppiness and luck.

Do not base your experimental uncertainties on the accepted values.

Determining the source of uncertainty and the magnitude of this uncertainty is often difficult. Some errors or uncertainties are caused by natural fluctuations or irregularities. These cannot be eliminated. To estimate these uncertainties one frequently uses mathematical methods similar to

those discussed in the section below titled, Averages and Deviation. Another method of estimating uncertainties is to assign an uncertainty to the measurement equal to the finest scale reading on the measuring instrument. For example, if a ruler may be marked in millimeters. then the uncertainty in any measurement with this ruler can be given as 1 mm. But with practice, one might be able to interpolate the scale and reduce the error to 0.25 mm.

EXPRESSING ERRORS:

For each measured value, **A**, there is an **estimated error, ΔA**. The complete result is given by **A ± ΔA**. This means that the “true value” probably lies between a maximum value of **A + ΔA** and a minimum value of **A - ΔA**. Sometimes the terms **relative error** and **percent error** is used, where:

$$\text{Relative Error} = \frac{\text{Estimated Error}}{\text{Estimated Value}} = \frac{\Delta A}{A}$$

$$\text{Percent Error} = \text{Relative Error} \times 100\% = \frac{\Delta A}{A} \times 100\%$$

Errors are expressed in graphs by using error bars. Consider a data point (**A,B**) and the associated Uncertainties are **ΔA** and **ΔB** respectively. The vertical error bar is drawn from **B-ΔB** to **B+ΔB**.

Similarly, the horizontal error bar is drawn from **A-ΔA** to **A+ΔA**. For example if **A = 4.2±0.8** and **B = 3.2±0.5** then this point would be plotted like:

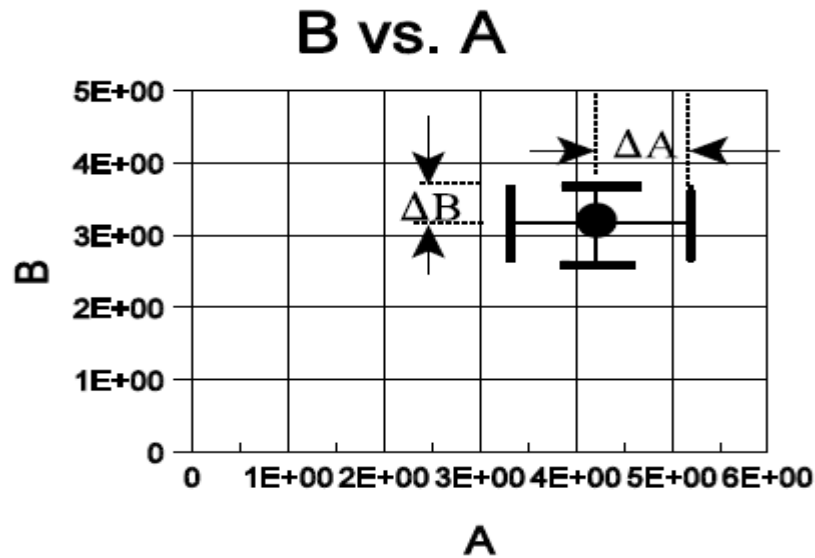


Figure 1.5: Expressing errors

7.4.2 PROPAGATION OF ERRORS

When measured values are used to calculate other values, the uncertainties in these measured values causes uncertainties in the calculated values. Calculating the uncertainties in the calculated values is

called *error propagation*. For the few simple cases that are discussed below, let **C** be a function **A** and **B** and the associated uncertainties are ΔC , ΔA , and ΔB , respectively.

1) Product with a Constant

Here $C = k A$ where **k** is a constant. Then

$$\Delta C = k \Delta A.$$

this rule can be applied if **k** is a measured quantity

with a relatively negligible uncertainty, for example, if **k** were the gravitational constant, **g**.

2) Addition and Subtraction

Here $C = A + B$ or $C = A - B$. In either case:

$$\Delta C = \sqrt{(\Delta A)^2 + (\Delta B)^2}$$

Note that ΔC is less than $\Delta A + \Delta B$. This is an expression

of the fact that the uncertainties in A and B are independent of each other. (In math lingo, one could say that ΔA , and ΔB are orthogonal or perpendicular to each other. Note how the calculation of ΔC is identical to the Pythagorean theorem for the sides of a right triangle.)

Why should the same formula work for addition and subtraction? Notice that the original uncertainties are squared.

3) Multiplication or Division

If $C = A B$ or $C = A/B$ then:

$$\frac{\Delta C}{C} = \sqrt{\left(\frac{\Delta A}{A}\right)^2 + \left(\frac{\Delta B}{B}\right)^2}$$

$$\Delta C = C \times \sqrt{\left(\frac{\Delta A}{A}\right)^2 + \left(\frac{\Delta B}{B}\right)^2}$$

Again the assumption is that the uncertainties in A and B are independent. An example of when this is not true is $C = A^2$. This brings up another rule.

4) Raised to a Power

This is the case where $C = A^n$ where n is a constant. In this case:

$$\frac{\Delta C}{C} = |n| \frac{\Delta A}{A}$$

5) Graphical Analysis of Uncertainties in Slopes and Intercepts.

If the slope or intercept of a line on a plot is the required calculated value (or the required value is calculated from these values) then the uncertainty of the slope and intercept will also be required. Graphically one can estimate these uncertainties. First, draw the best line possible, and then draw the two lines that just barely pass through the data. The differences of these slopes and intercepts from those of the best fit line provide an estimate of the uncertainties in these quantities.

In the graph below the **solid line** is a good fit to the data and the while the **dotted line** and the **dashed line** represent the extremes for lines that just fit the data.. The dashed line has the steepest slope and will be referred to as the max. line, while the dotted line will be referred to as the min. line. The slopes and y- intercepts for these three lines are:

	Slope	(dif.)	y-inter.	(dif.)
Best fit		1.0	2.0	
Max. line	1.16	(0.16)	-1.5	(-3.5)
Min. Line	0.81	(-0.19)	5.2	(3.2)

Thus if the data was expected to fit the equation:

$$y = a + bx,$$

then one would estimate the constants as:

$$a = 2.0 \pm 3.4$$

$$b = 1.0 \pm 0.2.$$

The uncertainties are based on averaging the absolute value of the differences (labeled (dif.)).

Graphical Error Analysis of Slope and y-intercept

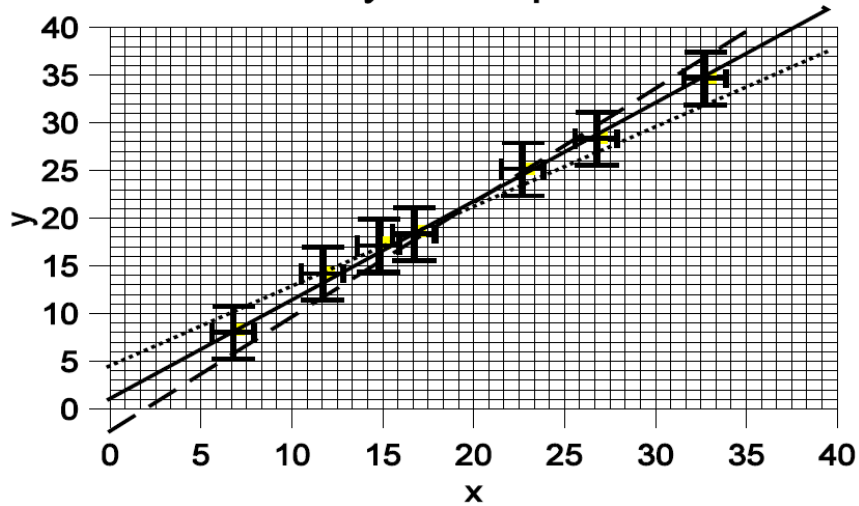


Figure 1.6: Graphical Error Analysis of Slope and y- intercept

6) General Formula (*Advanced*)

All of the above examples of propagation of errors are special cases of the general formula.

Consider a calculated variable z that is a function of two measured variables x and y , then one writes: If the uncertainties associated with x and y are Δx and Δy , respectively. The uncertainty, Δz in z is:

$$\Delta z = \sqrt{\left(\frac{\partial z}{\partial x}\right)^2 (\Delta x)^2 + \left(\frac{\partial z}{\partial y}\right)^2 (\Delta y)^2}$$

Here a symbol $\frac{\partial z}{\partial x}$ that may be somewhat unfamiliar. It is the partial derivative of z with respect to x . Partial derivatives is used when a variable is a function of more than one variable. To evaluate a partial

Derivative, one just takes the normal derivative but treats the variable y as a constant. That is one

$$\frac{dz}{dy} \frac{dy}{dx}$$

does not do any chain rule, kind of stuff. Note that the uncertainties are assumed to be independent of each other and add as if they are vectors at right angles to each other.

7.5 ACCURACY AND PRECISION

Accuracy vs. Precision: These two words do not mean the same thing. "Accuracy" deals with how close is a measured value to an accepted or "true" value.

"Precision" deals with how reproducible are a given measurement. Because, *precision* is a measure of how reproducible a measurement is, one can gain some knowledge of the precision simply by taking a number of measurements and comparing them. If the true value is not known, the *accuracy* of a measurement is more difficult to know. Here are a few things to consider when trying to estimate (or explain) the quality of a measurement:

1. How well does your equipment make the needed measurements? Two examples of problems: a meter stick with 2 mm worn off one end and trying to measure 0.01 mm with a meter stick.
2. How does the lack of precision, or *uncertainty* in any one measure the variable effect the final calculated value? For example, if the data is expected to fall in a straight line, some uncertainties may only shift the intercept and leave the slope unchanged.
3. Making a plot that shows each of your measurements, can help you "see" your uncertainty.

Also, certain uncertainties are sidestepped by extracting a slope from a plot and calculating the final value from this slope.

4. One can estimate the uncertainty after making multiple measurements. First, note that when plotted, about an of the data points will be outside of the *error bars* as they are normally drawn at one *standard deviation*. For example, if there are six measurements, one expects that 2 of the points (probably one too high and one too low) will be outside the normal error bars. One can draw a reasonable set of error bars based on this assumption, by drawing the upper part of the

error bar between the highest two points and the lower part of the error bar between the lowest two points.

Data analysis is seldom a straightforward process because of the presence of uncertainties.

Data can not be fully understood until the associated uncertainties are understood.

7.6 LEAST SQUARES ANALYSIS

Although extracting a slope or a y-intercept from a graph is relatively straightforward, the method has several limitations. The result is effected by the user's skill and bias. Also, precision may be lost in the process of graphing and extracting the slope and intercept. Evaluating the uncertainties using graphical methods is usually even more difficult. If data is only evaluated graphically, two different people evaluating the same data would get different slopes. Part of the scientific method is that the techniques must be reproducible. Usually, when a curve (e.g. a straight line) is fitted to some data, the method used is **least squares analysis**. Although this method is quite general, only the limited case of fitting a straight line to data is presented here. For this discussion the following assumptions are made:

- 1) The uncertainties are the same for all of the data points.
- 2) The uncertainties are only in the dependent variable, y.
- 3) The uncertainties are all random and the multiple measurements for a given value of the independent variable, x, would be distributed according to a Gaussian.

The data, in this case, would be pairs of measurements (x_i and y_i) and the goal is to determine the straight line ($y = a + bx$) that best models this data. One wants to make the minimize the difference between the measured value y_i , and $a + bx_i$ which the associated value calculated from the straight line. Explicitly one is looking for values of **a** and **b** the minimize

$$\chi^2 = \sum_{i=1}^n (y_i - (a + b x_i))^2$$

χ^2 is pronounced chi-squared and is an indicator of the goodness of fit.

Without any prove the following results of the linear least squares analysis of data yields a straight line, $y = a + bx$, are given:

$$a = \frac{1}{\Delta'} \left[\left(\sum (x_i^2) \right) \left(\sum y_i \right) - \left(\sum x_i \right) \left(\sum (x_i y_i) \right) \right]$$

$$b = \frac{1}{\Delta'} \left[n \times \sum (x_i y_i) - \left(\sum x_i \right) \left(\sum y_i \right) \right]$$

$$\Delta' = n \times \sum (x_i^2) - \left(\sum x_i \right)^2$$

Here all of the sums are from $i = 1$ to n .

Note that all of the individual sums can be built up a point at a time, which make calculation relatively simple. This is especially easy if, **one uses a spreadsheet**. With x and y in the two columns labeled **A** and **B**, one can have column **C** be x^2 , column **D** be y^2 , and column **E** be the product xy . Then sum each column and calculate the values.

Many calculators and spreadsheet programs provide access to least squares analysis routines.

You are encouraged to learn to use these programs If one needs to include uncertainties σ_i associated with the y_i in this calculation the above formula becomes:

$$a = \frac{1}{\Delta'} \left[\sum \left(\frac{x_i^2}{\sigma_i^2} \right) \sum \left(\frac{y}{\sigma_i^2} \right) - \sum \left(\frac{x_i}{\sigma_i^2} \right) \sum \left(\frac{x_i y_i}{\sigma_i^2} \right) \right]$$

$$b = \frac{1}{\Delta'} \left[\sum \left(\frac{1}{\sigma_i^2} \right) \sum \left(\frac{x_i y_i}{\sigma_i^2} \right) - \sum \left(\frac{x_i}{\sigma_i^2} \right) \sum \left(\frac{y_i}{\sigma_i^2} \right) \right]$$

$$\Delta' = \sum \left(\frac{1}{\sigma_i^2} \right) \sum \left(\frac{x_i^2}{\sigma_i^2} \right) - \left(\sum \left(\frac{x_i}{\sigma_i^2} \right) \right)^2$$

7.7 AVERAGE Vs. DEVIATION

The *average* of a series of measurements is one of the most common methods of analyzing data.

The average, or *arithmetic mean* for a series of n numbers: $x_1, x_2, x_3, \dots, x_n$ is defined as:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} = \sum_{i=1}^n \frac{x_i}{n}$$

Example: The length of a table was measured three

times with the following results in meters:

$x_1 = 1.42$, $x_2 = 1.45$, and $x_3 = 1.41$.

Then the average is:

$$\bar{x} = \frac{1.42 + 1.45 + 1.41}{3} = 1.43 \text{ meters}$$

Often one wants to compare individual measurements to the average. The *deviation* is a

simple quantity that is frequently used for this kind of comparison. The deviation δx_i of the measurement labeled i , from the average, is:

$$\delta x_i = x_i - \bar{x}.$$

Continued Example: The deviations for the above

example are:

$$\delta x_1 = 1.42 - 1.43 = -0.01 \text{ meters ,}$$

$$\delta x_2 = 1.45 - 1.43 = 0.02 \text{ meters ,}$$

$$\delta x_3 = 1.41 - 1.43 = -0.02 \text{ meters .}$$

7.8 STANDARD DEVIATION

If the deviations of a measurement were averaged, the result would be zero because of high and low values would cancel each other. Generally one expresses the fluctuation about the average measurement with by calculating and quoting the **standard deviation**, σ of the n measurements.

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

Continued Example: $\sigma = 0.0212$ meter

$$\sigma = \sqrt{\frac{(-0.01)^2 + (+0.02)^2 + (-0.02)^2}{3 - 1}} \quad (20)$$

Finally, the term, standard deviation of the mean defined as:

$$\sigma_m = \frac{\sigma}{\sqrt{n}} .$$

The standard deviation of the mean, is the standard measure for describing of the the precision of a measurement, i.e. how well a number of measurements agree with themselves. Thus:

$$\Delta x = \sigma_m$$

$$\Delta x = \frac{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}}{n}$$

Continued Example: $\Delta x = \sigma_m = 0.012$ meter

7.9 LET US SUM UP

- Data analysis, also known as analysis of data or data analytics.
- The term *data analysis* is sometimes used as a synonym for data modeling.
- The process of data analysis involves data requirement, data collection, data processing, data cleaning, exploratory data analysis, Modeling and algorithms, data product and communication.
- Initial data analysis phase involves quality of data, quality of measurements, initial transformation, characteristics of the data sample, the final stage of initial data analysis and analysis.
- In the main analysis phase analyses aimed at answering the research question are performed as well as any other relevant analysis needed to write the first draft of the research report.
- First and foremost, data analysis means understanding what your results mean. When analyzing your data, try to think through the physical processes which are occurring.
- All measurements have errors. Errors may arise from three sources: a) **Careless errors** b) **Systematic errors** c) **Random errors**
- Calculating the uncertainties in the calculated values is called *error propagation*
- "**Accuracy**" deals with how close is a measured value to an accepted or "true" value.
- "**Precision**" deals with how reproducible is a given measurement.
- When a curve is fitted to some data the method used is **least squares analysis**.

7.10 CHECK YOUR PROGRESS

Q1. State whether the following statements are true (T) or false (F):

- i- The term *data analysis* is sometimes used as a synonym for data modeling.
- ii- Initial data analysis phase involves quality of data.
- iii- Measurements have no errors.
- iv- Calculating the uncertainties in the calculated values is called error propagation.
- v- When a curve is not fitted to any data the method used is least squares analysis.

Ans. to Q. No. 1: i) True ii) True iii) False iv) True v) False

7.11 FURTHER READINGS

- Adèr, H.J. & Mellenbergh, G.J. (with contributions by D.J. Hand) (2008). *Advising on Research Methods: A Consultant's Companion*. Huizen, the Netherlands: Johannes van Kessel Publishing.
- ASTM International (2002). *Manual on Presentation of Data and Control Chart Analysis*, MNL 7A, ISBN 0-8031-2093-1
- Chambers, John M.; Cleveland, William S.; Kleiner, Beat; Tukey, Paul A. (1983). *Graphical Methods for Data Analysis*, Wadsworth/Duxbury Press. ISBN 0-534-98052-X
- Chekanov, S. (2016) *Numeric Computation and Statistical Data Analysis on the Java Platform*, Springer. ISBN 978-3-319-28531-3
- Fandango, Armando (2008). *Python Data Analysis, 2nd Edition*. Packt Publishers.

1.12 MODEL QUESTIONS

- Q.1. Explain the concept of Data Analysis.
- Q.2. What do you mean by the term error analysis? Explain.
- Q.3. What is the difference between initial data analysis and main data analysis?
- Q.4. Write short notes on
 - a. Accuracy and Precision

b. Average vs deviation

c. Least Square Analysis

d. Standard deviation

Q.5. What are the advantages of the data analysis?

Q.6. What do you mean by propagation of errors? Explain with the help of an example.

UNIT 8: HISTORY OF DEVELOPMENT OF COMPUTERS

CONTENTS:

- 8.1 Learning Objectives
- 8.2 Introduction
- 8.3 History of Computer
- 8.4 Types of computers
 - 8.4.1 According to the purpose of the computer.
 - 8.4.2 According to the operation of the computer.
 - 8.4.3 According to the size of the computer.
- 8.5 General awareness of Computer Hardware
 - 8.5.1 CPU
 - 8.5.2 Peripheral devices
- 8.6 Let Us Sum Up
- 8.7 Check Your Progress
- 8.8 Further Readings
- 8.9 Model Questions

8.1 LEARNING OBJECTIVES

After thoroughly studying this unit, you would be able to understand:

- Computers and history of computers
- Different types of computers
- Various definition and concepts of computers

- Computers and various peripheral devices of the computer

8.2 INTRODUCTION

We are living in the computer age. Most of our day to day jobs are being influenced by the use of computers. It is used increasingly in each and every field of our life. In the areas of science and technology improvements cannot be achieved without the use of computers. Hence it has become necessary to have basic knowledge about computers.

Strictly speaking, the computer is a calculating device having certain important characteristics like speed, storage capacity, accuracy etc. But, nowadays it is used for many more applications other than computing. It has become an indispensable tool in the field of communications.

8.3 HISTORY OF COMPUTER

The history of computer development is a computer science topic that is often used to reference the different generations of computing devices. Each one of the five generations of computers is characterized by a major technological development that fundamentally changed the way computers operate. Most major developments from the 1940's to present day have resulted in increasingly smaller, cheaper, more powerful and more efficient computing devices.

First Generation of Computers (1942-1955)

The beginning of commercial computer age is from UNIVAC (Universal Automatic Computer). It was developed by two scientists Mauchly and Eckert at the Census Department of United States in 1947. The first generation computers were used during 1942-1955. They were based on vacuum tubes. Examples of first generation computers are ENIVAC and UNIVAC-1.

Advantages:

- Vacuum tubes were the only electronic component available during those days.
- Vacuum tube technology made possible to make electronic digital computers.
- These computers could calculate data in a millisecond.

Disadvantages

- The computers were very large in size.

- They consumed a large amount of energy.
- They heated very soon due to thousands of vacuum tubes.
- They were not very reliable.
- Air conditioning was required.
- Constant maintenance was required.
- Non-portable.
- Costly commercial production.
- Limited commercial use.
- Very slow speed.
- Limited programming capabilities.
- Used machine language only.
- Used magnetic drums which provide very fewer data storage.
- Used punch cards for input.
- Not versatile and very faulty.

Second Generation Computers (1955-1964)

The second generation computers used transistors. The scientists at Bell laboratories developed transistor in 1947. These scientists include John Barden, William Brattain, and William Shockley. The size of the computers was decreased by replacing vacuum tubes with transistors. The examples of second generation computers are **IBM 7094 series**, **IBM 1400 series** and **CDC 164** etc.

Advantages

- Smaller in size as compared to the first generation computers.
- The 2nd generation Computers were more reliable
- Used less energy and were not heated.

- Wider commercial use
- Better portability as compared to the first generation computers.
- Better speed and could calculate data in microseconds
- Used faster peripherals like tape drives, magnetic disks, printer etc.
- Used Assembly language instead of Machine language.
- Accuracy improved.
-

Disadvantages

- The cooling system was required
- Constant maintenance was required
- Commercial production was difficult
- Only used for specific purposes
- Costly and not versatile
- Punch cards were used for input.

Third Generation Computers (1964-1975)

The Third generation computers used the integrated circuits (IC). Jack Kilby developed the concept of an integrated circuit in 1958. It was an important invention in the computer field. The first IC was invented and used in 1961. The size of an IC is about ¼ square inch. A single IC chip may contain thousands of transistors. The computer became smaller in size, faster, more reliable and less expensive. The examples of third generation computers are IBM 370, IBM System/360, UNIVAC 1108 and UNIVAC AC 9000 etc.

Advantages

- Smaller in size as compared to previous generations.
- More reliable.

- Used less energy
- Produced less heat as compared to the previous two generations of computers.
- Better speed and could calculate data in nanoseconds.
- Used fan for heat discharge to prevent damage.
- Maintenance cost was low because hardware failure is rare.
- Totally general purpose
- Could be used for high-level languages.
- Good storage
- Versatile to an extent
- Less expensive
- Better accuracy
- Commercial production increased.
- Used mouse and keyboard for input.

Disadvantages

- Air conditioning was required.
- The highly sophisticated technology required for the manufacturing of IC chips.

Fourth Generation Computers (1975-Present)

The fourth generation computers started with the invention of Microprocessor. The Microprocessor contains thousands of ICs. Ted Hoff produced the first microprocessor in 1971 for Intel. It was known as Intel 4004. The technology of integrated circuits improved rapidly. The LSI (Large Scale Integration) circuit and VLSI (Very Large Scale Integration) circuit was designed. It greatly reduced the size of the computer. The size of modern Microprocessors is usually one square inch. It can contain millions of electronic circuits. The examples of fourth generation computers are Apple Macintosh & IBM PC.

Advantages

- More powerful and reliable than previous generations.
- Small in size
- Fast processing power with less power consumption
- Fan for heat discharging and thus to keep cold.
- No air conditioning required.
- Totally general purpose
- Commercial production
- Less need for repair.
- Cheapest among all generations
- All types of High-level languages can be used in this type of computers

Disadvantages

- The latest technology is required for manufacturing of Microprocessors.

Fifth Generation Computers (Present & Beyond)

Scientists are working hard on the 5th generation computers with quite a few breakthroughs. It is based on the technique of Artificial Intelligence (AI). Computers can understand spoken words & imitate human reasoning. Can respond to its surroundings using different types of sensors. Scientists are constantly working to increase the processing power of computers. They are trying to create a computer with real IQ with the help of advanced programming and technologies. IBM Watson computer is one example that outsmarts Harvard University Students. The advancement in modern technologies will revolutionize the computer in future.

8.4 TYPES OF COMPUTERS

The computers are classified into various types depends on their purpose, cooperation, and size. In general, computers are classified into major categories based on.

(a) According to the purpose of the computer.

- (b) According to the operation of the computer.
- (c) According to the size of the computer.

8.4.1 Classification as per purpose of the computer

1. General purpose computers.
2. Special purpose computer.

1. General Purpose Computers:

These computers are theoretically used for any type of applications. These computers can be used in solving a business Problem and also used to solve a mathematical equation with same accuracy and consistency. Most of the computer now are general purpose digital computers. All the P.C's, which have become a household affair.

2. Special Purpose Computers :

These digital computers are designed, made and used for any specific job. These are usually used for those purposes which are critical and need great accuracy and response to Satellite launching, weather forecasting etc.

8.4.2 According to the operation of the computer.

According to the operational principle of computers, they are categorized as analog, digital and hybrid computers.

- 1. Analog Computers:** These are almost extinct today. These are different from a digital computer because an analog computer can perform several mathematical operations simultaneously. It uses continuous variables for mathematical operations and utilizes mechanical or electrical energy.
- 2. Digital Computers:** They use digital circuits and are designed to operate on two states, namely bits 0 and 1. They are analogous to states ON and OFF. Data on these computers is represented as a series of 0s and 1s. Digital computers are suitable for complex computation and have higher processing speeds. They are programmable. Digital computers are either general purpose computers or special purpose ones.

Specific purpose computers, as their name suggests, are designed for specific types of data processing while general purpose computers are meant for general use.

- 3. Hybrid Computers:** These computers are a combination of both digital and analog computers. In this type of computers, the digital segments perform process control by conversion of analog signals to digital ones.

8.4.3 According to the size of the computer.

According to the sizes of the computers, the computers are classified as follows.

- 1. Supercomputers:** A supercomputer is a computer that performs at or near the currently highest operational rate for computers. Traditionally, supercomputers have been used for scientific and engineering applications that must handle very large databases or do a great amount of computation (or both). Although advances like multi-core processors and GPGPUs (general-purpose graphics processing units) have enabled powerful machines for personal use (*see: desktop supercomputer, GPU supercomputer*), by definition, a supercomputer is exceptional in terms of performance. At any given time, there are a few well-publicized supercomputers that operate at extremely high speeds relative to all other computers. The term is also sometimes applied to far slower (but still impressively fast) computers. The largest, most powerful supercomputers are really multiple computers that perform parallel processing. In general, there are two parallel processing approaches, symmetric multiprocessing (SMP) and massively parallel processing (MPP).
- 2. Servers:** They are computers designed to provide services to client machines in a computer network. They have larger storage capacities and powerful processors. Running on them are programs that serve client requests and allocate resources like memory and time to client machines. Usually, they are very large in size, as they have large processors and many hard drives. They are designed to be fail-safe and resistant to crash.

- 3. Mainframe Computers:** A **mainframe** is a large, expensive, powerful computer that can handle hundreds or thousands of connected users simultaneously. Mainframes store huge amounts of data, instructions, and information. Most major corporations use mainframes for business activities. With mainframes, large businesses are able to bill millions of customers, prepare payroll for thousands of employees, and manage thousands of items in inventory. One study reported that mainframes process more than 83 percent of transactions around the world. Servers and other mainframes can access data and information from a mainframe. People also can access programs on the mainframe using terminals or personal computers.
- 4. Wearable Computers:** A record-setting step in the evolution of computers was the creation of wearable computers. These computers can be worn on the body and are often used in the study of behavior modeling and human health. Military and health professionals have incorporated wearable computers into their daily routine, as a part of such studies. When the users' hands and sensory organs are engaged in other activities, wearable computers are of great help in tracking human actions. Wearable computers do not have to be turned on and off and remain in operation without user intervention
- 5. Minicomputers:** A minicomputer is a type of computer that possesses most of the features and capabilities of a large computer but is smaller in physical size. A minicomputer fills the space between the mainframe and microcomputer and is smaller than the former but larger than the latter. Minicomputers are mainly used as small or mid-range servers operating business and scientific applications. However, the use of the term minicomputer has diminished and has merged with servers. A minicomputer may also be called a mid-range computer. Minicomputers emerged in the mid-1960s and were first developed by IBM Corporation. They were primarily designed for business applications and services that require the performance and efficiency of mainframe computers. Minicomputers are generally used as mid-range servers, where they can operate mid-sized software applications and support numerous users simultaneously. Minicomputers may contain one or more processors, support multiprocessing, and tasking, and are generally resilient to high workloads.

Although they are smaller than mainframe or supercomputers, minicomputers are more powerful than personal computers and workstations.

- 6. Microcomputers:** A microcomputer is a computer with a central processing unit (CPU) as a microprocessor. Designed for individual use, a microcomputer is smaller than a mainframe or a minicomputer. The term microcomputer is not as commonly used as it was during the 1970s-1980s. We now refer to microcomputers as, simply, computers, or personal computers (PC). A microcomputer's CPU includes random access memory (RAM), read-only memory (ROM) memory, input/output (I/O) ports, interconnecting wires and a motherboard. In 1970, Hewlett-Packard (HP) manufactured and released a microcomputer as a calculator with varying levels of programmability, although the Data point 2200 by Computer Terminal Corporation (CTC) is credited as the first microcomputer. Intel's x86 processor family can also be traced back to CTC's release.
- 7. Desktops:** A desktop is intended to be used in a single location. The spare parts of a desktop computer are readily available at relatively lower costs. Power consumption is not as critical as that in laptops. Desktops are widely popular for daily use in the workplace and households.
- 8. Laptops:** Similar in operation to desktops, laptop computers are miniaturized and optimized for mobile use. Laptops run on a single battery or an external adapter that charges the computer batteries. They are enabled with an inbuilt keyboard, touchpad acting as a mouse and a liquid crystal display. Their portability and capacity to operate on battery power have proven to be of great help to mobile users.
- 9. Notebooks:** They fall into the category of laptops, but are inexpensive and relatively smaller in size. They had a smaller feature set and lesser capacities in comparison to regular laptops, at the time they came into the market. But with passing time, notebooks too began featuring almost everything that notebooks had. By the end of 2008, notebooks had begun to overtake notebooks in terms of market share and sales.
- 10. Personal Digital Assistants (PDAs):** It is a handheld computer and popularly known as a palmtop. It has a touchscreen and a memory card for storage of data. PDAs can

also be used as portable audio players, web browsers and smartphones. Most of them can access the Internet by means of Bluetooth or Wi-Fi communication.

11. Tablet Computers: Tablets are mobile computers that are very handy to use. They use the touchscreen technology. Tablets come with an onscreen keyboard or use a stylus or a digital pen. Apple's iPad redefined the class of tablet computers.

8.5 GENERAL AWARENESS OF COMPUTER HARDWARE

8.5.1 CPU

The processor also called the central processing unit (CPU) interprets and carries out the basic instructions that operate a computer. The processor significantly impacts overall computing power and manages most of a computer's operations. On a personal computer, all functions of the processor usually are on a single chip. Some computer and chip manufacturers use the term microprocessor to refer to a personal computer processor chip. Most processor chip manufacturers now offer multi-core processors. A multi-core processor is a chip with two or more separate processors. Two common multi-core processors used today are dual-core and quad-core. A dual-core processor is a chip that contains two separate processors. Similarly, a quad-core processor is a chip with four separate processors. Each processor on a multi-core chip generally runs at a slower clock speed than a single-core processor, but multi-core chips typically increase overall performance. For example, although a dual-core processor does not double the processing speed of a single-core processor, it can approach those speeds. Multicore processors also are energy efficient, requiring lower levels of power consumption and emitting less heat in the system unit. Processors contain a control unit and an arithmetic logic unit (ALU). These two components work together to perform processing operations.

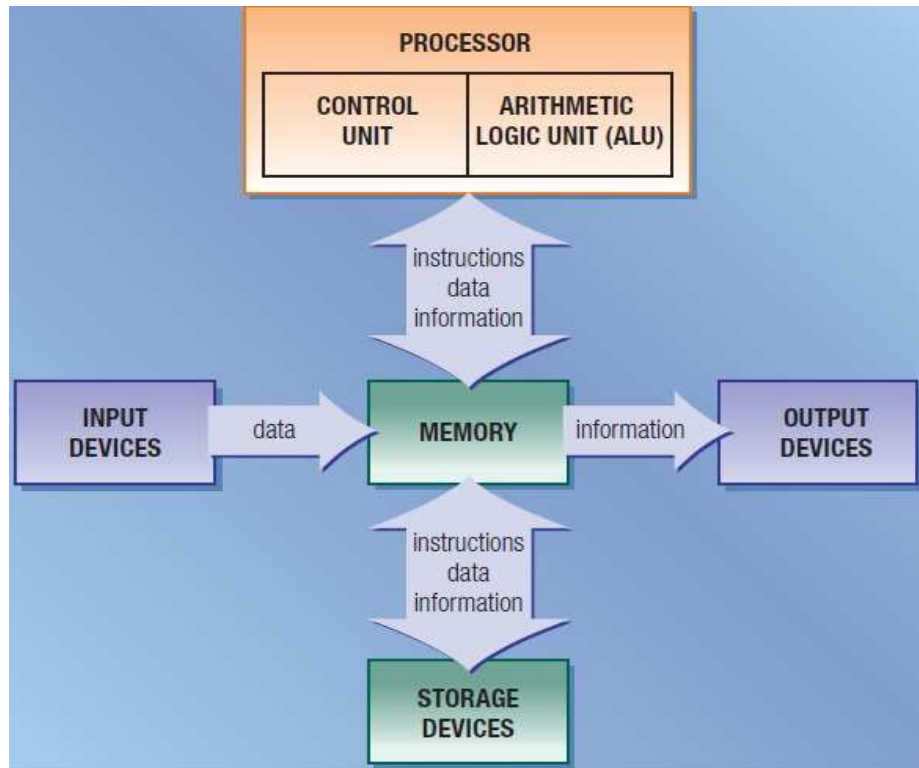


Figure 2.1: Central Processing Unit

The Control Unit

The **control unit** is the component of the processor that directs and coordinates most of the operations in the computer. The control unit has a role much like a traffic cop: it interprets each instruction issued by a program and then initiates the appropriate action to carry out the instruction. Types of internal components that the control unit directs include the arithmetic/logic unit and buses.

The Arithmetic Logic Unit

The **arithmetic logic unit (ALU)**, another component of the processor, performs arithmetic, comparison, and other operations. Arithmetic operations include basic calculations such as addition, subtraction, multiplication, and division. Comparison operations involve comparing one data item with another to determine whether the first item is greater than, equal to, or less than the other item. Depending on the result of the comparison, different actions may occur.

Machine Cycle

For every instruction, a processor repeats a set of four basic operations, which comprise a machine cycle : (1) fetching, (2) decoding, (3) executing, and, if necessary, (4) storing. Fetching is the process of obtaining a program instruction or data item from memory. The term decoding refers to the process of translating the instruction into signals the computer can execute. Executing is the process of carrying out the commands. Storing, in this context, means writing the result to memory (not to storage medium).

8.5.2 Peripheral devices

A **peripheral** is a device that connects to the system unit and is controlled by the processor in the computer. Examples of peripherals are modems, disk drives, printers, scanners, and keyboards. Peripheral devices are connected to the computer externally.

These devices are used for performing some specific functions.

Peripheral devices are as follows:

- I. Input Devices
- II. Output Devices
- III. Other Peripherals

1. INPUT DEVICES

Input devices accept data and instructions from the user. Following are the examples of various input devices, which are connected to the computer for this purpose.

- I. Keyboard
- II. Mouse
- III. Light Pen
- IV. Optical/magnetic Scanner
- V. Touch Screen
- VI. Microphone for voice as input
- VII. Track Ball

Keyboard

A keyboard is the most common input device. Several kinds of keyboards are available, but they resemble each other with minor variations. The keyboard in most common use is the QWERTY board. Generally, the standard keyboard has 104 keys. In these keyboards, the cursor control keys are duplicated to allow easier use of the numeric pad.

Mouse

A mouse is an electro-mechanical, hand-held device. It is used as a pointer. It can perform functions like selecting menu commands, moving icons and resizing windows, starting programs, and choosing options. The most common mouse uses an internal, magnetically coated ball, to detect the movement of the mouse across a flat surface, usually a desktop. Now a days Optical or laser mouse is used to detect the movement. All windows based applications today are designed to work with a mouse. A mouse is used to replace hard-to-remember key combinations with easier "Point and Click" actions. However, it cannot substitute all keyboard operations. It can be alternative for commands based operations.

Light pen

An input device that utilizes a light-sensitive detector to select objects on a display screen. A light pen is similar to a mouse except that with a light pen you can move the pointer and select objects on the display screen by directly pointing to the objects with the pen.

Optical Scanner

These devices are used for automatic data collection. The devices in this category completely eliminate manual input of data. For example, the bar-code reader is actually just a special type of image scanner. An image scanner translates printed images into an electronic format that can be stored in a computer's memory, and with the right kind of software, one can alter a stored image. Another example of the scanner is optical character recognition (OCR) device, used by banks to convert the scanned image of a typed or printed page into text that can be edited on the computer.

Touch Screen

Touch panel displays and pads are now being offered as alternatives to the keyboard. Here the input can be given through the computer screen, that accepts the input through monitor; users touch electronic buttons displayed on the screen or they may use a light pen.

Microphone

The microphone is an input device, which takes voice as input. The voice communication is more error-prone than information through the keyboard. There are two types of microphones available.

1. Desktop Microphone
2. Hand-held Microphone

Track Ball

Trackball, a pointing device, is a mouse lying on its back. To move the pointer, you rotate the ball with your thumb, your fingers, or the palm of your hand. There are usually one to three buttons next to the ball, which you use just like mouse buttons. The advantage of trackballs over a mouse is that the trackball is stationary so it does not require much space to use it. In addition, you can place a trackball on any type of surface, including your lap. For both these reasons, trackballs are popular pointing devices for portable computers.

2. OUTPUT DEVICES

Output devices return processed data that is information, back to the user. Some of the commonly used output devices are:

1. Monitor (Visual Display Unit)
2. Printers
3. Plotter
4. Speakers

1. Monitor

Out of all the output devices, the monitor is perhaps the most important output device because people interact with this device most intensively than others. Computer information is displayed,

visually with a video adapter card and monitor. Information processed within the CPU, that needs to be visually displayed, is sent to video adapter. The video adapter converts information from the format used, in the same manner as a television displays information sent to it by a cable service.

Two basic types of monitors are used with microcomputers, which are as follows:

1. CRT
2. LCD

Cathode Ray Tube (CRT):

CRT or Cathode Ray Tube Monitor is the typical monitor that you see on a desktop computer. It looks a lot like a television screen and works the same way. This type uses a large vacuum tube, called cathode ray tube (CRT).

Liquid Crystal Displays (LCD):

This type of monitors is also known as flat panel monitor. Most of these employ liquid crystal displays (LCDs) to render images. These days LCD monitor are very popular. When people talk about the capabilities of various monitors, one critical statistic is the resolution of the monitor. Most monitors have a resolution of at least 800 x 600 pixels. High-end monitors can have resolutions of 1024 x 768 pixels or even 1280 x 1024 pixels. Thus monitors are available either in low resolution or in high resolution.

2. Printer

After a document is created on the computer, it can be sent to a printer for a hard copy (printout). Some printers offer special features such as colored and large page formats. Some of the most commonly used printers are:

1. Laser Printer
2. Ink Jet Printer
3. Dot Matrix Printer
4. Line Printer

1. Laser Printer: A laser printer produces a high-quality print that one normally finds in publishing. It is extremely fast and quiet. Moreover, the operation of a laser printer is easy with automatic paper loading and no smudging or messing up of ink ribbons. The fastest laser printer can print up to 200 pages per minute in monochrome (black and white) and up to 100 pages per minute in color.

2. Ink-Jet Printer: An ink-jet printer creates an image directly on paper by spraying ink through as many as 64 tiny nozzles. Although the image it produces is not generally quite as sharp as the output of a laser printer, the quality of ink-jet images is still high. In general, the ink-jet printer offers an excellent middle ground between dot matrix and laser printer. Like laser printer, an ink-jet printer is quiet and convenient, but not particularly fast. Typically, an ink-jet printer is more expensive than a dot-matrix printer but costs only half as much as a laser printer.

3. Dot Matrix Printer: The dot matrix printer was very popular at one point in time. It is a very versatile and inexpensive output device. In dot matrix printer the print head physically "hits" the paper with the ribbon and produces text (or images) by combinations of dots; hence the name dot matrix printer. Its speed is measured in characters per second (CPS). Although it is less expensive, it is louder, slower and produces lower print quality.

4. Line Printer: A line printer is generally used with large computer systems to produce text-based data processing reports. Line printers are high-speed printers with speeds ranging anywhere from 100 to about 3800 lines per minute. In the past, print quality on line printers was not high. Developments in technology are improving the print quality on line printers. These are in the cost range of lacs of Rupees.

3. Plotter

A plotter is a special kind of output device that, like a printer, produces images on paper, but does so in a different way. Plotters are designed to produce large drawings or images, such as construction plans for buildings or blueprints for mechanical objects. A plotter can be connected to the port normally used by a printer. An array of different colored pens in a clip rack and a robotic arm is part of the plotter. The instructions that a plotter receives from a computer consist of a color, and beginning and ending coordinates for a line. With that information, the plotter picks up the appropriate pen through its arm, positions it at the beginning coordinates drops the

pen down to the surface of the paper and draws to the ending coordinates. Plotters draw curves by creating a sequence of very short straight lines. Plotters usually come in two designs:

1. Flat Bed: Plotters of small size to be kept on the table with restriction of paper size.
2. Drum: These plotters are of big size using rolls of paper of unlimited length.

4. Speaker

Speakers are another type of output device, which allows you to listen to a voice like music, and conversation with people.

8.6 LET US SUM UP

A computer is used only when it interacts with the external devices such as keyboard, mouse, monitor, speaker, which are called as peripheral devices. In this lesson, we have discussed in details about major peripheral devices of a computer system. You have learned that only hardware does not do anything on its own without the help of software. Hence we have explained in details about the software, its types and the computer languages used to prepare the software.

8.7 CHECK YOUR PROGRESS

1. Choose the correct answer.

(a) The task of performing arithmetic and logical operations is called:

- (i) ALU (ii) editing (iii) storage (iv) Output

(b) The ALU and CU jointly are known as

- (i) RAM (ii) ROM (iii) CPU (iv) None of above

(c) The process of producing results from the data for getting useful information is called:

- (i) output (ii) input (iii) processing (iv) storage

2. Write True or False for the following:

(a) The mouse is an output device.

(b) OCR stands for Optical Content Reader.

(c) LCD Monitor is used in a notebook computer.

(d) The speed of DOT Matrix Printer is measured in Characters per Second.

(e) Plotters are used to produce high-quality drawings and images, such as construction plans for buildings or blueprints for mechanical objects.

Ans: 1. (a) i (b) iii (c) i

2. (a) False (b) False (c) True (d) True (e) True

8.8 FURTHER READINGS

- Computer fundamentals by Anita Geol, Pearson Education
- Fundamentals of Computers by V, Rajaraman
- Fundamentals of Computers by P K Sinha

8.9 MODEL QUESTIONS

Q1. What is a computer? Draw a block diagram to illustrate the basic organization of a computer system and explain the functions of various units.

Q2. What is an input device? Briefly describe various important input devices.

Q3. What is a computer? Explain different types of computers in detail.

Q4. Explain the different generations of computers in detail.

UNIT 9: INTRODUCTION TO COMPUTERS AND COMPUTING

CONTENTS:

- 9.1 Learning Objectives
- 9.2 Introduction
- 9.3 Basic Structure and Function of Computer
 - 9.3.1 Advantages and Disadvantages of Using Computers
 - 9.3.2 Advantage of using computers
 - 9.3.3 Disadvantages of using computers
 - 9.3.4 Functional unit of computer
- 9.4 Memory
 - 9.4.1 Bytes and Addressable Memory
 - 9.4.2 Memory Sizes
 - 9.4.3 Types of Memory
- 9.5 Input/ Output Devices
- 9.6 Computer Languages
- 9.7 Operating System
- 9.8 Introduction to Windows
- 9.9 Data Processing
- 9.10 Principles of Programming
- 9.11 Algorithms and Flowcharts
- 9.12 Let Us Sum Up
- 9.13 Answer to Check Your Progress
- 9.14 Further Readings
- 9.15 Model Questions

9.1 LEARNING OBJECTIVES

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

- Understand the capabilities and limitations of computers and technology.
- Identify the components of the computer
- Define different types of memory
- Describe input/ output devices
- Define operating system
- Define data processing
- Define algorithms and flowcharts

9.2 INTRODUCTION

Computers are everywhere: at work, at school, and at home. Mobile devices, such as many cell phones, often are classified as computers. Computers are a primary means of local and global communication for billions of people. Employees correspond with clients, students with classmates and teachers, and family with friends and other family members. Through computers, society has instant access to information from around the globe. Local and national news, weather reports, sports scores, airline schedules, telephone directories, maps and directions, job listings, credit reports, and countless forms of educational material always are accessible. From the computer, you can make a telephone call, meet new friends, share photos and videos, share opinions, shop, book flights, file taxes, or take a course. In the workplace, employees use computers to create correspondence such as e-mail messages, memos, and letters; calculate payroll; track inventory; and generate invoices. Some applications such as automotive design and weather forecasting use computers to perform complex mathematical calculations. At school, teachers use computers to assist with classroom instruction. Students use computers to complete assignments and research. Instead of attending class on campus, some students take entire classes directly from their computer. Computers can do such a wide variety of things because they can be programmed. This means that computers are not designed to do just one job, but to do any job that their programs tell them to do. A program is a set of instructions that a computer follows to perform a task.

9.3 BASIC STRUCTURE OF COMPUTER

A computer is an electronic device, operating under the control of instructions stored in its own memory, that can accept data, process the data according to specified rules, produce results, and stores the results for future use. List of instructions are called programs & internal storage is called computer memory. Computers process data into information. Data is a collection of unprocessed items, which can include text, numbers, images, audio, and video. Information conveys meaning and is useful to people.

9.3.1 Advantages and Disadvantages of Using Computers

Society has reaped many benefits from using computers. Both business and home users can make well-informed decisions because they have instant access to information from anywhere in the world. A user is anyone who communicates with a computer or utilizes the information it generates. Students, another type of user, have more tools to assist them in the learning process.

9.3.2 Advantage of using computers

The benefits of computers are possible because computers have the advantages of speed, reliability, consistency, storage, and communications.

1. **Speed:** When data, instructions, and information flow along electronic circuits in a computer, they travel at incredibly fast speeds. Many computers process billions or trillions of operations in a single second.
2. **Reliability:** The electronic components in modern computers are dependable and reliable because they rarely break or fail.
3. **Consistency:** Given the same input and processes, a computer will produce the same results consistently. Computers generate error-free results, provided the input is correct and the instructions work.
4. **Storage:** Computers store enormous amounts of data and make this data available for processing anytime it is needed.

5. **Communications:** Most computers today can communicate with other computers, often wirelessly. Computers allow users to communicate with one another.

9.3.3 Disadvantages of using computers

Some disadvantages of computers relate to the violation of privacy, public safety, and the impact on the labor force, health risks, and the impact on the environment.

1. **Violation of Privacy:** In many instances, where personal and confidential records were not protected properly, individuals have found their privacy violated and identities were stolen.
2. **Public Safety:** Adults, teens, and children around the world are using computers to share publicly their photos, videos, journals, music, and other personal information. Some of these unsuspecting, innocent computer users have fallen victim to crimes committed by dangerous strangers.
3. **Impact on Labor Force:** Although computers have improved productivity and created an entire industry with hundreds of thousands of new jobs, the skills of millions of employees have been replaced by computers. Thus, it is crucial that workers keep their education up-to-date. A separate impact on the labor force is that some companies are outsourcing jobs to foreign countries instead of keeping their homeland labor force employed.
4. **Health Risks:** Prolonged or improper computer use can lead to health injuries or disorders. Computer users can protect themselves from health risks through proper workplace design, good posture while at the computer, and appropriately spaced work breaks. Another health risk, called computer addiction, occurs when someone becomes obsessed with using the computer.
5. **Impact on Environment:** Computer manufacturing processes and computer waste are depleting natural resources and polluting the environment. Strategies that can help protect the environment include recycling, regulating manufacturing processes, extending the life of computers, and immediately donating replaced computers.

9.3.4 Functional unit of the computer

A computer as shown in performs basically five major computer operations or functions irrespective of their size and make. These are

- 1) It accepts data or instructions by way of input.
- 2) It stores data.
- 3) It can process data as required by the user.
- 4) It gives results in the form of output.
- 5) It controls all operations inside a computer.

A computer consists of five functionally independent main parts input, memory, arithmetic logic unit (ALU), and output and control unit.

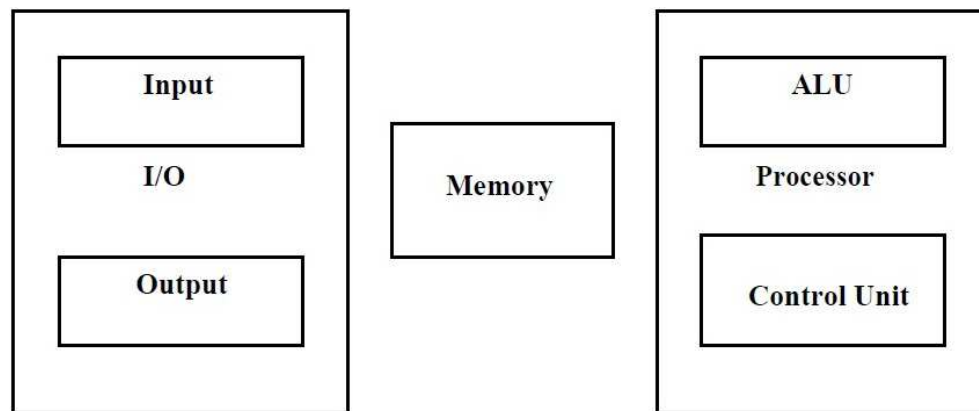


Figure 3.1: Functional Units of Computer

1. Input: This is the process of entering data and programs into the computer system. You should know that a computer is an electronic machine like any other machine which takes as inputs raw data and performs some processing giving out processed data. Therefore, the input unit takes data from us to the computer in an organized manner for processing. The source program high-level languages program/coded information/simple data is fed to a computer through input devices keyboard is a most common type. Whenever a key is pressed, one corresponding word or number is translated into its equivalent binary code over a cable & fed either to memory or processor. Joysticks, trackballs, mouse, scanners etc are other input devices.

2. Memory: The process of saving data and instructions permanently is known as storage. Data has to be fed into the system before the actual processing starts. It is because the processing speed of Central Processing Unit (CPU) is so fast that the data has to be provided to CPU with the same speed. Therefore the data is first stored in the storage unit for faster access and processing. This storage unit or the primary storage of the computer system is designed to do the above functionality. It provides space for storing data and instructions.

The storage unit performs the following major functions:

1. All data and instructions are stored here before and after processing.
2. Intermediate results of processing are also stored here.

3. Processing: The task of performing operations like arithmetic and logical operations is called processing. The Central Processing Unit (CPU) takes data and instructions from the storage unit and makes all sorts of calculations based on the instructions given and the type of data provided. It is then sent back to the storage unit. There are mainly three units used for processing: ALU, CU, and Memory unit.

Arithmetic logic unit (ALU):-

(a) Logical Unit: After you enter data through the input device it is stored in the primary storage unit. The actual processing of the data and instruction are performed by Arithmetic Logical Unit. The major operations performed by the ALU are addition, subtraction, multiplication, division, logic, and comparison. Data is transferred to ALU from the storage unit when required. After processing the output is returned back to the storage unit for further processing or getting stored. Most of the computer operators are executed in ALU of the processor like addition, subtraction, division, multiplication, etc. the operands are brought into the ALU from memory and stored in high-speed storage elements called to register. Then according to the instructions, the operation is performed in the required sequence.

(b) Control Unit: The next component of the computer is the Control Unit, which acts as the supervisor seeing that things are done in proper fashion. Control Unit is

responsible for coordinating various operations using time signal. The control unit determines the sequence in which computer programs and instructions are executed. Things like processing of programs stored in the main memory, interpretation of the instructions and issuing of signals for other units of the computer to execute them. It also acts as a switchboard operator when several users access the computer simultaneously. Thereby it coordinates the activities of computer's peripheral equipment as they perform the input and output. It effectively is the nerve center that sends signals to other units and senses their states. The actual timing signals that govern the transfer of data between input unit, processor, memory, and output unit are generated by the control unit.

The control and the ALU are many times faster than other devices connected to a computer system. This enables a single processor to control a number of external devices such as keyboards, displays, magnetic and optical disks, sensors and other mechanical controllers.

4. Output: This is the process of producing results from the data for getting useful information. Similarly, the output produced by the computer after processing must also be kept somewhere inside the computer before being given to you in human readable form. Again the output is also stored inside the computer for further processing. These actually are the counterparts of the input unit. Its basic function is to send the processed results to the outside world.

Examples: Printer, speakers, monitor etc.

9.4 MEMORY

A memory is just like a human brain. It is used to store data and instructions. Computer memory is the storage space in the computer, where data is to be processed and instructions required for processing are stored. The memory is divided into a large number of small parts called cells. Each location or cell has a unique address, which varies from zero to memory size minus one. Memory consists of electronic components that store instructions waiting to be executed by the processor, data needed by those instructions, and the results of processed data

(information). Memory usually consists of one or more chips on the motherboard or some other circuit board in the computer.

Memory stores three basic categories of items:

- (1) The operating system and other system software that control or maintain the computer and its devices;
- (2) Application programs that carry out a specific task such as word processing.
- (3) The data being processed by the application programs and resulting information. This role of memory to store both data and programs is known as the stored program concept.

9.4.1 Bytes and Addressable Memory

A byte (character) is the basic storage unit in memory. When application program instructions and data are transferred to memory from storage devices, the instructions and data exist as bytes. Each byte resides temporarily in a location in memory that has an address. An address simply is a unique number that identifies the location of the byte in memory. To access data or instructions in memory, the computer references the addresses that contain bytes of data.

9.4.2 Memory Sizes

Manufacturers state the size of memory chips and storage devices in terms of the number of bytes the chip or device has available for storage. Recall that storage devices hold data, instructions, and information for future use, while most memory holds these items temporarily. A kilobyte (KB or K) is equal to exactly 1,024 bytes. To simplify memory and storage definitions, computer users often round a kilobyte down to 1,000 bytes. For example, if a memory chip can store 100 KB, it can hold approximately 100,000 bytes (characters). A megabyte (MB) is equal to approximately 1 million bytes. A gigabyte (GB) equals approximately 1 billion bytes. A terabyte (TB) is equal to approximately 1 trillion bytes.

9.4.3 Types of Memory

Memory is primarily of three types –

- Cache Memory

- Primary Memory/Main Memory
- Secondary Memory

Cache Memory

Cache memory is a very high-speed semiconductor memory which can speed up the CPU. It acts as a buffer between the CPU and the main memory. It is used to hold those parts of data and program which are most frequently used by the CPU. The parts of data and programs are transferred from the disk to cache memory by the operating system, from where the CPU can access them. Memory cache helps speed the processes of the computer because it stores frequently used instructions and data. Most personal computers today have at least two types of memory cache: L1 cache and L2 cache.

L1 cache is built directly in the processor chip. L1 cache usually has a very small capacity, ranging from 8 KB to 128 KB.

L2 cache is slightly slower than L1 cache but has a much larger capacity, ranging from 64 KB to 16 MB. Current processors include advanced transfer cache, a type of L2 cache built directly on the processor chip. Processors that use advanced transfer cache perform at much faster rates than those that do not use it. Personal computers today typically have from 512 KB to 8 MB of advanced transfer cache. Cache speeds up processing time because it stores frequently used instructions and data. When the processor needs an instruction or data, it searches memory in this order: L1 cache, then L2 cache, then RAM with a greater delay in processing for each level of memory it must search. If the instruction or data is not found in memory, then it must search a slower speed storage medium such as a hard disk, CD, or DVD.

Advantages

The advantages of cache memory are as follows –

- Cache memory is faster than main memory.
- It consumes less access time as compared to main memory.
- It stores the program that can be executed within a short period of time.

- It stores data for temporary use.

Disadvantages

The disadvantages of cache memory are as follows –

- The cache memory has limited capacity.
- It is very expensive.

Primary Memory (Main Memory)

Primary memory holds only those data and instructions on which the computer is currently working. It has a limited capacity and data is lost when power is switched off. It is generally made up of the semiconductor device. These memories are not as fast as registers. The data and instruction required to be processed reside in the main memory. It is divided into two subcategories RAM and ROM.

Characteristics of Main Memory

- These are semiconductor memories.
- It is known as the main memory.
- Usually volatile memory.
- Data is lost in case power is switched off.
- It is the working memory of the computer.
- Faster than secondary memories.
- A computer cannot run without the primary memory.

Random Access Memory

RAM (Random Access Memory) is the internal memory of the CPU for storing data, program, and program result. It is a read/write memory which stores data until the machine is working. As soon as the machine is switched off, data is erased. Access time in RAM is independent of the address, that is, each storage location inside the memory is as easy to reach as other locations and takes the same amount of time. Data in the RAM can be accessed randomly but it is very expensive. RAM is volatile, i.e. data stored in it

is lost when we switch off the computer or if there is a power failure. Hence, a backup Uninterruptible Power System (UPS) is often used with computers. RAM is small, both in terms of its physical size and in the amount of data it can hold.



Figure 3.2: RAM

RAM is of two types –

- Static RAM (SRAM)
- Dynamic RAM (DRAM)

Static RAM (SRAM)

The word **static** indicates that the memory retains its contents as long as power is being supplied. However, data is lost when the power gets down due to volatile nature. SRAM chips use a matrix of 6-transistors and no capacitors. Transistors do not require power to prevent leakage, so SRAM need not be refreshed on a regular basis. There is extra space in the matrix, hence SRAM uses more chips than DRAM for the same amount of storage space, making the manufacturing costs higher. SRAM is thus used as cache memory and has very fast access.

Characteristic of Static RAM

- Long life
- No need to refresh
- Faster
- Used as cache memory

- Large size
- Expensive
- High power consumption

Dynamic RAM (DRAM)

DRAM, unlike SRAM, must be continually refreshed in order to maintain the data. This is done by placing the memory on a refresh circuit that rewrites the data several hundred times per second. DRAM is used for most system memory as it is cheap and small. All DRAMs are made up of memory cells, which are composed of one capacitor and one transistor.

Characteristics of Dynamic RAM

- Short data lifetime
- Needs to be refreshed continuously
- Slower as compared to SRAM
- Used as RAM
- Smaller in size
- Less expensive
- Less power consumption

ROM

ROM stands for Read Only Memory. The memory from which we can only read but cannot write on it. This type of memory is non-volatile. The information is stored permanently in such memories during manufacture. A ROM stores such instructions that are required to start a computer. This operation is referred to as bootstrap. ROM chips are not only used in the computer but also in other electronic items like a washing machine and microwave oven.



Figure 3.3: ROM

MROM (Masked ROM)

The very first ROMs were hard-wired devices that contained a pre-programmed set of data or instructions. These kinds of ROMs are known as masked ROMs, which are inexpensive.

PROM (Programmable Read Only Memory)

PROM is a read-only memory that can be modified only once by a user. The user buys a blank PROM and enters the desired contents using a PROM program. Inside the PROM chip, there are small fuses which are burnt open during programming. It can be programmed only once and is not erasable.



Figure 3.4: Programmable Read-Only Memory (PROM)

EPROM (Erasable and Programmable Read Only Memory)

EPROM can be erased by exposing it to ultra-violet light for the duration of up to 40 minutes. Usually, an EPROM eraser achieves this function. During programming, an electrical charge is trapped in an insulated gate region. The charge is retained for more than 10 years because the charge has no leakage path. For erasing this charge, ultra-violet light is passed through a quartz crystal window (lid). This exposure to ultra-violet light dissipates the charge. During normal use, the quartz lid is sealed with a sticker.



Figure 3.5: Erasable Programmable Only Memory (EPROM)

EEPROM (Electrically Erasable and Programmable Read Only Memory)

EEPROM is programmed and erased electrically. It can be erased and reprogrammed about ten thousand times. Both erasing and programming take about 4 to 10 ms (millisecond). In EEPROM, any location can be selectively erased and programmed. EEPROMs can be erased one byte at a time, rather than erasing the entire chip. Hence, the process of reprogramming is flexible but slow.

Advantages of ROM

The advantages of ROM are as follows –

- Non-volatile in nature
- Cannot be accidentally changed
- Cheaper than RAMs
- Easy to test
- More reliable than RAMs
- Static and do not require refreshing
- Contents are always known and can be verified

SECONDARY MEMORY

This type of memory is also known as external memory or non-volatile. It is slower than the main memory. These are used for storing data/information permanently. CPU directly does not access these memories; instead, they are accessed via input-output routines. The contents of

secondary memories are first transferred to the main memory, and then the CPU can access it. For example, disk, CD-ROM, DVD, etc.

Characteristics of Secondary Memory

- These are magnetic and optical memories.
- It is known as the backup memory.
- It is a non-volatile memory.
- Data is permanently stored even if power is switched off.
- It is used for storage of data in a computer.
- The computer may run without the secondary memory.
- Slower than primary memories.

Hard Disk Drive

The hard disk drive is made up of a series of circular disks called platters arranged one over the other almost $\frac{1}{2}$ inches apart around a spindle. Disks are made of non-magnetic material like aluminum alloy and coated with 10-20 nm of magnetic material. Standard diameter of these disks is 14 inches and they rotate with speeds varying from 4200 rpm (rotations per minute) for personal computers to 15000 rpm for servers. Data is stored by magnetizing or demagnetizing the magnetic coating. A magnetic reader arm is used to read data from and write data to the disks. A typical modern HDD has capacity in terabytes (TB).

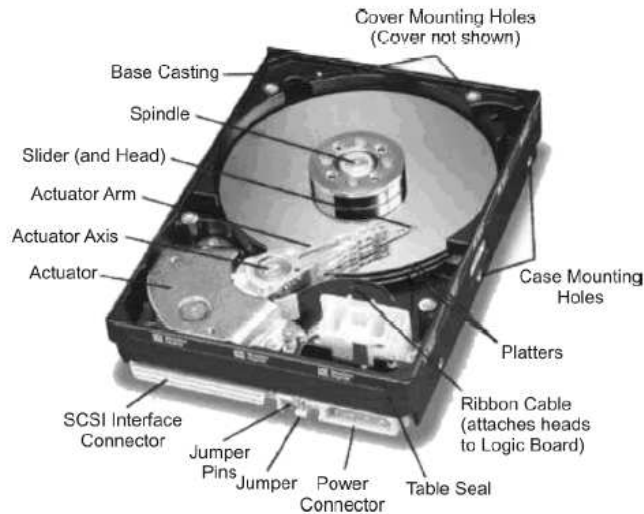


Figure 3.6: Hard Disks

CD Drive

CD stands for Compact Disk. CDs are circular disks that use optical rays, usually lasers, to read and write data. They are very cheap as you can get 700 MB of storage space for less than a dollar. CDs are inserted in CD drives built into CPU cabinet. They are portable as you can reject the drive, remove the CD and carry it with you. There are three types of CDs –

CD-ROM (Compact Disk – Read Only Memory) – The data on these CDs are recorded by the manufacturer. Proprietary Software, audio or video are released on CD-ROMs.

CD-R (Compact Disk – Recordable) – Data can be written by the user once on the CD-R. It cannot be deleted or modified later.

CD-RW (Compact Disk – Rewritable) – Data can be written and deleted on these optical disks again and again.

DVD Drive

DVD stands for **Digital Video Display**. DVD are optical devices that can store 15 times the data held by CDs. They are usually used to store rich multimedia files that need high storage capacity. DVDs also come in three varieties – read-only, recordable and rewritable.

Pen Drive

Pen drive is a portable memory device that uses solid-state memory rather than magnetic fields or lasers to record data. It uses a technology similar to RAM, except that it is nonvolatile. It is also called USB drive, key drive or flash memory.

Blu Ray Disk

Blu Ray Disk (BD) is an optical storage media used to store high definition (HD) video and other multimedia files. BD uses shorter wavelength laser as compared to CD/DVD. This enables writing arm to focus more tightly on the disk and hence pack in more data. BDs can store up to 128 GB data.

9.5 INPUT/OUTPUT DEVICES

Input is any data and instructions entered into the memory of a computer. An input device is any hardware component that allows users to enter data and instructions into a computer. The following pages discuss a variety of input devices.

Keyboard and Pointing Devices

Two of the more widely used input devices are the keyboard and the mouse. Most computers include a keyboard or keyboarding capabilities. The mouse is a pointing device because it allows a user to control a pointer on the screen. In a graphical user interface, a pointer is a small symbol on the screen whose location and shape change as a user moves a pointing device. A pointing device can select text, graphics, and other objects; and click buttons, icons, links, and menu commands. The following pages discuss the keyboard and a variety of pointing devices.

The Keyboard

Many people use a keyboard as one of their input devices. A **keyboard** is an input device that contains keys users press to enter data and instructions into a computer. All computer keyboards have a typing area that includes the letters of the alphabet, numbers, punctuation marks, and other basic keys. Many desktop computer keyboards also have a numeric keypad on the right side of the keyboard. Most of today's desktop computer keyboards are enhanced keyboards. An enhanced keyboard has 12 or more function keys

along the top and a set of arrow and additional keys between the typing area and the numeric keypad. Function keys are special keys programmed to issue commands to a computer. Keyboards with media control buttons allow you to control your media player program, access the computer's CD/DVD drive, and adjust speaker volume. Internet controls allow you to open an e-mail program, start a Web browser, and search the Internet. Some keyboards include buttons and other features specifically for users that enjoy playing games on the computer. Desktop computer keyboards often attach via a cable to a serial port, a keyboard port, or a USB port on the system unit. Some keyboards, however, do not have any wires connecting the keyboard to the system unit. A wireless keyboard, or cordless keyboard, is a battery-powered device that transmits data using wireless technology, such as radio waves or infrared light waves. Wireless keyboards often communicate with a receiver attached to a port on the system unit.

Mouse

A **mouse** is a pointing device that fits comfortably under the palm of your hand. With a mouse, users control the movement of the pointer. As you move a mouse, the pointer on the screen also moves. Generally, you use the mouse to move the pointer on the screen to an object such as a button, a menu, an icon, a link, or text. Then, you press a mouse button to perform a certain action associated with that object. A mechanical mouse, which was the first type of mouse used with personal computers, has a rubber or metal ball on its underside. Electronic circuits in the mouse translate the movement of the mouse into signals the computer can process. A mechanical mouse is placed on a **mouse pad**, which is a rectangular rubber or foam pad that provides better traction than the top of a desk. Most computer users today have some type of optical mouse, which has no moving mechanical parts inside. Instead, an **optical mouse** uses devices that emit and sense light to detect the mouse's movement. Some use optical sensors, and others use a laser. The latter type often is referred to as a **laser mouse**, which usually is more expensive than the former. An optical mouse is more precise than a mechanical mouse and does not require cleaning as does a mechanical mouse, but it also is more expensive.

Trackball

A **trackball** is a stationary pointing device with a ball on its top or side. To move the pointer using a trackball, you rotate the ball with your thumb, fingers, or the palm of your hand. In addition to the ball, a trackball usually has one or more buttons that work just like mouse buttons.

Touchpad

A **touchpad** is a small, flat, rectangular pointing device that is sensitive to pressure and motion. To move the pointer using a touchpad, slide your fingertip across the surface of the pad. Some touchpads have one or more buttons around the edge of the pad that works like mouse buttons. On most touchpads, you also can tap the pad's surface to imitate mouse operations such as clicking. Touchpads are found most often on notebook computers.

Pointing Stick

A **pointing stick** is a pressure-sensitive pointing device shaped like a pencil eraser that is positioned between keys on a keyboard. To move the pointer using a pointing stick, you push the pointing stick with a finger. The pointer on the screen moves in the direction you push the pointing stick. By pressing buttons below the keyboard, users can click and perform other mouse-type operations with a pointing stick.

Light Pen

A **light pen** is a handheld input device that can detect the presence of light. To select objects on the screen, a user presses the light pen against the surface of the screen or points the light pen at the screen and then presses a button on the pen.

Touch Screen

A **touchscreen** is a touch-sensitive display device. Users can interact with these devices by touching areas of the screen. Because touch screens require a lot of arm movements, you do not enter large amounts of data using a touchscreen. Instead, users touch words, pictures, numbers, letters, or locations identified on the screen. With some smartphones, portable media players, and other personal mobile devices, you can touch the screen to perform tasks such as dialing telephone numbers, entering text, and making

on-screen selections. Many handheld game consoles also have touch screens. Kiosks, which are freestanding computers, have touch screens.

Joysticks and Wheels

Users running game software or flight and driving simulation software often use a joystick or wheel to control an airplane, vehicle, or player. A **joystick** is a handheld vertical lever mounted on a base. You move the lever in different directions and press buttons to control the actions of the simulated vehicle or player. A **wheel** is a steering-wheel-type input device that you turn to simulate driving a car, truck or another vehicle. Most wheels also include foot pedals for acceleration and braking actions. Joysticks and wheels typically attached via a cable to a personal computer or game console.

Scanners and Reading Devices

Some input devices save users time by capturing data directly from a source document, which is the original form of the data. Examples of source documents include time cards, order forms, invoices, paychecks, advertisements, brochures, photos, inventory tags, or any other document that contains data to be processed. Devices that can capture data directly from a source document include optical scanners, optical readers, barcode readers, RFID readers, magnetic stripe card readers, and magnetic-ink character recognition readers.

OPTICAL SCANNERS

An optical scanner, usually called a **scanner**, is a light-sensing input device that reads printed text and graphics and then translates the results into a form the computer can process. A **flatbed scanner** works in a manner similar to a copy machine except it creates a file of the document in memory instead of a paper copy (Figure 5-19). Once you scan a picture or document, you can display the scanned object on the screen, modify its appearance, store it on storage medium, print it, fax it, attach it to an e-mail message, include it in another document, or post it to a Web site or photo community for everyone to see. Many scanners include OCR (optical character recognition) software, which can read and convert text documents into electronic files. OCR software converts a scanned image into a text file that can be edited, for example, with a word processing program.

OPTICAL READERS

An optical reader is a device that uses a light source to read characters, marks, and codes and then converts them into digital data that a computer can process. Two technologies used by optical readers are optical character recognition and optical mark recognition.

Optical character recognition (OCR)

It involves reading typewritten, computer-printed, or hand-printed characters from ordinary documents and translating the images into a form the computer can process. Most **OCR devices** include a small optical scanner for reading characters and sophisticated software to analyze what is read. OCR devices range from large machines that can read thousands of documents per minute to handheld wands that read one document at a time. Many companies use OCR characters on turnaround documents. A **turnaround document** is a document that you return (turn around) to the company that creates and sends it. For example, when consumers receive a bill, they often tear off a portion of the bill and send it back to the company with their payment. The portion of the bill they return usually has their payment amount, account number, and other information printed in OCR characters.

Optical mark recognition (OMR) devices read hand-drawn marks such as small circles or rectangles. A person places these marks on a form, such as a test, survey, or questionnaire answer sheet.

Barcode readers

A **barcode reader**, also called a barcode scanner, is an optical reader that uses laser beams to read barcodes. A **barcode** is an identification code that consists either of a set of vertical lines and spaces of different widths or a two-dimensional pattern of dots, squares, and other images. A 2-D bar code can store much more data than the traditional linear barcode. The barcode represents data that identifies the manufacturer and the item.

RFID READERS RFID (radio frequency identification)

It is a technology that uses radio signals to communicate with a tag placed in or attached to an object, an animal, or a person. RFID tags, which contain a memory chip and an antenna, are available in many shapes and sizes. An **RFID reader** reads information on the tag via radio waves. RFID readers can be handheld devices or mounted in a stationary object such as a doorway.

MAGNETIC STRIPE CARD READERS

A **magnetic stripe card reader**, often called a magistrate reader, reads the magnetic stripe on the back of credit cards, entertainment cards, bank cards, and other similar cards. The stripe contains information identifying you and the card issuer. Some information stored in the stripe includes your name, account number, the card's expiration date, and a country code.

MICR READERS MICR (magnetic-ink character recognition)

These devices read text printed with magnetized ink. A **MICR reader** converts MICR characters into a form the computer can process. The banking industry almost exclusively uses MICR for check processing. Each check in your checkbook has precoded MICR characters beginning at the lower-left edge.

The output is data that has been processed into a useful form. That is computers process data (input) into information (output). Users view or watch the output on a screen, print it, or hear it through speakers, headphones, or earphones. While working with a computer, a user encounters four basic categories of output: text, graphics, audio, and video. Very often, a single form of output, such as a Web page, includes more than one of these categories. An **output device** is any hardware component that conveys information to one or more people. Commonly used output devices include display devices; printers; speakers, headphones, and earphones; fax machines and fax modems; multifunction peripherals; data projectors; and interactive whiteboards.

Display devices

A **display device** is an output device that visually conveys text, graphics, and video information. Desktop computers typically use a monitor as their display device. A **monitor** is a display device that is packaged as a separate peripheral. Some monitors

have a tilt-and-swivel base that allows users to adjust the angle of the screen to minimize neck strain and reduce glare from overhead lighting. With some, you can rotate the screen. Monitor controls permit users to adjust the brightness, contrast, positioning, height, and width of images.

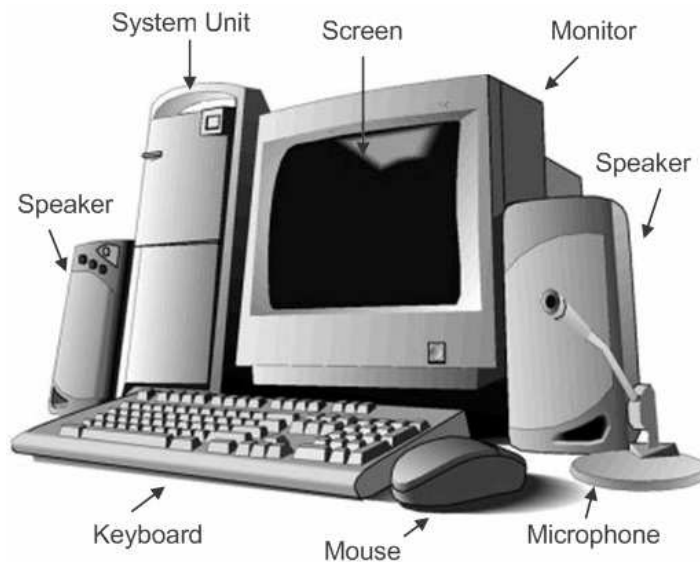


Figure 3.7: Video Display Terminal Consists of a Monitor and a Keyboard

LCD Monitors and LCD Screens

An **LCD monitor**, also called a flat panel monitor, is a desktop monitor that uses a liquid crystal display to produce images. These monitors produce sharp, flicker-free images. LCD monitors have a small footprint; that is, they do not take up much desk space. LCD monitors are available in a variety of sizes, with the more common being 17, 19, 20, 22, and 27 inches — some are 45 or 65 inches. Many are widescreen, which are much wider than they are tall. You measure a monitor the same way you measure a television, that is, diagonally from one corner to the other. Mobile computers, such as notebook computers and Tablet PCs, and mobile devices, such as Ultra-Mobile PCs, portable media players, and smartphones often have built-in LCD screens. Notebook computer screens are available in a variety of sizes, with the more common being 14.1, 15.4, 17, and 20.1 inches. Tablet PC screens range from 8.4 inches to 14.1 inches. Typical screen sizes of Ultra-Mobile PCs are 5 inches to 7 inches. Portable media players

usually have screen sizes from 1.5 inches to 3.5 inches. On smartphones, screen sizes range from 2.5 inches to 3.5 inches.



Figure 3.8: LCD Monitor

Plasma Monitors

A **plasma monitor** is a display device that uses gas plasma technology, which sandwiches a layer of gas between two glass plates. Plasma monitors offer screen sizes up to 60 inches wide and richer colors than LCD monitors but are more expensive. Like LCD monitors, plasma monitors can hang directly on a wall.

CRT Monitors

A **CRT monitor** is a desktop monitor that contains a cathode-ray tube. A cathode-ray tube (CRT) is a large, sealed glass tube. The front of the tube is the screen. CRT monitors for desktop computers are available in various sizes, with the more common being 15, 17, 19, 21, and 22 inches. In addition, to monitor size, advertisements also list a CRT monitor's viewable size. The viewable size is the diagonal measurement of the actual viewing area provided by the screen in the CRT monitor. A 21-inch monitor, for example, may have a viewable size of 20 inches. A CRT monitor usually costs less than an LCD monitor but also generates more heat and uses more power than an LCD monitor. To help reduce the amount of electricity used by monitors and other computer components, the United States Department of Energy (DOE) and the United States Environmental Protection Agency (EPA) developed the **ENERGY STAR program**. This program encourages manufacturers to create energy-efficient devices that require

little power when the devices are not in use. Monitors and devices that meet ENERGY STAR guidelines display an ENERGY STAR label.

Printers

A **printer** is an output device that produces text and graphics on a physical medium such as paper or transparency film. Many different printers exist with varying speeds, capabilities, and printing methods. There are mainly two types of printers:

- Impact printer
- Non Impact printer

Impact printers

An impact printer makes contact with the paper. It usually forms the print image by pressing an inked ribbon against the paper using a hammer or pins. Following are some examples of impact printers.

Dot-Matrix Printers

The dot-matrix printer uses print heads containing from 9 to 24 pins. These pins produce patterns of dots on the paper to form the individual characters. The 24 pin dot-matrix printer produces more dots than a 9 pin dot-matrix printer, which results in much better quality and clearer characters. The general rule is the more pins, the clearer the letters on the paper. The pins strike the ribbon individually as the print mechanism moves across the entire print line in both directions, i-e, from left to right, then right to left, and so on. The user can produce a color output with a dot-matrix printer (the user will change the black ribbon with a ribbon that has color stripes). Dot-matrix printers are inexpensive and typically print at speeds of 100-600 characters per second.



Figure 3.9: A Dot-matrix Printer

Daisy-wheel printers

In order to get the quality of type found on typewriters, a daisy-wheel impact printer can be used. It is called daisy-wheel printer because the print mechanism looks like a daisy; at the end of each “Petal” is a fully formed character which produces a solid-line print. A hammer strikes a “petal” containing a character against the ribbon and the character prints on the paper. Its speed is slow typically 25-55 characters per second.

Line printers

In a business where an enormous amount of material is printed, the character-at-a-time printers are too slow; therefore, these users need line-at-a-time printers. Line printers, or line-at-a-time printers, use a special mechanism that can print a whole line at once; they can typically print the range of 1,200 to 6,000 lines per minute. Drum, chain, and band printers are line-at-a-time printers.

Drum printer

A drum printer consists of a solid, cylindrical drum that has raised characters in bands on its surface. The number of print positions across the drum equals the number available on the page. This number typically ranges from 80-132 print positions. The drum rotates at a rapid speed. For each possible print position, there is a print hammer located behind the paper. These hammers strike the paper, along the ink ribbon, against the proper character on the drum as it passes. One revolution of the drum is required to print each line. This means that all characters on the line are not printed at exactly the same time, but the time required to print the entire line is fast enough to call them line printers. Typical speeds of drum printers are in the range of 300 to 2000 lines per minute.

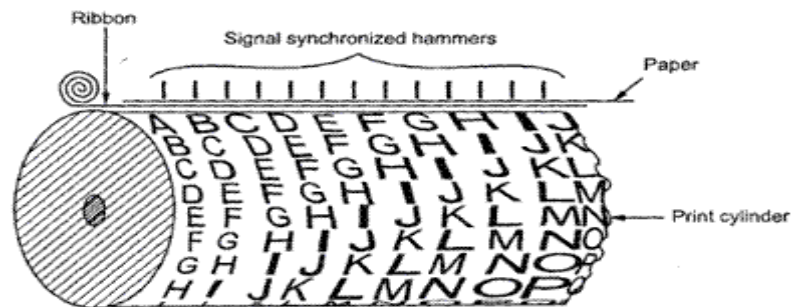


Figure 3.10: Drum Printer

Chain printers

A chain printer uses a chain of print characters wrapped around two pulleys. Like the drum printer, there is one hammer for each print position. The circuitry inside the printer detects when the correct character appears at the desired print location on the page. The hammer then strikes the page, pressing the paper against a ribbon and the character located at the desired print position. An impression of the character is left on the page. The chain keeps rotating until all the required print positions on the line have filled. Then the page moves up to print the next line. Speeds of chain printers range from 400 to 2500 characters per minute.

Band printers

A band printer operates similarly to chain printer except it uses a band instead of a chain and has fewer hammers. Band printer has a steel band divided into five sections of 48 characters each. The hammers on a band printer are mounted on a cartridge that moves across the paper to the appropriate positions. Characters are rotated into place and struck by the hammers. Font styles can easily be changed by replacing a band or chain.

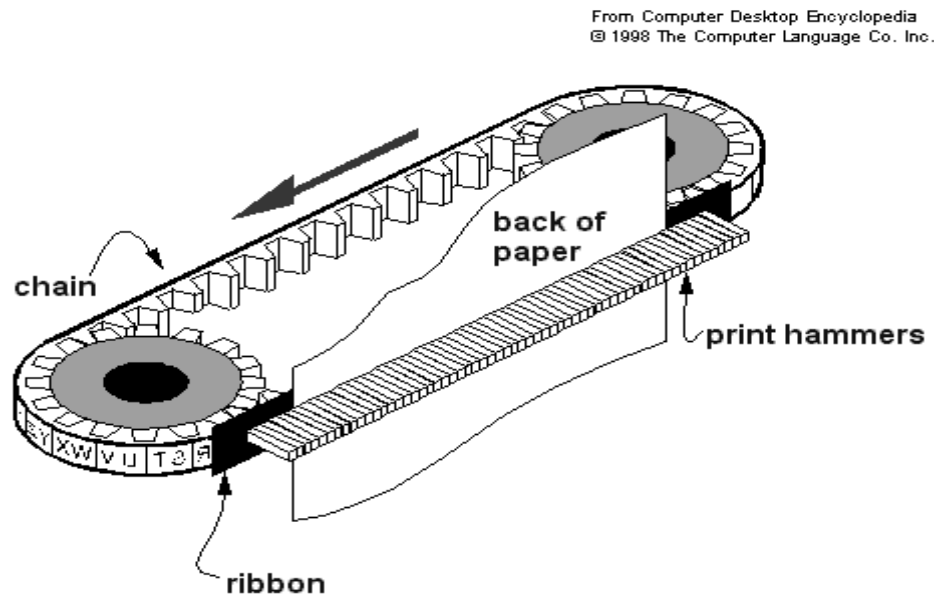


Figure 3.11: Printing Mechanism of a Chain/Band Printer

Non-impact printers

Non-impact printers do not use a striking device to produce characters on the paper, and because these printers do not hammer against the paper they are much quieter. Following are some non-impacted printers.

Ink-jet printers

Ink-jet printers work in the same fashion as dot-matrix printers in the form images or characters with little dots. However, the dots are formed by tiny droplets of ink. Ink-jet printers form characters on paper by spraying ink from tiny nozzles through an electrical field that arranges the charged ink particles into characters at the rate of approximately 250 characters per second. The ink is absorbed into the paper and dries instantly. Various colors of ink can also be used. One or more nozzles in the print head emit a steady stream of ink drops. Droplets of ink are electrically charged after leaving the nozzle. The droplets are then guided to the paper by electrically charged deflecting plates [one plate has a positive charge (upper plate) and the other has a negative charge (lower plate)]. A nozzle for black ink may be all that's needed to print text, but full-color printing is also possible with the addition of needed to print text, but full-color printing is also possible with the addition three extra nozzles for the cyan, magenta, and yellow primary colors. If a droplet isn't needed for the character or image being formed, it is recycled back to its input nozzle. Several manufacturers produce color ink-jet printer. Some of these printers come with all their color inks in a cartridge; if you want to replace on color, you must replace all the colors. Other color ink-jet printers allow you to replace ink individually. These printers are a better choice if the user uses one color more than other colors. These printers produce less noise and print in better quality with greater speed.



Figure 3.12: An Inkjet Printer

Laser printers

A laser printer works like a photocopier. Laser printers produce images on paper by directing a laser beam at a mirror which bounces the beam onto a drum. The drum has a special coating on it to which toner (an ink powder) sticks. Using patterns of small dots, a laser beam conveys information from the computer to a positively charged drum to become neutralized. From all those areas of the drum which become neutralized, the toner detaches. As the paper rolls by the drum, the toner is transferred to the paper printing the letters or other graphics on the paper. A hot roller bonds the toner to the paper. Laser printers use buffers that store an entire page at a time. When a whole page is loaded, it will be printed. The speed of laser printers is high and they print quietly without producing much noise. Many home-use laser printers can print eight pages per minute, but faster and print approximately 21,000 lines per minute, or 437 pages per minute if each page contains 48 lines. When high-speed laser printers were introduced they were expensive. Developments in the last few years have provided relatively low-cost laser printers for use in small businesses.



Figure 3.13: Laser Printer

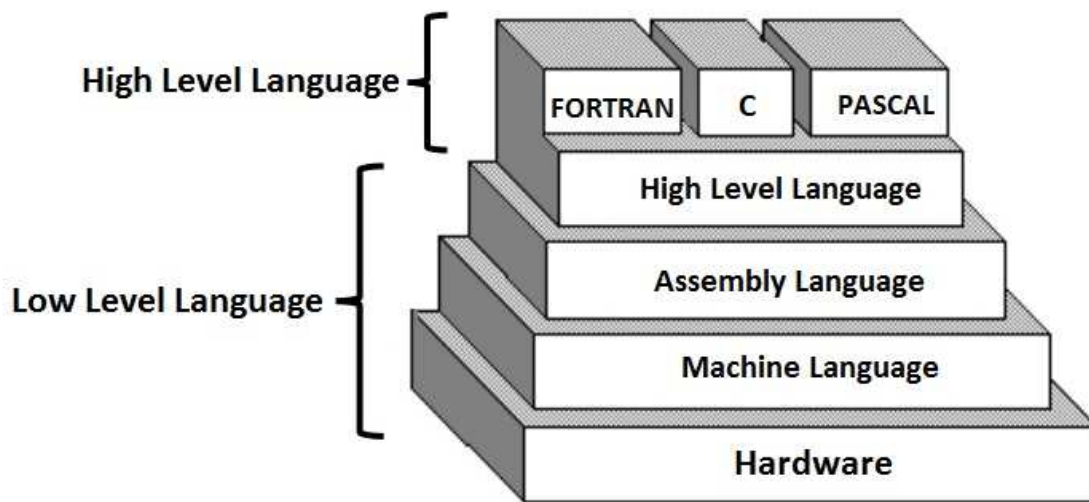
9.6 COMPUTER LANGUAGES

A computer program is a series of instructions that directs a computer to perform tasks. A computer programmer, sometimes called a developer, creates and modifies computer programs. To create a program, programmers sometimes write, or code, a program's instructions using a programming language. A programming language is a set of words, abbreviations, and symbols that enables a programmer to communicate instructions to a computer. Other times, programmers use a program development tool to create a program. A program that provides a user-friendly environment for building programs often is called a program development tool. Programmers use a variety of programming languages and tools to create programs. Several hundred programming languages exist today. Each language has its own rules for writing the instructions. Languages often are designed for specific purposes, such as scientific applications, business solutions, or Web page development. Two types of languages are low-level and high-level. A low-level language is a programming language that is machine dependent. A machine-dependent language runs on only one particular type of computer. Each instruction in a low-level language usually equates to a single machine instruction, discussed further in the next section. With a high-level language, by contrast, each instruction typically equates to multiple machine instructions. High-level languages often are machine independent. A machine-independent language can run on many different types of computers and operating systems.

Low-Level Languages

Two types of low-level languages are machine languages and assembly languages. Machine language, known as the first generation of programming languages, is the only language the computer directly recognizes. Machine language instructions use a series of binary digits (1s and 0s) or a combination of numbers and letters that represent binary digits. The binary digits corresponding to the on and off electrical states. As you might imagine, coding in machine language is tedious and time-consuming. With an assembly language, the second generation of programming languages, a programmer writes instructions using symbolic instruction codes. Assembly languages also use symbolic addresses. A symbolic address is a meaningful name that identifies a storage location. For example, a programmer can use the name RATE to refer to the storage location that contains a pay rate. Despite these advantages, assembly languages can be difficult to learn. In addition, programmers must convert an assembly language program into

machine language before the computer can execute, or run, the program. That is, the computer cannot execute the assembly source program. A source program is a program that contains the language instructions, or code, to be converted to machine language. To convert the assembly language source program into machine language, programmers use a program called an assembler.



Computer Language and its Types

Figure 3.14 Computer languages and types

High-Level Language:

The high-level languages are the most used and also more considered programming languages that help a programmer to read, write and maintain. It is also the third generation language that is used and also running till now by many programmers. They are less independent of a particular type of Computer and also require a translator that can convert the high-level language to machine language. The translator may be an interpreter and Compiler that helps to convert into binary code for a Computer to understand. There are various high-level programming languages like C, FORTRAN or Pascal that are less independent and also enables the programmer to write a program. The Compiler plays an important role on the Computer as it can convert to machine

language and also checks for errors if any before executing. There are several high-level languages that were used earlier and also now like COBOL, FORTRAN, BASIC, C, C++, PASCAL, LISP, Ada, Algol, Prolog, and Java. It is user-friendly as the programs are written in English using words, symbols, characters, numbers that need to be converted to machine code for processing.

9.7 OPERATING SYSTEM

An Operating System is a set of programs that controls and coordinates the use of computer hardware among various application programs. It provides an environment within which user can execute programs. A computer can be divided into four components: the hardware, the operating system, the applications programs, the users. It is the operating system that manages all the above components. The various functions of the operating system are :

- Controlling Input/Output devices (Keyboard, Mouse, Monitor, Printer, Plotter etc..)
- Memory and File storage management
- CPU Scheduling and controlling processes
- Loading, initiating, executing and supervising user applications programs
- Handling errors and restarting
- Providing command interface between user and computer system

Examples of the Operating system are UNIX (Solaris, IRIX, HPUnix, Linux, DEC Unix) Microsoft Disk Operating System (MS-DOS), WIN95/98, WIN NT, OS/2 etc.

MS-DOS

Short for **Microsoft Disk Operating System**, **MS-DOS** is a non-graphical command line operating system derived from 86-DOS that was created for IBM compatible computers. MS-DOS originally written by Tim Paterson and introduced by Microsoft in August 1981 and was last updated in 1994 when MS-DOS 6.22 was released. MS-DOS allows the user to navigate, open, and otherwise manipulate files on their computer from a command line instead of a GUI like Windows. Today, MS-DOS is no longer used; however, the command shell, more commonly known as the **Windows**

command line is still used by many users. The picture to the right is an example of what an MS-DOS window more appropriately referred to as the Windows command line looks like running under Microsoft Windows.

Most computer users are only familiar with how to navigate Microsoft Windows using the mouse. Unlike Windows, MS-DOS is a command-line and is navigated by using MS-DOS commands. For example, if you wanted to see all the files in a folder in Windows you would double-click the folder to open the folder in Windows Explorer. In MS-DOS, to view that same folder you would navigate to the folder using the `cd` command and then list the files in that folder using the `dir` command.

Types of DOS commands

Any instruction given to the computer to perform a specific task is called command. The DOS has several commands, each for a particular task and these are stored in DOS directory on the disk.

The commands are of two types :

(a) Internal Commands: These are inbuilt commands of MS-DOS i.e. these are stored in Command interpreter file (COMMAND.COM). These commands reside in the memory as long as the machine is at the system prompt(C:\>) level. To use these commands no extra /external file is required. E.g. DATE, TIME, DIR, VER etc.

(b) External commands: These are separate program (.com) files that reside in DOS directory and when executed behave like commands. An external command has predefined syntax. for e.g. HELP, DOSKEY, BACKUP, RESTORE, FORMAT etc.

Basic Dos Commands:

(a) Directory Commands :

DIR: To list all or specific files of any directory on a specified disk.

MD: To make directory or subdirectory on a specified disk/drive.

CD or CHDIR: Change DOS current working directory to specified directory on specified disk or to check for the current directory on the specified or default drive.

RMDIR or RD: Removes a specified sub-directory only when it is empty. This command cannot remove root directory (C:\) or current working directory.

TREE: Displays all of the directory paths found on the specified drive.

PATH: Sets a sequential search path for the executables files, if the same is not available in the current directory.

SUBST: Substitutes a string alias for the pathname and creates a virtual drive.

b) File Management Commands:

COPY: Copies one or more files from source disk/drive to the specified disk/drive.

XCOPY: Copies files and directories, including lower-level directories if they exist.

DEL: Removes specified files from specified disk/drive.

REN: Changes the name of a file (Renaming).

ATTRIB: Sets or shows file attributes (read, write, hidden, Archive).

BACKUP: Stores or back up one or more files/directories from source disk/drive to other destination disk/drive.

RESTORE Restores files that were backed up using the BACKUP command.

EDIT: Provides a full-screen editor to create or edit a text file.

FORMAT: Formats a disk/drive for data storage and use.

(c) General Commands:

TIME: sets or displays the system time.

DATE: Sets or displays system date.

TYPE: Displays the contents of at the specified file.

PROMPT: Customizes the DOS command prompt.

If a user requires help on any DOS commands he/she may type help and command name at the command prompt.

9.8 INTRODUCTION TO WINDOWS:

MS Windows is a **multitasking operating** system which uses a Graphical User Interface (GUI) to link the user to the computer. A GUI is an interface that helps users to interact with the computer by use of windows, icons, and menus. The interaction between the user and the computer heavily relies on the use of a mouse and keyboard which make it possible to enter data into the computer and also manipulate it the way we want. Windows (created by Microsoft) is the most dominant OS on the market today. The two most popular versions of Windows for the desktop are XP and Vista (Vista being the latest version). There is also a mobile version of Windows as well as a server version of Windows (the latest being Windows Server 2008). Windows is all proprietary, closed-source which is much different than Linux licenses. Most of the popular manufacturers make all of their hardware compatible with Windows which makes Windows operate and almost all kinds of new hardware.

Features Of Windows Graphical User Interface: Windows using graphical user interface environment that simplifies the user's interaction with the computer by representing programs, commands, files, and other options as visual elements, such as icons, pull-down menus, buttons, scrollbars, windows, and dialog boxes. **START BUTTON** On the very left corner a Start button is located in windows. By clicking it we can access many of the main programs within the computer. It is a very easy way to interact with computer programs just click on start button and access these programs like open a document, paint, calculator, manage files, maintain systems and many more. **TASKBAR** In the bottom section of the screen, there is (or there may be) a long bar. It is called the Taskbar:



Figure 3.15 windows layout

On the left side of the taskbar, there is an object labeled start. The appearance of this object depends on the type of Windows you are using but it plays the same role. On the right side of Start, there is a wide area that is empty when the computer starts. This is actually referred to as the taskbar. In some versions (Windows 95, 98, Windows Server 2003), the color of the taskbar may be gray (or Silver). In some other versions (Windows Millennium), the taskbar may be yellowish. In Windows XP, it may be blue. On the far right side of the taskbar is a section called the tray area. One of the things that this area display is the current time (as set on yourcomputer).

Icons are small graphical images that can represent your computer's programs, files, folders and printers amongst other things. To activate the program/file/folder that an icon represents you simply double click (two clicks in quick succession) on it with the left mouse button. This will activate the icon and either start a program or open a file/folder. The icons on your desktop can be renamed by right-clicking on them and selecting rename. They can be deleted by right clicking and selecting delete. It's possible to create your own Windows desktop icons for programs, files, folders, etc. Examples of icons are

- My computer
- My document
- Recycle bin
- Shortcut

My Computer: My computer icon has information displays all the components of the computer including disk drives and networks all parts of the computer can be accessed through My Computer's hierarchical structure double click on any icon to view the contents of a disk, folder or run a program.

My Document: My Documents is the name of a special folder on the computer's hard disk that the system commonly uses can store his documents, music, pictures, downloads, and other files.

Recycle Bin: When you delete a file, Windows XP will place the file into the recycle bin (instead of deleting it altogether), this allows you to restore the file in case you deleted it by mistake. For example, if you deleted a file by accident you could double click on the recycle bin icon to see its contents. Inside you will see the file/folder that you deleted, to restore the file back to where you deleted it from, right click on it and then select restore from the menu that appears.

Desktop After you have started the computer, the area you are looking at is called the desktop. The desktop is usually different from one computer to another. This is because some items get added as new programs are installed on a computer, and other items get deleted at will. Eventually, you will know how to change the way your desktop looks. The Microsoft Windows desktop is made of various parts. From the upper left to the lower left side of the screen, there are small pictures or images called icons.

9.9 DATA PROCESSING

Data processing refers to the transformation raw data into meaningful output. Data processing, Manipulation of data by a computer. It includes the conversion of raw data to a machine-readable form, the flow of data through the CPU and memory to output devices, and formatting or transformation of output. Any use of computers to perform defined operations on data can be included under data processing. In the commercial world, data processing refers to the processing of data required to run organizations and businesses.

Data processing methods:

1. Manual Data Processing:

In manual data processing, data is processed manually without using any machine or tool to get required results. In manual data processing, all the calculations and logical operations are performed manually on the data. Similarly, data is transferred manually from one place to another. This method of data processing is very slow and errors may occur in the output. Mostly,

is processed manually in many small business firms as well as government offices & institutions. In an educational institute, for example, marks sheets, fee receipts, and other financial calculations (or transactions) are performed by hand. This method is avoided as far as possible because of the very high probability of error, labor intensive and very time-consuming. This type of data processing forms the very primitive stage when technology was not available or it was not affordable. With the advancement in technology, the dependency on manual methods has drastically decreased

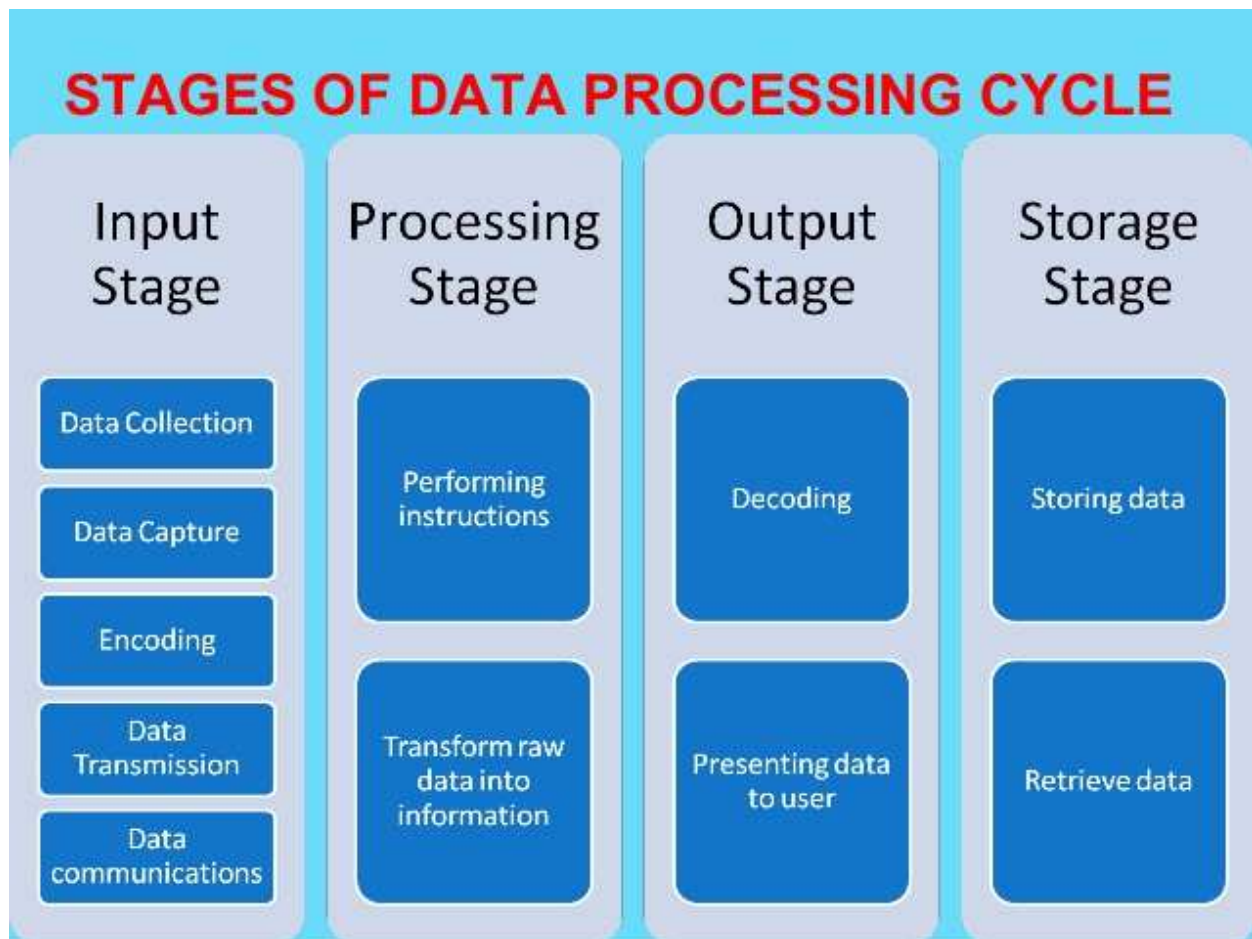


Figure 3.16 Data Processing Methods

2. Mechanical Data Processing:

In mechanical data processing method, data is processed by using different devices like typewriters, mechanical printers or other mechanical devices. This method of data processing is faster and more accurate than manual data processing. These are faster than the manual mode but

still form the early stages of data processing. With invention and evolution of more complex machines with better computing power this type of processing also started fading away. Examination boards and printing press use mechanical data processing devices frequently.

3. Electronic Data Processing

Electronic data processing or EDP is the modern technique to process data. The data is processed through the computer; Data and set of instructions are given to the computer as input and the computer automatically processes the data according to the given set of instructions. The computer is also known as electronic data processing machine. This method of processing data is very fast and accurate. For example, in a computerized education environment results of students are prepared through the computer; in banks, accounts of customers are maintained (or processed) through computers etc.

9.10 PRINCIPLES OF PROGRAMMING

Computer programming is the act of writing computer programs, which are a sequence of instructions written using a Computer Programming Language to perform a specified task by the computer. A computer program is a sequence of instructions written using a Computer Programming Language to perform a specified task by the computer.

The two important terms that we have used in the above definition are –

- Sequence of instructions
- Computer Programming Language

Uses of Computer Programs

Today computer programs are being used in almost every field, household, agriculture, medical, entertainment, defense, communication, etc. Listed below are a few applications of computer programs –

- MS Word, MS Excel, Adobe Photoshop, Internet Explorer, Chrome, etc., are examples of computer programs.
- Computer programs are being used to develop graphics and special effects in movie making.
- Computer programs are being used to perform Ultrasounds, X-Rays, and other medical examinations.
- Computer programs are being used in our mobile phones for SMS, Chat, and voice communication.

9.11 ALGORITHMS AND FLOWCHARTS

From a programming point of view, an **algorithm** is a step-by-step procedure to resolve any problem. An algorithm is an effective method expressed as a finite set of well-defined instructions. Thus, a computer programmer lists down all the steps required to resolve a problem before writing the actual code. Following is a simple example of an algorithm to find out the largest number from a given list of numbers –

1. Get a list of numbers $L_1, L_2, L_3, \dots, L_N$
2. Assume L_1 is the largest, Largest = L_1
3. Take next number L_i from the list and do the following
4. If Largest is less than L_i
5. Largest = L_i
6. If L_i is the last number from the list then
7. The print value stored in Largest and come out
8. Else repeat the same process starting from step 3

The above algorithm has been written in a crude way to help beginners understand the concept. You will come across more standardized ways of writing computer algorithms as you move on to advanced levels of computer programming.

Properties of algorithm

Donald Ervin Knuth has given a list of five properties for an algorithm, these properties are:

(1) Finiteness: An algorithm must always terminate after a finite number of steps. It means after every step one reaches closer to the solution of the problem and after a finite number of steps, algorithm reaches an endpoint.

(2) Definiteness: Each step of an algorithm must be precisely defined. It is done by well-thought actions to be performed at each step of the algorithm. Also, the actions are defined unambiguously for each activity in the algorithm.

(3) Input: Any operation you perform need some beginning value/quantities associated with different activities in the operation. So the value/quantities are given to the algorithm before it begins.

(4) Output: One always expects output/result (expected value/quantities) in terms of output from an algorithm. The result may be obtained at different stages of the algorithm. If some result is from the intermediate stage of the operation then it is known as intermediate result and result obtained at the end of the algorithm is known as the end result. The output is expected value/quantities always have a specified relation to the inputs.

(5) Effectiveness: Algorithms to be developed/written using basic operations. Actually, operations should be basic, so that even they can in principle be done exactly and in a finite amount of time by a person, by using paper and pencil only.







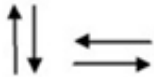
FLOWCHART:

The flowchart is a diagram which visually presents the flow of data through processing systems. This means by seeing a flow chart one can know the operations performed and the sequence of these operations in a system. Algorithms are nothing but a sequence of steps for solving problems. So a flow chart can be used for representing an algorithm. A flowchart will describe the operations (and in what sequence) are required to solve a given problem. You can see a flowchart as a blueprint for a design you have made for solving a problem.

For example, suppose you are going for a picnic with your friends then you plan for the activities you will do there. If you have a plan of activities then you know clearly when you will do what activity. Similarly, when you have a problem to solve using a computer or in another word you need to write a computer program for a problem then it will be good to draw a flowchart prior to writing a computer program. The flowchart is drawn according to defined rules.

Flowchart Symbols

There are 6 basic symbols commonly used in flowcharting of assembly language Programs: Terminal, Process, and input/output, Decision, Connector and Predefined Process. This is not a complete list of all the possible flowcharting symbols; it is the ones used most often in the structure of Assembly language programming.

Symbol	Name	Function
	Process	Indicates any type of internal operation inside the Processor or Memory
	Input/Output	Used for any Input / Output (I/O) operation. Indicates that the computer is to obtain data or output results
	Decision	Used to ask a question that can be answered in a binary format(Yes/No, True/False)
	Connector	Allows the flowchart to be drawn without intersecting lines or without a reverse flow.
	Predefined Process	Used to invoke a subroutine or an Interrupt program.
	Terminal	Indicates the starting or ending of the program, process, or interrupt program
	Flow Lines	Shows direction of flow.

General Rules for flowcharting

1. All boxes of the flowchart are connected with Arrows. (Not lines)
2. Flowchart symbols have an entry point on the top of the symbol with no other entry points. The exit point for all flowchart symbols is on the bottom except for the Decision

symbol.

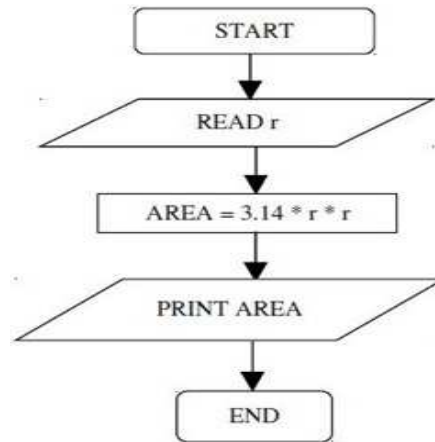
3. The Decision symbol has two exit points; these can be on the sides or the bottom and one side.
4. Generally, a flowchart will flow from top to bottom. However, an upward flow can be shown as long as it does not exceed 3 symbols.
5. Connectors are used to connect breaks in the flowchart. Examples are:
 - From one page to another page.
 - From the bottom of the page to the top of the same page.
 - An upward flow of more than 3 symbols
6. Subroutines and Interrupt programs have their own and independent flowcharts.
7. All flowcharts start with a Terminal or Predefined Process (for interrupt programs or subroutines) symbol.
8. All flowcharts end with a terminal or a contentious loop.

Flowcharting uses symbols that have been in use for a number of years to represent the type of operations and/or processes being performed. The standardized format provides a common method for people to visualize problems together in the same manner. The use of standardized symbols makes the flowcharts easier to interpret; however, standardizing symbols is not as important as the sequence of activities that make up the process.

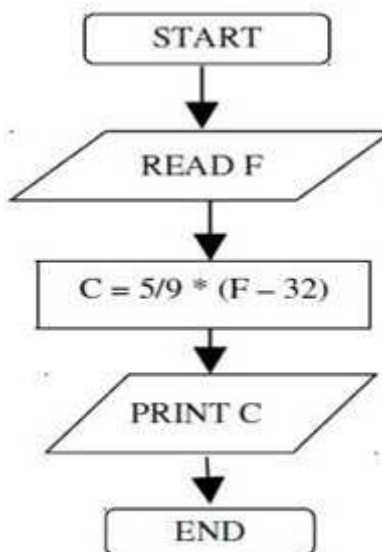
Some examples of Flowcharts

Now, we will discuss some examples on flowcharting. These examples will help in the proper understanding of flowcharting technique. This will help you in program development process in next unit of this block.

Problem1: Find the area of a circle of radius r .



Problem 2: Convert temperature Fahrenheit to Celsius.



Advantages of using Flowcharts

As we discussed flowchart is used for representing an algorithm in pictorial form. This pictorial representation of a solution/system is having many advantages. These advantages are as follows:

1. **Communication:** A Flowchart can be used as a better way of communication of the logic of a system and steps involved in the solution, to all concerned particularly to the client of the system.
2. **Effective analysis:** A flowchart of a problem can be used for effective analysis of the problem.

3. **Documentation of Program/System:** Program flowcharts are a vital part of a good program documentation. Program document is used for various purposes like knowing the components in the program, complexity of the program etc.
4. **Efficient Program Maintenance:** Once a program is developed and becomes operational it needs time to time maintenance. With help of flowchart, maintenance becomes easier.

9.12 LET US SUM UP

- A **computer** is an electronic device, operating under the control of instructions stored in its own memory, that can accept data, process the data according to specified rules, produce results, and stores the results for future use. The electric, electronic, and mechanical components of a computer, or **hardware**, including input devices, output devices, a system unit, storage devices, and communications devices.
- An **input device** allows you to enter data or instructions into a computer.
- An **output device** conveys information to one or more people.
- The **system unit** is a case that contains the electronic components of a computer that are used to process data.
- **Storage device** records and/or retrieves items to and from storage media.
- The system unit contains volatile and nonvolatile memory. **Volatile memory** loses its contents when the computer's power is turned off. **Nonvolatile memory** does not lose its contents when the computer's power is turned off. RAM is the most common type of volatile memory. ROM, flash memory, and CMOS are examples of nonvolatile memory. **RAM** consists of memory chips that can be read from and written to by the processor and other devices. **ROM** refers to memory chips storing permanent data and instructions that usually cannot be modified. **Flash memory** can be erased electronically and rewritten. **CMOS** technology uses battery power to retain information even when the power to the computer is turned off.

- Any hardware component that allows users to enter data and instructions into a computer is an **input device**. A **keyboard** is an input device that contains keys users press to enter data and instructions into a computer. Computer keyboards have a typing area that includes letters of the alphabet, numbers, punctuation marks, and other basic keys. An enhanced keyboard also has function keys programmed to issue commands, a numeric keypad, arrow keys, and additional keys and buttons.
- The operating system provides a user interface, manages programs, manages memory, coordinates tasks, configures devices, establishes an Internet connection, and monitors performance. The **user interfaces** control how data and instructions are entered and how information is displayed. Two types of user interfaces are a **command-line interface** and a **graphical user interface (GUI)**. Managing programs refers to how many users, and how many programs, an operating system can support at one time. An operating system can be single user/single tasking, single user/multitasking, multiuser, or multiprocessing.
- Algorithm and flowchart are two types of tools to explain the process of a program.
- To write a logical step-by-step method to solve the problem is called algorithm, in other words, an algorithm is a procedure for solving problems. In order to solve a mathematical or computer problem, this is the first step of the procedure. An algorithm includes calculations, reasoning and data processing. Algorithms can be presented by natural languages, pseudocode, and flowcharts, etc.
- A flowchart is the graphical or pictorial representation of an algorithm with the help of different symbols, shapes, and arrows in order to demonstrate a process or a program. With algorithms, we can easily understand a program.

9.13 CHECK YOUR PROGRESS

Q1. CD-ROM stands for

- a. Compactable Read Only Memory
- b. Compact Data Read Only Memory
- c. Compactable Disk Read Only Memory
- d. Compact Disk Read Only Memory

Q2. ALU is

- a. Arithmetic Logic Unit
- b. Array Logic Unit
- c. Application Logic Unit
- d. None of above

Q3. MICR stands for

- a. Magnetic Ink Character Reader
- b. Magnetic Ink Code Reader
- c. Magnetic Ink Cases Reader
- d. None

Q4. Which of the following are two main components of CPU?

- a. Control unit and registers
- b. Register and main memory
- c. Control unit and ALU
- d. ALU and Bus

Q5. What command would you type to list all of the files in the current directory?

- a. DIR
- b. LIST ALL
- c. SHOW
- d. SHOW ME

Q6. A hard disk is divided into tracks which are further subdivided into:

- a. Clusters
- b. Sectors
- c. Vectors
- d. Head

Ans. 1. d 2. a 3. a 4. c 5. a 6. b

FURTHER READINGS

- Computer fundamentals by Anita Geol, Pearson Education
- Fundamentals of Computers by V. Rajaraman, PHI publication
- Computing Fundamentals and C Programming by E Balagurusamy, TMH publication
- Operating system concepts by Silberschatz, Galvin, Gagne, Willy Publication

POSSIBLE QUESTIONS

Q.1. What is computer and explain the basic structure of computer?

Q.2. Explain the functional units of the computer.

Q.3. Explain memory and different types of memory in details.

Q. 4. What is the operating system? And state some functions of the operating system.

Q5. What is data processing? Explain the various methods used for data processing.

Q 6. Define algorithm and explain the properties of the algorithm.

Q7. What is flowchart and explain the advantages of using a flowchart?

Q 8. Write short notes on:

- | | |
|----|-------------------|
| a. | Computer language |
| b. | DOS |
| c. | Windows |
| d. | Operating system |
| e. | Algorithm |
| f. | Flowchart |

UNIT 10: PROGRAMMING IN CHEMISTRY

CONTENTS:

- 10.1 Learning Objectives
- 10.2 Introduction
- 10.3 Development of small computer codes
- 10.4 VANDER WAALS equation
- 10.5 pH titration
- 10.6 Kinetics
- 10.7 Radioactive Decay
- 10.8 Evaluation of Lattice Energy
- 10.9 ionic radii
- 10.10 Let Us Sum Up
- 10.11 Answer to Check Your Progress
- 10.12 Further Readings
- 10.13 Model Questions

10.1 LEARNING OBJECTIVES

After complete study and understanding of this unit, you should be able to:

- Understand the technique of writing programs for some chemistry applications
- Able to write computer codes for many more chemistry applications.

10.2 INTRODUCTION

The increasing power of computers, which has become available at constantly decreasing cost, has made computational methods an invaluable tool for chemists to understand and design new complex chemical systems. One great reason for using

computational methods in chemistry is that they can in principle provide an understanding of chemical processes at a level of detail which is often not available experimentally. Transition state modeling provides one such example. While transition states and reaction paths can be studied computationally almost as simply as ground states, experimental information about these unstable points on the energy surface can often only be obtained indirectly, for example by measuring the activation parameters of reactions. Since the lifetimes of transition states are only on the order of a chemical vibration, spectroscopy with a fem to second-time resolution is needed in order to study transition states directly experimentally.

There are many uses of computers in chemistry. Broadly, they are divided into the following categories and sub-categories.

Programming:

- A. **Time-dependent differential equations**, in which the initial conditions are known, and one wants to know how they evolve forward in time: these include:
 - Chemical Reactions (evolution of concentrations with time)
 - Classical dynamics (evolution of atomic positions with time)
 - Time-dependent Quantum dynamics (atoms under lasers)
- B. **Time-independent differential equation**: primarily used for the (time independent) Schrodinger equation, which is solved in two main fields of chemical interest:
 - Structure of molecules and their energies (electronic structure, or ab-initio).
 - Vibrational frequencies
- C. **Random variables and stochastic processes** (important for diffusion, noise, and similar processes which affect chemical reactions, among other places).
- D. **Data-analysis**, including especially Fourier (and today Wavelet), transforms, 3-D data plotting algorithms, etc. A related area is: **Bioinformatic tools** are emerging for describing DNA genomics, etc.

10.3 DEVELOPMENT OF SMALL COMPUTER CODES

Developing computer codes or programs requires acquiring programming skills. These skills can be acquired by studying existing programs for typical applications. Computational chemistry is a branch of chemistry that uses computer simulation to assist in solving chemical problems. It uses methods of theoretical chemistry, incorporated into efficient computer programs, to calculate the structures and properties of molecules and solids. It is necessary because, apart from relatively recent results concerning the hydrogen molecular ion (dihydrogen cation, see references therein for more details), the quantum many-body problem cannot be solved analytically, much less in closed form. While computational results normally complement the information obtained by chemical experiments, it can in some cases predict hitherto unobserved chemical phenomena. It is widely used in the design of new drugs and materials.

10.4 VANDER WAALS EQUATION

The van der Waals equation was proposed in the year 1873. It was a first step towards taking into account interaction forces which are acting between real gases molecules.

Van der Waal's equation

Though the equation $PV = nRT$ was arrived at first experimentally and then theoretically, yet it fails to explain the behaviour of real gases. Van der Waals attributed the deviation of real gases from the gas equation to the following faulty assumptions of the kinetic theory:

- (i) The actual volume of the gas molecules is negligible as compared with the total volume of the gas.
- (ii) The gas molecules do not exert any appreciable attraction on each other.

Both these assumptions are not true at high pressure and low temperature. At high pressure, the volume is reduced to a great extent and the actual volume of the molecules cannot be supposed as negligible under such conditions. Moreover, the molecules come closer and the attractive force between them should also be taken into consideration.

Van der Waals introduced the necessary corrections as follows:

(a) **Volume correction:** At higher pressure, the volume is much reduced and at this state, the volume of gas molecules no longer remains negligible in comparison with the total volume V occupied by the gas. Thus the free space available for the motion of molecules is reduced. If v be the volume of the molecules at rest, effective volume when they are in motion must be greater. It has been calculated to be approximately four times the actual volume of the molecules.

I. e.

$4v$ and is generally denoted by b . Therefore the actual volume available in which the molecules are free to move will be

= Total volume — Effective volume.

i.e., correct volume ($V - b$)

(b) **Pressure correction:** The pressure of a gas is due to the hits of the molecules on the walls of the containing vessel. The attractive force between the molecules comes into play when the molecules are brought close together by compressing the gas. A molecule in the body of the gas is attracted in all the directions when forces acting in opposite directions cancel out, but a molecule, the boundary of the gas is subjected to an inward pull due to the unbalanced molecular attraction.

In this way, some of the energy of the molecule about to strike the wall of the vessel is used up in overcoming this inward pull. Therefore, it will not strike the opposite wall with the same force. The observed pressure consequently will be less than the ideal pressure. Therefore the ideal pressure, P_i will be equal to observed pressure P plus a pressure correction, P_a depending upon the attractive forces,

I.e. $P_i = P + P_a$

Now, the force of attraction \propto number of molecules in the interior which are attracting (i.e., density)

The total force of attraction $\propto (\text{density})^2 \propto \frac{1}{(\text{Volume})^2}$

Or, pressure due to the force of attraction $\propto \frac{1}{(\text{Volume})^2}$

i.e.,
$$Pa = \frac{a}{V^2}$$

(where V is the volume of the gas and 'a' is the coefficient of attraction)

$$\therefore \text{ideal pressure} = P + \frac{a}{V^2}$$

Introducing, both these corrections, the equation becomes

It is known as van der Waals equation. This equation explains the behavior of real gases with great accuracy and also the deviations of gas laws from ideal behavior.

The units of 'a' and 'b' are expressed in the units of pressure and volume as the unit of 'a' is $\text{atm } L^2 \text{mol}^{-2}$ that of 'b' is $L \text{mol}^{-1}$ and

The following is the C language program which calculates the pressure of a gas, when all the other quantities are known:

```
#include<stdio.h> main () {  
  
float p,n,r,t,v,a,b;  
  
printf (" enter the value of volume, temperature and moles of the gas\n");  
  
scanf ("%f%f%f", &v, &t, &n);  
  
printf (" Enter the values for Gas Law constant, Van Der Waal constant a and van der  
wall constant b\n");  
  
scanf ("%f%f%f", &r, &a, &b);
```

$p = (n \cdot r \cdot t_0 / (v - n \cdot b_0 - (n \cdot n \cdot a) / (n \cdot v))); \text{ printf} (" \text{ the pressure of the gas} = \%f \text{ atm} ", p); \}$

10.5 *pH TITRATION*

The word titration comes from the Latin word "titulus", which means inscription or title. The French word title means rank. Therefore, Titration means the determination of concentration or rank of a solution with respect to water with a pH of 7.

The standard solution is usually added from a graduated vessel called a burette. The process of adding standard solution until the reaction is just complete is termed as titration and the substance to be determined is said to be titrated. All chemical reactions cannot be considered as titrations. A reaction can serve as a basis for a titration procedure only if the following conditions are satisfied:

1. The reaction must be a fast one.
2. It must proceed stoichiometrically.
3. The change in free energy (ΔG) during the reaction must be sufficiently large for spontaneity of the reaction.
4. There should be a way to detect the completion of the reaction.

We will look at four types of titrations:

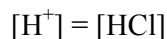
- I. A strong acid with a strong base
- II. A weak acid with a strong base
- III. A strong acid with a weak base
- IV. Titration of a polyprotic acid

. A strong acid titrated with a strong base

We will calculate the pH of 25 mL of 0.1 M HCl titrated with 0.1 M NaOH. At each point in the titration curve, we will need to determine two quantities, the **concentration of H^+** remaining in the solution, and the **volume** of the solution. From these, we can calculate the $[H^+]$ and, from that, the **pH**.

1. **The initial pH**, before the addition of a strong base.
-

In dealing with strong acids and bases, assume that they dissociate completely in water.



In this example, we have a 0.1 M solution of HCl, so the $[\text{H}^+] = 0.1 \text{ M}$

$$-\log(0.1) = 1 = \text{pH}$$

II. A weak acid with a strong base

We will calculate the pH of 25 mL of 0.1 M CH_3COOH titrated with 0.1 M NaOH. At each point in the titration curve, we will need to determine two quantities, the **concentration of H^+** remaining in the solution, and the **volume** of the solution. From these, we can calculate the $[\text{H}^+]$ and, from that, the **pH**. This is a weak acid with a $K_a = 1.8 \times 10^{-5}$.

1. **The initial pH**, before the addition of a strong base.

In dealing with weak acids and bases, there will be an incomplete dissociation in water. The degree of dissociation will be determined by K_a , the acid dissociation constant in this case.



$$K_a = \frac{[\text{H}^+][\text{CH}_3\text{COO}^-]}{[\text{CH}_3\text{COOH}]}$$

In this example, we have a 0.1 M solution of CH_3COOH , so the $[\text{H}^+]$ can be calculated using the equilibrium expression:

$$K_a = 1.8 \times 10^{-5} = \frac{[x][x]}{[0.1-x]}, \text{ where } x \text{ is the amount of weak acid that dissociates.}$$

$$x = 1.34 \times 10^{-3} = [\text{H}^+] \quad \text{pH} = 2.87$$

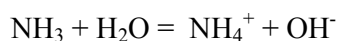
The pH is higher in this problem than in the strong acid titration because we have partial dissociation of the acid, and consequently, a lower $[\text{H}^+]$.

III. A weak base with a strong acid

We will calculate the pH of 25 mL of 0.1 M NH_3 titrated with 0.1 M HCl . At each point in the titration curve, we will need to determine two quantities, the **concentration of H^+** remaining in the solution, and the **volume** of the solution. From these, we can calculate the **$[\text{H}^+]$** and, from that, the **pH**. This is a weak base with a $K_b = 1.8 \times 10^{-5}$.

1. The initial pH, before the addition of strong acid.

In dealing with weak acids and bases, there will be an incomplete dissociation in water. The degree of dissociation will be determined by K_b , the base dissociation constant in this case.



$$K_b = \frac{[\text{NH}_4^+][\text{OH}^-]}{[\text{NH}_3]}$$

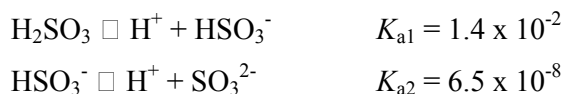
In this example, we have a 0.1 M solution of NH_3 , so the $[\text{OH}^-]$ can be calculated using the equilibrium expression:

$$K_b = 1.8 \times 10^{-5} = \frac{[x][x]}{[0.1-x]}, \text{ where } x \text{ is the amount of weak base that dissociates.}$$

$$x = 1.34 \times 10^{-3} = [\text{OH}^-] \qquad \text{pOH} = 2.87 \quad \text{pH} = 14.00 - 2.87 = 11.13$$

IV. Titration of a polyprotic acid

Polyprotic acids contain more than one acidic proton and more than one $\text{p}K_a$. An example is a sulfurous acid, H_2SO_3 .



It will require **2 equivalents of base** to neutralize **one equivalent of the acid**.

We can **approximate** the titration curve of 40 mL of 0.10 M sulfurous acid with 0.10 M NaOH fairly easily, by making some simplifying assumptions.

1. The pH of the acid in water will be determined solely by the K_{a1} .

In this example, the initial pH is given by:

$$1.4 \times 10^{-2} = \frac{[\text{H}^+][\text{HSO}_3^-]}{[\text{H}_2\text{SO}_3]} = \frac{x^2}{(0.1 - x)} \quad x = [\text{H}^+] = 0.031$$

[Note: In this case, the 5% approximation is not good, and you have to solve for x either using the quadratic equation or the method of successive approximations.]

$$\text{pH} = -\log 0.031 = \mathbf{1.51}$$

10.6 KINETICS

Chemical kinetics is the study and discussion of chemical reactions with respect to reaction rates, the effect of various variables, re-arrangement of atoms, the formation of intermediates etc. There are many topics to be discussed, and each of these topics is a tool for the study of chemical reactions. By the way, the study of motion is called **kinetics**, from Greek kinesis, meaning movement.

At the macroscopic level, we are interested in amounts reacted, formed, and the rates of their formation. At the molecular or microscopic level, the following considerations must also be made in the discussion of the **chemical reaction mechanism**.

- Molecules or atoms of reactants must collide with each other in chemical reactions.
- The molecules must have sufficient energy (discussed in terms of activation energy) to initiate the reaction.

- In some cases, the orientation of the molecules during the collision must also be considered.

rates of reactions affected by four factors

- concentrations of reactants
- the temperature at which reaction occurs
- presence of a catalyst
- surface area of solid or liquid reactants and/or catalysts

The first order rate constant K is generally expressed as

$$K = \frac{2.303}{t} \log \frac{a}{a-x}$$

Where k= first-order rate constant

t= temperature

a= initial concentration of reactant a

x= moles of a already converted to the product after time t

the following is the C program, which calculates the first order rate constant from the kinetic data

```
#include<stdio.h>
main()
{
double t[10], u[10], K, vmax;
int i,n;
printf(" enter the number of observation\n");
scanf("%d",&n);
printf(" Enter the values of Temperature and Volume respectively\n");
for(i=0;i<n;i++)
{
scanf("%f",&t[i], &v[i]);
}
vmax= v[n-1];
printf("temp \t k\n\n");
for(i=0;i<(n-1);i++)
{
K=(2.303/t[i]*log(vmax/(vmax-v[i]));
```

}
}

10.7 RADIOACTIVE DECAY

Radioactive decay is the spontaneous breakdown of an atomic nucleus resulting in the release of energy and matter from the nucleus. Remember that a radioisotope has unstable nuclei that do not have enough binding energy to hold the nucleus together. Radioisotopes would like to be stable isotopes so they are constantly changing to try and stabilize. In the process, they will release energy and matter from their nucleus and often transform into a new element. This process, called **transmutation**, is the change of one element into another as a result of changes within the nucleus. The radioactive decay and transmutation process will continue until a new element is formed that has a stable nucleus and is not radioactive. Transmutation can occur naturally or by artificial means.

10.8 EVALUATION OF LATTICE ENERGY

Lattice Energy is a type of potential energy that may be defined in two ways. In one definition, the lattice energy is the energy required to break apart an ionic solid and convert its component atoms into gaseous ions. This definition causes the value of the lattice energy to always be positive since this will always be an endothermic reaction. The other definition says that lattice energy is the reverse process, meaning it is the energy released when gaseous ions bind to form an ionic solid. As implied in the definition, this process will always be exothermic, and thus the value for lattice energy will be negative. Its values are usually expressed with the units kJ/mol.

Lattice Energy is used to explain the stability of ionic solids. Some might expect such an ordered structure to be less stable because the entropy of the system would be low. However, the crystalline structure allows each ion to interact with multiple opposite charged ions, which cause a highly favorable change in the enthalpy of the system. A lot of energy is released as the oppositely charged ions interact. It is this that causes ionic solids to have such high melting and boiling points. Some require such high temperatures that they decompose before they can reach a melting and/or boiling point.

Born and Lande provided the following formula for the calculation of lattice energies of crystals.

$$E = -\frac{N_A M z^+ z^- e^2}{4\pi\epsilon_0 r_0} \left(1 - \frac{1}{n}\right)$$

- N_A = Avogadro constant;
- M = Madelung constant, relating to the geometry of the crystal;
- z^+ = charge number of cation
- z^- = charge number of anion
- e = elementary charge, 1.6022×10^{-19} C
- ϵ_0 = permittivity of free space
 $4\pi\epsilon_0 = 1.112 \times 10^{-10}$ C²/(J·m)
- r_0 = distance to closest ion
- n = Born exponent, typically a number between 5 and 12, determined experimentally by measuring the compressibility of the solid, or derived theoretically.

The following is the C language program, which calculates the lattice energy of crystals, when all the quantities on the right hand side of the above expression are known.

```
#include<stdio.h>
main()
{
double lattice, na, a, z , e, ro, n;
printf(" Enter the values of Avagadro Number, Modeling constant, Dist of separation of ions
in the crystal \n"0);
scanf("%f %f %f, &na,&a, &ro);
printf(" Enter the val;ues of Born Exponent, Electronic charge and ionic charge \n");
scanf(" %f %f %f , &n, &e, &z);
lattice=(na*a*z*z*e*e)*(1-1/n)/((1.112e-10)*ro);
print("Lattice Energy of the given crystal=%f Kjmol-1",lattice);
}
```

10.9 IONIC RADII

Ionic radii follow similar trends to atomic radii with one critical difference. Cations have very different radii than anions. Cations are all smaller than their neutral analogs while anions are all larger. This is easy to understand since cations have lost electrons. As a result, they have both fewer electrons in the highest energy atomic orbitals farthest from the nuclei and the remaining electrons feel a stronger pull from the nucleus. Look at Na^+ which is isoelectronic with Ne. It goes from being one of the "largest" atoms on the left-hand side of the periodic table to effectively one of the smallest with an electron configuration that is the same as neon (all the way on the right-hand side). Moreover, Ne has a nuclear charge of $Z=10$ and Na^+ has a $Z=11$. Thus the Na^+ should be smaller than Ne. However, as atomic radii and ionic radii are often defined differently then this comparison is difficult to make.

Conversely adding an electron to F to make F^- also generates an ion that isoelectronic with Ne. However, now you have added electrons and kept the number of protons constant. Thus F^- will be larger than Ne (and larger than Na^+).

The general trends will continue to hold. From top to bottom of the periodic table ions will increase in radii. However, now left to right the radius is more of a function of the number of electrons. Mg^{2+} is smaller than Na^+ . Both have 10 electrons but $Z=12$ for Mg and $Z=11$ for Na. Similarly, O^{2-} will be larger than F^- as both have 10 electrons but $Z=8$ for oxygen and $Z=9$ for fluorine.

10.10 LET US SUM UP

Computer availability has increased to the point where most general chemistry students have easy access to them. Many of these students also know something about programming. It may be time to take advantage of this and have students start programming computers to solve problems as an integral part of their general chemistry course.

10.11 CHECK YOUR PROGRESS

1. Which one of the following has the greatest Lattice Energy?
 - a) NaCl
 - b) CaCl₂
 - c) AlCl₃
 - d) KCl
2. From top to bottom in a group of periodic table atomic radii
 - a) Decrease
 - b) Increase
 - c) Remain same
 - d) None
3. From left to right in a period atomic radii
 - a) Decrease
 - b) Increase
 - c) Remain same
 - d) None
4. An increase in the number of electrons increases the
 - a) Van Der Waal forces
 - b) Enthalpy change of vaporization
 - c) Boiling points of noble gases
 - d) The melting point of metalloids only
5. Many ionic compounds have some covalent ability due to
 - a) Ion polarization
 - b) Charge polarization
 - c) Proton polarization
 - d) Electron polarization

Ans. 1. c 2. b 3. a 4. b 5. a

10.12 FURTHER READINGS

- Computational Chemistry by A. C. Norris, John Wiley
- Numerical Recipes in FORTRAN/C by W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, Cambridge University Press, 2nd Ed. 1996.

10.13 MODEL QUESTIONS

1. Explain Van der Waals equation and generate a C code that calculates the pressure of a gas when other parameters are given.
2. How we evaluate lattice energy explain in detail and write a C program to evaluate lattice energy.
3. Write a short note on:
 - a) Kinetics
 - b) Ionic Radii
 - c) Radioactive decay
 - d) pH titration

UNIT 11: USE OF COMPUTER PROGRAMMES

CONTENTS:

- 11.1 Learning Objectives
- 11.2 Introduction
- 11.3 Operate a PC
 - 11.3.1 Getting started with a computer
 - 11.3.2 Using a Computer
 - 11.3.3 Running/Installation Standard program and packages
- 11.4 Linear Regression
 - 11.4.1 Introduction
 - 11.4.2 Uses of Regression
- 11.5 X-Y plot
- 11.6 MATLAB
 - 11.6.1 Starting and stopping
 - 11.6.2 Windows and such
 - 11.6.3 Using MATLAB as a calculator
 - 11.6.4 Using MATLAB's built-in functions
 - 11.6.5 On-line help
 - 11.6.6 Suppressing output with the semicolon
- 11.7 EASYPLOT
 - 11.7.1 Introduction
 - 11.7.2 Launching Easyplot
- 11.8 DBASE
 - 11.8.1 Introduction
 - 11.8.2 dBase Programming language
 - 11.8.3 File formats
- 11.9 Word Processing Software
 - 11.9.1 Introduction
 - 11.9.2 MS Word
 - 11.10 ChemDraw
- 11.10 Let Us Sum Up
- 11.11 Answer to Check Your Progress

11.1 LEARNING OBJECTIVE

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

- Operate a personal computer. Running/Installation Standard programs and packages in the computer
- Introduction and uses of Linear Regression

- Students understand basic concepts of computer application tools such as X-Y plot, MATLAB, EASYPLOT, DBASE, Word processing software like MS-Word, ChemDraw.

11.2 INTRODUCTION

A computer is more than just another household appliance. The vast amount of information and possibilities can be overwhelming. But you can accomplish a lot with a computer, and using one can be a good experience. Let's walk through getting started with your first computer.

11.3 OPERATE A PC

11.3.1 GETTING STARTED WITH A COMPUTER

A computer is more than just another household appliance. The vast amount of information and possibilities can be overwhelming. But you can accomplish a lot with a computer, and using one can be a good experience. Let's walk through getting started with your first computer.

Turning on a computer for the first time can be different from one computer to the next. Your experience could be different from this lesson. It's OK to ask someone for help.

If you're using a desktop computer, you'll need to make sure that the keyboard, mouse, and monitor are plugged into the computer case before you continue.

Turning on a computer

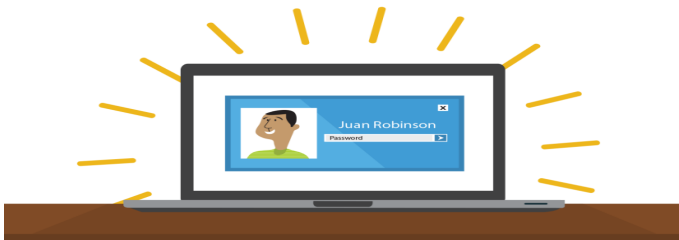
The very first step is to **turn on** the computer. To do this, locate and press the **power** button. It's in a different place on every computer, but it will have the universal power button symbol (shown below).



Figure 5.1 Turning point

Once turned on, your computer takes time before it's ready to use. You may see a few different displays flash on the screen. This process is called **booting up**, and it can take anywhere from 15 seconds to several minutes.

Once the computer has **booted up**, it may be ready to use, or it may require you to **log in**. This means identifying yourself by typing your username or selecting your profile, then typing your password. If you've never logged in to your computer before, you may need to **create an account**.



The keyboard and mouse

You interact with a computer mainly by using the **keyboard** and **mouse**, or a **trackpad** on laptops. Learning to use these devices is essential to learning to use a computer. Most people find it comfortable to place the keyboard on the desk directly in front of them and the mouse to one side of the keyboard.



Figure 5.2 keyboard and mouse

The mouse controls the **pointer** on the screen. Whenever you move the mouse across the desk, the pointer will move in a similar manner. A mouse usually has two buttons, which are referred to as the left button and the right button. You will often interact with the computer by moving the mouse pointer over something on the computer screen, then clicking one of the buttons.



On laptops, you can use the **trackpad**, located below the keyboard, instead of a mouse. Simply drag your finger across the trackpad to move the **pointer** on the screen. Some trackpads do not have buttons, so you'll either press or tap the trackpad to click.

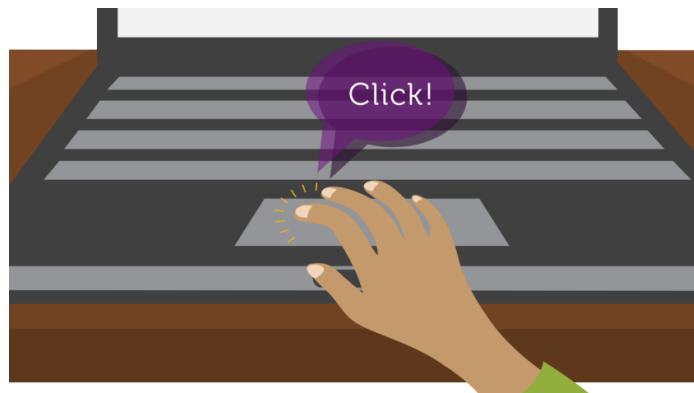
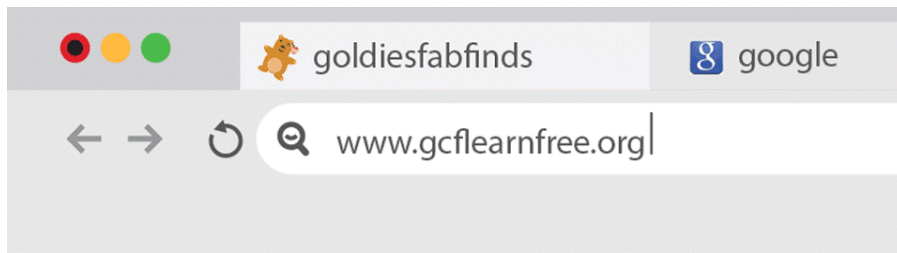


Figure 5.3 trackpad,

The keyboard allows you to type letters, numbers, and words into the computer. Whenever you see a flashing vertical line—called the **cursor**—you can start typing.



Note that the **mouse pointer** is also called a **cursor**, but it is shaped differently. The **keyboard cursor** is also called the **insertion point**.

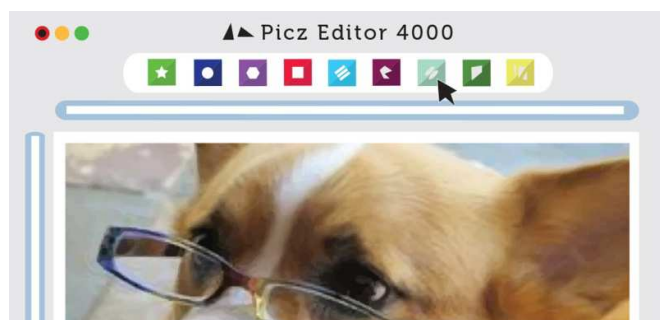
11.3.2 USING A COMPUTER

The main screen you'll start from is the **desktop**. This is sort of like the main menu or a table of contents. From here, you can access the programs and features you need to use your computer.

Icons are used to represent the different files, applications, and commands on your computer. An icon is a small image that's intended to give you an idea at a glance of what it represents, like a logo. Double-clicking an icon on the desktop will open that application or file.



A **button** is a command that performs a specific function within an application. The most commonly used commands in a program will be represented by buttons.



Menus are organized collections of commands and shortcuts. Click a **menu** to open it and display the commands and shortcuts within. Then click an item in the **menu** to execute it.



When you open an application or folder, it is displayed in its own **window**. A **window** is a contained area—like a picture within a picture—with its own menus and buttons specific to that program. You can rearrange multiple **windows** on the desktop and switch between them.

11.3.3 RUNNING/INSTALLATION STANDARD PROGRAM AND PACKAGES

Installation (or **setup**) of a computer program (including device drivers and plugins), is the act of making the program ready for execution. Because the process varies for each program and each computer, programs (including operating systems) often come with an *installer*, a specialized program responsible for doing whatever is needed for their installation. Installation may be part of a larger software deployment process.

Installation typically involves code being copied/generated from the installation files to new files on the local computer for easier access by the operating system. Because the code is generally copied/generated in multiple locations, uninstallation usually involves more than just erasing the program folder. For example, registry files and other system code may need to be modified or deleted for a complete uninstallation.

11.3.3.1 TYPES OF INSTALLATION

Attended installation

On Windows systems, this is the most common form of installation. An installation process usually needs a user who attends it to make choices, such as accepting or declining an end-

user license agreement (EULA), specifying preferences such as the installation location, supplying passwords or assisting in product activation. In graphical environments, installers that offer a wizard-based interface are common. Attended installers may ask users to help mitigate the errors. For instance, if the disk in which the computer program is being installed was full, the installer may ask the user to specify another target path or clear enough space in the disk.

Silent installation

An installation that does not display messages or windows during its progress. "Silent installation" is not the same as "unattended installation" (see below): All silent installations are unattended but not all unattended installations are silent. The reason behind a silent installation may be convenience or subterfuge. Malware is almost always installed silently.

Unattended installation

An installation that is performed without user interaction during its progress or with no user present at all. One of the reasons to use this approach is to automate the installation of a large number of systems. An unattended installation either does not require the user to supply anything or has received all necessary input prior to the start of installation. Such input may be in the form of command line switches or an *answer file*, a file that contains all the necessary parameters. Windows XP and most Linux distributions are examples of operating systems that can be installed with an answer file. In an unattended installation, it is assumed that there is no user to help mitigate errors. For instance, if the installation medium was faulty, the installer should fail the installation, as there is no use to fix the fault or replace the medium. Unattended installers may record errors in a computer log for later review.

Headless installation

Installation performed without using a computer monitor connected. In attended forms of headless installation, another machine connects to the target machine (for instance, via a local area network) and takes over the display output. Since a headless installation does not need a user at the location of the target computer, unattended headless installers may be used to install a program on multiple machines at the same time.

Scheduled or automated installation

An installation process that runs at a preset time or when a predefined condition transpires, as opposed to an installation process that starts explicitly on a user's command. For instance,

a system administrator willing to install a later version of a computer program that is being used can schedule that installation to occur when that program is not running. An operating system may automatically install a device driver for a device that the user connects. (See plug and play.) Malware may also be installed automatically. For example, the infamous Conficker was installed when the user plugged an infected device to their computer.

Clean installation

A clean installation is one that is done in the absence of any interfering elements such as old versions of the computer program being installed or leftovers from a previous installation. In particular, the clean installation of an operating system is an installation in which the target disk partition is erased before installation. Since the interfering elements are absent, a clean installation may succeed where an unclean installation may fail or may take significantly longer.

Network installation

Network installation, shortened net install, is an installation of a program from a shared network resource that may be done by installing a minimal system before proceeding to download further packages over the network. This may simply be a copy of the original media but software publishers which offer site licenses for institutional customers may provide a version intended for installation over a network.

An *installation program* or *installer* is a computer program that installs files, such as applications, drivers, or other software, onto a computer. Some installers are specifically made to install the files they contain; other installers are general-purpose and work by reading the contents of the software package to be installed.

The differences between a package management system and an installer are:

Criterion	Package manager	Installer
Shipped with	Usually, the operating system	Each computer program
Location of installation	One central installation database	It is entirely at the discretion of the installer. It could be a file within the app's folder, or among the operating

information		system's files and folders. At best, they may register themselves with an uninstallers list without exposing installation information.
Scope of maintenance	Potentially all packages on the system	Only the product with which it was bundled
Developed by	One package manager vendor	Multiple installer vendors
Package format	A handful of well-known formats	There could be as many formats as the number of apps
Package format compatibility	Can be consumed as long as the package manager supports it. Either newer versions of the package manager keep supporting it or the user does not upgrade the package manager.	The installer is always compatible with its archive format if it uses any. However, installers, like all computer programs, may be affected by software rot.

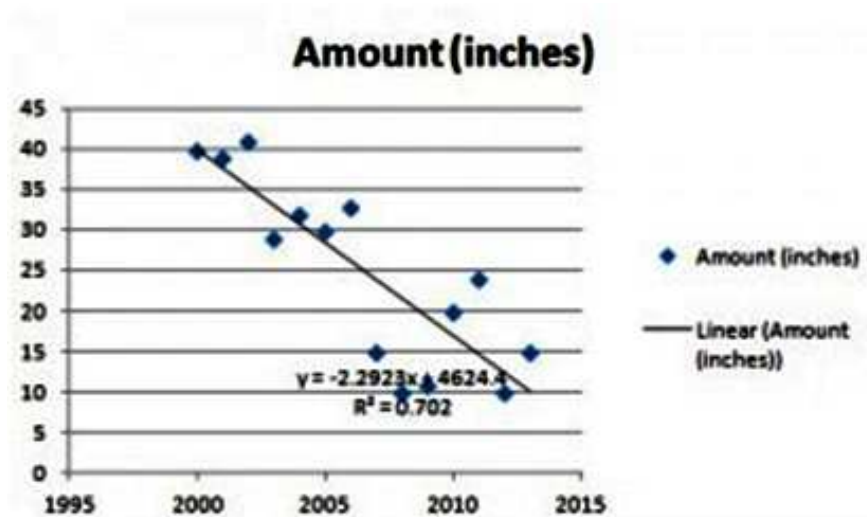
11.4 LINEAR REGRESSION

11.4.1 INTRODUCTION

In statistics, it's hard to stare at a set of random numbers in a table and try to make any sense of it. For example, global warming may be reducing average snowfall in your town and you are asked to predict how much snow you think will fall this year. Looking at the following table you might guess somewhere around 10-20 inches. That's a good guess, but you could make a *better* guess, by using regression.

Year	Amount (inches)
2000	40
2001	39
2002	41
2003	29
2004	32
2005	30
2006	33
2007	15
2008	10
2009	11
2010	20
2011	24
2012	10
2013	15

Essentially, regression is the “best guess” at using a set of data to make some kind of prediction. It’s fitting a set of points to a graph. There’s a whole host of tools that can run a regression for you, including Excel, which I used here to help make sense of that snowfall data:



Just by looking at the regression line running down through the data, you can fine-tune your best guess a bit. You can see that the original guess (20 inches or so) was way off. For 2015, it looks like the line will be somewhere between 5 and 10 inches! That might be “good

enough”, but regression also gives you a useful equation, which for this chart is:
 $y = -2.2923x + 4624.4$.

What that means is you can plug in an x value (the year) and get a pretty good estimate of snowfall for any year. For example, 2005:

$y = -2.2923(2005) + 4624.4 = 28.3385$ inches, which is pretty close to the actual figure of 30 inches for that year.

Best of all, you can use the equation to make predictions. For example, how much snow will fall in 2017?

$y = 2.2923(2017) + 4624.4 = 0.8$ inches.

Regression also gives you an R squared value, which for this graph is 0.702. This number tells you how good your model is. The values range from 0 to 1, with 0 being a terrible model and 1 being a perfect model. As you can probably see, 0.7 is a fairly decent model so you can be fairly confident in your weather prediction!

the regression equation that shows how the set of predictor variables can be used to predict the outcome. The simplest form of the equation with one dependent and one independent variable is defined by the formula $y = c + b*x$, where y = estimated dependent score, c = constant, b = regression coefficients, and x = independent variable.

11.4.2 USES OF REGRESSION

Three major uses for regression analysis are (1) causal analysis, (2) forecasting an effect and (3) trend forecasting. Other than correlation analysis, which focuses on the strength of the relationship between two or more variables, regression analysis assumes a dependence or causal relationship between one or more independent variables and one dependent variable.

Firstly, the regression might be used to identify the strength of the effect that the independent variable(s) have on a dependent variable. Typical questions are what is the strength of the relationship between dose and effect, sales and marketing spend, age and income.

Secondly, it can be used to forecast effects or impact of changes. That is, the regression analysis helps us to understand how much the dependent variable change with a change in one or more independent variables. Typical questions are, “how much additional Y do I get for one additional unit X?”

Thirdly, regression analysis predicts trends and future values. The regression analysis can be used to get point estimates. Typical questions are, “what will the price for gold be in 6 months from now?” “What is the total effort for a task X?”

11.5 X-Y PLOT

The XY Graph block displays an X-Y plot of its inputs.

The block has two scalar inputs. The block plots data in the first input (the x -direction) against data in the second input (the y -direction). This block is useful for examining limit cycles and other two-state data. Data outside the specified range does not appear.

Parameters

x-min

Specify the minimum x -axis value.

x-max

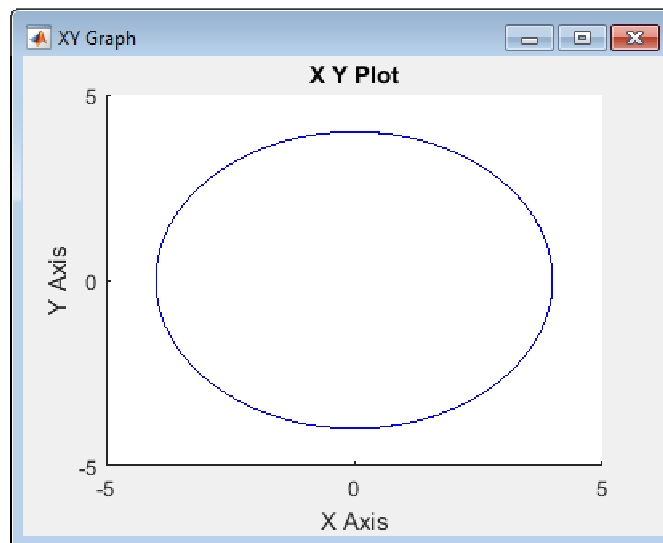
Specify the maximum x -axis value.

y-min

Specify the minimum y -axis value.

y-max

Specify the maximum y -axis value.



11.6 MATLAB

11.6.1 INTRODUCTION

MATLAB is very useful for making scientific and engineering plots. You can create plots of known, analytical functions, you can plot data from other sources such as experimental measurements, you can analyze data, perhaps by fitting it to a curve, and then plot a comparison. MATLAB also has powerful built-in routines for drawing contour and three-dimensional surface plots $z = f(x,y)$ data.

This section is an introduction to MATLAB by way of example. First, some screenshots of MATLAB are shown to help you get oriented to using MATLAB on your computer. Then some simple MATLAB commands are presented that you are encouraged to try. The syntax and general use of these commands are described in detail later.

Starting and stopping

If you are using a DOS or Unix computer, start MATLAB by typing "Matlab" at the command prompt. If you are using a Macintosh, or a computer running Windows, double-click on the MATLAB icon. Refer to the MATLAB User's guide for additional information to help you get started on your computer.

To stop MATLAB type "quit" at the MATLAB prompt, i.e.

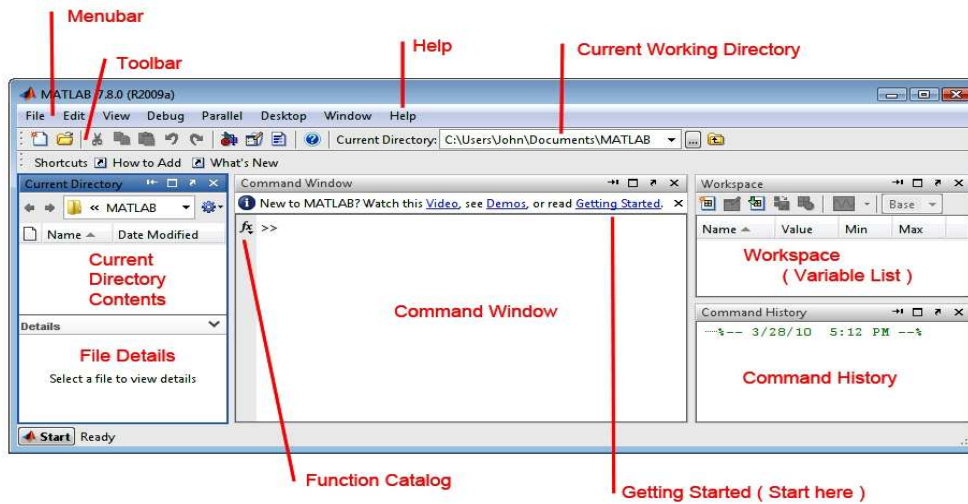
```
>> quit
```

Windows and such

MATLAB evolved as an interface to a powerful library of linear algebra routines programmed in FORTRAN. Though it is a very modern computer package, MATLAB's roots are still evident in the way that you interact with it. The most basic way to get MATLAB to do work is to type a command in the Command Window.

- Command window
- Figure window(s)
- Editor window(s)

The MATLAB Work Environment



Using MATLAB as a calculator

MATLAB can be used as an expression evaluator. To do this you simply type a mathematical expression into the command window. The command window prompt is `>>`. To enter an expression, type it after the prompt (correct any mistakes by backspacing) and press return. MATLAB prints the result back to the command window. For example, to "evaluate" pi type

```
>> pi
```

MATLAB responds with

```
Answer =
```

```
3.1416
```

pi is a built-in MATLAB variable. The preceding command is actually a request for MATLAB to print the value of this variable.

Try the following examples

```
>> sin(pi/4)
```

```
>> 2^(log2(4))
```

```
>> sqrt(9)
```

Assigning Values to Variables

MATLAB allows you create variables on the fly. To create a variable just use it on the left-hand side of an equal sign. The following examples show how to assign values to three variables, x, y, and z. It also shows the MATLAB response to the assignment statements.

```
>> x = 5
```

```
x = 5
```

```
>> y = pi/4
```

```
y = 0.7854
```

```
>> z = y + x^0.25
```

```
z = 2.2807
```

Note that z exists only as a numerical value, not as the formula with which it was created.

Use the who command to list the currently active variables. For the preceding session, this results in

```
>> who
```

Your variables are:

```
answer    x      y      z
```

```
leaving 7433616 bytes of memory free.
```

To print the value of a variable that is already defined just type its name at the prompt.

```
>> z
```

```
z = 2.2807
```

The variables section describes how to work with MATLAB variables

Built-in Functions and Variables

MATLAB has many very powerful built-in functions. As the preceding example demonstrated these functions include the familiar set found on many hand-held calculators. Here is just a tiny sample of other MATLAB functions

Function Name	Meaning
bessel	Bessel function of the first kind
erf	error function
gamma	gamma function
chol	Cholesky factorization of a matrix
det	determinant of a matrix
lu	LU decomposition of a matrix
svd	singular value decomposition of a matrix
plot	plot $y = f(x)$ data
surf	3-D shaded surface plot
contour	contour plot of $z = f(x,y)$ data
movie	play recorded movie frames

A sample of built-in MATLAB functions

You can also create your own functions. To do this you type MATLAB commands into a plain text file. The file must have the extension ".m", and because of this MATLAB functions are often referred to as "m-files". Creating and using functions is described in the Programming Matlab section.

MATLAB also has some special variables and functions. You have already encountered the "ans" variable which is automatically created whenever a mathematical expression is not assigned to another variable.

The built-in eps variable is used as a tolerance for internal calculations. Initially, eps is equal to machine epsilon. You can reset the value of eps, but this is not recommended unless you are very sure you know what you are doing.

The real max, real min, Inf and NaN built-in variables are used to handle floating point exceptions. All numerical variables in MATLAB are stored as 32-bit floating point numbers. This corresponds to "double precision" on most computers (supercomputers and some high-end workstations being the exception). The Inf and NaN values appear if a floating point exception occurred during the calculations.

In general, you should not try to (re)assign the values of any built-in variables listed in the table below.

Variable Name	Meaning
ans	value of an expression
eps	floating point precision
pi	π (3.141592...)
realmax	largest floating point number
realmin	smallest floating point number
Inf	∞ (a number larger than realmax)
NaN	not a number

Special variables and constants in MATLAB

On-line Help (!)

MATLAB provides on-line help for all built-in functions and commands. To browse a list of help topics type

```
>> help
```

To get help on a specific topic, for example, the cosine function, type

```
>> help cos
```

Using help is one way to learn about the variety of built-in functions. The help command is most useful when you know the name of the function, but are unsure how to use it. The MATLAB manuals provide cross-referenced and indexed descriptions of all aspects of using MATLAB.

Suppressing Output with Semicolon

In the preceding examples, it was useful to have MATLAB print the results of the calculations. This is not always the case. MATLAB will print the result of every assignment operation unless the expression on the right-hand side is terminated with a semicolon. Unlike C the semicolon is not required by MATLAB syntax. Rather, it is a convenience that allows multiple statements to be executed without printing the intermediate results. For example, most lines in m-files end with semicolons because only the final results of the calculations are of interest.

To see how the semicolon works enter the following statements exactly as shown, ending each line with a carriage return. Do not enter the prompt character, >>.

```
>> x = pi/3;  
>> y = sin(x)/cos(x);  
>> z = y - tan(x);  
>> z  
>> y
```

The last two lines do not end in semicolons so MATLAB prints the results of evaluating z and y. The the values of z and y were **assigned** in the y = and z = statements. The last two lines merely printed the values of the expressions z and y.

11.7 EASYPLOT

11.7.1 INTRODUCTION

Easyplot allows you to do 2-D, 3-D, polar and multiple graphs and contour plots. You can also fit curves to your data or plot mathematical expressions. Easyplot is available on all the Physics Department computers and to Physics students across the campus via the Physics Department software. It is fast and easy to use and generates publication-quality graphs.

11.7.2 LAUNCHING EASYPLOT

Click on the Easyplot icon (blue square with lines on it). It comes up with a menu giving you a choice of plotting data or an equation. Assuming you want to plot data:

Either: type in the name of a file containing the data (they should be in columns of numbers which represent x, y, z, error bar, r, or θ . The error bar corresponding to a particular column must be in the next column. You can have as many columns as you like. Missing data points can be written as //miss)

Or: choose enter data and enter your data (in columns) in the data table. (You can use the data table to edit the data file later on if you make a mistake in entering data). Click 'plot' when you have finished.

Defining what the columns should be

Easyplot assumes that the first column is an x and all the others are y. To change this, select TOOLS then DEFINE DATA. Type in a string like “xyexyeixye” (i is for ignore). The menu explains how to define columns. You should now see the graph you are expecting.

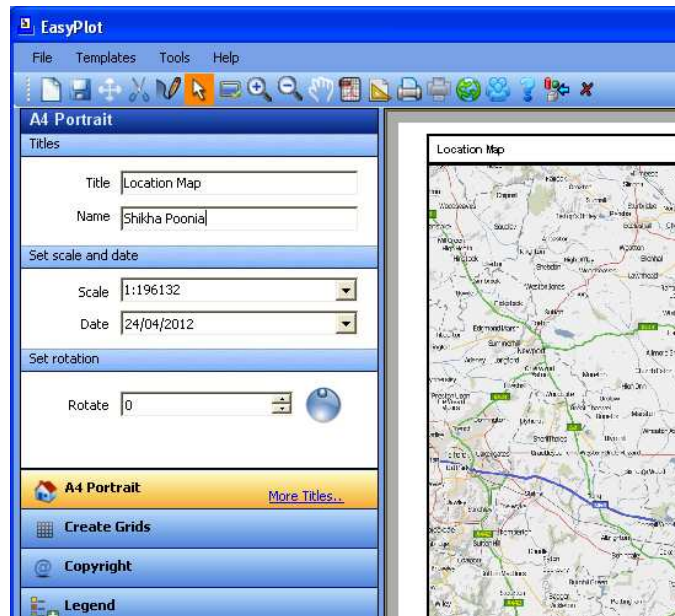
Changing the symbols :

To change the size, colour or type of a marker or tick mark, go to STYLE then RESTYLE DATA.

Tic marks scaling the axes and choosing log Axes

To change where you want tick marks and labels on the axes, or to change the scale on the axes, or select log axes, select STYLE, then AXIS SETUP (you can also click on the axes). Note that the tic marks point OUT on the screen but IN when you print - see a print preview. To change the size of annotations, use the FILE, PRINT OPTIONS menu. Click on the little squares next to the axes to enter axis titles.

Error bars.



To get error bars you must first enter a column of error bars in your data, and tell Easyplot what it is (e.g. under TOOLS, DEFINE DATA will define four columns, with errors on x and y). A column of errors must be placed immediately after the relevant data column. To show error bars, switch them to using STYLE, ERROR BARS then STYLE, RESTYLE DATA to

see them. Note that Easyplot should be set up to display error bars in the normal way, i.e. 51 means an error bar of 1 either side of 5. However, the program can interpret error bars in other ways, so always check that it is working correctly for your application. You can do this either by looking at the graph or by checking FILE, PREFERENCES and ERROR BARS. If the buttons 'error value' and 'half of entire bar' are on then the errors will be displayed in the normal way.

Multiple curves and graphs:

ADD then CURVE allows you to plot a new data set on an existing graph which means you can have several data lines on one graph. ADD then GRAPH allows you to define a new graph with new axes, title etc. You can change the size of a graph with the mouse (move one of the corners while holding the left-hand button down).

Thus, you can have several graphs on the screen at once (these will also print at the same time if you wish to do so).

Modifying data:

You may want to re-scale a curve, or mathematically transform it. Use TOOLS, TRANSFORM, CURVE, then supply an equation (e.g. $y=y*10$ or $y=\log(y)$). Under EDIT, DATA you can edit your data set - either numerically or by moving points around on the graph. You can also transform the whole graph, but if you do this you will find that Easyplot retains both the original and modified graph (the original will be hidden behind the modified one: make it smaller to see both). Then any style changes you may have made will be found only in the original.

Printing:

FILE PRINT - modify plot using OPTIONS if required, and use s to see what the graph will look like (or use print preview). Then type P to send to the printer. If you need to prepare lots of high-quality graphs, printing to a Postscript printer is very much faster. To get colour plots, select an appropriate printer and if necessary choose the colour output print option. ALWAYS check your plot using the print preview before sending to the colour printer. NOTE: If you would like to include an Easyplot graph for putting into a word processing package, just copy the whole graph to the clipboard (EDIT, COPY) and paste into the document.

Entering text, lines, and arrows:

Easyplot allows you to type text directly on the screen. Use the ADD ANNOTATION menu for text. Once the text is written you can move it by 'dragging' with the mouse. Note that you change the size of lettering on the graph using the print options. Use ADD LINE of lines. For lines, the cursor will change to a pencil shape. To return to the normal cursor press Esc. Lines can be adjusted by clicking and 'dragging' one end of a line with the mouse while the cursor is in its normal (+) mode. Hold the ALT key down as you draw the line to put an arrow at the end where you release the mouse button.

Changing the type of graph:

STYLE menu. You can select a 3-D option for instance - see help pages for more info. on this.

Panic

There are ways in which Easyplot will get into a tangle. You may accidentally zoom or select a curve for instance and find you can't do anything. If EDIT-UNDO doesn't work, then try OPTIONS AUTORANGE to recover control. The Esc key can also get you out of trouble. Of course, you should always save your file at regular intervals.

Save and restore:

Save a plot using FILE SAVE. To call it up again use ADD GRAPH.

Greek characters, subscripts, superscripts:

When you're typing text (e.g. axis titles, annotations) Easyplot comes up with a menu which you can use to select various options. Strangely, once you have chosen superscript (for instance) you can't go back to normal text. So it's easier to type everything in normally then go through selecting text with the mouse and converting to whatever you want. Alternatively (and more quickly) use the keyboard strokes: Greek: type \a to get alpha etc. It does upper and lower case. Superscripts: x^2 gives x²; to get more figures in the exponent you have to use the menu method. Subscripts: O_3 gives O₃.

Getting a table to label the different lines in multi-line plots:

This is called a legend. Use OPTIONS LEGEND to introduce a box and click on curves to put something in it.

Highlighting curves:

Use the right mouse button. You will see a little box in the lower left-hand corner of the screen with a sample of the curve in it. Keep hitting the right mouse button until you see the curve you want. If you keep the button down, the curve glimmers. To delete a curve, highlight and hit delete.

Multiple axes:

ADD X-AXIS or Y-AXIS. If you choose multiple y-axes the current axis will have a little y above it. To choose a particular axis (e.g. to edit it) click on it using the mouse. To plot data against a given axis, highlight and, while holding down the right mouse button, type y or x as appropriate. To enter the second title, choose EDIT, TITLE after selecting the second axis.

Zoom:

You may zoom in (and zoom in only) on any part of a graph by clicking and 'dragging' with the mouse while the cursor is in normal (+) mode. The position of the cursor when you press the left mouse button and its position when you leave go define opposite corners of a rectangle for zooming in. In order to zoom out, you must choose OPTIONS RANGE/AUTORANGE.

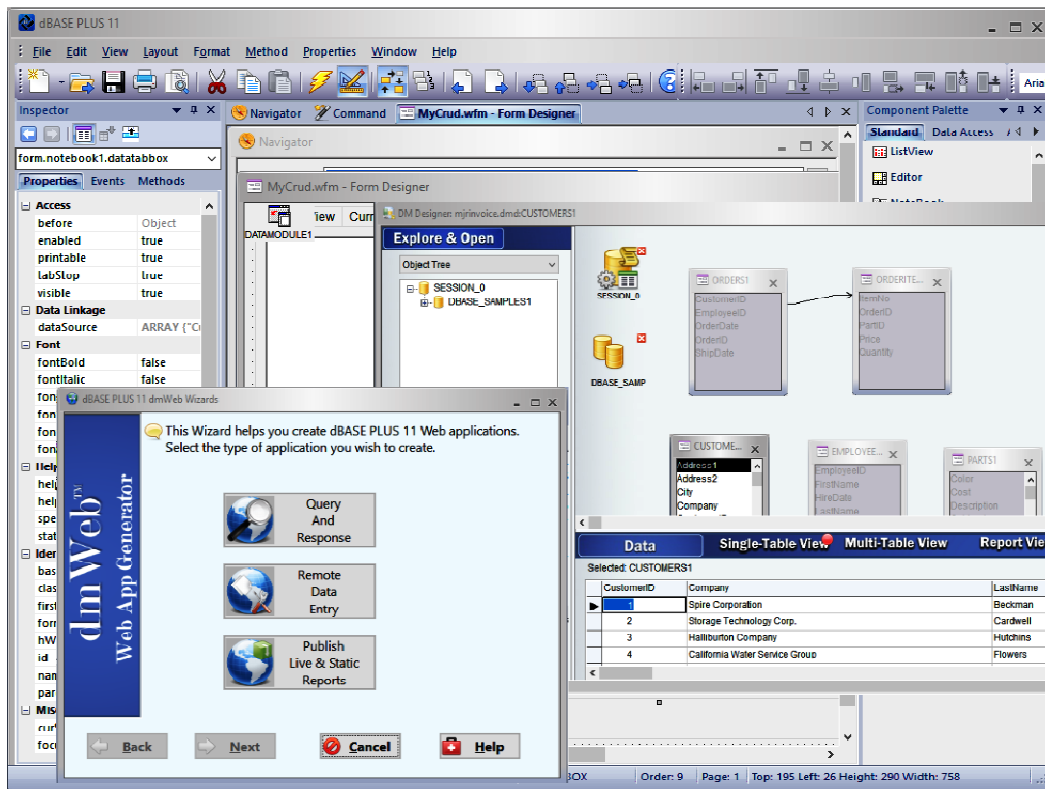
11.8 DBASE

11.8.1 INTRODUCTION

dBase (also stylized **dBASE**) was one of the first database management systems for microcomputers and the most successful in its day. The dBase system includes the core database engine, a query system, a forms engine, and a programming language that ties all of these components together. dBase's underlying file format, the .dbf file, is widely used in applications needing a simple format to store structured data.

dBase was originally published by Ashton-Tate for microcomputer operating system CP/M in 1980 and later ported to Apple II and IBM PC computers running DOS. On the PC platform, in particular, dBase became one of the best-selling software titles for a number of years. A major upgrade was released as dBase III, and ported to a wider variety of platforms, adding UNIX, and VMS. By the mid-1980s, Ashton-Tate was one of the "big three" software

publishers in the early business software market, the others being Lotus Development and WordPerfect.



dBase has evolved into a modern object-oriented language that runs on 32-bit Windows. It can be used to build a wide variety of applications including web apps hosted on a Windows server, Windows rich client applications, and middleware applications. dBase can access most modern database engines via ODBC drivers. DOS versions can still run on the latest Windows/Linux machines using the Virtual Machine (VM) DOSbox, or the later variant dbDOS. In 2015, dBase, LLC. Introduced a set of new utilities called dbfUtils. The utilities include: dbfExport(TM) to export data from .dbf files to CSV, Microsoft(R) Excel, and XML, dbf Import(TM) allows users to import CSV, Microsoft(R) Excel, and XML files into a .dbf. dbf Compare(TM) allows 2 .dbf tables to compare differences and gives an easy way to synchronize tables. Finally, the new dbfInspect(TM) allows users to deep-dive into the .dbf files and their contents.

11.8.2 DBASE PROGRAMMING LANGUAGE

For handling data, dBase provided detailed procedural commands and functions to open and traverse records in data files (e.g., USE, SKIP, GO TOP, GO BOTTOM, and GO recno), manipulate field values (REPLACE and STORE), and manipulate text strings (e.g., STR() and SUBSTR()), numbers, and dates. Its ability to simultaneously open and manipulate multiple files containing related data led Ashton-Tate to label dBase a "relational database" although it did not meet the criteria defined by Dr. Edgar F. Codd's relational model; it could more accurately be called an application development language and integrated navigational database management system that is influenced by relational concepts.

The dBase product used a runtime interpreter architecture, which allowed the user to execute commands by typing them in a command line "dot prompt." Upon typing a command or function and pressing the return key, the interpreter would immediately execute or evaluate it. Similarly, program scripts (text files with PRG extensions) ran in the interpreter (with the DO command), where each command and variable was evaluated at runtime. This made dBase programs quick and easy to write and test because programmers didn't have to first compile and link them before running them. (For other languages, these steps were tedious in the days of single- and double-digit megahertz CPUs.) The interpreter also handled automatically and dynamically all memory management (i.e., no preallocating memory and no hexadecimal notation), which more than any other feature made it possible for a business person with no programming experience to develop applications.

Conversely, the ease and simplicity of dBase presented a challenge as its users became more expert and as professional programmers were drawn to it. More complex and more critical applications demanded professional programming features for greater reliability and performance, as well as greater developer productivity.

Over time, Ashton-Tate's competitors introduced so-called clone products and compilers that had more robust programming features such as user-defined functions (UDFs) to supplement the built-in function set, scoped variables for writing routines and functions that were less likely to be affected by external processes, arrays for complex data handling, packaging features for delivering applications as executable files without external runtime interpreters, object-oriented syntax, and interfaces for accessing data in remote database management systems. Ashton-Tate also implemented many of these features with varying degrees of

success. Ashton-Tate and its competitors also began to incorporate SQL, the ANSI/ISO standard language for creating, modifying, and retrieving data stored in relational database management systems.

Today, implementations of the dBase language have expanded to include many features targeted for business applications, including object-oriented programming, manipulation of remote and distributed data via SQL, Internet functionality, and interaction with modern devices.

The following example opens an employee table ("empl"), gives every manager who supervises 1 or more employees a 10-percent raise, and then prints the names and salaries.

```
USE empl
REPLACE ALL salary WITH salary * 1.1 FOR supervisors > 0
LIST ALL fname, lname, salary TO PRINT
* (comment: reserved words shown in CAPITALS for illustration purposes)
```

11.8.3 FILE FORMATES

A major legacy of dBase is its .dbf file format, which has been adopted in a number of other applications. For example, the shapefile format, developed by ESRI for spatial data in its PC ArcInfo geographic information system, uses .dbf files to store feature attribute data.

Microsoft recommends saving a Microsoft Works database file in the dBase file format so that it can be read by Microsoft Excel.^[28]

A package is available for Emacs to read xbase files.

LibreOffice and OpenOffice Calc can read and write all generic dbf files.

dBase's database system was one of the first to provide a header section for describing the structure of the data in the file. This meant that the program no longer required advance knowledge of the data structure, but rather could ask the data file how it was structured. There are several variations on the .dbf file structure, and not all dBase-related products and .dbf file structures are compatible. VP-Info is unique in that it can read all variants of the dbf file structure.

A second filetype is the .dbt file format for memo fields. While character fields are limited to 254 characters each, a memo field is a 10-byte pointer into a .dbt file which can include a much larger text field. dBase was very limited in its ability to process memo fields, but some other xBase languages such as Clipper treated memo fields as strings just like character fields for all purposes except permanent storage.

dBase uses .ndx files for single indexes, and .mdx (*multiple-index*) files for holding between 1 and 48 indexes. Some xBase languages such as VP-Info include compatibility with .ndx files while others use different file formats such as .ntx used by Clipper and .idx/.cdx used by FoxPro or FlagShip. Later iterations of Clipper included drivers for .ndx, .mdx, .idx and .cdx indexes.

11.9 WORD PROCESSING SOFTWARE

11.9.1 INTRODUCTION

Word processing software is used to manipulate a text document, such as a resume or a report. You typically enter text by typing, and the software provides tools for copying, deleting and various types of formatting. Some of the functions of word processing software include:

- Creating, editing, saving and printing documents.
- Copying, pasting, moving and deleting text within a document.
- Formatting text, such as font type, bolding, underlining or italicizing.
- Creating and editing tables.
- Inserting elements from other software, such as illustrations or photographs.
- Correcting spelling and grammar.

Word processing includes a number of tools to format your pages. For example, you can organize your text into columns, add page numbers, insert illustrations, etc. However, word processing does not give you complete control over the look and feel of your document. When design becomes important, you may need to use desktop publishing software to give you more control over the layout of your pages.

Word processing software typically also contains features to make it easier for you to perform repetitive tasks. For example, let's say you need to send a letter to all your customers regarding a new policy. The letter is the same for all customers except for the name and address at the top of the letter. A mail merge function allows you to produce all the letters using one template document and a table with customer names and addresses in the database.

Text editors shouldn't be confused with word processing software. While they do also allow you to create, edit and save text documents, they only work on plain text. Text editors don't use any formatting, such as underlined text or different fonts. Text editors serve a very different purpose from word processing software. They are used to work with files in plain text format, such as source code of computer programs or configuration files of an operating system. An example of a text editor would be Notepad on the Windows platform.

11.9.2 MS-WORD

Microsoft Word is a word processor developed by Microsoft. It was first released on October 25, 1983, under the name *Multi-Tool Word* for Xenix systems. Subsequent versions were later written for several other platforms including IBM PCs running DOS (1983), Apple Macintosh running Classic Mac OS (1985), AT&T Unix PC (1985), Atari ST (1988), OS/2 (1989), Microsoft Windows (1989), SCO Unix (1994), and macOS (2001). Commercial versions of Word are licensed as a standalone product or as a component of Microsoft Office, Windows RT or the discontinued Microsoft Works suite. Microsoft Word Viewer and Office Online are freeware editions of Word with limited features.

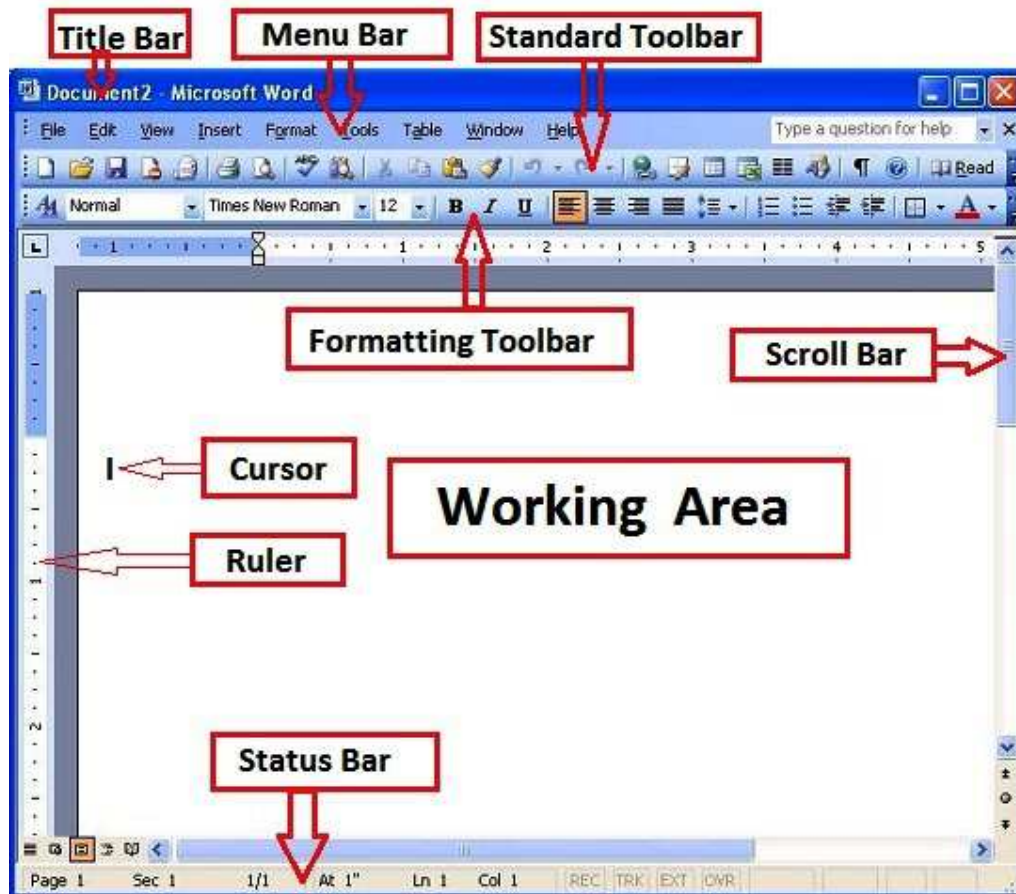
The word for Windows is available stand-alone or as part of the Microsoft Office suite. The word contains rudimentary desktop publishing capabilities and is the most widely used word processing program on the market. Word files are commonly used as the format for sending text documents via e-mail because almost every user with a computer can read a Word document by using the Word application, a Word viewer or a word processor that imports the Word format (see Microsoft Word Viewer). Word 6 for Windows NT was the first 32-bit version of the product, released with Microsoft Office for Windows NT around the same time as Windows 95. It was a straightforward port of Word 6.0. Starting with Word 95, releases of Word were named after the year of its release, instead of its version number.

Word 2010 allows more customization of the Ribbon, adds a Backstage view for file management, has improved document navigation, allows creation and embedding of screenshots, and integrates with Word Web App.

File formats

File extensions

Microsoft Word's native file formats are denoted either by a .doc or .docx filename extension.



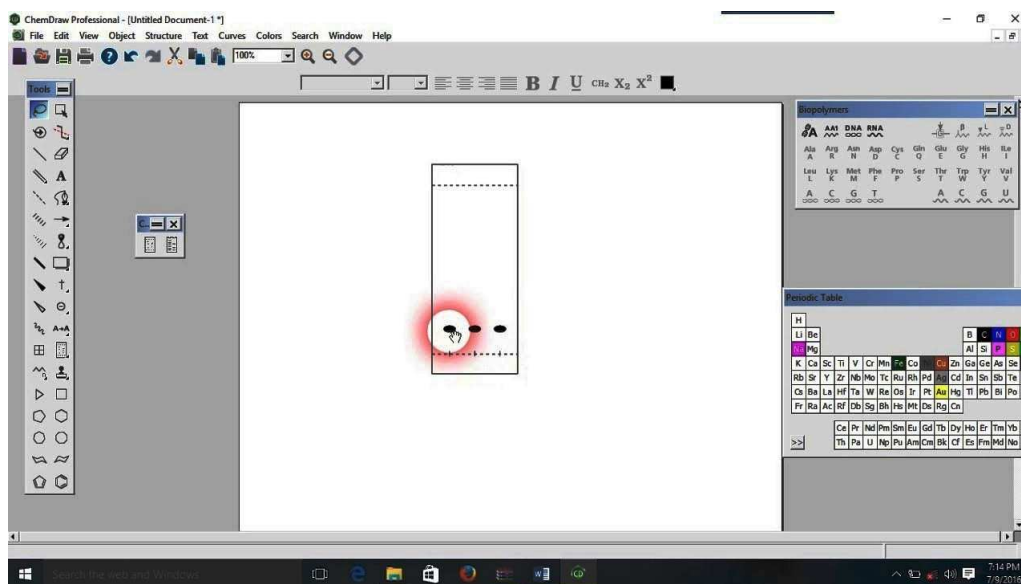
Although the .doc extension has been used in many different versions of Word, it actually encompasses four distinct file formats:

1. Word for DOS
2. The word for Windows 1 and 2; Word 3 and 4 for Mac OS
3. Word 6 and Word 95 for Windows; Word 6 for Mac OS
4. Word 97 and later for Windows; Word 98 and later for Mac OS

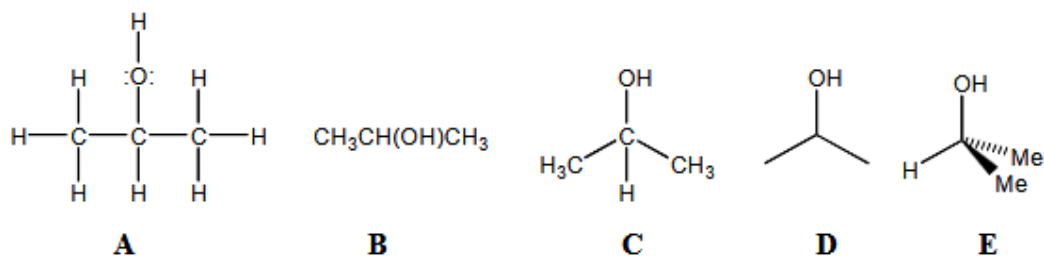
The newer .docx extension signifies the Office Open XML international standard for Office documents and is used by Word 2007 and later for Windows, Word 2008 and later for macOS, as well as by a growing number of applications from other vendors, including OpenOffice.org Writer, an open source word processing program.

11.9.3 ChemDraw

Chem Draw can be used to draw simple two-dimensional representations of organic molecules. These structures must precisely describe the connectivity of atoms. Five different



representations of 2-propanol (isopropanol or "rubbing alcohol") are shown below. While structure. Shows all bonds and lone pairs of electrons, a number of short-cuts are often used when drawing molecules. For example, -CH is used to represent the methyl groups in (B) and (C).

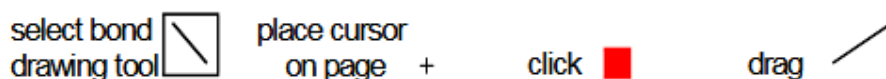


Bond-line structures (D) are often the simplest representation of organic molecules. A chain of carbon atoms is shown as a series of lines; each angle between lines and each line end is a carbon atom. The carbon atoms are not labeled. All heteroatoms (i.e., N, O, S, halogens) are labeled. Hydrogens on carbons are not shown; although it is assumed that they are present to account for the tetravalency of carbon. Since sp^3 carbon atoms are tetrahedral it is often useful to show the structure of organic molecules using wedge-and-dash structures (E). The wedge indicates a bond coming out of the page toward; the dashed bond is going into the page. In structure E the methyl groups are abbreviated as “Me”. Other common abbreviations for substituents are Et (ethyl), i-Pr (isopropyl), t-Bu (tert-butyl). Chem Draw can also be used to draw Newman and Fischer projections of organic molecules (see your textbook).

Getting Started

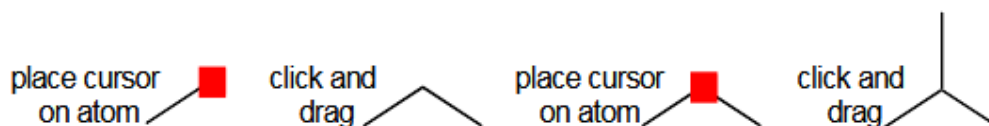
The ChemDraw screen has a vertical palette of tools on the left, a series of pull-down Menus at the top of the screen, and a blank window on which to draw. The following instructions illustrate how to draw molecules. Work through the step-by-step instructions to draw the structure of 2-propanol shown on the right.

- i. Launch ChemDraw by clicking on the program icon (or use the start menu).
- ii. First, you will draw a bond: Click on the bond drawing tool in the palette on the left-hand side of the screen. Move the cursor to the blank window, click-hold-drag- and release the mouse to draw a bond.



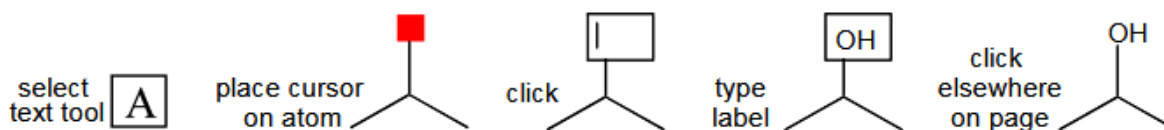
Each end of the bond represents a carbon atom until you change it to another type of atom (see below).

iii. You will now draw the other bonds: Place the cursor on one end of the bond; a black square will appear. Click-and-drag the mouse to draw a second bond. Click- and-drag from the middle atom to draw a third bond.



iv. One of the carbon atoms should be changed to a hydroxyl group: Click on the text tool; place the cursor on the atom you want to label, a black square will appear.

Click the mouse when the end of a bond is highlighted with a black square, a text box will appear. Type OH using the keyboard, and then click the mouse elsewhere on the window. Numbers entered in atom labels are automatically written as subscripts.



Moving and Rotating Structures

Draw a molecule and then move, shrink or enlarge, and rotate it using the following instructions.

i. Click on the Lasso tool. Click on the window close to a molecule you want to move; hold down the mouse, circle around the molecule and release. The circled molecule will be highlighted in a highlight box.

ii. To move the molecule: Move the cursor into the middle of the box until a hand appears. Click and-drag to move the highlighted molecule.

iii. To shrink or enlarge the molecule: Move the cursor to the bottom right corner of the highlight box (a double-headed arrow will appear). Click- and drag to shrink and enlarge the highlighted molecule.

iv. To rotate the molecule: Put the cursor at the top right corner of the highlight box (a curved arrow will appear). Click and drag to rotate the highlighted molecule.

11.10 LET US SUM UP

- Once turned on, your computer takes time before it's ready to use. You may see a few different displays flash on the screen. This process is called **booting up**.
- **Installation** (or **setup**) of a computer program (including device drivers and plugins), is the act of making the program ready for execution.
- An *installation program* or *installer* is a computer program that installs files, such as applications, drivers, or other software, onto a computer.
- Regression is the “best guess” at using a set of data to make some kind of prediction. It’s fitting a set of points to a graph.
- Regression also gives you an R squared value. This number tells you how good your model is. The values range from 0 to 1, with 0 being a terrible model and 1 being a perfect model.
- Three major uses for regression analysis are (1) causal analysis, (2) forecasting an effect and (3) trend forecasting.
- The XY Graph block displays an X-Y plot of its inputs.
- MATLAB is very useful for making scientific and engineering plots. You can create plots of known, analytical functions, you can plot data from other sources such as experimental measurements, you can analyze data, perhaps by fitting it to a curve, and then plot a comparison
- Easyplot allows you to do 2-D, 3-D, polar and multiple graphs and contour plots. You can also fit curves to your data or plot mathematical expressions.
- **dBase** (also stylized **dBASE**) was one of the first database management systems for microcomputers and the most successful in its day. The dBase system includes the core database engine, a query system, a forms engine, and a programming language that ties all of these components together.
- **Word processing software** is used to manipulate a text document, such as a resume or a report. You typically enter text by typing, and the software provides tools for copying, deleting and various types of formatting.
- **Microsoft Word** is a word processor developed by Microsoft. It was first released on October 25, 1983, under the name *Multi-Tool Word* for Xenix systems.

- ChemDraw can be used to draw simple two-dimensional representations of organic molecules. These structures must precisely describe the connectivity of atoms.

11.11 CHECK YOUR PROGRESS

Q1. State whether the following statements are true (T) or false (F):

- vi- An *installation program* or *installer* is a computer program that installs files.
- vii- Regression is the “best guess” at using a set of data to make some kind of prediction.

Q1. State whether the following statements are true (T) or false (F):

- i. The term *data analysis* is sometimes used as a synonym for data modeling.
- ii. Initial data analysis phase involves quality of data.
- iii. Measurements have no errors.
- iv. Calculating the uncertainties in the calculated values is called error propagation.
- v. When a curve is not fitted to any data the method used is least squares analysis.
- vi. Regression gives you an R squared value. The values range from 2 to 3.
- vii. MATLAB is very useful for making scientific and engineering plots.
- viii. **dBase** was one of the first payroll management systems for microcomputers.

Ans. to Q. No. 1: i) True ii) True iii) False iv) True v) False

FUTURE READINGS

- Uchenick, Gordon. "Apparatus and method for preventing unauthorized use of computer programs." U.S. Patent No. 4,458,315. 3 Jul. 1984
- Weitzman, Eben, and Matthew B. Miles. *Computer programs for qualitative data analysis*. Sage, 1995
- Pfleeger, Charles P., and Shari Lawrence Pfleeger. *Security in computing*. Prentice Hall Professional Technical Reference, 2002.

- Seber, George AF, and Alan J. Lee. *Linear regression analysis*. Vol. 936. John Wiley & Sons, 2012.
- Grant, Michael, Stephen Boyd, and Yinyu Ye. "CVX: Matlab software for disciplined convex programming." (2008).
- Forrest, John Raymond. *Easyplot, a high-level plotting system*. Diss. Oklahoma State University, 1972.
- Byers, Robert A., and B. Johnson-Grau. "Everyman's database primer. Featuring dBase III plus." (1986).
- Bangert-Drowns, Robert L. "The word processor as an instructional tool: A meta-analysis of word processing in writing instruction." *Review of Educational Research* 63.1 (1993): 69-93.

MODEL QUESTIONS

- Q.1. Explain the concept of Booting.
- Q.2. What do you mean by the term Linear Regression? Explain.
- Q.3. What is the use of X-Y plot?
- Q.4. Write short notes on
- a. MATLAB
 - b. EASYPLOT
 - c. DBASE
 - d. ChemDraw
- Q.5. What are the advantages of word processing software?
- Q.6. What do you mean by installation or setup of a computer program?