UNIT – I

LESSON – 1

Elements of Crystallography

Aim: To introduce elements of crystallography.

Objectives:

- To learn about the classification of solids based on order in the arrangement of atoms or molecules.
- To gain insight on the concept of lattice and periodicity in lattices and the symmetry considerations.
- To know about different types of unit cells of a crystal.
- To learn about Bravais lattices and classification of these in to crystal classes.

Structure:

1.1 Classification of Solid materials:

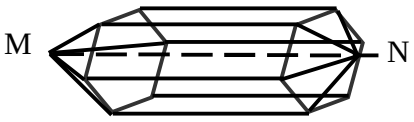
Matter is usually regarded to exist in the Solid State or in Fluid State. The fluid state is further divided into liquid and gaseous states. However, on the basis of modern concept, the matter is more conveniently divided into the condensed state and the gaseous state. The former state is subdivided into solid and liquid states. The word "Solid" means any material whose constituent particles are relatively fixed in position. (Except for thermal vibration). A solid appears as a continuous rigid body.

Experiments on solids proved they are composed of discrete basic units (atoms). In some solids the atoms are not distributed randomly, but are arranged in a highly ordered manner relative to each other. Such group of ordered atoms is known as "crystal". There are several types of crystalline structures, depending on the geometry of the atomic arrangement. So, solids are classified on the basis of their degree and type of order. There are 3 classes.

1. Crystalline
2. Semi-crystalline (poly-crystalline)
3. Non-crystalline (Amorphous).

1.2 **Single and Poly crystalline materials**: -

In general a solid is said to be crystal if the constituent particles (i.e. atoms, ions (or) molecules) are arranged in a three dimensional periodic manner. The regularity in the appearance of the crystal found in nature (or) grown in laboratory has led to believe that crystals are formed by a regular repetition of identical building blocks in three dimensional shape. Shown in fig (1.1, 1.2).



Quartz

Fig 1.1 External forms of the crystal found in nature

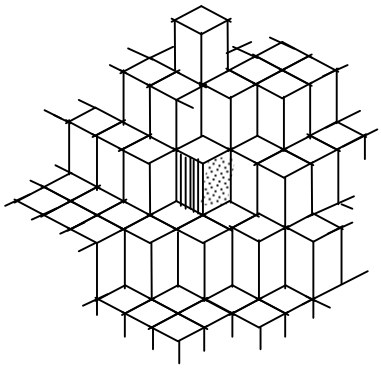


Fig 1.2 Formation of crystals by regular repetition of identical building blocks

Observations on crystals showed, that, it is bound by optically plane faces , sharp straight edges and interfacial angles. In general the relation ship is expressed as,

$$f + c = e + 2 \text{ ---------} (1.1)$$

Where "f" is the number of faces, "c" is the number of angles & "e" is the number of edges.

Example: Determine the number of edges in a quartz crystal if there are 18 faces and fourteen angles.

Solution: Given f = 18, c = 14; using equation 1.1, number of edges e= 18+14-2 = 30

Single Crystal: -

In a single crystal periodicity extends through out the material. Crystals that occur in nature or grown from solution show regular faces and sharp edges and corners. Quartz, diamond ,some gem stones and ice can be cited as natural crystals.In fact the word crystal has its root from latin which means 'ice like'. Common salt, alum, sugar and copper sulphate are some materials which can be grown very easily from solution.

Single crystals of some compounds cannot be grown easily from solution. Some such materials are grown from melt using methods like Bridgeman method. Other materials which cannot be grown efficiently by these methods are grown using vapour deposition techniques. The single crystals grown from melt or from vapour normally do not show external crystalline features but have internal periodic arrangement. Crystals of semiconducting matrials are grown using these methods.

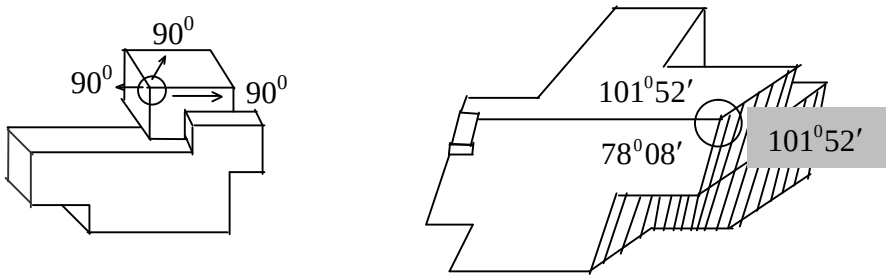


Fig. 1.3 Examples of monocrystals (single crystals).
(a) Common salt (b) Quartz crystal

A single crystal has a typical feature i.e. its anisotropy. It means the difference in its physical properties in different directions. It has a sharp melting point.

Poly crystalline (or) Semi crystalline solids: -

A solid consisting of many crystallites grown together in the form of an interlocking mass, oriented randomly, and separated by well defined boundaries is said to be a poly crystalline solid. (Fig – 1.4)

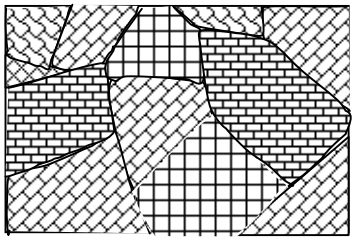


Fig 1.4 Polycrystalline aggregates separated by well defined boundaries.

Ex: rock, sand, metals, salts etc.

Due to random distribution of crystallites, a poly crystalline solid is isotropic, i.e. its properties are same in all directions.

In case of amorphous solids the order is limited to a few molecular distances. Glass is a good example for amorphous materials. It may be noted that glass has some other features and all amorphous materials are not glasses.

Now we discuss the some important definitions regarding crystal physics.

We know that crystal consist of a microscopic particles arranged in a three dimensional periodic manner. In order to describe the periodicity in crystals, in 1848 Bravais introduced the concept of space lattice.

In a perfect crystal, there is a regular arrangement of atoms. This periodicity in the arrangement generally varies in different directions. It is very convenient to imagine points in space about which these atoms are located. These imagined points in space are called “lattice points”. The totality of such points forms a “crystal lattice” (or) “space lattice”. The two dimensional (2d) lattice points are shown in fig 1.5.

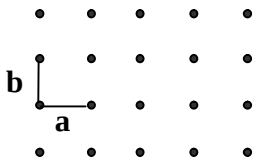


Fig 1.5 Two-Dimensional array of lattice points.

1.3 Symmetry consideration:

The definite ordered arrangement of the faces and edges of a crystal is known as “crystal symmetry”. It is a powerful tool for the study of the internal structure of crystal.

A crystal possesses different symmetries (or) symmetry elements. They are described by certain operations. A symmetry operation is one that leaves the crystal and its environment invariant, i.e. the body becomes indistinguishable from its initial configuration after symmetry operation. The geometrical locus about which a group of finite operations act is known as "symmetry element". A crystalline solid can have the following symmetry elements.

(i) Pure translation
(ii) Proper rotation
(iii) Reflection
(iv) Inversion
(v) Improper rotation

(i) Pure translation: -

A two dimensional space lattice is shown in fig1.6(a) The distance between any two nearest neighbours along the x direction is "a" and along the y direction is "b". A perfect crystal maintains this periodicity in both x and y directions from $-\infty$ to ∞ i.e. the periodicity of atoms A, B and C are equivalent. In other words, to an observer located at any of the atomic sites, the crystal appears exactly the same. From the above point we can say that a crystal possesses a translation symmetry. If the crystal is translated by any vector joining two atoms say R in fig.1.6 (a), the crystal appears exactly the same as before the translation. Therefore crystal remains invariant under any translation. A translation operation is a displacement vector represented in terms of the basis vectors a, b and c as

$\mathbf{T} = n_1\mathbf{a} + n_2\mathbf{b} + n_3\mathbf{c}$ where n_1, n_2 and n_3 are integers.

(ii) Proper rotation (through an angle ϕ): -

The proper rotation is shown in fig1.6(b). Let us imagine a line (or) axis passing through the center & normal to the fig.1.6(b), so that the J's are represented by a rotation through any angle $\phi = 2\pi/n$ about the axis of rotation, the axis is said to have n-fold symmetry.

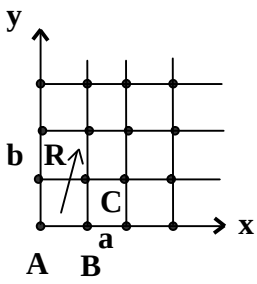


Fig 1.6(a) Translation in two dimensional lattice

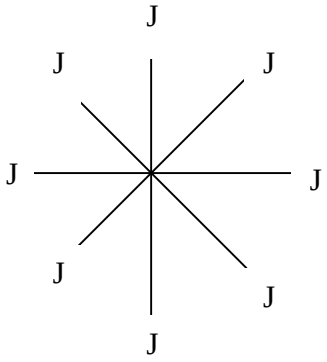


Fig 1.6(b) Rotation

Because of reticular structure of crystals, only 1-,2-,3-,4- & 6- fold rotational symmetries are possible. Crystalline solid cannot possess either 5-fold (or) any other rotational symmetry higher than 6-fold.

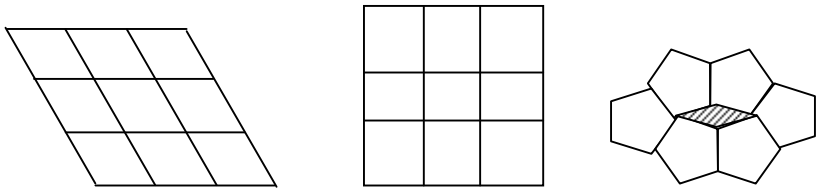


Fig 1.6(c) Some Possible and Non-existent Symmetry axes

(iii) Reflection: -

The proper reflection is shown in fig 1.6(d). In the fig, we find that a plane transforms left-handed object into a right-handed one and vice-versa. The element of symmetry in this case is known as a symmetry plane (or) a mirror plane and symbolically represented by the letter "m". i.e. A plane in a cell such that, when a mirror reflection in this plane is performed, the cell remains invariant.

iv) Inversion center (center of symmetry):

The proper inversion is shown in fig 1.6(e). A cell has an inversion center if there is a point at which the cell remains invariant when the mathematical transformation $\gamma \rightarrow -\gamma$ is performed on it. This is similar to reflection, with the difference that reflection occurs in a plane of mirror, while inversion is equivalent to reflection through point called inversion center (or) center of symmetry. The inversion center has the property of inverting all space through points.

Ex: All bravais lattices have inversion symmetry. A Non-bravais lattice may or may not have an inversion center, depending on the symmetry of the basis.

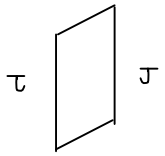


Fig 1.6(d) reflection

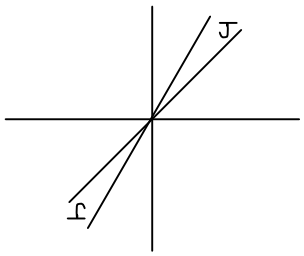


Fig 1.6(e) Inversion

1.4 Two dimensional crystalline lattices:

Introduction to Space lattice

In general infinite number of lattices are possible because there is no restriction on the length a, b of the lattice translations and on angle ϕ between them. Such a lattice is known as oblique lattice and is shown in fig 1.7(a).

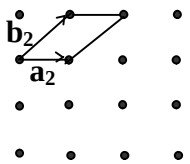


Fig 1.7 (a) Obliqe

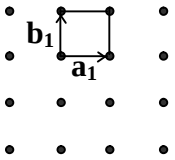


Fig 1.7 (b) square

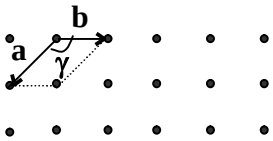


Fig1.7(c) Hexagonal lattice

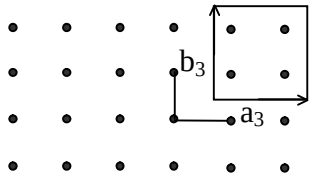


Fig1.7(d) Rectangular lattice

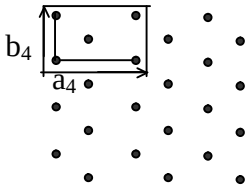


Fig 1.7(e) centered rectangle

Oblique lattice is invariant under the rotation 2π /n (n= 1 & 2) about any lattice point. However, this can also be invariant under the rotation 2π /n with n=3,4,6 (or) mirror reflection if some suitable restrictions are imposed on a, b & ϕ. These symmetry elements in turn put restrictions on the shape of the lattice. The resulting lattices are known as special lattices. They are

1.Square lattice $|a| = |b|$, ϕ **(or)** γ **= 90^0** **2. Hexagonal lattice** $|a| = |b|$, γ **= 120^0**

3. Rectangular lattice $|a| \neq |b|$, γ **= 90^0** **4. Centered rectangular**

Lattice axes are shown for both the primitive cell and the rectangular unit cell for which $|a| \neq |b|, \gamma = 90^0$

Table 1.1 Bravais lattice in two dimensions

S.No.	Lattice type	Conventional unit cell	Axes & angles	Point group symmetry about lattice point
1.	Oblique	Parallelogram	a ≠ b; ϕ (or) $\gamma \neq 90^0$	2mm
2.	Square	Square	a = b; $\gamma = 90^0$	4mm
3.	Hexagonal	Rhombus	a = b; $\gamma = 120^0$	6mm
4.	Primitive rectangular	Rectangle	a ≠ b; $\gamma = 90^0$	2mm
5.	Centered rectangular	Rectangle	a ≠ b; $\gamma = 90^0$	2mm

There are in all five Bravais lattices, ten point groups and seventeen space groups in two-dimensions. Out of the five Bravais lattices, one is general and other four are obtained by exhausting the feasible axial relationships between a and b and the relative orientations of the two. The general lattice is termed as oblique lattice.

1.5 Three dimensional crystalline lattices:

In previous section, we observe that in two dimensions suitable restrictions on lattice translations & angles allow only five types of lattices. By extending the same idea to a 3-dimensional case and applying the similar restrictions on the lattice translations a, b &c and angles α, β and γ we obtain a Bravais lattice.

A three-dimensional unit cell is defined by vectors a, b &c representing its edges (or) crystal axes and the angles α, β &γ. The numbers of Bravais lattices is 14 with 32 points groups and 230 space groups. Based on relationships between a, b & c in terms of magnitude and relative orientations α, β &γ, the 14 types of unit cells are grouped into seven different classes of crystal lattices. They are Triclinic, monoclinic, orthorhombic,

tetragonal, cubic, trigonal & hexagonal. The table 1.2 gives the 7 classes of crystal lattices.

Table 1.2 Crystal classes

Crystal system	Restriction on conventional cell, axes & angles	Associate lattice		Characteristic symmetry element
		Number	Symbol	
Triclinic	$a \neq b \neq c$; $\alpha \neq \beta \neq \gamma \neq 90^0$	1	P	None
Monoclinic	$a \neq b \neq c$; $\gamma = \alpha = 90^0 \neq \beta$	2	P, C	One 2-fold-rotation axis.
Orthorhombic	$a \neq b \neq c$; $\alpha = \beta = \gamma = 90^0$;	4	P, C, F, I	Three 2-fold rotation axis
Tetragonal	$a=b \neq c$; $\alpha = \beta = \gamma = 90^0$;	2	P, I	One 4-fold rotation axis
Cubic	a=b=c; $\alpha = \beta = \gamma = 90^0$;	3	P, I, F	Four 3-fold rotation axis
Trigonal	a=b=c; $\alpha = \beta = \gamma < 120^0 \neq 90^0$;	1	P	One 3-fold rotation axis
Hexagonal	$a=b \neq c$; $\alpha = \beta = 90^0$; $\gamma = 120^0$;	1	P	One 6-fold-rotation axis.

P stands for Primitive, C for Side centered (or) base centered, I for Body centered and F for Face centered

1.6 Periodicity in crystals – Basic definitions:

The study of crystal physics aims to interpret the macroscopic properties in terms of properties of the microscopic particles of which the solid is composed. The geometric form & physical properties of crystalline solids can be determined by using X-rays, electron beams & neutron beams.

The main difference between the crystals & other solids is that there is a arrangement of atoms, ions (or) molecule in three dimensional periodic manner in crystals and this is absent in other solids. In order to explain this periodicity Bravais introduced the concept of space lattice in 1848. Space lattice is obtained by simply considering the translation of an object (J) to a finite distance (a) & when repeated systematically along the three crystallographic directions i.e. x, y & z. The one-dimensional periodic array of object is shown in below fig. 1.8a.

It is easy to represent this periodicity by replacing each object in the array with a point. It should be remembered that a point is an imaginary infinitesimal spot in space i.e. lattice points are imaginary i.e.Fig 1.8a represents real objects whereas Fig 1.8b shows the same array interms of imaginary lattice points.

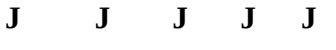


Fig 1.8(a)

Fig 1.8(b)

If translation "a" is combined with another non collinear translation "b" then a two dimensional array is obtained as shown in 1.8 c and 1.8d

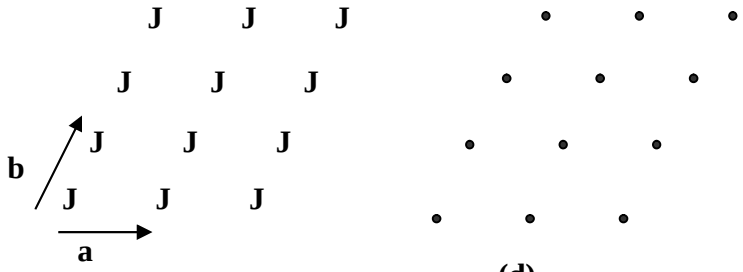


Fig 1.8 Two dimensional array of: (c) Objects, (d) Points; a plane lattice

Similarly if "a" & "b" is combined with a third non-coplanar translation c, then a three dimensional array is obtained, which is shown in fig1.9.

The characteristic feature of the space lattice is that the environment around any one point is identical. The location of any lattice point can be defined as, $T=n_1a+n_2b+n_3c$. Where n_1, n_2, n_3 are arbitrary integers.

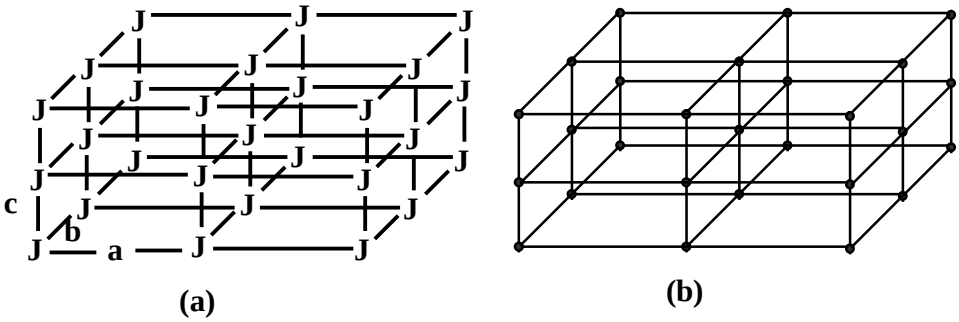


Fig 1.9Three dimensional array of: (a) Objects, (b) Points; a space lattice

Basis:

In preceding section we studied the periodicity & lattice points. Now it is essential to distinguish a lattice from a crystal. A crystal structure is formed only when a group of atoms (or) molecules are attached identically to each lattice points. This group of atoms (or) molecules is called basis. Basis is identical in composition, arrangement and orientation, which is repeated periodically in space to form the crystal structure.

∴ Lattice + Basis = crystal structure

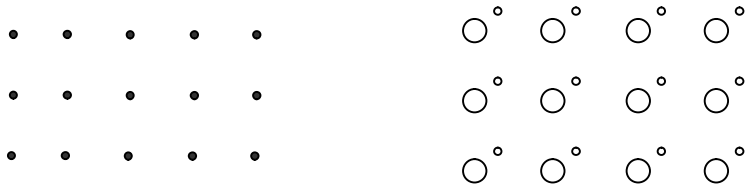


Fig 1.10(a) space lattice

Fig 1.10 (b) Basis containing two different atoms

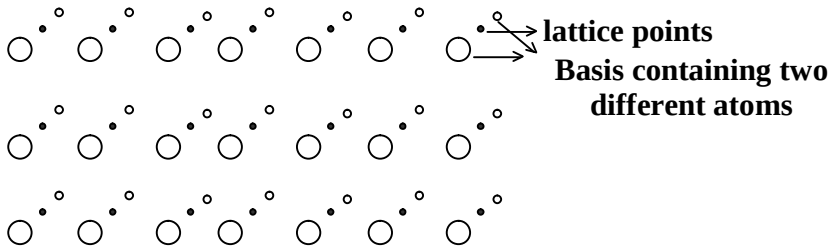


Fig 1.10 (c) Crystal structure (lattice + Basis)

1.7 Unit cell: -

The atomic order in crystalline solids indicates that the small groups of atoms form a repetitive pattern. Thus, in describing crystal structure, it is often convenient to subdivide the

structure into small repeat entities called unit cells. I.e. in every crystal some fundamental grouping of particles is repeated. Such collection of particles is called a "Unit Cell". Unit cells for most crystals are parallelopiped (or) cubes having 3 set of parallel faces.

It can be used to represent the crystal symmetry. It is a "building block" of the crystal structure. The fig 1.11(a) shows a unit cell of a 3 dimensional crystal lattice. The lattice is made-up of a repetition of unit cells, and a unit cell can be completely described by the three vectors **a**, **b**, **c** & the angles between them (α, β, γ). If the values of these intercepts & interfacial angles are known, we can easily determine the form & actual size of the unit cell.

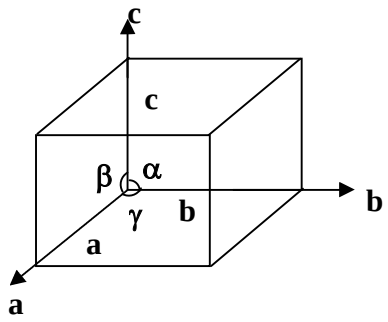


Fig 1.11(a) Lattice parameter of a unit cell

Table1.3 unit cell volume of different lattice types

Lattice type	Volume
Cubic	a^3
Orthorhombic	abc
Tetragonal	a^2c
Hexagonal	$\sqrt{3}\,a^2c/2$
Rhombohedral	$a^3\sqrt{1-3\cos^2\alpha+2\cos^3\alpha}$
Monoclinic	$abc\sin\beta$
Triclinic	$abc\sqrt{1-\cos\alpha-\cos\beta-\cos\gamma+2\cos\alpha\cos\beta\cos\gamma}$

1.8 Seven Crystal Systems:

In 3 dimensional case by applying the restrictions on the lattice translations a, b, c and angles α, β, γ, one can verify that only 7 crystal groups (or) basic systems are possible. They are collectively known as Bravais lattices. In order to specify the given

arrangement of points in a space lattice (or) of atoms in a crystal, it is customary to define its co-ordinates with reference to a set of axes chosen with its origin at a lattice point. The three axes & angles are defined as shown in fig 1.11(b). Each space lattice has a convenient set of axes, however only seven different systems of axes have been found to be sufficient for representing all Bravais lattices.

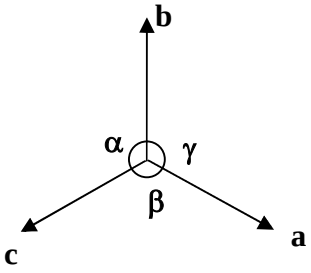


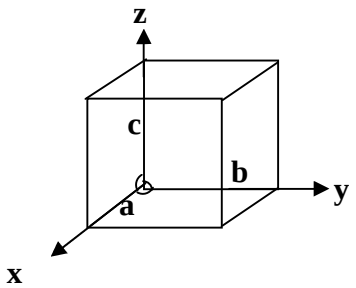
Fig 1.11(b) The crystallographic axes and the corresponding angles

They are 1.Triclinic 2.Monoclinic 3.Orthorhombic 4.Tetragonal
5.Trigonal 6.Hexagonal 7.Cubic

Explanation of each crystal system is given below.

(1) Cubic crystal system: -

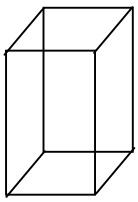
The crystal axes are perpendicular to one another & the repetitive interval is the same along all the 3 axes. Cubic crystallites may be simple (or) body centered (or) face centered.



$a = b = c$; $\alpha = \beta = \gamma = 90^0$;

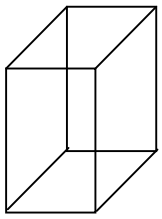
(2) Tetragonal crystal system:-

The crystal axes are perpendicular to one another. The repetitive intervals along two axes are the same, but the interval along the third axis is different. Tetragonal lattices may be simple (or) body centered.



$a = b \neq c$; $\alpha = \beta = \gamma = 90^0$;

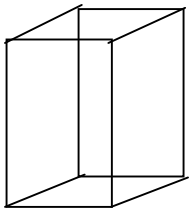
(3) Orthorhombic crystal system:-



The crystal axes are perpendicular to one another, but the repetitive intervals are different along all the three axes. Orthorhombic lattices may be simple, base centered, body centered (or) face centered.

$a \neq b \neq c; \quad \alpha = \beta = \gamma = 90^0;$

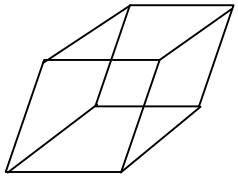
(4) Monoclinic Crystal system:-



Two of the crystal axes are perpendicular to each other, but the third is obliquely inclined. The repetitive intervals are different along all the three axes. Monoclinic lattices may be simple (or) base centered.

$a \neq b \neq c; \quad \alpha = \beta = 90^0 \neq \gamma$

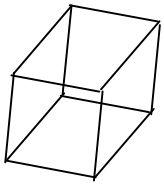
(5) Triclinic crystal system:-



None of the crystal axes is perpendicular to any of the others, and the repetitive intervals are different along all the three axes.

$a \neq b \neq c; \quad \alpha \neq \beta \neq \gamma \neq 90^0$

(6) Trigonal (or) rhombohedral crystal system:-



The three axes are equal in length and are equally inclined to each other at an angle other than 90^0.

(7) Hexagonal crystal system:-

Two of the crystal axes are 60^0 apart while the third is perpendicular to both of them. The repetitive intervals are the same along the axes that are 60^0 apart, but the interval along the third axis is different. The properties of seven are shown in table 1.4.

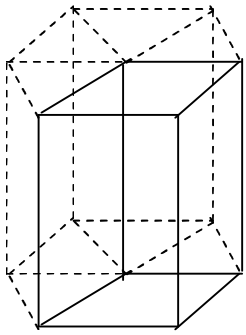


Table 1.4

S. No.	Crystal system	Axial length of unit cell (a,b&c)	Inter axial angles (α, β & γ)	Example
1	Cubic	a=b=c	$\alpha=\beta=\gamma=90^0$	Au, Cu, NaCl
2	Tetragonal	a=b≠c	$\alpha=\beta=\gamma=90^0$	TiO_2, SnO_2
3	Orthorhombic	a ≠b ≠c	$\alpha=\beta=\gamma=90^0$	KNO_3, $BaSO_4$
4	Monoclinic	a ≠b ≠c	$\alpha=\beta=90^0\neq\gamma$	$CaSO_4$, $2H_2O$,$FeSO_4$
5	Triclinic	a ≠b ≠c	$\alpha\neq\beta\neq\gamma\neq90^0$	$K_2Cr_2O_7$
6	Trigonal	a=b=c	$\alpha=\beta=\gamma\neq90^0$	As, Sb, Bi, Calcite
7	Hexagonal	a=b≠c	$\alpha=\beta=90^0$&$\gamma=120^0$	SiO_2, Zn, Mg, Cd

1.9 The Fourteen Bravais Lattices:-

A three-dimensional unit cell is defined by vectors a, b, &c representing its edges (or) crystal axes and the angles α, β & γ defined as shown in fig 1.12

The number of bravais lattices is 14 with 32-point groups & 230 space groups. Based on the relationship between a, b &c in terms of magnitude and relative orientation α, β & γ. The 14 types are grouped into seven different classes of crystal lattices. A description of the 14 bravais lattices of 3 dimensions along with the axial relationship for the class of crystal lattices to which each belongs is entered in table 1.5.

The simple cube (SC):-

The unit cell is cube having one atom (or) molecule at each corner. So there are eight atoms (or) molecules at eight corners of the cube. Since each corner atom is shared by eight surrounding cubes, share of each cube comes to one eighth of an atom, shown in fig 1.13a.

$\therefore$ Total No. of atoms = $\frac{1}{8}\times 8$ = 1 atom.

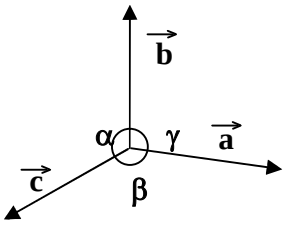


Fig 1.12 Notation for angles between the crystal axis

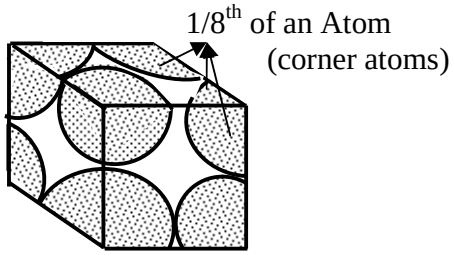


Fig 1.13(a) The simple cube

The Body Centered cube (BCC):-

The unit cell of body centered cube structure is shown in fig 1.13(b). It has eight corner atoms & one center atom. The number of atoms belonging to this cube are (a) One center atom &

(b) $\frac{1}{8}\times 8$ = 1 corner atom

$\therefore$ Total No. of atoms per cube =1+1=2

The Face centered cube (FCC): -

The unit cell of face centered cube structure is shown in fig 1.13c. It has 6 face centered and eight corner atoms. The six face centered atoms at six faces of the cube are shared by their adjacent cubes. Hence, a total of 6/2=3 such atoms belong to the cube. As each corner atom is shared by eight surrounding cubes, the share of each cube comes to one eighth of an atom.

$\therefore$ Total no. of atoms per cube =3 + $\frac{1}{8}\times 8$=3+1=4.

Similarly monoclinic lattice has two types simple and base centered. Orthorhombic has four types three as mentioned above for cubic and base centered, Tetragonal has two types of lattices. These are shown in fig 1.14.

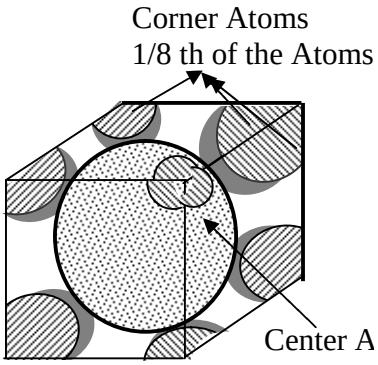


Fig 1.13(b) Body centered cube

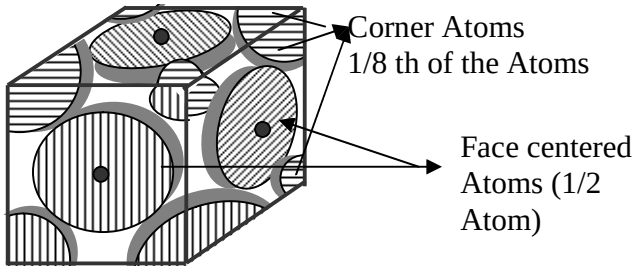


Fig 1.13(c) Face centered cube

Table 1.5

Class	Bravais lattice	No. of lattices	Unit cell Characteristics
Triclinic	Simple	1	$a \neq b \neq c$ & $\alpha \neq \beta \neq \gamma \neq 90^0$
Monoclinic	Simple Base centered	2	$a \neq b \neq c$ & $\alpha = \beta = 90^0 \neq \gamma$
Orthorhombic	Simple base-centered body-centered	4	$a \neq b \neq c$ & $\alpha = \beta = \gamma = 90^0$
Tetragonal	Simple Body centered	2	$a = b \neq c$ & $\alpha = \beta = \gamma = 90^0$
Cubic	Simple body-centered face-centered	3	$a = b = c$ & $\alpha = \beta = \gamma = 90^0$
Trigonal	Simple	1	$a = b = c$ & $\alpha = \beta = \gamma \neq 90^0$
Hexagonal	Simple	1	$a = b \neq c$ & $\alpha = \beta = 90^0$ $\gamma = 120^0$

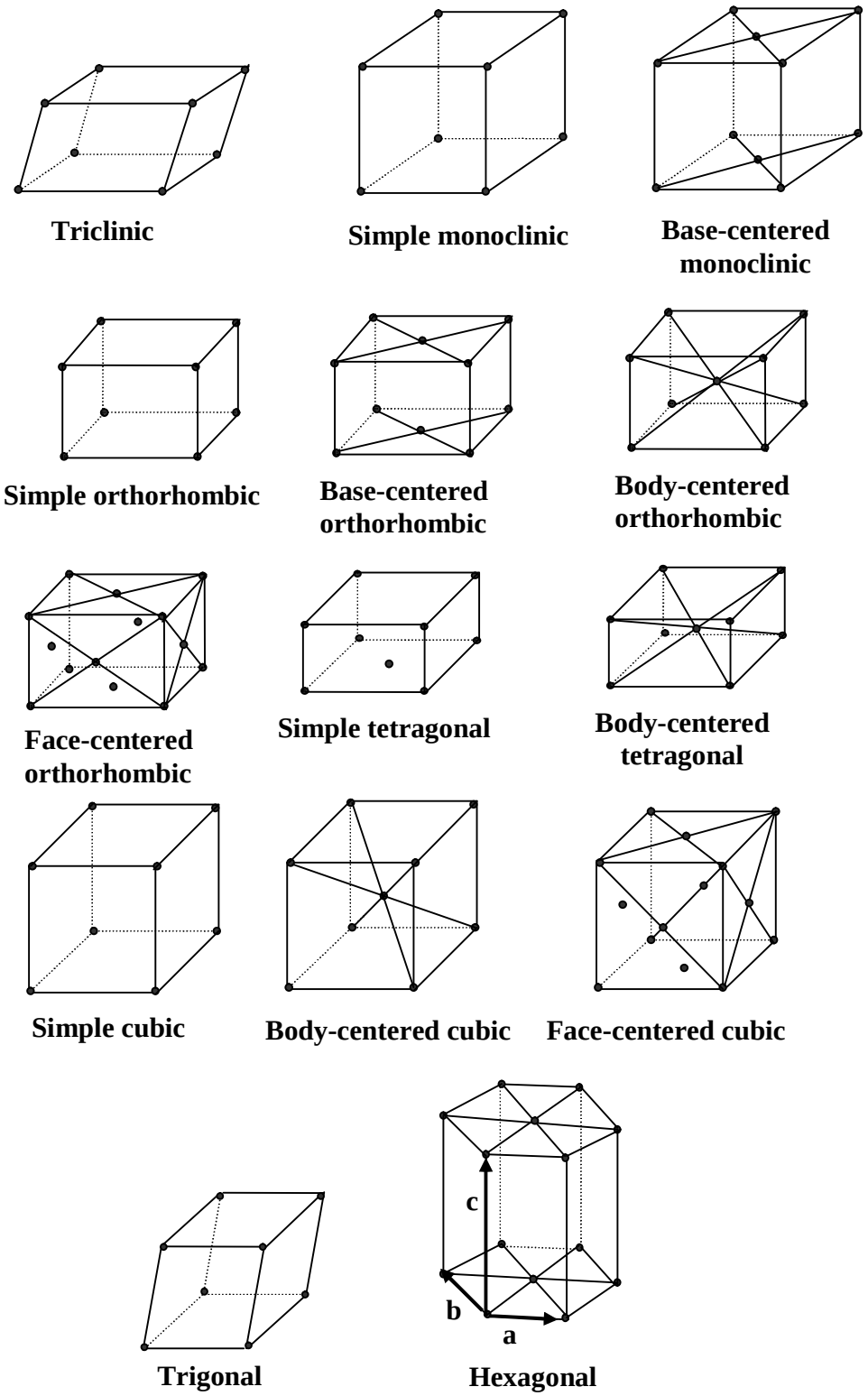


Fig 1.10 The 14 Bravais lattices grouped into the 7 crystal systems

1.10 Summary:

1. Crystal = Space lattice + Basis

Space lattice: An infinite periodic array of points in space

Basis: An atom or an identical group of atoms attached to every lattice point. It is identical for every lattice point in terms of composition, relative orientation and separation.

2. Lattice points are connected to each other by the translation vector $t = n_1a + n_2a + n_3c$, where n_1, n_2, n_3 are integers and a, b, c are primitive translation vectors (the axes of a primitive cell).

3. Primitive cell: A minimum volume unit cell having only one lattice point.

4.

Crystal types	Bravais lattices	Point groups	Space groups
Two-dimensional crystals	5	10	17
Three-dimensional crystals	14	32	230

5. Identical atoms may crystallize into the closest-packed structure either with HCP or with FCC symmetry, having the pattern of layers as:

HCP: AB AB AB AB …..
FCC: ABC ABC ABC …..

For both the structures:(i) The number of nearest neighbors (coordination number) = 12.

(ii) The packing fraction = 0.743

1.11 Key words:

Crystal – Amorphous – polycrystalline – basis – lattice – Bravais lattice – space lattice – lattice point – unit cell primitive unit cell – non primitive unit cell – crystal systems – point group – space group – symmetry operations – translation – rotation – reflection – inversion – center of symmetry – proper – rotation – improper rotation – Triclinic – monoclinic – orthorhombic – tetragonal – rhombohedral –trigonal – hexagonal close packed – cubic – simple – face centered – body centered – coordination number.

1.12 Review questions:

1.What are the essential differences between crystalline and non-crystalline materials'. Explain the concepts of lattice, basis and crystal structure. How are they related?

2. Mention and explain with examples the types of lattices in cubic system. How many lattice points are there in each of these lattices?

3. Explain the following terms used in crystallography

a. Primitive cell
b. Unit cell
c. Packing factor
d. Coordination number.

4. Define space lattice. Name the seven types of crystal systems and give the relation of length of axes and relation of angles between the axes of a unit cell in each type.

5. Explain the physical basis of classifying crystals into: (a) 14 Bravais lattice (b) 7 systems and (c) 32 point groups.

6. What is meant by 'Symmetry elements' in crystals? Discuss the various types of symmetry elements present in a cube.

7. What are symmetry operations? Explain the meaning of a 'n-fold rotation axis and a 'n-fold screw axis'. Prove that five fold axis of rotation is not compatible with a lattice.

8. What is Bravais lattice? What is the maximum number of Bravais lattices possible? How will you account for the existence of thousands of structures from these lattices?

9. The end-centered orthorhombic is one of the Bravais lattice but the end-centered tetragonal is not. Give reasons.

10. The primitive cell of fcc lattice is rhombohedral. Why then is the rhombohedral lattice included separately in the Bravais list?

11. What are point group and space group? Give their number for two and three dimensional lattices. List all the point groups of a two-dimensional lattice.

12. Show that the base centered and face centered tetragonal does not give any new Bravais lattice.

1.13 Text and Reference Books:

1. Elements of Solid state Physics by J.P.Srivatsava (PHI)
2. Elements of Solid state Physics by Ali Omar (Pearson Education)
3. Solid state Physics by S.L. Kakani and C. Hemrajani (S Chand)
4. Solid state Physics by M.A.Wahab (Narosa)
5. Solid state Physics by C.Kittel
6. Solid State Physics by C.J.Dekker

UNIT – I
LESSON – 2

Elements of crystallography-II

Aim: To introduce the concept of point groups and space groups and space groups and discuss about some common crystal structures.

Objectives:

- To acquaint with the nomenclature of crystal directions and planes.
- To acquaint with the concept of point groups and their nomenclature.
- To learn about the common crystal forms under cubic class.

Structure:

2.1 Nomenclature of crystal directions and planes:-

In a crystal there exists directions and planes, which contain a large concentration of atoms. For the crystal analysis, it is necessary locate these directions and planes.

Crystal directions:-

Consider the straight line passing through the lattice points A, B in fig.2.1(a).

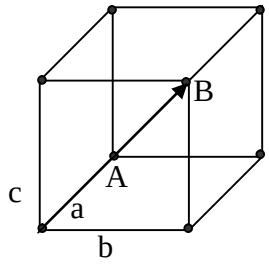


Fig 2.1(a) The [111] direction is a cubic lattice.

To specify the direction of straight line we proceed as follows. We choose one lattice point on the line as an origin, the point A. Then we choose the lattice vector joining A to any point on the line, say point B. This vector can be written as

$$\mathbf{R} = n_1\mathbf{a} + n_2\mathbf{b} + n_3\mathbf{c}$$

Generally square brackets are used to indicate a direction. The direction is now specified by the integral triplet [n_1 n_2 n_3]. If the numbers n_1, n_2 & n_3 have a common factor, this factor is removed, i.e the triplet [n_1 n_2 n_3] is the smallest integer of the same relative ratios. In the above fig the direction shown is the [111] direction

When we speak of a direction, we do not mean one particular straight line, but a whole set of parallel straight lines, which are completely equivalent by virtue of the translational symmetry.

Crystal planes:-

The crystal plane may be regarded as made-up of an aggregate of a set of parallel equidistant planes, passing through the lattice points. For example for a given lattice, the lattice planes can be chosen in different ways as shown in fig 2.1(b).

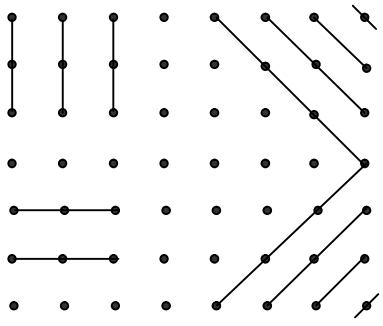


Fig 2.1(b) Different crystal planes

2.2 Miller indices:-

It is difficult to designate the different planes in a crystal. To avoid this difficulty Miller evolved a method to designate a set of parallel planes in a crystal by three numbers (h, k, l) know as Miller indices.

The steps for the determination of Miller indices of a set of parallel planes are illustrated as follows

(1) Determine the intercepts made by the plane along the three crystallographic axes (x, y, z)

i.e x y z

2a 3b c

pa qb rc

where p = 2; q = 3; r = 1;

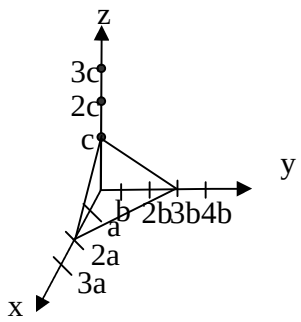


Fig 2.2(a)

(2) Express the intercepts as multiples of the unit cell dimensions, (or) lattice parameters along the axes i.e.

$$\frac{2a}{a} \quad \frac{3b}{b} \quad \frac{c}{c} \Rightarrow 2 \quad 3 \quad 1$$

(3) Determine the reciprocals of these numbers i.e. $\rightarrow \frac{1}{2} \quad \frac{1}{3} \quad \frac{1}{1}$

(4) Reduce the reciprocals to the smallest set of integral number and enclose them in brackets $\Rightarrow \frac{6}{2} \quad \frac{6}{3} \quad \frac{6}{1} \Rightarrow (3 \quad 2 \quad 6)$

Thus Miller indices may be defined as the reciprocal of the intercepts made by the plane on the crystallographic axes when reduced to smallest number.

Important features of Miller indices of crystal planes are,

a) All the parallel equidistant planes have the same Miller indices.
b) A plane parallel to one of the Co-ordinate axes has an intercept of infinity.
c) If the Miller indices of two planes have the same ratio i.e.(844) and (422) or (211) then the planes are parallel to each other.
d) If (h, k, l) are the Miller indices of a plane, then the plane cuts the axes into h, k & l equal segments respectively.

The Miller indices define a set of parallel planes or a set of parallel planes. If (h, k, l) are the Miller indices of a plane, then, the plane cuts the axes into h, k and l equal segments respectively. If a plane cuts an axis on the negative side of the origin, the corresponding index is negative and is indicated by placing a bar above the index; $(h\bar{k}l)$. The cube faces of a cubic crystal are (100); (010); (001); $(\bar{1}00); (0\bar{1}0)$ and $(00\bar{1})$. Planes equivalent by symmetry are denoted by curly

brackets around miller indices; the cube faces are {100}. Regarding direction the x axis is the [100] direction; the –y axis is the $[0\bar{1}0]$ direction. A full set of equivalent directions is denoted this way: <uvw>. In cubic crystals a direction [uvw] is perpendicular to a plane (uvw) having the same indices, but this is not generally true in other crystal systems.

The positions of points in a unit cell are specified in terms of lattice coordinates, in which each coordinate is a fraction of the axial length, a, b or c, in the direction of the coordinate, with the origin taken at the corner of unit cell. Thus the coordinates of the central point of a cell are $\frac{1}{2}\frac{1}{2}\frac{1}{2}$, and the face center positions are $\frac{1}{2}\frac{1}{2}0$; $0\frac{1}{2}\frac{1}{2}$; $\frac{1}{2}\frac{1}{2}0$.

2.3 Spacing between planes of same Miller indices: -

It is necessary to know the inter planar distance between planes labeled by the same Miller indices for the x-ray diffraction from the crystal. Let us call this distance d_{hkl}.

Now, we shall derive a formula for the spacing between two parallel planes in a given cell with the help of figure shown in **fig 2.3.**

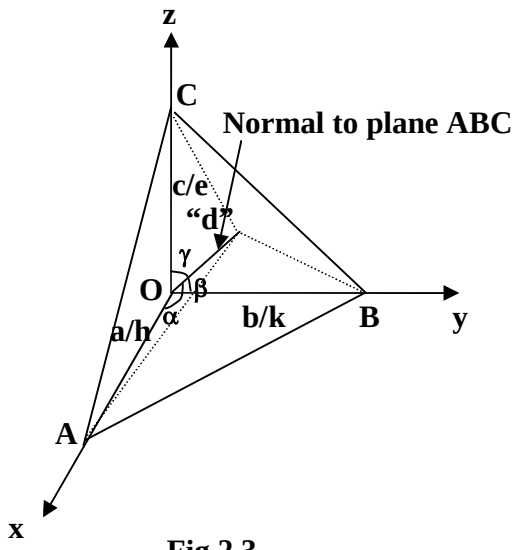


Fig 2.3

For convenience, we shall take a simple unit cell in which co-ordinate axes are orthogonal, therefore they are mutually perpendicular (cubic, tetragonal and orthorhombic cells), for the calculation of inter-planer spacing using cartesian co-ordinates.

In the fig ox, oy & oz are orthogonal axes the origin O is taken at any lattice point. Now we consider any set of crystal planes defined by the Miller indices, (h, k, l). Suppose the reference plane passes through the origin and the next plane makes intercepts a/h, b/k & c/l on x, y & z-axes respectively.

A normal is drawn on to the plane ABC from the origin. The length "d" of this normal plane will be the distance between the adjacent planes.

Now we have to find an expression for d in terms of a, b, c & h, k, l.

Since d is normal to the plane ABC, we write

$$\text{Cos } \alpha = \left(\frac{ON}{OA}\right); \quad \text{Cos}\beta = \left(\frac{ON}{OB}\right); \quad \text{Cos } \gamma = \left(\frac{ON}{OC}\right);$$

Where $\angle\alpha = \angle$NOA, $\angle\beta = \angle$NOB, & $\angle\gamma = \angle$NOC

According to law of direction co-sines,

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$$

$$\Rightarrow \left(\frac{ON}{OA}\right)^2 + \left(\frac{ON}{OB}\right)^2 + \left(\frac{ON}{OC}\right)^2 = 1$$

$$\Rightarrow \left(\frac{d}{a/h}\right)^2 + \left(\frac{d}{b/k}\right)^2 + \left(\frac{d}{c/l}\right)^2 = 1$$

$$\Rightarrow d^2\left[\frac{h^2}{a^2} + \frac{k^2}{b^2} + \frac{l^2}{c^2}\right] = 1$$

$$\Rightarrow d = \frac{1}{\sqrt{\frac{h^2}{a^2} + \frac{k^2}{b^2} + \frac{l^2}{c^2}}}$$

The above relation is applicable to the primitive lattice in cubic, orthorhombic & tetragonal systems. For tetragonal crystal a = b

$$\Rightarrow d = \left[\frac{h^2k^2}{a^2} + \frac{l^2}{c^2}\right]^{-1/2}$$

For cubic system a = b = c

$$\Rightarrow d = \frac{1}{\sqrt{h^2 + k^2 + l^2}}$$

Note:- In non-orthogonal lattice, calculation of inter planer spacing is more complex.

2.4 Point groups:-

In previous lesson we discussed the symmetry operations. We know that the symmetry operation is one that takes the crystal into it self. There is a possibility that the symmetry

elements may be combined with one another in different ways; but the symmetry elements in each such combination are required to satisfy certain conditions.

For example, when one symmetry operation A is followed by another operation B, the resultant effect is capable of being produced by a third symmetry operation C.

$$A \times B = C.$$

where A, B & C are members of combination of certain symmetry elements permitted by the symmetry of the concerned crystal lattice.

The symmetry elements in a combination should also satisfy the following mathematical conditions.

$(A \times B) \times C = A \times (B \times C)$ (The associative law)

$A \times I = A$ I is the Identity operator

$A^{-1} \times A = I$ Where A^{-1} is an operation which is the inverse of operation A. Since A is also the inverse of A^{-1} we have the relation $(A^{-1})^{-1} = A$.

A set of symmetry elements that give the above condition is said to form a group commonly known as the "Point group". It can also be shown that if A and B are elements of a group $(AB)^{-1} = B^{-1}A^{-1}$. If A is a symmetry operation that brings a framework into coincidence with itself and B is some other such operation. The effect of applying first A followed by B may result in a symmetry operation C that makes the framework to coincide with itself. This may be written as BA = C. The operation AB need not be equal to C but may be equal to some other symmetry operation D which is also a member of the same group in which A and B are members. If AB = BA we say that the operations A and B commute. Thus for example, two successive rotations about the same axis will commute since they result in a rotation about the same axis through an angle, which is the sum of two separate angles of rotation. The result of two successive rotations about different axes however does depend on the order in which they are performed.

In the case of a cube it can be shown that $C_4(z)\ C_4(z) = C_2(z)$

$C_4(z)\ C_4(y) = C_3$(about body diagonal).

All the operations contained in the group are performed at a point , which leave the body invariant. The body may be for example, a molecule, a group of atoms, a lattice, or a crystal structure. The word `point' group can be more appreciated if we consider molecular symmetry. A symmetry operation which rotates or reflects a molecule into itself must leave the center of mass of the molecule unmoved. If the molecule has a plane or axis of a symmetry, the center of

the mass must lie on this playing or axis. It follows that all the axes and planes of the symmetry of a molecule must intersect in at least one common point, and that point remains fixed under all the symmetry operations of the molecule. For this reason the symmetry group of a molecule is generally referred to as its point group. These conclusions are valid for lattice points and there are in total 32 distinct combinations of symmetry operations performed on a point in various crystals. This number exhausts all the possible ways of combining the symmetry elements such that the symmetry elements in each combination meet the requirements imposed by the mathematical conditions (A x I = A). This gives 32 point groups in 3 dimensional lattices, in which 27 are non-cubic and five (5) are cubic point groups. There exist two standard nomenclature systems for point group. They are

1. The Schonflies 2. International.

Table 2.4a shows the 27 Non-cubic crystal point groups.

Table 2.4b shows the compositions of the 5 cubic point groups

2.5 Space groups:-

In a crystal, point group symmetry operations are combined with translational elements, provided they are compatible. Such combinations are called space groups. The description of the real crystal structures requires precise knowledge about the composition and symmetry of the basis of the atoms attached to each lattice point. A crystal is regarded to have been constructed by translating its basis through vectors of respective Bravais lattice. The translational symmetry element exhibited by crystals is

a) Glide plane b) Screw axis

Table 2.4(a) Twenty-seven (27) non-cubic crystal point groups

Schönflies	*International*	*Meaning in terms of symmetry elements*	*Number*
C_n	n	These groups have only an n-fold axis of rotation (n = 1, 2, 3, 4, 6)	5
C_{nv}	n m m (n even) n m (n odd)	The groups have an n-fold rotation axis, a mirror plane containing the rotation axis, and as many additional mirror planes as may be required by the presence of the n-fold axis (n = 2, 3, 4, 6).	4
C_{nh}		In addition to the n-fold rotation axis, there is a single mirror plane perpendicular to the rotation axis.	
	$\left(\frac{n}{m}\right)$	n = 2, 4, 6 (even)	
	$\bar{n}$	n = 1, 3 (odd)	5
S_n	$\bar{n}$	These groups have only an n-fold roto- reflection axis (n = 2, 4, 6).	3
D_n	n 2 2 (n even) n 2 (n odd)	These groups contain an n-fold axis of rotation, a 2-fold rotation axis perpendicular to the n-fold axis, and as many additional 2-fold axes as the presence of the n-fold axis demands (n = 2, 3, 4, 6).	4
D_{nh}	$\frac{n}{m}\frac{2}{m}\frac{2}{m}$	In addition to all the elements of D_n, there is a mirror plane perpendicular to the n-fold rotation axis.	4
	$\left(\frac{n}{mmm}\right)$	n = 2, 4, 6 (even)	
	$\bar{n}$ 2 m	n = 3 (odd)	
D_{nd}		In addition to all the elements of D_n, there are mirror planes containing the n-fold rotation axis such that the planes bisect the angles between the 2-fold axes of rotation.	2
	$\bar{n}$ 2 m	n = 2 (even)	
	$\bar{n}\,\frac{2}{m}$	n = 3 (odd)	

Table 2.4(b) The five (5) cubic crystal point groups

Schönflies	*International*	*Meaning in terms of symmetry elements*
T	2 3	The tetrahedral point group (in a crystal the alternate corners of the cube are vacant). It contains three mutually perpendicular 2-fold axes, plus four 3-fold axes, angles between which are bisected by the 2-fold axes.
T_d	$\bar{4}$ 3 m	In addition to all the elements of ***T***, there is a mirror plane through each pair of 3-fold axes (i.e. two mutually perpendicular mirror planes through each 2-fold axis. In all, there are thus six mirror planes.
T_h	$\frac{2}{m}\ \bar{3}$ (m 3)	The groups has all symmetry elements of ***T***, plus a centre of inversion ***i***.
O	4 3 2	The octahedral point group. It contains three mutually perpendicular 4-fold axes and four 3-fold axes which have the same orientation with respect to one another as the 2-fold and 3-fold axes of the tetrahedral point group.
O_n	$\frac{4}{m}\ \bar{3}\ \frac{2}{m}$ (m 3 m)	The group has all symmetry elements of ***O***, plus a centre of inversion ***i***.

a) Glide plane:

When a mirror plane is combined with simultaneous translation operation in a crystal, one obtains glide plane. The glide plane is always parallel to the mirror plane.

b) Screw axis:

In a crystal the rotation axis coupled with the translation parallel to the rotation axis will give rise to a new symmetry element called screw axis. The symbol, which is generally used for screw axis in n_m. This is performed by rotation of $\frac{2\pi}{n_m}$ and a translation of $\frac{m}{n}$ times the translation vector parallel to the rotation axis in a crystal.

The complete pattern of symmetry element is known as a space group, which in addition to having the translational symmetry elements contains all the elements of the respective point

group. There are in all 230 space groups in three dimensional crystals. It is surprising that this number would be reduced to 14 in the case of a completely symmetric basis.

All possible crystal structures are characterized by different space groups. These space groups are obtained by introducing a basis with symmetries of each of the point groups of the system into each of the Bravais lattices.

For example there are 5 point groups & 3 Bravais lattices (P, P, C) for a cubic system. So, with these possibilities 15 different crystal structures can be produced, represented by 15 different space groups. The space group of a crystal is usually represented on a plane by a repetitive pattern of the allowed point symmetry elements, using different standard symbols for different symmetry elements. For the representation of space group both the *Schönflies* & international systems are used.

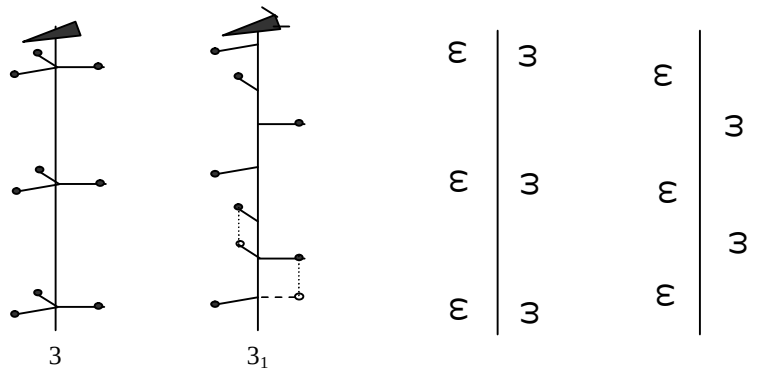


Fig 2.13 (a) & (b) Illustration of Screw operation
(c) Mirror reflection; (d) gride reflection.

2.6 Common Crystal Structures:-

A large percentage of metallic structures crystallize in hcp, fcc, bcc. Simple cubic is very rare in metals.

Simple Cubic:-

The simplest crystal structure that we can think of is that of simple cubic symmetry with a basis of one atom. In this structure the atoms are situated at the corners of the cube touching each other along the edges, which can be shown in fig 2.6(a).

Each atom surrounded by 6 nearest neighbors. So, that the Co-ordination number is 6.

The atomic radius is a/2 (r = a/2). Where a is the cube edge.

The number of atoms per unit cell is 1. Polonium is the lone known example of this class in nature.

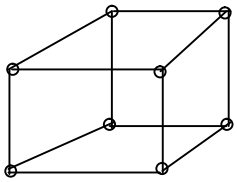


Fig 2.6 (a)

The CsCl Structure:-

The structure of Cesium Chloride is shown in fig 2.6(b). The space lattice is simple cube.

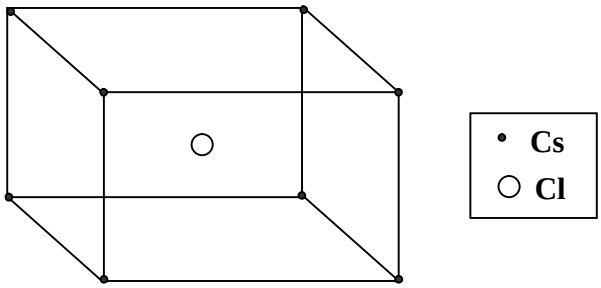


Fig 2.6(b) The unit cell structure of CsCl crystal

The basis has one Cs^+ ion of 000 and one Cl^- ion at ½ ½ ½. The central atom is surrounded by eight atoms of the other type at the corners, the co-ordination number is thus eight. It may be noticed that this structure cannot be interpreted as body centred cubic (BCC). The lattice points of CsCl are two interpenetrating simple cubic lattices,

Crystal	Lattice parameter A^0 ($A^0 = 10^{-10}$m)
CsCl	4.11
CsBr	4.29
CSI	4.56
TICI	3.84
TIBr	3.97
TII	4.18

the corner of one sub-lattice is the body centre of the other. One sub-lattice is occupied by Cs^+ ions and the other by Cl^- ions. The lattice parameter of some ionic crystal having this structure is given below.

Crystal of Alkali Metals:-

Crystals of alkali metals (Li, Na, K, Rb, Cs) are typical representatives of the body centered cubic (bcc) structure. The unit cell is non-primitive with 2 lattice points and the basis of one atom. The cell consist of one atom of each corner & one atom in the centre of the cube which is shown in fig 2.6(c). Each atom has only 8 nearest neighbors. Therefore e Co-ordination number is 8.

(1) Total No. of atoms = 1 + 1 = 2

(2) Atomic radius r = $\frac{\sqrt{3}a}{4}$

(3) Volume of atoms in unit cell u = $2\times\frac{4}{3}\pi r r^3 = \frac{\pi\sqrt{3}a^3}{8}$

(4) Volume of a unit cell V = a^3.

(5) Packing factor $\simeq \frac{u}{V} = \frac{\pi\sqrt{3}a^3}{8\times a^3} \simeq 0.68$.

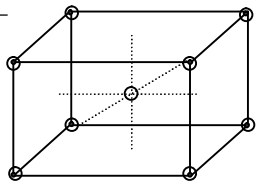


Fig 2.6(c)

Crystals of Noble Metals:-

The Noble metals Cu, As, Au, crystallize in another structure, the face centred cubic (fcc). The unit cell has four lattice points with the basis of one atom. The positions of the atoms in the unit cell are 000, ½ ½ 0, 0 ½ ½ , ½ 0 ½. The Co-ordination number is 12. The structure is also close packed. The lattice parameters are given below (in $\overset{0}{A}$).

Cu	3.61
Ag	4.08
Au	4.07

The NaCl Structure:-

Bravais lattice is fcc. An atom on the edge contributes $1/4^{th}$ of the atom. Thus, the unit cell gets three sodium atom from those on the edges $\left(\frac{12}{4}=3\right)$ & one from that at the centre. The total cell consists of 4 Na atoms. Similarly the contribution of chlorine atoms to the unit cell comes to 4 [8/8 + 6/2 = 4]. A single unit cell accommodates four formula unit cells of NaCl. The positions of atoms in the unit cell are,

Na: $\frac{1}{2}\ \frac{1}{2}\ \frac{1}{2}$ $\quad 0\ 0\ \frac{1}{2}$ $\quad 0\ \frac{1}{2}\ 0$ $\quad \frac{1}{2}\ 0\ 0$

Cl: $0\ 0\ 0$ $\quad \frac{1}{2}\ \frac{1}{2}\ 0$ $\quad \frac{1}{2}\ 0\ \frac{1}{2}$ $\quad 0\ \frac{1}{2}\ \frac{1}{2}$.

The Sodium Chloride (NaCl) structure is shown in fig 2.6(d).

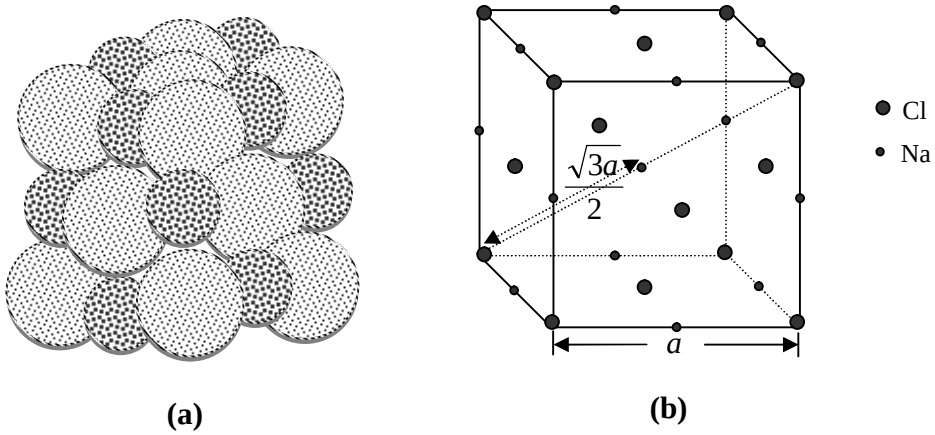


Fig 2.6(d): (a) The arrangement of Na and Cl (shown bigger) atoms on the unit cell of NaCl crystal. (b) Locations of Na and Cl (shown bigger) atoms in the cube representing the unit cell. The separation between the basis partners is half the length of the body diagonal $\frac{\sqrt{3}a}{2}$.

Since each ion has six nearest neighbors of opposite kind, the co-ordination number is 6. Some of the crystals representative of NaCl arrangement, along with their lattice parameter are shown in table.

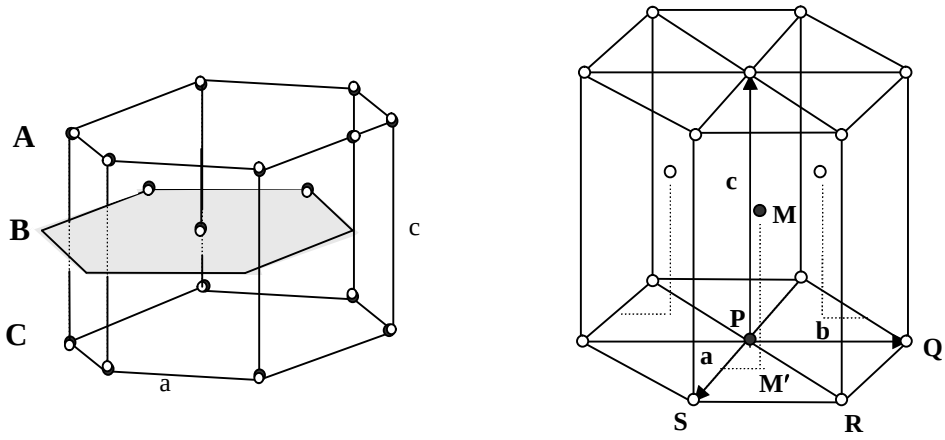
Crystal	Lattice parameter A^0
NaCl	5.63
LiH	4.08
KBr	6.50
RbI	7.33
NH_4I	4.37
NiO	4.17
UO	4.92
PbS	5.92

Hexagonal close-packed structure(hcp):-

Fig (a), in fig (2.6e) represent the hexagonal close-packed structure and fig(b) represent the primitive axes of the hcp crystal. The unit cell contains one atom at each corner, one

atom each at the centre of the Hexagonal faces and 3 more atoms within the body of the cell. Each atom touches 3 atoms in the layer below its plane, six atoms in its own plane, and 3 atoms in the layer above. Hence the Co-ordination number of this structure is 12. The top layer contains seven atoms. Each atom is shared by 6 surrounding hexagon cells and the centre atom is shared by 2 surrounding cells. The three atoms within the body of the cell are carefully contributing to the cell.

Thus the total number of atoms in a unit cell is = $\frac{3}{2}+\frac{3}{2}+3=6.$



(a) **(b)**

Fig 2.6(e): (a) The hexagonal close-packed structure. (b) The primitive axes of the hcp crystal. The c is normal to the plane of a and b. The two atoms of one basis are shown as solid circles.

Diamond Structure:-

The diamond lattice can be considered to be formed by inter penetrating two fcc lattice along the body diagonal by $(1/4)^{th}$ cube edge. One sub-lattice has its origin at the point 0, 0, 0 and the other at a point quarter of the way along the body diagonal $\left(\frac{a}{4},\frac{a}{4},\frac{a}{4}\right)$. The basic diamond lattice and the atomic positions in the cubic cell of diamond projected on a cube face are shown in fig 2.6(f).

The fractions denote height about the base in units of cube edge. The print at 0 and ½ are on the fcc lattice, those at ¼ and ¾ are on a similar lattice displace among the body diagonals by ¼ of the cube edge.

The packing factor is $(XZ)^2 = (XY)^2 + (YZ)^2 = \frac{a^2}{8}+\frac{a^2}{16}=\frac{3a^2}{16}$

But, XZ = 2r

$\therefore (2r)^2 = \frac{3a^2}{16} \quad \Rightarrow a = \frac{8r}{\sqrt{3}}.$

Packing factor = $\frac{U}{V} = \frac{8 \times \frac{4}{3}\pi r^3}{a^3} = \frac{\pi\sqrt{3}}{16} = 0.34\,(or)\,34\%$

Thus it is a loosely packed structure. Carbon, Silicon and Germanium crystallize in this structure.

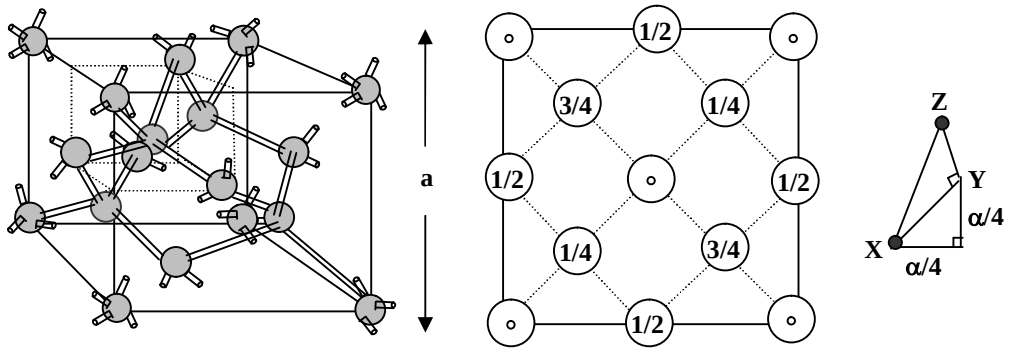


Fig 2.6(f) Diamond structure.

Zincblende structure:

In diamond structure consists of two fcc lattice displaced from each other by one-quarter of a body diagonal. The cubic zinc sulfide structure results from the diamond structure when Zn atoms are placed on one fcc lattice and S atoms on the other fcc lattice. The coordinates of Zn atoms are 000; there are four atoms of ZnS per unit cell.

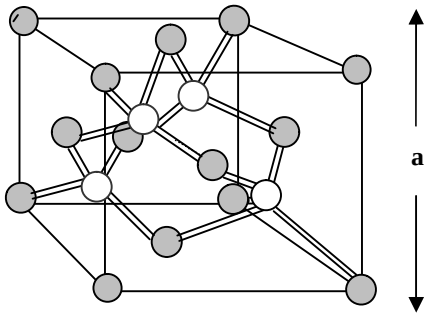


Fig 2.7 Zincblende Structure

Each atom has about it four equally distant atoms of the opposite kind arranged at the corners of a regular tetrahedron. The diamond structure possesses a center of symmetry at the mid point of each line connecting neighbor atoms; the ZnS structure does not have inversion symmetry. Examples of the cubic zinc sulfide structure are CuF, CuCl, ZnS, ZnSe, CdS, InAs, InSb, SiC, AlP etc.

2.7 Summary:

When we speak of a direction, we do not mean one particular straight line, but a whole set of parallel straight lines, which are completely equivalent by virtue of the symmetry. Square brackets are used to indicate a direction (line joining two lattice points). Each direction is specified by the triplet [n_1 n_2 n_3], where n_1, n_2 and n_3 are three integers having no common factor. It is difficult to designate the different planes in a crystal. To avoid this difficulty Miller evolved a method to designate a set of parallel planes in a crystal by three numbers (h, k, l) know as Miller indices.

The Miller indices define a set of parallel planes. Distance between a plane passing through the origin and having Miller indices (hkl) and another consecutive parallel plane is given by $d = \dfrac{1}{\sqrt{\dfrac{h^2}{a^2} + \dfrac{k^2}{b^2} + \dfrac{l^2}{c^2}}}$

The above relation is applicable to the primitive lattice I cubic, orthorhombic and tetragonal systems. It is also

For tetragonal crystal a = b and $d = \left[\dfrac{h^2k^2}{a^2} + \dfrac{l^2}{c^2}\right]^{-1/2}$

For cubic system a = b = c and $d = \dfrac{1}{\sqrt{h^2 + k^2 + l^2}}$

In a non-orthogonal lattice, calculation of inter planer spacing is more difficult. A set of symmetry elements that satisfy group properties are said to form a group commonly known as the “Point group” as all the operations contained in the group are performed at a point in the crystal lattice. There are in total 32 point groups in crystals.

The crystal structure is complete only when its point group & space group are known. The description of the real crystal structures requires precise knowledge about the composition and symmetry of the basis of the atoms attached to each lattice point. A crystal is regarded to have been constructed by translating its basis through vectors of respective Bravais lattice. A space group, which in addition to having the translational symmetry elements contains all the elements of the respective point group, symbolizes the complete array thus obtained. There are in all 230 space groups in crystals. These space groups are obtained by introducing a basis with symmetries of each of the point groups of the system into each of the Bravais lattices.

For example there are 5 point groups and 3 Bravais lattices (P, P, C) for a cubic system. So, with these possibilities 15 different crystal structures can be produced, represented by 15 different space groups. The space group of a crystal is usually represented on a plane by a repetitive pattern of the allowed point symmetry elements, using different standard symbols for different symmetry elements. For the representation of space group both the *Schönflies* & in international systems are used.

Several diatomic crystals and metals form cubic crystals with non-primitive unit cells. Some follow Sodium chloride structure, some follow Cesium chloride structure some follow Zinc blend and some other follow diamond structure.

In CsCl structure the basis has one Cs^+ ion at 000 and one Cl^- ion at ½ ½ ½. The lattice points of CsCl are two interpenetrating simple cubic lattices the corner of one sub-lattice is the body centre of the other. One sub-lattice is occupied by Cs^+ ions and the other by Cl^- ions.

Crystals of alkali metals (Li, Na, K, Rb, Cs) are typical representatives of the body centered cubic (bcc) structure. The unit cell is non-primitive with 2 lattice points and the basis of one atom. The cell consist of one atom of each corner & one atom in the centre of the cube.

The Noble metals Cu, As, Au, crystallize in another structure, the face centered cubic (fact). The unit cell has four lattice points with the basis of one atom. The positions of the atoms in the unit cell are 000, $\frac{1}{2}\frac{1}{2}0, 0\frac{1}{2}\frac{1}{2}, \frac{1}{2}0\frac{1}{2}$. The Co-ordination number is 12. For rock salt structure Brava's lattice is fact. An atom on the edge contributes $1/4^{th}$ of the atom. Thus, the unit cell gets three sodium atom from those on the edges $\left(\frac{12}{4}=3\right)$ & one from that at the centre. The total cell consist of 4 Na atoms. Similarly the contribution of chlorine atoms to the unit cell comes to 4[8/8 + 6/2=4]. A single unit cell accommodates four formula unit cells of NaCl. The positions of atoms in the unit cell are,

Na: $\frac{1}{2}\frac{1}{2}\frac{1}{2}$ $00\frac{1}{2}$ $0\frac{1}{2}0$ $\frac{1}{2}00$

Cl: 000 $\frac{1}{2}\frac{1}{2}0$ $\frac{1}{2}0\frac{1}{2}$ $0\frac{1}{2}\frac{1}{2}$

Since each ion has six nearest neighbors of opposite kind, the co-ordination number is 6. in hcp crystal the unit cell contains one atom at each corner, one atom each at the

centre of the Hexagonal faces and 3 more atoms within the body of the cell. Each atom touches 3 atoms in the layer below its plane, six atoms in its own plane, and 3 atoms in the layer above. Hence the Co-ordination number of this structure is 12. The top layer contains seven atoms. Each atom is shared by 6 surrounding hexagon cells and the center atom is shared by 2 surrounding cells. The three atoms within the body of the cell are contributing to the cell. Thus the total number of atoms in a unit cell is 6.

The diamond lattice can be considered to be formed by inter penetrating two fcc lattice along the body diagonal by ($1/4^{th}$)cube edge. One sub-lattice has its origin at the point 0, 0, 0 and the other of a point quarter of the way along the body diagonal $\left(\frac{a}{4},\frac{a}{4},\frac{a}{4}\right)$. The fraction denote height about the base in units of cube edge. The point at 0 and ½ are on the fcc lattice, those ¼ and ¾ are on a similar lattice displaced among the body diagonals by ¼ of the cube edge.

The cubic zinc sulfide structure results from the diamond structure when Zn atoms are placed on one fcc lattice and S atoms on the other fcc lattice. The coordinates of Zn atoms are 000; there are four atoms of ZnS per unit cell. Each atom has about it four equally distant atoms of the opposite kind arranged at the corners of a regular tetrahedron. The diamond structure possesses a center of symmetry at the mid point of each line connecting neighbor atoms; the ZnS structure does not have inversion symmetry. Examples of the cubic zinc sulfide structure are CuF, CuCl, ZnSl, ZnSe, CdS, InAs, InSb, SiC, AlP etc.

2.8 Keywords:

Miller indices – crystal planes – inter-planar distance – symmetry elements – center of symmetry – point groups – Translational symmetry – glide plane – Screw axis – Bravais lattices – space groups – common crystal structures.

2.9 Review questions:

1. What do you understand by Miller indices of a lattice plane. Find the Miller indices of a plane that makes an intercept of 3a, 3b and c along the three crystallographic axes, where a, b, c being the primitive vectors of the lattice.
2. Determine the Miller indices of a plane that makes intercept of 3A, 4A and 5A on the coordinate axes of an orthorhombic crystal with a: b: c = 1:2:5.

3. Draw a (110) plane in the unit cell of a cubic crystal.
4. Show that the (hkl) plane is perpendicular to the [hkl] direction in a cubic lattice.
5. Calculate the number of atoms per unit cell for a rock-salt crystal. Given a = 5.63A; Mol.wt. of NaCl = 58.5 and the density is 2180kg/m^3.
6. The density of bcc iron is 7.9×10^3kg/m^3 and the atomic wt. is 56. Calculate the size of the unit cell.
7. Show that maximum packing ratio in the diamond structure is $\frac{\pi\sqrt{3}}{16}$
8. The diamond lattice may be considered to be a combination of two interpenetrating fcc sub-lattices. One sub-lattices has its origin at the point 0, 0, 0 and the other at a point one-quarter of the way along the body diagonal. a) Write the position of all the atoms in the unit cell. b) How many atoms are contained in the unit cell.
9. How many planes of type {hkl} are found in the cubic system. How many of the type {hk0}.
10. Explain zinc blend structure and give two examples of compounds, which crystallize in this structure.
11. Describe the diamond, sodium chloride and cesium chloride structures.
12. Show that for a simple cubic lattice $d_{100} : d_{110} : d_{111} = \sqrt{6} : \sqrt{3} : \sqrt{2}$
13. Derive the relationship between atomic radius and lattice parameter for fcc crystal and calculate molecular packing.
14. What point groups does each of the following structures belong a) fc cubic b) diamond c) bc cubic.
15. Explain clearly the three space group, point group and Bravais lattice. Show that base centered and face centered tetragonal does not give any new Bravais lattice.

2.10 Text and Reference Books:

1. Elements of Solid State Physics by J.P.Srivastava (PHI)
2. Elements of solid State Physics by Omar (Person education)
3. Solid State Physics by S.L.Kakani and C.Hemrajani (S.Chand)
4. Solid State Physics by M.A. Wahab (Narosa)
5. Solid State Physics by C.Kittel (Asia Publishing house)
6. Solid State Physics by C.J. Dekker.

UNIT – I

LESSON – 3

RECIPROCAL LATTICE

Aim: To learn about reciprocal lattices and Brillouin zones

Objectives:

- To introduce the concept of reciprocal lattice
- To show the relationship between real and reciprocal lattices
- To write Braggs law in terms of reciprocal lattice vectors
- To apply the reciprocal lattice concept to some important crystal lattices.

Structure:

Introduction:

In the earlier lessons we learnt about the crystal structures in terms of the lattice symmetry, basic composition, and various intersecting families of planes as visualized in the space defined by the three crystal edges. The common experimental technique employed to investigate crystal structures is the x-ray diffraction method. The crystal photograph recorded on the X- ray film represents the diffraction pattern of the crystal in which each of the spots corresponds to the diffraction maximum of a different family of parallel planes, with a certain orientation and common spacing. It is worthwhile to find some correlation between structure of a crystal and its planar X-ray photograph. This can be done conveniently by introducing the concept of reciprocal lattice.

3.1 Reciprocal Lattice:-

It is necessary to consider sets of planes in a crystal. This can be done in terms of their normal. Geometrically we have an advantage, that is, the planes are of 2-dimensions while normal lines are of one-dimensional nature. But it is not enough if we consider the orientation of planes alone to study the diffraction of x-rays by crystals but it is also necessary to know the inter-planar spacing d since they only determine the reflection angles θ. These inter-planar spacing may also be represented in the normal to the planes by appropriately limiting their lengths.

From the above discussion it is clear that one can indicate the orientation of a set of parallel planes by their common normal and the inter planar spacing by restricting the lengths of the normal proportionately.

Consider any given space lattice and apply the following.

1. From a common origin draw a normal to each crystal plane.
2. Set the length of each normal equal to (or) 2π times the reciprocal of the inter planar spacing d_{hkl}.
3. Mark a point at the end of each normal, which represents the crystal plane.

A collection of points obtained in this way corresponding to various crystal planes form a lattice array and is known as "reciprocal lattice". The points in the reciprocal lattice are called reciprocal lattice points. These points in 3 dimensional space form the reciprocal lattice space. This is also called k-space. From the concept of reciprocal lattice it may be understood that the "Co-ordinates of Points" in the reciprocal lattice

space are defined by (h k l), the Miller indices. The concept of reciprocal lattice plays a very important role in the field of x-ray crystallography.

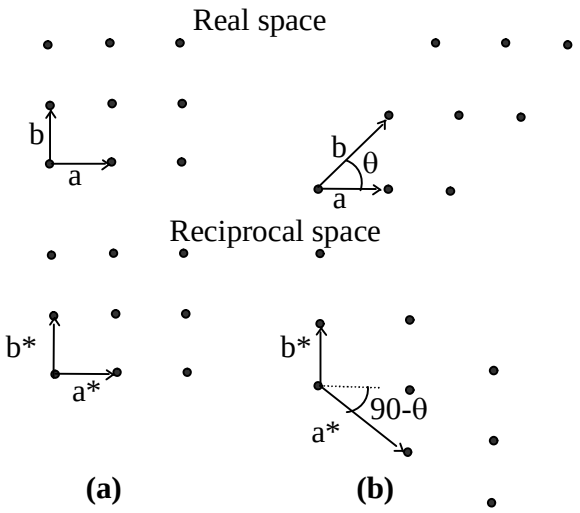


Fig 3.1 Direct and reciprocal lattice for a rectangular system

3.2 Construction of Reciprocal Lattice:-

For the construction of reciprocal lattice first fix the origin at any arbitrary lattice point O in the direct lattice and then identify three crystal axes (a, b, c) and shown in fig 3.2a

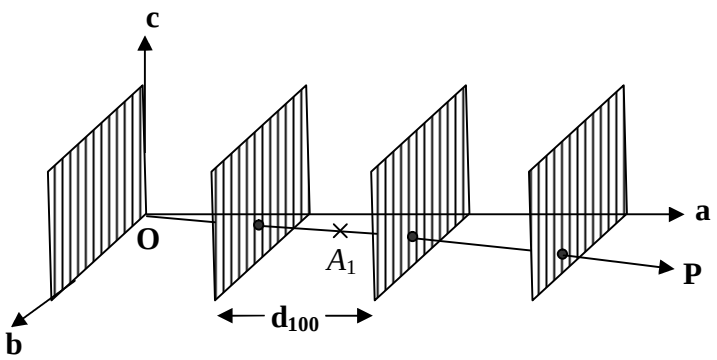


Fig 3.2a Construction of the reciprocal lattice, a, b, c represent the primitive axes of the direct lattice. A_1 is the reciprocal lattice point referred to d_{100} planes and lies on the normal to d_{100} planes. The normal OP passing through the origin O gives the direction of the a^* axis of the reciprocal lattice ($|OA_1| = |a^*|$).

Let us now draw a family of d_{100} planes where d-stands for the inter planar separation and (100) are the Miller indices. Drop a normal on **OP** to these set of planes from the origin *O*. The vector **OP** in this case represents the direction of $\mathbf{a}^*$. The complete set of reciprocal lattice vectors along this direction are given by

$$\mathbf{g} = \frac{2\pi n}{\mathbf{d}_{100}}, \quad (n = 1, 2, 3,\ldots) \text{ ---------- } (3.1)$$

The tip of these vectors will be located at the respective reciprocal lattice points A_1, A_2, A_3,…. The smallest of these $\mathbf{OA}_1$, is denoted by $\mathbf{a}^*$ and expressed as

$$\mathbf{OA_1} = \mathbf{a}^* = \frac{2\pi}{\mathbf{d}_{100}} = \frac{2\pi\hat{e}}{d_{100}} \text{ ---------- } (3.2)$$

where $\hat{e}$ stands for the unit vector along **OP** or $\mathbf{a}^*$.

Thus A_1 represents the reciprocal lattice point corresponding to d_{100} planes and A_2 will correspond to d_{200} planes which are parallel to d_{100} planes with half common spacing.

Therefore, it is reasonable to write

$$\mathbf{OA_2} = \frac{2\pi}{\mathbf{d}_{200}} = \frac{2\pi\hat{e}}{d_{200}} \text{ ---------- } (3.3)$$

It is also very clear that a reciprocal lattice point truly symbolizes a family of parallel planes because its position is defined implicitly by the orientation and explicitly by the common spacing of planes. And the general expression for reciprocal lattice vectors along $\mathbf{a}^*$, identified as one of the three primitive translation vectors, is

$$g_{h00} = h\mathbf{a}^*, \quad (h = 1, 2, 3, \ldots) \text{ ---------- } (3.4a)$$

We can similarly write vectors along the other two primitive vectors $\mathbf{b}^*$ and $\mathbf{c}^*$ as

$$g_{0k0} = k\mathbf{b}^*, \quad (k = 1, 2, 3, \ldots) \text{ ---------- } (3.4b)$$

and

$$g_{00l} = l\mathbf{c}^*, \quad (l = 1, 2, 3, \ldots) \text{ ---------- } (3.4c)$$

Following this procedure for all the planes of a crystal lattice, its reciprocal lattice may be mapped. The general reciprocal lattice vector should read as

$$g = h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^* \text{ ---------- } (3.5)$$

As also inferred towards the end of the preceding section, the reciprocal vector **g** given by (3.5) is normal to the family of planes (*hkl*). Suppose one such plane passes through the origin *O* in a crystal system. Another member of the family and closest to this plane is drawn in Fig 3.2b. Vectors **p, q** and **r** lie in this plane. Now, it is a simple exercise to show that **g** is normal to planes (*hkl*). All we need is to show that

$$\mathbf{g} \cdot \mathbf{t} = 0 \quad \text{with} \quad \mathbf{t \equiv p, q, r.}$$

From the definition of the Miller indices, it follows that the intercepts of the plane, shown in Fig.3.2b, with **a, b** and **c** axis vectors of the direct crystal are vectorially written as **a**/*h*, **b**/*k* and **c**/*l*, respectively. This enables us to write

$$\mathbf{p} = \left(\frac{\mathbf{a}}{\mathrm{h}} - \frac{\mathbf{b}}{\mathrm{k}}\right); \quad \mathbf{q} = \left(\frac{\mathbf{c}}{l} - \frac{\mathbf{b}}{\mathrm{k}}\right); \text{ and } \quad \mathbf{r} = \left(\frac{\mathbf{c}}{l} - \frac{\mathbf{a}}{\mathrm{h}}\right)$$

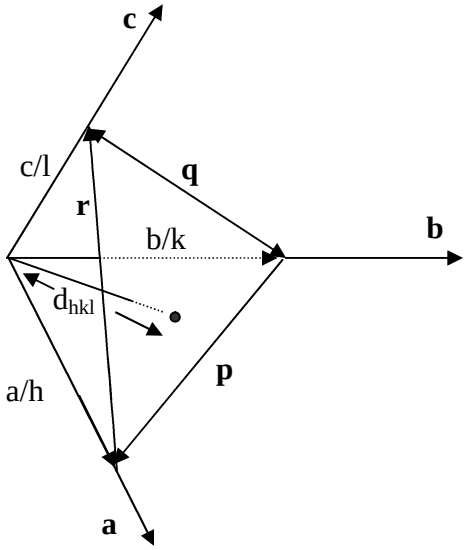


Fig 3.2b The figure shows intersections of a plane (*hkl*) with crystal axes a, b, c at general orientations. The three intercepts measure a/h, b/k and c/l, respectively. Vectors p, q and r lie in the plane (*hkl*).

It can be checked that the scalar product of any of these vectors with **g** vanishes, proving thereby that every reciprocal vector is normal to a family of parallel planes with common inter plane spacing.

One point of caution: *h, k* and *l* in (3.5) need not always represent the Miller indices and in principle they can have a common factor. In view of this fact it is safer to use the form

$$\mathbf{g} = m_1\mathbf{a}^* + m_2\mathbf{b}^* + m_3\mathbf{c}^* \quad \text{----------(3.6)}$$

where m_1, m_2 and m_3 are any integers.

3.3 Relationship between a, b, c and $\mathbf{a}^*$, $\mathbf{b}^*$, $\mathbf{c}^*$:

Consider a direct crystal lattice as shown in Fig. 3.3. Draw a normal *OA* from *O* on the face opposite to be face.

Draw a normal *OA* from *O* on the face opposite to bc face volume of the unit cell,

$$V = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = OA \times (\text{Area of the face bc}) = OA \times |\mathbf{S}| \quad \text{----------} \quad (3.7)$$

$$\therefore \frac{|\mathbf{S}|}{V} = \frac{1}{OA} \quad \text{(or)} \quad \frac{1}{d_{100}} = \frac{|\mathbf{S}|}{V} \qquad (\because OA = d_{100})$$

$$\text{Thus } \frac{\hat{e}}{d_{100}} = \frac{\mathbf{S}}{V}$$

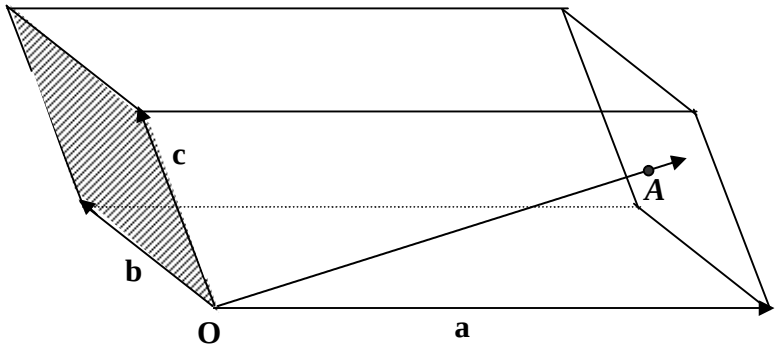


Fig 3.3 A general direct crystal lattice

Where $\hat{e}$ is the unit vector in the direction of the normal *OA*.

$$\therefore \frac{2\pi\hat{e}}{d_{100}} = \frac{2\pi\mathbf{S}}{V} = 2\pi \cdot \frac{\mathbf{b}\times\mathbf{c}}{\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})} \quad \text{----------} \quad (3.8)$$

$$\Rightarrow \mathbf{a}^* = 2\pi \cdot \frac{\mathbf{b}\times\mathbf{c}}{\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})} \quad \text{----------} \quad (3.9)$$

similarly $$\Rightarrow \mathbf{b}^* = 2\pi \cdot \frac{\mathbf{c}\times\mathbf{a}}{\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})} \quad \text{----------} \quad (3.10)$$

$$\Rightarrow \mathbf{c}^* = 2\pi \cdot \frac{\mathbf{a}\times\mathbf{b}}{\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})} \quad \text{----------} \quad (3.11)$$

From the above equations we observe that $\mathbf{a}^*$, $\mathbf{b}^*$ and $\mathbf{c}^*$ is orthogonal to two crystal axis vectors.

This can be compactly stated as, $\mathbf{i}\cdot\mathbf{j} = 2\pi\delta_{ij}$ ---------- (3.12)

Where **i**, **j** can be **a**, **b** and **c** and δ_{ij} is the usual delta function.

3.4 Bragg's Law in terms of reciprocal lattice vectors:-

We learn more about x ray diffraction in the forth coming lesson and for the present the Bragg's law derived for the X ray diffraction condition can be expressed suitably in the reciprocal lattice, We now see how to do this.

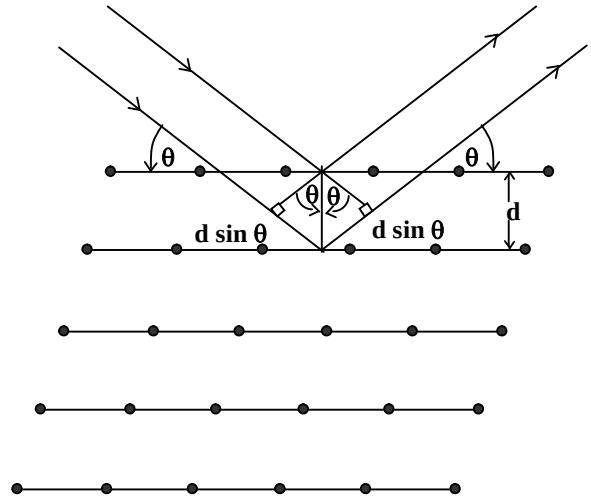


Fig 3.4a Bragg reflection from a family of planes with inter planar spacing d. Note that the incident beam is deflected by twice the Bragg angle θ.

In the X ray diffraction phenomenon, the crystal planes are believed to act like plane mirrors. Radiations reflected from two successive parallel planes under certain conditions may interfere constructively to produce a diffraction maximum. The Bragg diffraction shown in fig3.4a. Occurs for specular reflections (angle of incidence = angle of reflection). The constructive interference occurs when the path difference (2d sin θ) between the interfering rays equals an integral multiple of the X-ray wavelength λ. That is,

$$2d \sin \theta = n\lambda \quad \text{---------- (3.13)}$$

where d is the inter planar spacing θ is the angle of the incident radiation with the plane

n = 1, 2, 3 …… (order of diffraction)

The relation (3.13), which is a mathematical statement of the Bragg law, shows that the diffraction effects cannot be observed from a family of planes for any arbitrary angle of incidence.

Bragg's diffraction condition in terms of reciprocal lattice:

The Bragg condition can be expressed as a relation between vectors in the reciprocal lattice. The Bragg condition can be expressed as

$$\sin\theta_{hkl} = \frac{\lambda}{2d_{hkl}} \quad \text{----------} \quad (3.14)$$

in which the order of reflection is already included. The equation (3.14) can also be written as

$$\sin\theta_{hkl} = \frac{1/d_{hkl}}{1/\lambda} \quad \text{----------} \quad (3.15)$$

A geometrical interpretation of eqn.(3.15) is given in fig (3.4b). ***SO*** is a vector whose length is 1/λ. This vector is drawn in the direction of incident X-ray beam and ending at the origin of reciprocal lattice. Now a sphere of radius 1/λ is constructed about a point ***S*** as centre. Let this sphere intersect some point (***h′, k′, l′***) of the reciprocal lattice at ***P***.

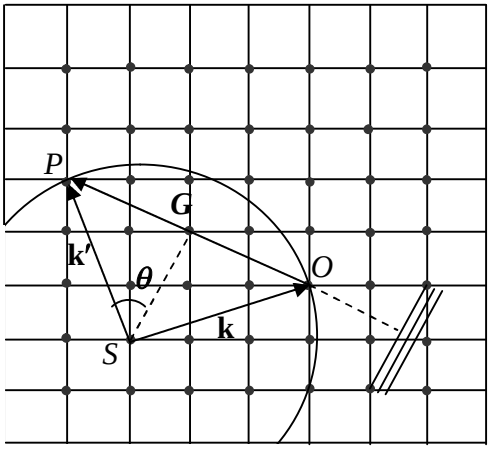


Fig 3.4b Vector geometry of Bragg reflection in the reciprocal lattice.

The vector ***OP*** then represents a vector counting the origin of the reciprocal lattice and a point (***h′, k′, l′***) of that lattice. The vector is normal to ***hkl*** plane of direct lattice and its length is $\mathbf{1/d_{hkl}}$. From the figure the length of the vector ***OP*** can be calculated from fig. which is $2\sin\theta/\lambda$.

Hence $2\sin\theta/\lambda = \frac{1}{d_{hkl}}$ ---------- (3.16)

i.e. $\lambda = 2d\sin\theta$

and the Bragg condition is satisfied. Thus the vector ***OP*** represents a normal to the reflecting planes **(*hkl*)** and the vector ***SP*** is in the direction of diffracted beam. The direction of the diffracted beam is shown in figure (3.4c). For any experimental set up, the direction of X-ray beam is defined as ***AO***.

Diffraction occurs only when the orientation of the crystal is such that a reciprocal lattice point ***P*** comes to lie on the circumference of a circle ***S*** of radius $1/\lambda$. When this occurs, a diffracted beam is developed in the direction ***SP***.

It is customary to imagine all the vectors of fig (3.4c) to be multiplied by a constant factor of 2π as represented in figure (3.4d). Here the vector **K** is 2π times the vector ***SO*** and the vector **G** is 2π times the vector ***OP***. Again the disposition of vector is such that vector ***SP*** must be vector sum of **K** and **G**.

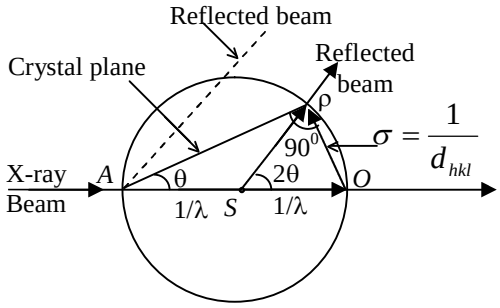


Fig 3.4c Showing the direction of diffracted beam.

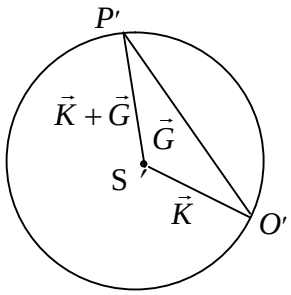


Fig 3.4d Vector diagram of the figure expanded by a scale factor 2π

Now the magnitude of the vector ***S′P′*** i.e., **K**+**G** and the magnitude of the incident beam vector ***S′O′*** i.e., **K** must be equal. Applying this condition the Bragg condition is satisfied ; the Bragg condition must imply that

$$\mathbf{(K+G)}^2 = (\mathbf{K+G}).(\mathbf{K+G}) = \boldsymbol{K^2}$$

$$\textbf{i.e.,}\quad \mathbf{2K \cdot G + G^2 = 0} \qquad \text{----------} \quad (3.17)$$

Equation (3.17) represents the vector form of Bragg equation.

Here $\boldsymbol{G} = \boldsymbol{h}^*\mathbf{a} + \boldsymbol{k}^*\mathbf{b} + \boldsymbol{l}^*\mathbf{c}$ **(*h, k, l* =** integers**)**

Brillouin Zones:

If we consider a parallelepiped formed by reciprocal lattices $\boldsymbol{a}^*$, $\boldsymbol{b}^*$ and $\boldsymbol{c}^*$, then this may be taken as the primitive cell of the reciprocal lattice. It can be observed that the eight corner points are shared among eight parallelepiped or we can say that one parallelepiped contains one-eight of each of eight corner points. In this way the parallelepiped contains one reciprocal lattice point. But in solid state physics, a primitive cell of a reciprocal lattice is taken as the smallest volume bounded by planes normal to each of (shorter) **G**'s at its midpoint.

Each of the new cell contains on lattice point and the point is at the centre of the cell as shown in fig 3.4e. The primitive cell formed in this way in the reciprocal lattice is called the first Brillouin zone. The same procedure adopted for real crystal lattice results in a real primitive cell called Wigner-Seitz cell.

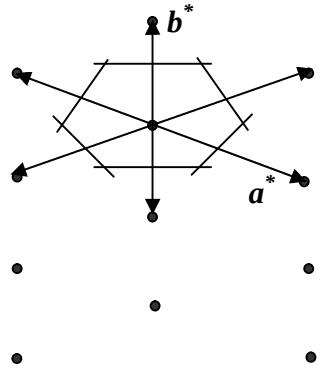


Fig 3.4e Construction of first Brillouin zone.

3.5Application to some crystal lattices:-

Reciprocal lattice to (sc) simple cubic lattice:-

The primitive translation vectors of a sc lattice may be written as,

$\mathbf{a} = a\hat{i}$; $\mathbf{b} = b\hat{j}$; $\mathbf{c} = c\hat{k}$; where the volume of the unit cell is $\mathbf{a} \cdot \mathbf{b} \times \mathbf{c} = a^3$. The Primitive translation vectors of the reciprocal lattice of the sclattice will be,

$$\mathbf{a}^* = 2\pi \cdot \frac{\mathbf{b}\times\mathbf{c}}{\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})} = \frac{2\pi}{a}\hat{i}(or)\frac{2\pi}{a}\hat{x}$$

$$\mathbf{b}^* = 2\pi \cdot \frac{\mathbf{c}\times\mathbf{a}}{\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})} = \frac{2\pi}{a}\hat{j}(or)\frac{2\pi}{a}\hat{y}$$

and $$\mathbf{c}^* = 2\pi \cdot \frac{\mathbf{a}\times\mathbf{b}}{\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})} = \frac{2\pi}{a}\hat{k}(or)\frac{2\pi}{a}\hat{z}$$

From these equations, it is evident that the reciprocal lattice to sc lattice is itself a sc lattice with a lattice const $2\pi/a$. The boundaries of the resulting primitive unit cell are the planes normal to the six reciprocal lattice vectors $\pm \mathbf{a}^*, \pm \mathbf{b}^*, \pm \mathbf{c}^*$ at their mid points,

$$\therefore \qquad \pm\frac{\mathbf{a}^*}{2} = \pm\frac{\pi}{a}\hat{i} \;;\quad \pm\frac{\mathbf{b}^*}{2} = \pm\frac{\pi}{b}\hat{j} \;;\quad \pm\frac{\mathbf{c}^*}{2} = \pm\frac{\pi}{c}\hat{k}\,.$$

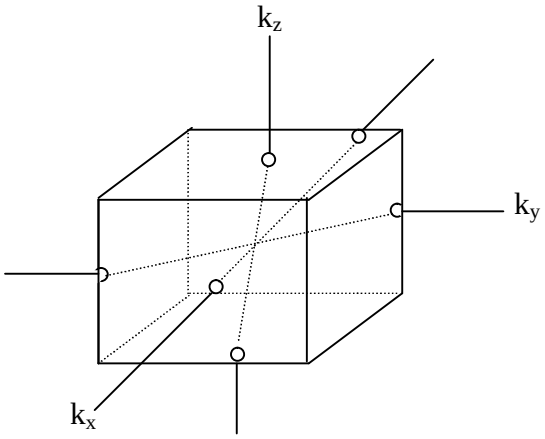


Fig 3.5a The first Brillouin zone of a simple cubic lattice

The space bounded by these Six planes is a cube of side $2\pi/a$ and volume $\left(2\pi/a\right)^3$. This cube is known as the first Brillouin Zone of the sc lattice is shown in fig 3.5a.

3.5.2 Reciprocal lattice to bcc lattice:-

The primitive translation vectors or the bcc lattice shown in fig 3.11 are given by,

$$\mathbf{a}^1 = \frac{a}{2}\;(\mathbf{i}+\mathbf{j}-\mathbf{k})$$

$$\mathbf{b}^1 = \frac{a}{2}\;(-\mathbf{i}+\mathbf{j}+\mathbf{k})$$

$$\mathbf{c}^1 = \frac{a}{2}\;(\mathbf{i}-\mathbf{j}+\mathbf{k})$$

Volume of the unit cell is $\frac{a^3}{2}$.

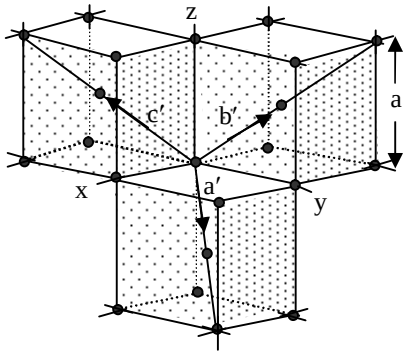


Fig 3.5b Primitive translation vectors of the bcc lattice

The primitive translation vectors $\mathbf{a}^*$, $\mathbf{b}^*$ and $\mathbf{c}^*$ of the reciprocal lattice are defined by

$$\mathbf{a}^* = \frac{2\pi}{a}\;(\mathbf{i}+\mathbf{j})$$

$$\mathbf{b}^* = \frac{2\pi}{a}\;(\mathbf{j}+\mathbf{k}) \qquad \text{----------} \quad (3.18)$$

$$\mathbf{c}^* = \frac{2\pi}{a}(\mathbf{k}+\mathbf{i})$$

We observe that, these reciprocal lattice vectors are just the primitive vectors of the fcc lattice, showing that an fcc lattice is the reciprocal lattice of the bcc lattice. The rhombohedron formed by $\mathbf{a}^*$, $\mathbf{b}^*$ and $\mathbf{c}^*$ represents the primitive cell of volume V expressed as

$$\mathbf{V} = \mathbf{a}^* \cdot (\mathbf{b}^* \times \mathbf{c}^*) = \frac{16\pi^3}{a^3}$$

Similarly, it can be shown that the reciprocal lattice of the fcc lattice is a bcc lattice. There are in all 12 shortest vectors for the lattice described by (3.18):

$$\frac{2\pi}{a}(\pm\mathbf{x} \pm \mathbf{y}); \quad \frac{2\pi}{a}(\pm\mathbf{y} \pm \mathbf{z}); \quad \frac{2\pi}{a}(\pm\mathbf{z} \pm \mathbf{x})$$

with the choices of signs being independent.

Planes perpendicular to these vectors at their mid-points enclose the volume of the first Brillouin zone which turns out to be a rhombododecahedron (Fig. 3.5c).

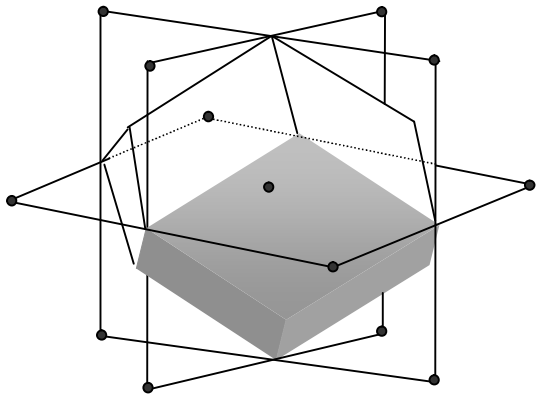


Fig 3.5c The first Brillouin zone of a bcc crystal. It is rhombododecahedral in shape.

Face-Centered Cubic Lattice:

The primitive translation vectors of the fcc lattice as shown in fig.3.5d may be taken as

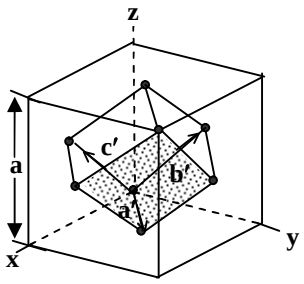


Fig 3.5d Primitive translation vectors of the fcc lattice.

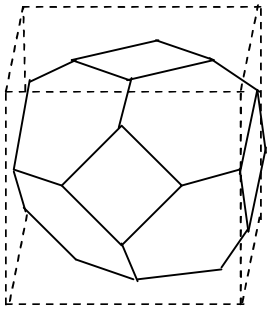


Fig 3.5e The first Brillouin zone in the fcc lattice.

$\mathbf{a} = \frac{a}{2}(\mathbf{i} + \mathbf{j})$;

$\mathbf{b} = \frac{a}{2}(\mathbf{i} + \mathbf{k})$; ---------- (3.19)

$\mathbf{c} = \frac{a}{2}(\mathbf{j} + \mathbf{k})$.

The volume of the primitive cell is $\mathbf{a} \cdot \mathbf{b} \times \mathbf{c} = \frac{1}{4}a^3$. Using the primitive translations of the reciprocal lattice are found to be

$\mathbf{a}^* = \frac{2\pi}{a}(\mathbf{i} + \mathbf{j} - \mathbf{k})$

$\mathbf{b}^* = \frac{2\pi}{a}(-\mathbf{i} + \mathbf{j} + \mathbf{k})$; ---------- (3.20)

$\mathbf{c}^* = \frac{2\pi}{a}(\mathbf{i} - \mathbf{j} + \mathbf{k})$

These are the primitive translations of a bcc lattice. We have now

$\mathbf{G} = (2\pi/a)[(h - k + l)\,\mathbf{i} + (h + k - l)\,\mathbf{j} + (-h + k + l)\,\mathbf{k}]$ ---------- (3.21)

The shortest non-zero G's are the eight vectors

$(2\pi/a)(\pm\mathbf{i} \pm\mathbf{j} \pm\mathbf{k})$. ---------- (3.22)

The zone boundaries are determined for the most part by the eight planes normal to these vectors at their midpoints, but it may be seen that the corners of the octahedron thus formed are truncated by the planes which are the perpendicular bisectors of the six vectors.

$(2\pi/a)\ (\pm 2\mathbf{i})$; $(2\pi/a)\ (\pm 2\mathbf{j})$; $(2\pi/a)\ (\pm 2\mathbf{k})$. ---------- (3.23)

The first zone is then the truncated octahedron shown in fig 3.5e.

3.5.3 Hexagonal space Lattice:-

The primitive translation vectors $\mathbf{b}_1$, $\mathbf{b}_2$ and $\mathbf{b}_3$ are drawn in Fig 3.5f. In relation to the directions of x, y, z axes as shown in the figure, the primitive vectors of the direct lattice are given by

$$\mathbf{b}_1 = \frac{\sqrt{3a}}{2}\mathbf{i} + \frac{a}{2}\mathbf{j}$$

$$\mathbf{b}_2 = -\frac{\sqrt{3a}}{2}\mathbf{i} + \frac{a}{2}\mathbf{j} \quad \text{---------- (3.24)}$$

$$\mathbf{b}_3 = c\mathbf{k}$$

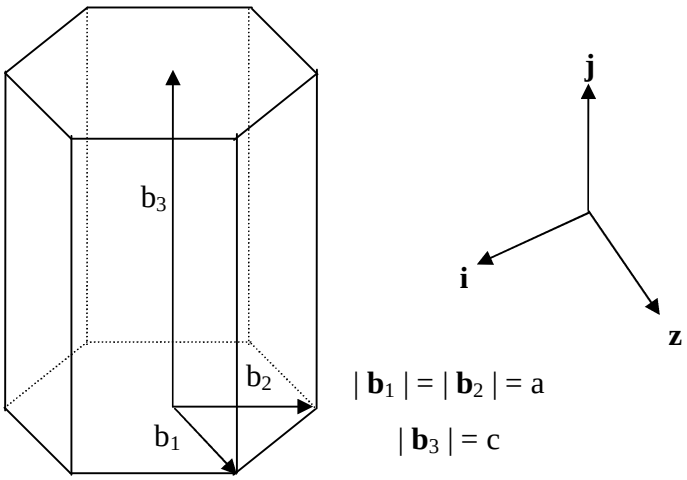


Fig 3.5f Hexagonal space lattice –b_1, b_2 and b_3 are the primitive vectors and the angle between b_1 and b_2 (on the hexagonal face) is 60^0.

Then the primitive vectors of the corresponding reciprocal lattice are

$$\mathbf{b}_1^* = \frac{2\pi}{a}\left(\frac{1}{\sqrt{3}}\hat{x} + \hat{y}\right)$$

$$\mathbf{b}_2^* = \frac{2\pi}{a}\left(-\frac{1}{\sqrt{3}}\hat{x} + \hat{y}\right) \quad \text{----------} \quad (3.25)$$

$$\mathbf{b}_3^* = \frac{2\pi}{c}\hat{z}$$

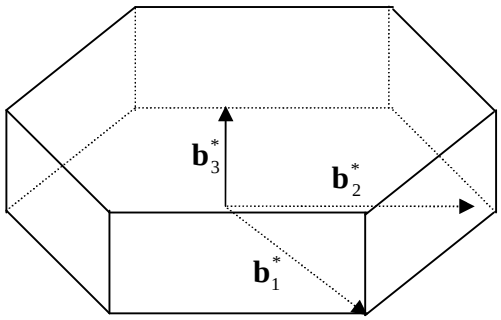


Fig 3.5g The reciprocal lattice of the hexagonal space lattice shown in Fig 3.5f

The unit cell formed by these vectors is again a hexagonal prism as shown by Fig.3.5g. The significant point about this prism is that

$|\mathbf{b}_3^*| < |\mathbf{b}_1^*|$ or $< |\mathbf{b}_2^*|$

3.6 Summary:

One can indicate the orientation of a set of parallel planes in a crystal by their common normal and the inter planer spacing by restricting the lengths of the normals proportionately by applying the following rules

1. From a common origin draw a normal to each crystal plane.
2. Set the length of each normal equal to (or) 2π times the reciprocal of the inter planer spacing d_{hkl}.
3. Mark a point at the end of each normal which represents the crystal plane.

A collection of points obtained corresponding to various crystal planes form a lattice array and this is known as "reciprocal lattice". The points in the reciprocal lattice are called reciprocal lattice points. These points in 3 dimensional space form the reciprocal lattice space. This is also called k-space. It may be noted that the reciprocal lattice to sc lattice is itself a sc lattice with a lattice const $\frac{2\pi}{a}$. An fcc lattice is the reciprocal lattice of the bcc lattice and vice versa.

A reciprocal lattice vector is expressed as $\mathbf{G} = h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^*$ where h,k,l are integers.

The fundamental translation vectors **a**, **b**, and **c** of direct lattice and $\mathbf{a}^*$, $\mathbf{b}^*$ and $\mathbf{c}^*$ of reciprocal lattice are mutually related as

$$\mathbf{a}^* = \frac{\mathbf{b}\times\mathbf{c}}{\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}},\ \mathbf{b}^* = \frac{\mathbf{c}\times\mathbf{a}}{\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}},\ \mathbf{c}^* = \frac{\mathbf{a}\times\mathbf{b}}{\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}}$$

$$\text{(or)}\ \mathbf{a}^* = 2\pi\frac{\mathbf{b}\times\mathbf{c}}{\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}},\ \mathbf{b}^* = 2\pi\frac{\mathbf{c}\times\mathbf{a}}{\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}},\ \mathbf{c}^* = 2\pi\frac{\mathbf{a}\times\mathbf{b}}{\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}}$$

The Bragg's law is also expressed as $\mathbf{2k.\ G + G^2 = 0;}$ D$\mathbf{k = G}$

A Brillouin zone is the locus of all those **k**-values in the reciprocal lattice which are Bragg reflected.

Properties of the reciprocal lattice

(a) Each vector of the reciprocal lattice is normal to a set of lattice planes of the direct lattice.

(b) If the components of G have no common factor, then |G| is inversely proportional to the spacing of the lattice planes normal to G, i.e. $|\mathbf{G}| = \frac{\mathbf{1}}{\mathbf{d_{hkl}}}$.

(c) $(\mathbf{a}^*\cdot\mathbf{b}^*\times\mathbf{c}^*) = \frac{(2\pi)^3}{\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}} = \frac{(2\pi)^3}{v_0}$.

(d) The direct lattice is the reciprocal of its own reciprocal lattice.

(e) The unit cell of the reciprocal lattice need not to be a parallelepiped.

Reciprocal lattice concept and Brillouin zone concept are useful not only in X-ray diffraction, but also in the band theory particularly in explaining the properties of materials.

3.7 Key words: Reciprocal lattice – k-space – Brillouin zone

3.8 Review questions:

1. Explain the concept of reciprocal lattice. Obtain reciprocal lattice vectors in terms of direct lattice vectors.
2. Show that the reciprocal lattice must belong to the same system as the original lattice.

3 .Derive reciprocal lattice vectors to fcc lattice.

4. Explain about reciprocal lattice and Brillouin zones in crystalline solids.

5. Derive Bragg's law in terms of reciprocal lattice.

6. Explain the properties and usefulness of reciprocal lattice.

7. Show that the reciprocal lattice of a hexagonal lattice is a hexagonal lattice with a rotation of axes.

8. Sketch the first three Brillouin zones of of a 2-dimensional square lattice.

9. A particular two dimensional lattice has the basic vectors $\mathbf{a} = 2\mathbf{x}$, $\mathbf{b} = \mathbf{x} + 2\mathbf{y}$. Find the basis vectors of the reciprocal lattice.

10. The primitive translation vectors of the horizontal space lattice are

$$\mathbf{a} = \frac{a}{2}\hat{x} + \frac{\sqrt{3}a}{2}\hat{y}\ ,\ \mathbf{b} = -\frac{a}{2}\hat{x} + \frac{\sqrt{3}a}{2}\hat{y} \text{ and } \mathbf{c} = c\hat{z}$$

(a) Show that the volume of primitive cell is $\frac{\sqrt{3}}{2}a^2c$

(b) Show that the primitive translations of the reciprocal lattice are

$\mathbf{a}^* = \frac{2\pi}{a}\hat{x} + \frac{2\pi}{\sqrt{3}a}\hat{y}, \mathbf{b}^* = -\frac{2\pi}{a}\hat{x} + \frac{2\pi}{\sqrt{3}a}\hat{y},\ \mathbf{c}^* = \frac{2\pi}{c}\hat{z}$. So that the lattice is its own reciprocal but with a rotation of axes.

11. What are Brillouin zones? Determine he reciprocal lattice vectors, which define the Brillouin zones of bcc and fcc lattices.

12. Show that the volume of the unit cell of the reciprocal lattice is inversely proportional to the volume of unit cell in direct lattice.

13. Deduce Bragg's law in terms of reciprocal lattice.

3.9 Text and reference books:

1. Elements of Solid State Physics by J.P. Srivastava (PHI)
2. Elements of solid State Physics by A. Omar (Pearson education)
3. Solid State Physics by C. Kittel (Asia Publishing house)
4. A Text Book of Solid State Physics by S.L. Kakani and C.Hemrajani (S.Chand)
5. Fundamentals of Solid State Physics by Saxena Gupta Saxena (Pragati Prakashan)

UNIT – I

LESSON: 4

Diffraction Methods

Aim: To learn about the principles of various diffraction methods

Objectives:

- To understand the phenomenon of diffraction of X rays by crystals
- To study the various experimental methods of X ray diffraction
- To understand the principles and uses of Electron and Neutron diffraction methods

Structure of the Lesson:-

Introduction:

Strong x-ray diffraction produced by crystalline solids is based on Bragg's Law. Crystals are composed of various intersecting planes each of which itself contains a number of atoms. Since the inter atomic separations in crystals are of the same order of magnitude as the wavelength of X-ray (~A), crystals act as diffraction gratings and produce diffraction on being irradiated by X-rays. W.H. Bragg and his son W.L. Bragg observed in 1913 such characteristic effects in the X-ray radiation reflected from crystals. In this phenomenon, the crystal planes are believed to act like plane mirrors. Radiations reflected from two successive parallel planes under certain conditions may interfere constructively to produce a diffraction maximum. The Bragg diffraction, shown in fig.4.1a, occurs for specular reflections (angle of incidence = angle of reflection). The constructive interference occurs when the path difference (2d sin θ) between the interfering rays equals an integral multiple of the X-ray wavelength λ. That is,

$2d \sin \theta = n\lambda$ -------- (4.1a)

where d is the inter planar spacing

θ is the angle of the incident radiation with the plane

n = 1, 2, 3 …… (order of diffraction)

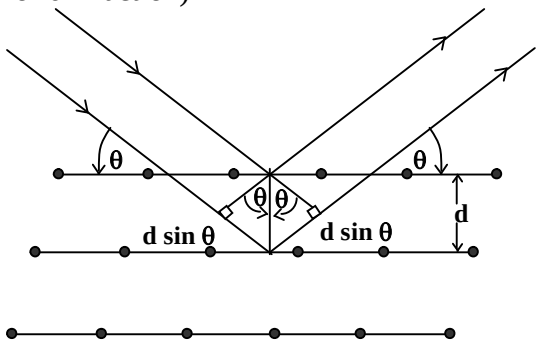


Fig 4.1a Bragg reflection from a family of planes with inter planar spacing d. Note that the incident beam is deflected by twice the Bragg angle θ.

The relation (4.1a), which is a mathematical statement of the Bragg law, shows that the diffraction effects cannot be observed from a family of planes for any arbitrary angle of incidence. Even if monochromatic X-rays are used, the match among d, θ and λ has to be sought to satisfy (4.1a) and get the Bragg diffraction. So, it is clear that Bragg diffraction is very much different from ordinary diffraction which generally puts no restriction on the incident angle.

Further, since $\sin \theta \leq 1$,

$$\lambda \leq 2d \quad \text{---------} \quad (4.1b)$$

The condition expressed by (4.1b) clearly explains why Bragg diffraction cannot occur for the visible radiation. The Bragg law as expressed by (4.1a) is essentially the consequence of the periodicity with only the elastic scattering of radiation taken into consideration .

4.1 The Diffraction of X-rays by Simple Lattice Arrays of Atoms:

When an electron comes in the path of *X*-ray wave, it is set into forced vibrations by the periodically changing electric field of the *X*-ray waves passing by it. Due to such oscillations, the electron is accelerated and decelerated. We also know that an accelerated electrically charged particle behaves as a source of electromagnetic disturbance and hence the electron acts as a source of electromagnetic wave of the same frequency and wave-length as the original *X*-ray wave. By this interaction the electron is said to scatter the original *X*-ray wave .

Let us consider one dimensional regular array of points, spaced a_0 apart as shown in fig 4.1. The atoms scatter *X*-radiations and produce about itself a new set of spherical wave envelopes. Any line up of envelopes constitute a combined wave moving in the direction of common tangent. The cooperative combination of scattered wavelets is known as diffraction. When the tangent is parallel to the original wavefront, the diffraction is known as zero order diffraction. When the tangent starts from the inner most spherical envelope of one atom, continuing through the second nearest envelope of next atom, through the third nearest envelope of next atom and so on, the wave built up along this front is known as first order diffraction. In this case the envelope difference is one. Similarly, when the envelope difference is two, we have second order diffraction i.e., the tangent connects envelope differences of two between neighboring atoms.

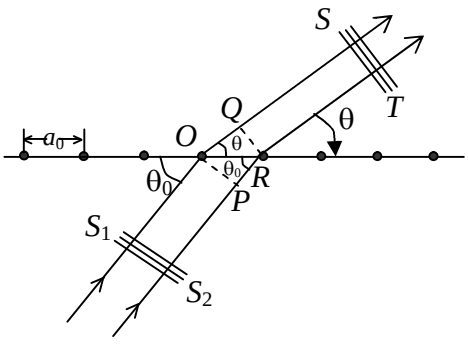


Fig 4.1 Diffraction of plane waves by one-dimensional point grating.

Now let us consider the condition of diffraction. Let the rays S_1O and S_2O represent distances in the incoming beam, and *OS* and *RT* distances in the scattered radiation.

There will be a maximum of intensity in the direction of *OS* and *RT* when the following condition is satisfied.

$S_1OS - S_2RT = m_1\lambda$ --------- (4.2a)

Where m_1 is an integer and λ is the wavelength of the radiation. The path difference ($S_1OS - S_2RT$) can be calculated with the help of fig 4.1.

$S_1OS - S_2RT = OQ - PR = a_0 \cos\theta - a_0 \cos\theta_0$ ---------- (4.2)

From equations (4.2a) and (4.2), we get

$a_0 \cos\theta - a_0 \cos\theta_0 = a_0 (\cos\theta - \cos\theta_0) = m_1\lambda$ ---------- (4.3)

The equation (4.3) shows that except zero order, no diffraction maxima can occur for wavelengths longer than $2a_0$ because the maximum value of ($\cos\theta - \cos\theta_0$)can be equal to 2. The atoms along one dimensional array scatter secondary radiations in all directions and hence the three dimensional representation of diffraction pattern may be obtained by the rotation about the line of grating. The diffraction maxima lie along cones whose mutual axis is the line of grating. The appearance of spectra of different orders on a plane parallel to line of atom array is shown in figure 4.2(a). When the incident radiation comes directly below grating i.e., $\theta_0 = \pi/2$ and $\cos\theta_0 = 0$, the spectrum of zero order appears as straight line upon this plane and other orders as hyperbolae.

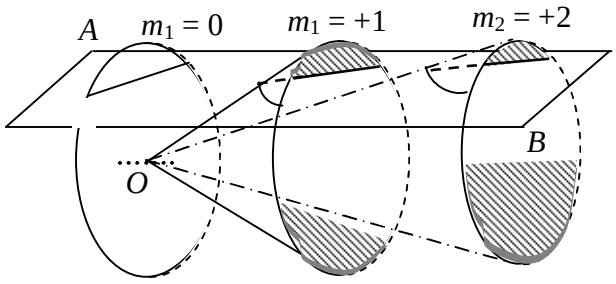


Fig 4.2(a)Diffraction pattern formed on the plane *AB* by one dimensional points array at *O*.

Now we consider the three dimensional case consisting of one-dimensional grating of equal interval a_0 along the *X*, *Y*, and *Z* directions. Corresponding to these axes, let α_0, β_0, and γ_0 be the direction cosines of the incident radiation and α, β and γ the corresponding quantities for a line in direction maxima in scattered radiation. In this case, we have the following equation (known as Laue's equations):

$a_0 (\alpha - \alpha_0) = m_1\lambda$

$a_0 (\beta - \beta_0) = m_2\lambda$ ---------- (4.4)

$a_0 (\gamma - \gamma_0) = m_3\lambda$

where m_1, m_2 and m_3 are integers.

The figure 4.2(b) illustrates the condition of diffraction in three dimensional gratings. Let the wave-fronts on the incoming monochromatic radiation be perpendicular to *Z*-axis. The axes of the diffraction cones from the component gratings parallel to *Z*-axis are vertical and their sections by the plane *AB* are circles. These circles are shown in figure 4.2(b). The maxima in the diffraction pattern will be a common intersection of two hyperbolae and circle. As no such intersection can take plane there is no monochromatic diffraction pattern at all produced for a given value of α_0, β_0, and γ_0. Since, the angle θ_0 between the beam and lattice rows is variable, it is possible to arrange this variable such that the three cos θ_0's define an identical direction. In this case diffraction occurs.

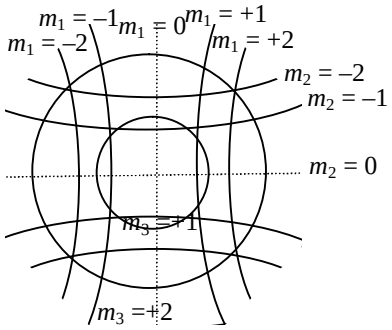


Fig 4.2(b) Illustrating the condition of diffraction in three dimensions.

4.2 Bragg's Law:

Schuster pointed out that *X*-rays are electromagnetic waves with wavelengths thousand times smaller than the visible light. In order to measure the wavelengths, a grating of corresponding dimensions is required and hence simple grating can not be used. Moreover, it is rather impossible to prepare artificially a grating of such fine dimensions. Laue and his collaborators showed that atoms in crystals are arranged in a regular manner and the spacing between them is comparable to the wavelength of *X*-rays and hence the crystal could act as suitable natural grating for diffracting the *X*-rays.

W.L.Bragg presented a simple explanation of the observed angles of the diffracted beams from a crystal. Consider a series of parallel rows in which the atoms are arranged in a given plane of the crystal. Suppose a parallel beam of X-rays is incident in a direction making a glancing angle θ with the surfaces of the planes. Each atom acts as a centre of disturbance and sends spherical wave-fronts by Huygens construction. As X-rays are much more penetrating than ordinary light, it is necessary to consider the rays

reflected not from a single layer but from several layers together. There will, however, be no diffracted beam unless the waves reflected from different planes are exactly in phase. Now they will reinforce themselves and an intense reflected beam will result. The condition that the reflected wavefronts be in same phase, is that path difference between the reflected wave from one layer and that from the next must be an exact wave length or an integral multiple of it. If there is a smallest disagreement in phase between the beams reflected from successive planes, it causes destructive interference.

Let us consider two parallel rays *LMN* and *PQR*, which are reflected by two atoms *M* and *Q* in adjacent layers as shown in figure 4.3(a). The atom *Q* is vertically below *M*. The length of the path *PQR* is greater than the length of the path *LMN*. The path difference is (*AQ* + *QB*) and according to the condition of reflection, we have

$$(AQ + QB) = n\lambda \quad \text{----------} \quad (4.5)$$

But from figure $AQ = BQ = d \sin \theta_0$

$$\therefore\ 2d \sin \theta_0 = n\lambda \quad \text{----------} \quad (4.6)$$

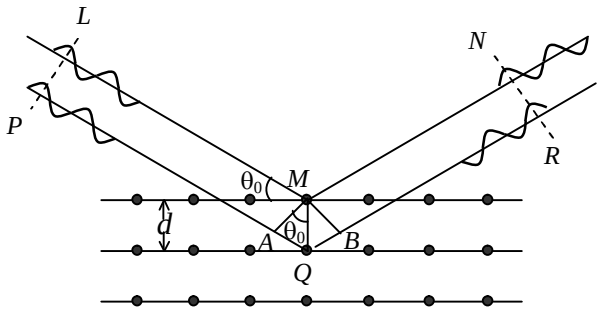


Fig 4.3(a) Illustrating the derivation of the Bragg's Law.

This is known as Bragg's law and gives the condition for the reflection of *X*-rays from series of atomic layers in a given plane. For a given set up with monochromatic radiation, the wavelength λ is fixed and hence equation (4.6) has only a particular set of solutions namely.

$$\theta_0 = \sin^{-1} 1 \left(\frac{\lambda}{2d}\right),$$
$$= \sin^{-1} 2 \left(\frac{\lambda}{2d}\right), \quad \text{----------} \quad (4.7)$$
$$= \sin^{-1} 3 \left(\frac{\lambda}{2d}\right) \ldots\ldots \text{etc.}$$

These are known as the first, second, third, etc. reflection according to *n* is 1, 2, 3 etc., this shows that the crystal cannot given rise to reflections at any angle but only at those discrete angles indicated by equation (4.7). Bragg reflection can occur only for wavelength $\lambda \leq 2d$. Due to this fact, the visible light wavelength cannot be used in diffraction.

The Bragg's law is a consequence of the periodicity of the space lattice. The law does not refer to arrangement or basis of atoms associated with each lattice point. The composition of the basis determines the relative intensity of the various order *n* of diffraction from a given set of parallel planes. As stated earlier, Bragg reflection can only occur for wavelength $\lambda \leq 2d$ and due to this fact the visible wavelength cannot be used in diffraction.

4.2.1 Bragg's law in Three Dimensions:

The three dimensional diffraction grating of a crystal may be regarded as composed of three sets of uni-dimensional gratings, each one of which consists of a row of atoms parallel to one of the three axes of crystal. Let x_1, y_1, z_1 represent the inter atomic distances along the three axes. Let, α_0, β_0 and γ_0 be the direction cosines of the incident radiation and α, β and γ the corresponding quantities for the scattered radiation. Then the conditions for diffraction along these three sets of linear grating are

$x_1 (\alpha - \alpha_0) = m_1\lambda$

$y_1 (\beta - \beta_0) = m_2\lambda$

$z_1 (\gamma - \gamma_0) = m_3\lambda$ ---------- (4.8)

where m_1, m_2 and m_3 are integers representing the order of the diffracted beam from each of the gratings. Considering the case of a cube crystal $x_1 = y_1 = z_1 = a_0$, the edge of the unit cube. Squaring and adding equations (4.8), we get

$$(\alpha^2 + \beta^2 + \gamma^2) + (\alpha_0^2 + \beta_0^2 + \gamma_0^2) - 2(\alpha\alpha_0 + \beta\beta_0 + \gamma\gamma_0) = \left(\frac{\lambda}{a_0}\right)^2 \left\{m_1^2 + m_2^2 + m_3^2\right\} \text{ ---------- (4.9)}$$

But we know that

$$\alpha^2 + \beta^2 + \gamma^2 = \alpha_0^2 + \beta_0^2 + \gamma_0^2 = 1 \quad \text{and}$$

$$\alpha\alpha_0 + \beta\beta_0 + \gamma\gamma_0 = \cos\phi$$

according to well know theorem in trigonometry. Here ϕ is the angle between incident and diffracted beams.

Applying these results to equation (4.9), we have

$$2 - 2\cos\phi = \left(\frac{\lambda}{a_0}\right)^2 \left\{m_1^2 + m_2^2 + m_3^2\right\}$$

$$4\sin^2\phi/2 = \left(\frac{\lambda}{a_0}\right)^2 \left\{m_1^2 + m_2^2 + m_3^2\right\}$$

$$2\sin\phi/2 = \left(\frac{\lambda}{a_0}\right) \sqrt{\left[\left\{m_1^2 + m_2^2 + m_3^2\right\}\right]}. \quad \text{---------- (4.10)}$$

The angle of deviation ϕ can be calculated with the help of equation (4.10). The equation (4.10) may be regarded as typical in the viewpoint of Laue and in Bragg's viewpoint, it can be expressed in the following manner : We know that a plane of atoms in a crystal is most commonly defined by its 'Miller indices', which are the reciprocals of the intercepts of the plane upon the *X*, *Y* and *Z* axes, respectively. It is customary to express these reciprocal intercepts in terms of their lowest prime numbers so that if the reciprocal intercepts are given as *h, k, l,* the actual reciprocal intercepts will be *nh, nk* and *nl* wher *n* is any integer. If, therefore, we choose such a plane of atoms that

$m_1 = nh$; $m_2 = nk$: $m_3 = nl$,

we may regard the diffracted beam of equation (4.10) to be a diffracted beam of the nth order from the plane (*h, k, l*). Hence

$$2\sin\phi/2 = \frac{\lambda}{a_0}\left\{\sqrt{\left(h^2 + k^2 + l^2\right)}\right\}n.$$

Or $$n\lambda = 2\frac{a_0}{\sqrt{(h^2 + k^2 + l^2)}}\sin\phi/2 \quad \text{---------- (4.11)}$$

But $\frac{a_0}{\sqrt{(h^2 + k^2 + l^2)}}$ is equal to the distance between successive *hkl* planes, and it is denoted by *d*. The grazing angle of incidence (or the angle of diffraction is ϕ / 2 and it is equal to θ_0. Hence

$n\lambda = 2d\sin\theta$

If instead of considering the case of cubic crystal, we consider the case of any other crystal, we arrive at a new expression in the denominator of equation (4.11) which would

represent the distance *d* between successive planes for that crystal system. Equation (4.11) is, therefore, perfectly general and applies to all types of crystals.

4.2.2 The Von Laue Treatment:

Consider the radiation scattered by two identical scattering centers *A* and *B* separated at a distance $\vec{\rho}$ as shown in fig 4.3 (b) Here the incident radiation is assumed to be a parallel beam and the scattered beam is assumed to be detected at a long distance away.

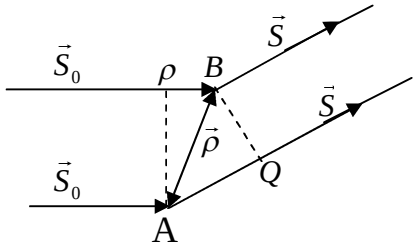


Fig 4.3 (b)

Let $\vec{S}_0$ and $\vec{S}$ be the unit vectors along the direction of incident beam and scattered beam respectively. The path difference between the radiation scattered from radiating center *B* and that scattered from *A* is given by

$$BP - AQ = \vec{S}_0 \cdot \vec{\rho} - \vec{S} \cdot \vec{\rho} = (\vec{S}_0 - \vec{S}) \cdot \vec{\rho}$$

Where *BP* and *AQ* are the projections of $\vec{\rho}$ on incident and scattered beams respectively. The vector $\vec{S}_0 - \vec{S}$ represents the direction normal to the plane that reflects the incident direction to the scattered direction as shown in fig 4.3 (c).

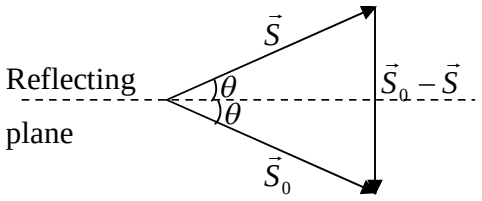


Fig 4.3(c)

The plane may be regarded as the reflecting plane in the sense of Bragg treatment.

If 2θ be the angle which $\vec{S}$ makes with $\vec{S}_0$, then the magnitude of $(\vec{S}_0 - \vec{S})$ i.e., $|\vec{S}_0 - \vec{S}| = 2 \sin\theta$ because $\vec{S}_0$ and $\vec{S}$ are unit vectors. The phase difference between the radiation scattered from the points *A* and *B* is given by

$$\phi_\rho = \frac{2\pi}{\lambda} \times \text{ Path difference } = \frac{2\pi}{\lambda}(\vec{S}_0 - \vec{S}) \cdot \rho \quad \text{----------} \quad (4.12)$$

The condition that there be a diffraction maxima in the direction $\vec{S}$ is that the phase difference between the scattering contributions from *A* and *B* must be an integer multiple of 2π. This condition can be satisfied in nearest–neighbor atoms if the radiation scattered by them add in phase. We know that nearest neighbor atoms are separated by primitive translation distances *a*, *b* and *c* hence replacing $\vec{\rho}$ by $\vec{a}, \vec{b}$ or $\vec{c}$ in equation (4.12) and applying the condition of maximum diffraction we have

$$\frac{2\pi}{\lambda}(\vec{S_0} - \vec{S}) \cdot \vec{a} = 2\pi h' = 2\pi nh$$
$$\frac{2\pi}{\lambda}(\vec{S_0} - \vec{S}) \cdot \vec{b} = 2\pi k' = 2\pi nk \quad \text{---------- (4.13)}$$
$$\frac{2\pi}{\lambda}(\vec{S_0} - \vec{S}) \cdot \vec{c} = 2\pi l' = 2\pi nl$$

Here *h′*, *k′*, and *l′* may be any three integers. These integers may form a smallest set of integers in such a way that they may be identical with Miller indices of a crystal plane or they may contain a common integer factor greater than unity. Considering the common factor as *n*, we can write them as $h' = nh$, $k' = nk$ and $l' = nl$. Now *h*, *k* and *l* are now three smallest integers identical with Miller indices. In a particular case *n* may be unity and in this case either set of integers may have the same meaning. Let α, β and γ be the angles between the scattering normal $(\vec{S_0} - \vec{S})$ and the $\vec{a}$-, $\vec{b}$- and $\vec{c}$- axes of the crystal respectively.

Now $(\vec{S_0} - \vec{S}) \cdot \vec{a} = 2 \sin\theta \cdot a \cdot \cos\alpha = 2a \sin\theta \cdot a \cdot \cos\alpha$

$(\because |\vec{S_0} - \vec{S}| = 2 \sin\theta)$

Similarly

$(\vec{S_0} - \vec{S}) \cdot \vec{b} = 2b \sin\theta \cos\beta$ ---------- (4.14)

and $(\vec{S_0} - \vec{S}) \cdot \vec{c} = 2c \sin\theta \cos\gamma$

Applying eqs. (4.14) to eqs. (4.13), we have

$2a \sin\theta \cos\alpha = nh\lambda$

$2b \sin\theta \cos\beta = nk\lambda$ ---------- (4.15)

$2c \sin\theta \cos\gamma = nl\lambda$

These conditions are known as Laue equations. The equations (4.15) serve to determine a unique value of θ, thus defining a scattering direction. From these equations we note that the direction cosines of the scattering normal ($\vec{S}_0 - \vec{S}$) are proportional to $\frac{h}{a}$, $\frac{k}{b}$ and $\frac{l}{c}$. We also know that direction cosines of the normal to *hkl* family of planes are proportional to $\frac{h}{a}$, $\frac{k}{b}$ and $\frac{l}{c}$. Therefore, we conclude that the scattering normal ($\vec{S}_0 - \vec{S}$) is identical to the normal to (*hkl*) planes. Hence the (*hkl*) planes may be regarded as the reflecting planes in the sense of Bragg treatment. In this way the Laue equations provide the validity of Bragg treatment.

If *d* be the spacing between two adjacent planes of a family (*hkl*) we have

$$d = \frac{a}{h}\cos\alpha = \frac{b}{k}\cos\beta = \frac{c}{l}\cos\gamma .$$

Then from either of equations (4.15), we have

$$2d \sin\theta = n\lambda \quad \text{----------} \quad (4.16)$$

which is the Bragg equation.

Here *n* is the order of the diffraction and it is the greatest common factor among the integers h', k', and l' enclosed in paranthesis i.e., ($h'k'l'$). Thus the first order diffraction maxima for the (111) planes is referred as (111) reflection. Similarly the second order diffraction maxima for the family of planes as (222) reflection and so on.

4.3 Experimental method in X-ray Diffraction:

4.3.1 The Laue Method:

The experimental equipment here is relatively simple and is shown in figure 4.4. The crystal is held stationary in the beam of X– rays. The rays after passing through the crystal are diffracted and are recorded on the photographic plate place at a certain distance from the crystal. Before passing through the crystal, the X– rays are limited to a fine pencil by a slit system. The diameter of the pinhole is importance from the stand point of detail in diffraction pattern. The smaller is the diameter the sharper is the interference. The crystal is set on a holder to adjust its

orientation. The Laue method employs white radiation (X– rays beam of continuous range of wavelength) which is usually obtained from a tungsten target at about 60,000 volts.

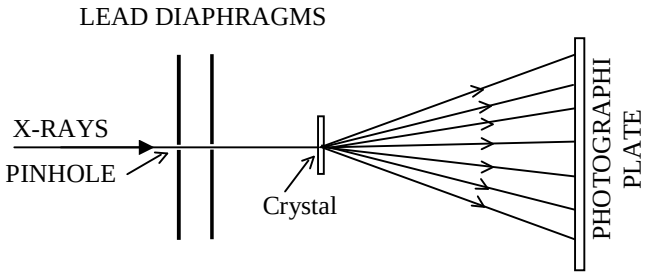
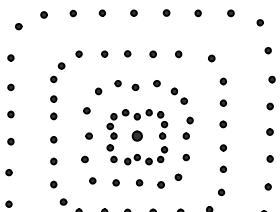


Fig 4.4 Diffraction of white X-rays by a single stationary crystal

We have seen that if a beam of X– rays of a given wavelength λ is passed in a given direction through a crystal, the diffraction is not in general to be expected. This is because very few sets of planes would be in a favourable position to meet the requirements of the Bragg equation and reflections would of course be rare. But since there is a whole range of wavelengths in the continuous spectrum, there will be discrete values of λ which satisfy the Bragg condition-no matter what may be the orientation of the lattice planes. In another words, for any values of d and θ_0, there will be found in the beam some value of λ such that diffraction can occur. We know that atoms of crystal have an orderly arrangement in all the these dimensions in space, hence the diffraction of X-ray will occur from many families of atomic planes at once, each family picking out the wavelength which it can diffract at the angle at which it finds itself. The sort of diffraction pattern obtained is illustrated in figure 4.5.

Examination of the Laue photograph shows that the spots do actually occur at the positions to be expected from the reflection law. When the primary beam passes along the axis symmetry of the crystal, the Laue pattern consists of a series of spots whose loci are ellipses which pass through the *central image* made by primary beam. The spots on any one ellipse are produced by planes belonging to the same zone i.e. planes which are parallel to one common direction.

The Laue pattern can be used to orient crystal for solid state experiments. Let us consider the case of a crystal with four fold axial symmetry which is oriented with the axis parallel to the beam. Each reflecting plane then selects a wavelength satisfying the



Bragg equation from the incident beam. The Laue pattern obtained in this case shows the four fold symmetry.

Practically this method is never used for crystal structure determination. In this case several wavelengths may be reflected in different orders from a single plane, so that different orders of reflection may superpose on a single spot. Due to this fact the determination of a reflected intensity is difficult and thus the determination of the basis.

4.3.2 Rotating Crystal Method:

The rotating crystal method is shown in fig 4.6. The X rays are generated in the X-ray tube and then the beam is made nearly monochromatic by a filter. The beam is now passed through collimating system, which permits a fine pencil of parallel X-rays. Now a crystal is mounted on a shaft which is arranged perpendicular to the incident beam. This shaft is rotated at a uniform angular rate by a small motor. Here one thing should be remembered that the crystal dimensions should not be greater than one mm, so that the crystal may completely be bathed by the incident radiation.

When the crystal is set into slow rotation about a fix axis, the sets of plane come successively into their reflecting position, i.e., the value of θ satisfies the Bragg equation. Beams from all planes parallel to the vertical rotation axis will lie in the horizontal plane. The diffraction pattern may be registered either upon a photographic plate perpendicular to X-rays beam pr upon a film a cylindrical camera, the axis of which coincides with the axis of rotation of the crystal.

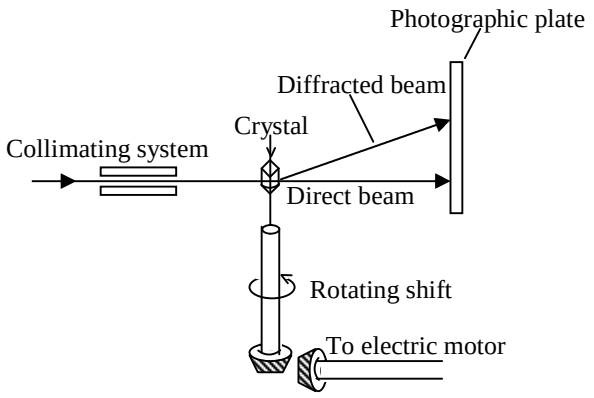
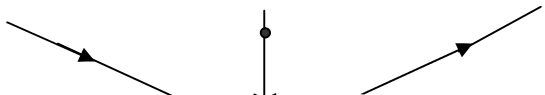


Fig 4.6. Rotating crystal method.

The two types of photographs may be obtained. The first is *complete rotation* photograph in which the turning of the crystal through a series of complete revolutions takes place. It is observed that each set of planes in the crystal diffracts four times during the rotation. The four diffracted beams are distributed in a rectangular pattern about the center point of the photograph. The second is *oscillation* photograph. In this case, instead of being given a continuous rotation, the crystal is made to oscillate back and



forth with a constant angular speed through a chosen angular range. It is observed that the oscillation photograph lacks the symmetry, which is shown by the complete rotation photograph. The limited range in oscillation photograph reduces the possibility of overlapping reflections.

Now we shall discuss the theory of the rotation photograph. Suppose the crystal to be rotated in the *X*-ray beam is mounted, in such a way that one of the crystallographic axes coincides with the axis be the *Z*-axis. Now consider the two neighbouring atoms *A* and *B* at lattice poiot along the *Z*-axis. The distance between *A* and *B* will be primitive translation *c* as shown in fig 4.7(a). In this case the condition of diffraction is given by

$$c(\cos\delta + \cos\phi) = n\lambda \quad \text{----------} \quad (4.17)$$

Where *n* is an integer representing the order of diffraction by the line grating. When the rotation axis is perpendicular to the beam $\phi = 90^0$ and the condition (4.17) now becomes

$$c \cos \delta = c \sin \mu = n\lambda \quad \text{----------} \quad (4.18)$$

For various order of diffraction n assumes values such as 0, 1, 2, 3, ….etc giving rise to the series of equations

$$\cos \delta_0 = \sin \mu_0 = 0$$
$$\cos \delta_1 = \sin \mu_1 = \lambda/c$$
$$\cos \delta_2 = \sin \mu_2 = 2\lambda/c$$
$$\cos \delta_3 = \sin \mu_3 = 3\lambda/c \quad \text{----------} \quad (4.19)$$
$$\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots$$
$$\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots$$
$$\cos \delta_n = \sin \mu_n = n\lambda/c$$

When λ is constant, such equations give the loci of all possible diffracted rays as the crystal turns. These loci are elements of a series of cones, shown in figure 4.7(b) of

which the half apex angles are given by δ_1 , δ_2 ,δ_3 etc. Any element of each of the cones makes angles μ_1, μ_2 , μ_3 etc. respectively with the horizontal plane.

The central horizontal plane contains all diffracted beams of order of zero. The Miller indices of all planes giving diffracted beams in horizontal plane must be presented by (*h, k, l*). similarly *X*-rays diffracted by planes of indices (*h, k, l*) lie on the first order cone defined by

$\mu_1 = \sin^{-1} \lambda/c$. ---------- (4.20)

Hence in general, all the beams diffracted from planes of indices (*hkl*) lie on the n th order cone defined by

$$\mu_n = \sin^{-1} \frac{n\lambda}{c} \quad \text{---------- (4.21)}$$

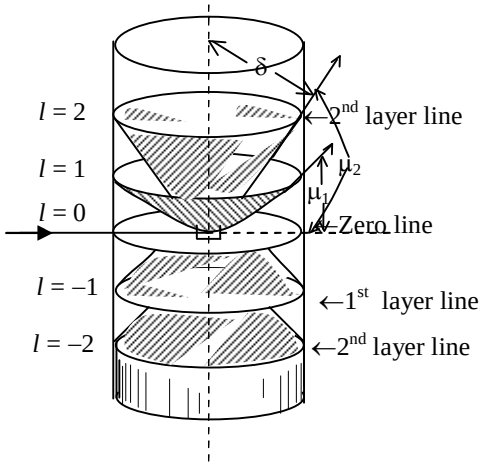


Fig 4.7(b) Formation of layer lines on a cylindrical film

Figure (4.7(b)) shows that the cones intersect the cylindrical film in a series of circles lying in planes perpendicular to the axis of rotation, which becomes a series of parallel straight lines, known as layer lines when the film is unrolled. The cone for $n = 0$ is plane perpendicular to the axis and containing the direction of incidence and the intersection of this with the film is called the equatorial or zero layer line. The succeeding lines are termed the first, second, third, etc. layer lines.

The dimensions of the unit cell of the structure are determined in the following way. The spacings of layer lines give lattice translations. If the distance of the film from the crystal is known, the distances of the layer lines from the equatorial line give the values of angle δ and λ is known, the value of *C* can be calculated. Now the rotation photographs are taken separately about all three axes. in this way translations *a, b, c* are calculated which give the dimensions of unit cell of structure. The rotating crystal method is this very powerful of giving the size of unit cell.

A powerful method which is much used now a days is due to Weissenberg. In the Weissenberg method, the crystal is continuously rotated through 180^0 and back and at the same time the cylindrical camera containing the film moves at a constant speed backward

and forward parallel to the axis of rotation. The movements of film and camera are so synchronized that a given position of camera corresponds accurately to a definite angular position of the crystal in its rotation. With the help of co-ordinates of the spot on the film, the angle of reflection as well as the position of the reflecting plane can be known. To allow only the spots due to one layer line, a metallic cylinder having an equatorial slit of a few millimeters is introduced between the crystal and the film. The photograph records all the spots belonging to that layer line spread pit into a characteristic pattern on the single film and can readily be indexed.

4.3.3 The Powder-Photograph Method:

The powder method is the only method which can be used with that large class of substances which can not be obtained easily in the form of perfect crystals of appreciable size. This class includes not only the most metals and their alloys but also a large number of compounds. The method was devised independently by Debye and Scherrer in Germany and by Hill in America. In this method a monochromatic X-rays beam is used and instead of using a single crystal, fine powders of crystalline aggregates of all kinds, having random of chaotic orientations are used. Such a powder requires no rotation because every atomic plane is present in every possible orientation and hence the diffraction depends upon the fact that in a fine powder the grains are arrange in an entirely chaotic manner. The entirely random orientations of the grains with respect to the beam means that some of them will be in a position to reflect the radiation from an important set of planes. Now diffracted rays go out from individual crystallites, which happen to be oriented with planes making an angle θ with the beam satisfying the Bragg equation. Any fragment in which the normal to the plane in questions makes an angle $(90 - \theta^0)$ with the incident beam will be in a position to reflect and since all orientations of the fragment are equally likely, the reflected rays will form a cone, concentric with the original beam and whose semi vertical angle is 2θ. There is such a cone of diffracted rays for each set of planes. The cones intercept the film in a series of concentric circular halves, from the radii of which the angle θ and hence the spacing of the planes can be deduced. The formation of powder photograph is shown in fig (4.8).

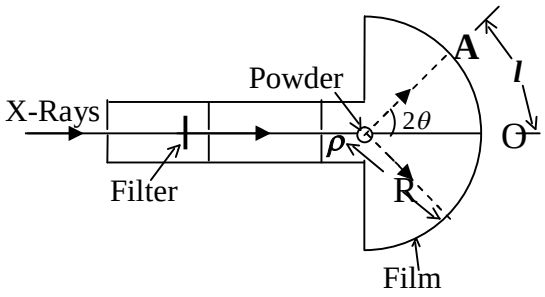


Fig 4.8(a) Design of an arrangement for taking powder photographs.

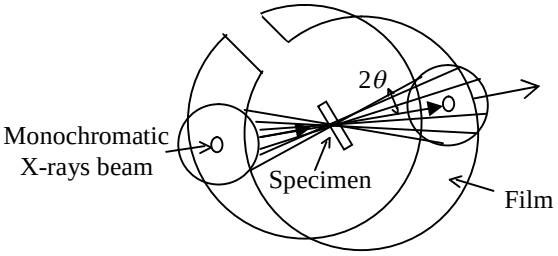


Fig 4.8(b) X-ray powder diffraction camera.

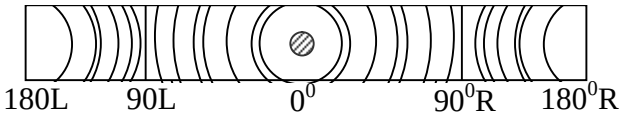


Fig 4.8(c) Arrangement of lines in a powder photograph

The radiation is made approximately monochromatic with the help of filter as shown in fig 4.8(a). *P* is the powder and *O* is the point where the direct beam would have struck the film. Point *A* on the film corresponds at which a spectrum with glancing angle θ is formed. The diffracted maxima lie on cones coaxial with the direct beam, and if a photographic plate is mounted normal to the direct beam, and if a photographic plate is mounted normal to the direct beam, concentric circles are registered upon it as shown in fig 4.8(b). Usually a plate or film in the form of cylindrical shame is employed whose axis is perpendicular to the beam. There appear arcs of the circles as shown in fig 4.8(c).

From fig. 4.8(c) it is observed that when rays are different through small angles, they make arcs around the central spot on the film, when the rays are diffracted through 90^0, the cones become flat and the corresponding trace is a straight line. When the diffracted angle increases above 90^0, the curvature is reversed and when the angle approaches to 180^0, the traces are nearly circular. Thus the curvature of lines changes from the centre to the outside of the film.

Now considering fig 4.8(a), if *l* is the distance from *O* to *A*, measured on the film and R is the radius of camera, then $\theta = l/2R$. In this way by measuring *l*, the value of θ can be calculated.

This method is very useful in investigating the structures of simple crystals particularly belonging to cubic system of which spacings *a, b, c* of unit cell are all equal. In these crystals, there are certain definite relationship between the angles at which spectra can occur. The spacings of all planes parallel to faces of the same form {*hkl*} are equal and therefore produce spectra at the same–angle. In the most general case, in which all *h, k, l* are different, there are 48 faces in the form and 24 sets of planes all having the same spacing. These all co-operate to produce one line on the film. When the three axes of different length are at right angles to one another, the general form {*h, k, l*} corresponds to six different spacing and hence there are six different lines on the film.

4.4 Electrons Diffraction:

In 1924 de Broglie suggested that a material particle like electron in motion is always associated with a wave whose wavelength is given by $\lambda = h/mv = h/p$, where *h* be the Plank's constant and *p*, the momentum of the particle. The experimental evidence of this fact was provided by Davisson and Germer and G.P. Thomson. Since the material particle (electron) can be described by wave, it can be diffracted by crystals like *X*-rays. The amount of matter required to produce electron diffraction is small as compared with *X*-rays diffraction and the time of exposure required for a photographic record is measured in a fraction of a second. Thus we here introduce the diffraction of electrons in connection with the exploration of the crystal structure. The discussion would be confined to comparison between the electron diffraction and *X*-rays diffraction.

First of all we shall calculate the wavelength of an electron beam appropriate to the crystal diffraction work. According to de Broglie idea, the wavelength λ associated with the particle is given by

$$\lambda = \frac{h}{mv} \quad \text{----------} \quad (4.22)$$

If the electrons are accelerated by an applied electric potential V, then the kinetic energy $\frac{1}{2}.mv^2$ of the electron is given by

$\frac{1}{2} mv^2 = eV$ ---------- (4.23)

where e is the electronic charge.

From equation (4.22) and (4.23) we have

$$\lambda = \frac{h}{\sqrt{(2meV)}} = \sqrt{\left(\frac{150}{V}\right)} \quad \text{----------} \quad (4.24)$$

where λ is in Å and V in volts.

It is evident from equation (4.24) that only 150 volts are required to produce electrons of a wavelength of one Å suitable for diffraction work. The X-rays suitable for diffraction work require approximately 12000 volts for the same wavelength.

The diffraction of electrons takes place in a similar way as in case of X-rays and the diffraction pattern obtained can be interpreted exactly in the same way as X-rays diffraction pattern.

Some special features of electron diffraction are:

1. In contrast with X-rays, electrons are scattered by the nucleus as well as by the electrons in atoms.
2. The scattering factor for electrons decreases with increasing value of Bragg's angle θ as in X-rays. However, the scattering efficiency of atoms is considerably greater for electrons than X-rays. Due to this reason the diffraction of electrons by gases requires much shorter exposure times than does X-rays diffraction.
3. Electrons are charged and interact strongly with matter as compare with X-rays.
4. Electrons penetrate a relatively short distance into a crystal. At normal incidence, an electron of about 50keV has a penetration depth of only about 500Å, whilest for small angles of incidence this may be only about 50 Å measured perpendicular to the surface. Thus the electron diffraction is particularly useful in investigating the structure of thin surface layers such as oxide layers on metals. These layers are not detected by X-rays diffraction because they penetrate deep into the solids and produce a pattern, which is the characteristic of the interior of the solid.
5. Electron diffraction can be used for the studies of orientation, lattice parameter and perfection of evaporated thin films.
6. By electron diffraction, the dislocation patterns in thin metallic foils can be see.

4.5 Neutrons Diffraction:

As neutrons are associated with wave and hence they can be diffracted from crystals like electrons. Here we shall consider the neutron diffraction by crystal in connection with the exploration the crystal structure, of course, the discussion will be confined to a comparison between X-ray, electron and neutron diffraction.

First of all we shall calculate the wavelength of suitable neutrons. We know that the wavelength in case of neutrons can be expressed as

$$\lambda = \frac{h}{p} = \frac{h}{M_n v} \quad \text{----------} \quad (4.25)$$

$$\text{E} = \frac{p^2}{2m} = \frac{\hbar^2 K^2}{2m} \qquad (\because\ p = \hbar K)$$

$$= \frac{10^{-68} \times 10^{21}}{3 \times 10^{-27}} \, Jouls$$

$$= 3 \times 10^{-21} \, jouls$$

$$= 200\text{K} = 0.02 \text{ eV}$$

where M_n is the mass of the neutrons. The mass of a neutron is 2000 times large in comparison to an electron i.e., the wavelength associated with a neutron is about 1/2000 times that for an electron of the same velocity. The energy $\in$ of the suitable neutrons may be calculated by the following expression

$$\lambda = \frac{h}{\sqrt{(2\text{M}_n v)}} \quad \text{----------} \quad (4.26)$$

By equation (4.26), the energy $\in$ of neutrons is approximately 0.1 e.V. for λ = 1Å which is required for diffraction work. It is important to note that for X-rays of 1Å the energy required is 10^4 e.V. while for electron it is about 10^2 e.V. In case of neutrons, the diffraction patterns are formed in a similar manner a in case of X-rays.

Some special features of the neutrons diffraction are :

1. Neutrons are scattered chiefly by the nuclei of the atoms.
2. As the wavelength of the neutrons is much greater than the dimensions of the scattering nucleus ($\cong 10^{-13}$ cm), the atomic scattering factor is nearly independent of the scattering angle.

3. The scattering power does not vary in a regular manner with atomic number. Due to this fact the light elements such as hydrogen and carbon produce relatively strong scattering than X-rays scattering because the X-rays scattering is done by electrons. This feature of the neutron diffraction enable us to deduce the positions of hydrogen and carbon atoms in a number of organic crystals.
4. The scattering from neighboring elements in the periodic system may differ appreciably. Hence neutron diffraction allows to detect with relative ease, ordered phases of an alloy, such as FeCo, where as their detection by X-rays is difficult.
5. The neutrons possess magnetic moments and these moments interact with the magnetic moments of the scattering atoms of the solid. This gives an additional scattering mechanism for neutrons, which often out-weights the nuclear scattering. Thus neutron diffraction methods are exceedingly valuable in structural studies of magnetic crystals.

In paramagnetic substances in which the atomic moments are randomly oriented, the magnetically scattered neutrons are incoherent in phase resulting in a diffuse background. This diffuse background of magnetic scattering is then superimposed on the lines produced by the nuclear scattering.

In ferromagnetic substances in which the magnetic moments within a domain are linked up in parallel, the diffuse background is absent.

In an antiferromagnetic solid the magnetic moments are aligned antiparallel and hence from the point of view of the neutron such atoms would appear to be different.

Figure (4.9) shows neutron diffraction patterns for Mno (Mn ion has a permanent moment), which is known to be an anti-ferromagnetic solid below and above its curie temperature (120^0).

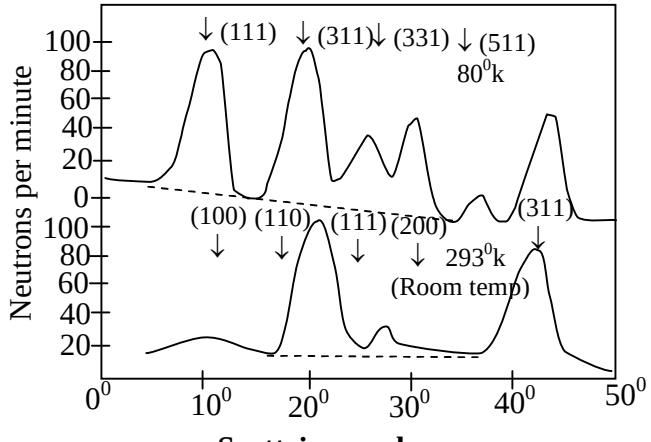


Fig 4.9 The neutron diffraction patterns of MnO at 80^0K and 293^0K (below & above the curie temperature).

At room temperature (293^0K) the pattern shows coherent diffraction peaks, like those of X-ray diffraction, and in the positions expected from a lattice of NaCl structure. The diffuse background of magnetic scattering is also a visible, which is inductive of no magnetic order at all. At low temperatures, in a addition to these peaks certain other peaks are located at positions which one can not expect on the basis of chemical structure of a unit cell.

4.6 Laue Derivation of Amplitude of Scattered Wave:

Let us consider the case of a plane wave which is incident on a small crystal. Again let in the free space at point x the amplitude be F, then

$$F(x) = F_0 e^{i(\mathbf{K}\cdot\mathbf{x}-\omega t)} \text{ ---------- } (4.27)$$

Referred to an origin at x = 0. Equation (4.27) represents a traveling wave having wave vector K, angular frequency ω and the wavelength λ = 2π/K, Now we place the crystal in the beam with origin O chosen anywhere within the crystal. Here it is assumed that the incident beam is not greatly disturbed by crystal i.e., neither by the refractive index nor by the loss of energy through scattering. At a point $\vec{\rho}$, the amplitude of the incident wave is given by

$F(\vec{\rho}) = F_0 e^{i(\mathbf{K}.\vec{\rho})}$ [at instant of time t = 0] ---------- (4.28)

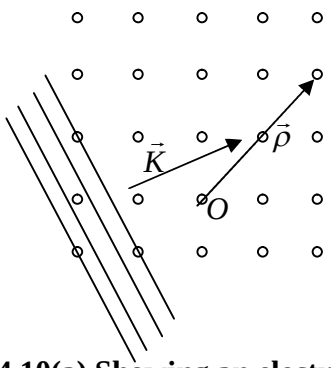


Fig 4.10(a) Showing an electromagnetic wave incident upon a small crystal.

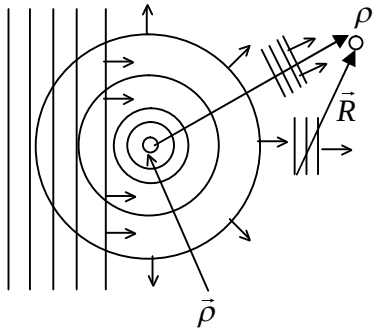


Fig 4.10(b) Showing the radiation scattered $\vec{\rho}$

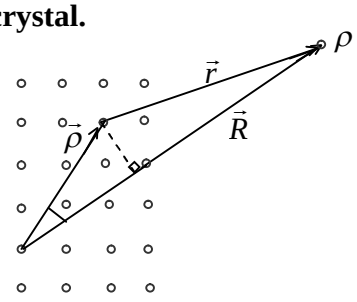


Fig 4.10(c) Showing the wave scattered at *O* as received at R

The atom at $\vec{\rho}$ will scatter some of the radiation out of the incident beam. As in fig 4.10 (a) and (b), the amplitude of the scattered radiation as seen at point P is R = $\vec{\rho}$ + r i.e., at a point distant r from $\vec{\rho}$ outside the crystal will be proportional to

$$(F_0 e^{iK\vec{\rho}})\left(\frac{e^{iKr}}{r}\right) \quad \text{---------- (4.29)}$$

where the first parenthesis contains the amplitude and phase factor of the incident beam and the second parenthesis describes the spatial variation of the radiation scattered from a point atom at $\vec{\rho}$. The total phase factor at R is

$$e^{i\mathbf{K}.\vec{\rho}}.\, e^{i\mathbf{K}r} = e^{(i\mathbf{K}.\vec{\rho} + i\mathbf{K}r)} \quad \text{---------- (4.30)}$$

From fig 4.10(c), we have

$$r^2 = (R - \vec{\rho})^2 = R^2 + \rho^2 - 2\rho R\cos(\vec{\rho}, R) = R^2\left[1 + \frac{\rho^2}{R^2} - \frac{2\rho}{R^2}\cos(\vec{\rho}, R)\right]$$

when R is at a large distance so that $\rho/R << 1$, we have

$$\cong R^2\left[1-\frac{2\rho}{R^2}\cos(\vec{\rho},R)\right]$$

or $r \cong R[1-(2\rho/R)\cos(\vec{\rho},R)]^{1/2}$

$$\cong R-\rho\cos(\vec{\rho},R)$$

now from equation (4.30), the total phase factor of the scattered wave on arriving at R is

$e^{[i\mathbf{K}\rho+i\mathbf{K}R-i\mathbf{K}\rho\cos(\rho,R)]}$ ---------- (4.31)

Now it can be assumed that the amplitude of the wave scattered from an element of volume of the crystal is proportional to the electron concentration $n(\vec{\rho})$ in the volume element. Hence the amplitude of the scattered radiation at R will be proportional to the integral

$$\int dV\,.\,n(\vec{\rho})\,.\,e^{[i\mathbf{K}\cdot\rho-i\mathbf{K}\cdot\rho\cos(\rho R)]} \quad \text{---------- (4.32)}$$

The factor e^{iKR} is omitted, being constant over the volume. Equation (4.32) can be written in more compact form i.e.

$$\int dV\,.\,n(\vec{\rho})\,.\,e^{-i\rho\cdot\Delta R} \quad \text{---------- (4.33)}$$

where

$iK\,.\,\vec{\rho}-iK\rho\cos(\vec{\rho},R)\equiv i\vec{\rho}\,(K-K')\equiv -i\vec{\rho}\,.\,\Delta K,$

K' is the wave vector in scattering direction R and

$$\Delta K\equiv K'-K.$$

Equation (4.33) gives the amplitude of the scattered wave.

Scattering from lattice of point atoms.

Consider a finite crystal and let all points be scattering centers.

The lattice points are defined by

$\vec{\rho} = ma + nb + pc$ ---------- (4.34)

where m, n, p are integers. The amplitude a of radiation by entire crystal seen at R will be proportional to

$$a \equiv \sum_{\rho} e^{-i\vec{\rho}.\Delta K}$$

$$= \sum_{mnp} e^{[-i(ma+nb+pc).\Delta K]}$$

$$= (\sum_{m} e^{[-im(a.\Delta K)]}) \ (\sum_{n} e^{[-in(b.\Delta K)]}) \ (\sum_{p} e^{[-ip(c.\Delta K)]}) \quad \text{---------- (4.35)}$$

We know that the intensity is the square of the amplitude, hence

$$\text{Intensity} = |\sum_{m} \exp.[-\text{im}\,(a\,.\,\Delta K)]|^2 \, |\sum_{n} \exp.[-\text{in}\,(b\,.\,\Delta K)]|^2 \, |\sum_{p} \exp.[-\text{ip}\,(c\,.\,\Delta K)]|^2$$

---------- (4.36)

Now we shall consider the value of one sum out of the three from equation (4.36). Let us consider the crystal of dimension Ma is the direction a where M is integer. We may let that m, n and p run from 0 to (M – 1), the crystal will have M^3 primitive cells because the total volume will be M^3 abc; abc is the volume of one cell. Thus

$$|\sum_{m=0}^{M-1} \exp.[-im(a\,.\,\Delta K)]|^2$$

This is a geometric series and the summation therefore will be

$$\left[\frac{1-\exp.[-iM(a.\Delta K)]}{1-\exp.[-i(a.\Delta K)]}\right]^2 \quad \text{---------- (4.37)}$$

Now using the series.

$$\sum_{m=0}^{M-1} x^m = \sum_{m=0}^{\infty} x^m - \sum_{m=M}^{\infty} x^m = \frac{1}{1-x} - \frac{x^M}{1-x} \text{ with } x \equiv \exp\,.\,[-i\ (a.\ \Delta K)] \quad \text{---------- (4.38)}$$

Thus sum in equation (4.37) can be written as

$$\left[\frac{\exp.\left[-\frac{1}{2}iM(a.\Delta K)\right]}{\exp.\left[-\frac{1}{2}i(a.\Delta K)\right]} \cdot \frac{\exp.\left[\frac{1}{2}iM(a.\Delta K)\right] - \exp.\left[-\frac{1}{2}iM(a.\Delta K)\right]}{\exp.\left[\frac{1}{2}i(a.\Delta K)\right] \ - \ \exp.\left[-\frac{1}{2}i(a.\Delta K)\right]}\right]^2$$

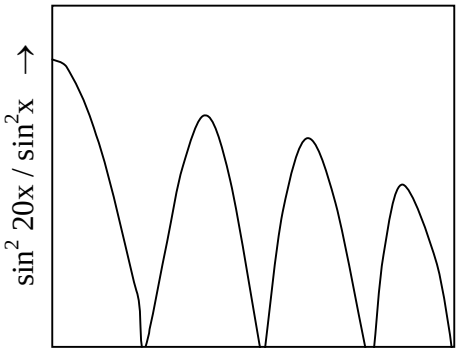
Multiplying by its complex conjugate, we get

$$\frac{\sin^2 \frac{1}{2}M(a.\Delta K)}{\sin^2 \frac{1}{2}(a.\Delta K)}$$

$$\therefore \left|\sum_{m} \exp.[-im(a\,.\,\Delta K)]\right|^2 = \frac{\sin^2 \frac{1}{2}M(a.\Delta K)}{\sin^2 \frac{1}{2}(a.\Delta K)}, \quad \text{---------- (4.39)}$$

A plot of the function in equation (4.39) is shown in figure (4.11) for M = 20. The intensity will be maximum when each term in the sum on left hand side is unity i.e.

$a\,.\,\Delta K = 2\pi q$ ---------- (4.40)



where q is an integer. At these values equation (4.39) has a value M^2.,

Now we shall consider the width of the maxima as the value of $a \cdot \Delta K$ is slightly changed. Let it be changed by $\in$ where $\in$ is the smallest non-zero number i.e.

$a \cdot \Delta K = 2\pi q + \in$.

When $\in = 2\pi/M$, we have

$\text{Sin } ½ M (a \cdot \Delta K) = \text{Sin } ½ M (2\pi q + 2\pi/M) = \sin \phi (Mq + 1) = 0$

In this way if we choose $\in = 2\pi/M$, the width of maxima is proportional to $2\pi/M$ or $1/M$. This shows that larger is the length of the crystal, smaller will be the width of maxima.

The area under the central maxima of equation (4.39) is given by the height ($\propto M^2$) times the width ($\propto$, 1/M), so the area is proportional to M, the number of atoms in the line. If the crystal in three dimensions has M^3 atoms, the scattered intensity will be directly proportional to M^3.

4.7 Summary:

The Bragg law, shows that the diffraction effects cannot be observed from a family of planes for any arbitrary angle of incidence. Even if monochromatic X-rays are used, the match among d, θ and λ has to be sought using the relation $2d \sin \theta = n\lambda$, to get the Bragg diffraction. Bragg diffraction is very much different from ordinary diffraction which generally puts no restriction on the incident angle.

Further, since $\sin \theta \leq 1$,

The condition expressed by $\lambda \leq 2d$ explains why Bragg diffraction cannot occur for the visible radiation.

Laue and his collaborators showed that atoms in crystals are arranged in a regular manner and the spacing between them is comparable to the wavelength of *X*-rays and hence the crystal could act as suitable natural grating for diffracting the *X*-rays.

The Laue method employs white radiation (X– rays beam of continuous range of wavelength) which is usually obtained from a tungsten target at about 60,000 Volts.

We know that atoms of crystal have an orderly arrangement in all the these dimensions in space, hence the diffraction of X-ray will occur from many families of atomic planes at

once, each family picking out the wavelength which it can diffract at the angle at which it finds itself. The Laue pattern can be used to orient crystal for solid state experiments. Practically this method is never used for crystal structure determination.

In the Weissenberg method, the crystal is continuously rotated through 180^0 and back and at the same time the cylindrical camera containing the film moves at a constant speed backward and forward parallel to the axis of rotation. The movements of film and camera are so synchronized that a given position of camera corresponds accurately to a definite angular position of the crystal in its rotation. With the help of co-ordinates of the spot on the film, the angle of reflection as well as the position of the reflecting plane can be known.

The powder method is the only method which can be used with that large class of substances which can not be obtained easily in the form of perfect crystals of appreciable size. This class includes not only the most metals and their alloys but also a large number of compounds. This method was devised independently by Debye and Scherrer in Germany and by Hill in America. In this method a monochromatic X-rays beam is used and instead of using a single crystal, fine powders of crystalline aggregates of all kinds, having random of chaotic orientations are used. Such a powder requires no rotation because every atomic plane is present in every possible orientation and hence the diffraction depends upon the fact that in a fine powder the grains are arrange in an entirely chaotic manner. The entirely random orientation of the grains with respect to the beam means that some of them will be in a position to reflect the radiation from an important set of planes. Now diffracted rays go out from individual crystallites, which happen to be oriented with planes making an angle θ with the beam satisfying the Bragg equation.

ELECTRON DIFFRACTION: The wave length associated with an electron is given by the relation

$$\lambda = \frac{h}{\sqrt{(2meV)}} = \sqrt{\left(\frac{150}{V}\right)}$$ where λ is in Å and V in volts.

It is evident from the equation that only 150 volts are required to produce electrons of a wavelength of one Å suitable for diffraction work. The X-rays suitable for diffraction work require approximately 12000 volts for the same wavelength.

The diffraction of electrons takes place in a similar way as in case of X-rays and the diffraction pattern obtained can be interpreted exactly in the same way as X-rays diffraction pattern. Electron diffraction is particularly useful in investigating the structure of thin surface layers such as oxide layers on metals. These layers are not detected by X-rays diffraction because they penetrate deep into the solids and produce a pattern, which is the characteristic of the interior of the solid.

Neutron diffraction : The mass of a neutron is 2000 times large in comparison to an electron i.e., the wavelength associated with a neutron is about 1/2000 times that for an electron of the same velocity. The energy $\in$ of the suitable neutrons may be calculated by the following expression

$\lambda = \frac{h}{\sqrt{(2M_n \nu)}}$ By this equation, the energy $\in$ of neutrons is approximately 0.1 e.V. for λ = 1Å which is required for diffraction work. It is important to note that for X-rays of 1Å the energy required is 10^4 e.V. while for electron it is about 10^2 e.V. In case of neutrons, the diffraction patterns are formed in a similar manner a in case of X-rays.

Some special features of the neutrons diffraction are :

Neutrons are scattered chiefly by the nuclei of the atoms.

As the wavelength of the neutrons is much greater than the dimensions of the scattering nucleus ($\cong 10^{-13}$ cm), the atomic scattering factor is nearly independent of the scattering angle.

The scattering power does not vary in a regular manner with atomic number. Due to this fact the light elements such as hydrogen and carbon produce relatively strong scattering than X-rays scattering because the X-rays scattering is done by electrons. This feature of the neutron diffraction enables us to deduce the positions of hydrogen and carbon atoms in a number of organic crystals.

The scattering from neighboring elements in the periodic system may differ appreciably. Hence neutron diffraction allows to detect with relative ease, ordered phases of an alloy, such as FeCo, where as their detection by X-rays is difficult.

The neutrons possess magnetic moments and these moments interact with the magnetic moments of the scattering atoms of the solid. This gives an additional scattering

mechanism for neutrons, which often out-weights the nuclear scattering. Thus neutron diffraction methods are exceedingly valuable in structural studies of magnetic crystals.

In paramagnetic substances in which the atomic moments are randomly oriented, the magnetically scattered neutrons are incoherent in phase resulting in a diffuse background. This diffuse background of magnetic scattering is then superimposed on the lines produced by the nuclear scattering.

In ferromagnetic substances in which the magnetic moments within a domain are linked up in parallel, the diffuse background is absent.

In an antiferromagnetic solid the magnetic moments are aligned antiparallel and hence from the point of view of the neutron such atoms would appear to be different.

4.8 Key words :

Bragg's law - Bragg's diffraction - Wave front – Grating - Laue's method - Weissenberg's method - Rotating crystal method - Oscillating crystal method - White x-rays - Electron diffraction - Neutron diffraction - Elastic scattering - Structure factor.

4.9 Review questions:

1. Obtain an expression for the intensity of the diffraction of X rays by a simple three dimensional lattice of scattering points and hence derive Bragg's law.

An orthorhombic crystal has cell edges a = 4A, b= 2A, c= 6A . X rays of wave length 1.5A suffer a Bragg's reflection in the first order from the plane (2,1,3) . Find the glancing angle of the incident X ray.

2. Derive the Bragg's law. Use it to explain the principle of X-ray powder diffraction pattern. How can a simple cubic and fcc cubic crystals be distinguished by X ray diffraction.

3. What are the various techniques used of x ray diffraction. State in each case

(i) The nature of X-rays used (ii) The nature of sample (iii) The information obtainable or application.

4. Develop Laue conditions for diffraction of waves by a crystal. Out line the procedure to determine the lattice parameters by the powder method.

5. Derive the condition for a systematic absence of reflections in bcc and fcc crystals. What are the absent reflections in CsCl. Why is the x ray diffraction pattern of NaCl different from that of KCl and in what way.

6. Powder diffraction patterns of three monoatomic cubic crystals with bcc and fcc structures are recorded using Debye-Scherrer camera. The angles of diffraction in degrees of the first four diffraction lines for the two samples marked as and B are as under

A	B
42.9	28.8
49.2	41.0
72.0	50.8
87.3	59.6

Determine the structure type of each sample. What is the size of the cubic cell in each case?. Take the wavelength of incident X-rays as 1.5A.

7. Describe briefly the methods for crystal structure determination by X-ray diffraction. Explain the importance of geometrical structure factor taking the example of cubic crystals.

8. The first reflection using copper K-alpha radiation from a sample of copper powder (fcc) has value of 86.7 mm. Compute the camera radius. (57.3mm)

9. Calculate Bragg angle at which electrons accelerated from rest through a potential difference of 80V will be diffracted from (111) planes of an fcc crystal of lattice parameter 3.5A.

10. A beam of thermal neutrons emitted from the opening of a reactor diffracted by the (111) planes of nickel crystal at an angle of $28^{o}\ 30^{'}$. Calculate the effective temperature of the neutrons, Nickel has fcc structure and its lattice parameter is 3.52A.

11. (a) Discuss the theory relevant to the crystal structure determination by X-ray diffraction technique apply the theory for a cubic crystal.

(b) Compare the relative advantages between electron, X-ray and neutron diffraction techniques.

12. (a) Explain the Bragg's law employed in X-ray diffraction in crystalline materials

(b) Explain how the Bragg's law or coalition is satisfied for laws diffraction and

powder deffraction.

(c) Distinguish between electron and neutron diffraction methods.

4.10 Text and Reference books:

1. Elements of Solid State Physics by J.P.Srivastava (PHI)
2. Solid State Physics by M.A. Wahab (Narosa)
3. Elements of Solid State Physics by A.Omar (Pearson education)
4. Solid State Physics by S.O. Pillai (New Age)
5. Solid State Physics by C.Kittel (Asia Publishing house)
6. A Text Book of Solid State Physics by S.L.Kakani and C.Hemrajani (S.Chand)
7. Fundamentals of Solid State Physics by Saxena Gupta Saxena (Pragati Prakasan)

UNIT – II

LESSON: 5

Nature of Binding in crystals

Aim: To know about the various kinds of binding forces that result in different types of crystals.

Objectives of Lesson:

- To know different types of bonds involved between atoms in crystals and also knowing the origin of bonds.
- To know about the cohesive energy, the electrostatic and repulsive overlap energy.
- To know about various types of crystals classified on the basis of nature of bonding.
- To know the relation between atomic radii and lattice constants.

Structure of the Lesson:

5.1 Cohesion of atoms.

5.2 Primary bonds.

5.2.1 Covalent bond.

5.2.2 The metallic bond

5.2.3 The ionic bond.

5.2.4 Mixed bonding.

5.3 Secondary bond

5.3.1 The van der Waal's bond

5.3.2 The hydrogen bond

5.4 Cohesive energy

5.4.1 Ionic crystals

5.4.2 Covalent Crystals

5.4.3 Metallic Crystals

5.4.4 Molecular crystals

5.4.5 Hydrogen bonded Crystals

Introduction:

Crystalline materials can be classified in many ways. We have already seen one type of classification based on lattice structures. Crystals can also be classified according to their properties like mechanical, thermal, electrical, magnetic etc,. The most convenient basis for classification of crystals is the character of the inter-atomic binding forces in various types of crystalline materials. According to this scheme of classification of crystals, all solids fall into one of five general categories; (*i*) molecular, (*ii*) ionic, (*iii*) covalent, (*iv*) metallic and (*v*) hydrogen bonded crystals. The distinction in the said five categories is not a sharp one, because some materials may belong to more than one class.

From the very existence of solids we may draw two general conclusions (*i*) there must act attractive forces between the atoms or molecules in a solid which keep them together; and (*ii*) there must be repulsive forces acting between the atoms as well, since large external pressures are required to compress a solid to appreciable extent. In order to understand the importance of these two types of forces, let us consider the simplest system e.g., a single pair of atoms.

5.1 Cohesion of Atoms:-

Consider two atoms kept at a finite separation. The nature of forces between atoms or the binding energy depends primarily on the distribution of positive charge over the inner ion core and of negative charge over the outer space within the atoms. Their potential energy is zero when the atoms are infinite distance apart. When they are brought closure to a finite separation, there will be a potential energy for the system whose sign depends on the relative order of magnitude of the repulsive and attractive forces. The attractive part of the potential energy is conventionally treated as negative since the atoms themselves do the work of attraction. The positive sign for the repulsive potential energy comes from the concept that in binding the atoms closer, the work is done against the

repulsive force by an external energy. The total potential energy may be represented by a general potential of the form.

$$U(\mathrm{r}) = -\frac{A}{r^n} + \frac{B}{r^m} \quad \text{----------} \quad (5.1)$$

Where r is inter-atomic distance and A and B are constants. The net force is expressed as

$$\mathrm{F(r)} = -\frac{dU(r)}{d\mathbf{r}} = -\frac{nA}{\mathbf{r}^{n+1}} + \frac{mB}{\mathbf{r}^{m+1}} \quad \text{----------} \quad (5.2)$$

The first term represents the attractive force and the second repulsive. The variation of potential energy and force as a function of the inter-atomic separation is traced in fig. 5.1. The net force at the separation r_0, when the atoms are in stable equilibrium forming a diatomic molecule, is zero [Fig 5.1(b)]. Obviously the potential energy at this position is minimum with the negative sign [Fig 5.1(a)].

So at the interatomic seperation r_0 , F(r) in equation 5.2 is zero. From this it follows that

$$(r_0)^{m-n} = \frac{B}{A} \cdot \left(\frac{m}{n}\right) \quad \text{----------} \quad (5.3)$$

$$\text{At } r = r_0 \qquad \mathrm{U}(\mathrm{r}_0) = -\frac{A}{r_0^n} + \frac{B}{r_0^m} = -\frac{A}{r_0^n}\left(1 - \frac{n}{m}\right) \quad \text{----------} \quad (5.3)$$

The minimum energy conditions require that $\left.\frac{d^2U}{dr^2}\right|_{r=r_0} > 0$. This condition leads to m > n.

It means the repulsive forces be of shorter range than the attractive forces. Here, we must note that although the attractive and repulsive forces are equal in equilibrium, the attractive and repulsive energies are not equal, since n ≠ m. We also must note that although the energy cannot, in general be represented accurately by a relation of the type shown in equation 5.1 the above treatment provides some useful qualitative conclusions about the bonding of atoms in the solids.

From the above argument it is evident that the atoms are brought closure the repulsive force increases faster than the attractive force at short distance. When the two atoms stay in stable equilibrium at a certain separation they are said to have formed a chemical bond between them. The atoms spend a part of their energy in bond formation or we can say that a bond stores a part of the energy of the atoms. Therefore, the electron energy of atoms decreases when a large number of them collect together to form a solid. This

suggests that when a solid is heated to break it finally forms isolated atoms, the energy stored in bonds must be carried away by the individual isolated atoms. How strongly the atoms are bound together in a solid is represented by its cohesive energy (or binding energy). A proper discussion on cohesive energy will be taken up later, after dealing with the types of bonding in solids. When on solidification a pronounced lowering of electron energy takes place, the bonds formed are strong and called ***primary bonds***. On the other hand, if the lowering of electron energies is small weak bonds are formed and referred to as ***secondary bonds***.

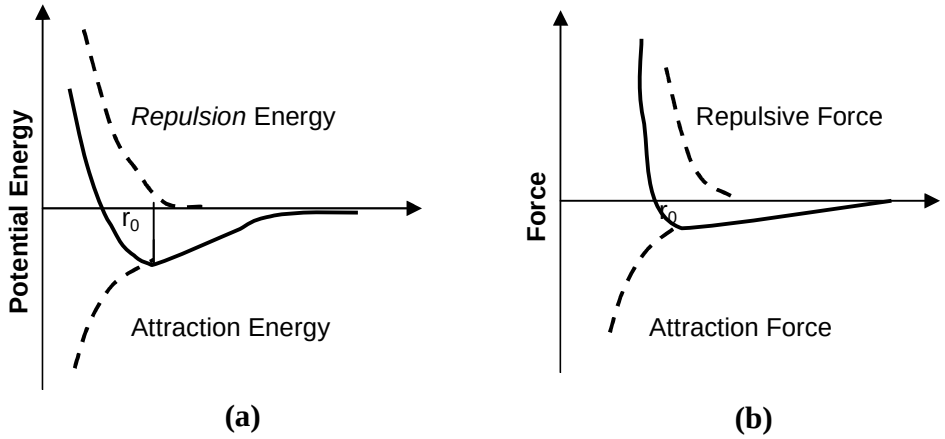


Fig.5.1 (a): Potential energy of a system of two atoms as a function of their inter atomic separation. (b): Force between two atoms as a function of their inter atomic separation. The r_0 denotes the separation at which the bond formation occurs.

The quantum mechanical interpretation of the probability function comes handy in understanding that some bonds are strong and directional while the others are weak and non-directional. It can be done qualitatively by acquiring information about the energies and the location of the bonding electrons with respect to positively charged ion cores. The probability that an electron confined in the small volume dV, may lie in the orbital ψ is given by $|\psi|^2 dV$. For the large value of the probability amplitude $|\psi|^2$ in a certain direction the bond formed is strong and concentrated along this direction and it behaves as directional. The small value of $|\psi|^2$ corresponds to the weaker non-directional bonds. Further more, if $|\psi|^2$ is spherically symmetric non-directional bonds result. The spherical symmetry means that the chance of finding a bonding electron in all directions is equal

and an atom can approach the other atom from any direction over the angle 2π for making the bond. Such a bonding gives rise to close packing exhibited by metals.

Now, we set out to outline the description of various forms of binding mechanisms in solids. Based on their comparative strength the bonds formed in various binding processes, they may be broadly be classified as **(a)** primary bonds **(b)**secondary bonds.

5.2 Primary Bonds:

Three limiting cases of primary bonds are identified as (a) covalent bonds (b) metallic bonds and (c) ionic bonds

5.2.1 The Covalent Bond:

A covalent bond involves the mutual sharing of a pair of electrons between a pair of atoms. The spins of the two electrons are oriented in opposite directions. In solid state, the most stable covalent bonds are formed between non-metallic atoms like those of N, O, C, F and Cl. Some of the other elements that are well known to form covalent bonded crystals include Si, Ge, As and Se. But the nature of bonds in these solids is only partly covalent. Strong covalent bonds are formed when each atom has at least one half filled orbital. In such a situation only there will be substantial lowering in the electrons energy when each of the bonding electrons occupies the orbitals the two atoms simultaneously. The lowering of electronic energy is proportional to the degree of overlap of the bonding orbitals. The more the overlap, the stronger the bond. Either the electrostatic repulsion or the Pauli exclusion principle controls the overlap. The molecular hydrogen offers itself as the simplest candidate for the explanation of the covalent bonding. Hydrogen has a single electron, which occupies the 1s orbital in the ground state. The orbital is half-filled and the Pauli principle allows it to accommodate one more electron with the opposite sign. Thus when two hydrogen atoms are brought closer, their electron charge distribution overlap and the covalent bond is formed by having the two electrons with opposite spins in the 1s orbital belonging to both the atoms. This configuration being lower in energy than the one in which the second electron goes into the 2s orbital with the probability of the same spin orientation belongs to the hydrogen molecule in the ground state.

The effect of repulsive interaction as a compulsion of Pauli principle is best demonstrated in rare gas solids whose atoms have completely filled orbitals with little

chance of overlap. This results in a large inter atomic distance (3.76 $\overset{o}{A}$ in A_r) compared to the bond length in a covalent-bonded molecule such as Cl_2 (2 $\overset{o}{A}$). There is a deficiency of one electron in the outermost shell of chlorine ($3p^5$) to saturate it. Because of this the outermost orbital of a chlorine atom has the tendency to overlap with that of the neighboring chlorine atom in search of the deficit electron. Since this forms the basis of covalent bond formation, the covalent bond is also known as ***saturable bond***.

The discussion on the covalent bonding remains in complete without a few remarks on elements in group IV of the periodic table (C, Si, Ge and grey Sn). These are extremely covalent elements and crystallize in the tetrahedrally coordinated diamond structure. Each atom participates in four covalent bonds with its four neighbours. These bonds are extremely directional and difficult to tilt. This provides the material with unusual strength and hardness. The wide use of diamond for making cutting tools is a well-acknowledged fact. Also the strong directional nature of the covalent bonds rules out close packing and increases emptiness in covalent bonded crystals. For example the fraction of volume of a diamond crystal actually occupied by atoms is a mere 0.34 which is 46 percent of the value for the close-packed fcc and hcp structures. In view of this fact the diamond lattice is called the empty lattice.

5.2.2 The Metallic Bond:

In the modern theory of metals the valence electrons being loosely bound to the respective parent atoms are treated as common to the whole assembly of atoms comprising the metal. The valence electrons generally one or two per atom, are considered free and, therefore, allowed to move freely over the whole volume of the metal. The metal is pictured as an assembly of positive metal ions embeded in a sea of free electron gas. The attraction between the metal ions and the electron gas gives rise to a strong cohesive force. The free electron gas serves as the glue. The concept of free electrons contributes significantly to the success of this model in explaining most of the properties of metals such as high electrical and thermal conductivities, high reflectivity and opacity. In reality the motion of valence electrons also called ***conduction electrons*** with reference to metals are affected, no matter how slightly, because of the presence of other particles including the fellow electrons. A complete theory does not take care of

this effect, which is of crucial advantage in explaining some of the dramatic properties (e.g. superconductivity) of solids in general.

Since the metallic bonds involve loosely bound electrons there can be a relatively large number of probable orientations for a metal bond. In other words, the bonds are weak and non-directional. This characterizes metals as having the tendency to crystallize in relatively closed packed structures with a large number of nearest neighbours. This fact is in agreement with the observed structures of metals–hcp, fcc and bcc. In all these structures the metallic bond is not saturated. For example in the bcc structure of lithium each bond has only $1/4^{th}$ of an electron since the only outermost electron ($3s^1$) is involved in eight bonds with atoms at the corners of the cube. Still the modern picture assumes that the metallic bond is more closely related to the electron–pair or covalent bond then to an ionic–type bond, which is discussed next.

5.2.3 The Ionic Bond:

The formation of an ionic bond is based on an electrostatic attraction between the positive and negative ions that are derived from the free atoms by the loss or gain of electrons. The ionic bond is responsible for the binding of the salts that comprise the combinations of the elements located on the right– and left–hand sides of the periodic table. Alkali halides are the typical representatives of such salts. The electronic configuration of the alkali and halogen atoms is nearly closed-shell configuration. The bond formation is facilitated by the ease with which the alkali atom looses one electron and the halogen atom accepts it to acquire closed-shell configuration. In another way we say that the origin of binding lies in the low ionization energy of the alkali atom and high electron affinity of the halogen atom. Take the example of NaCl in which the sodium atom has one electron more ($1s^2$, $2s^2$, $2p^6$, $3s^1$) than the neon ($1s^2$ $2s^2$ $2p^6$) and the chlorine atom is one electron short ($1s^2$, $2s^2$, $2p^6$ $3s^2$ $3p^5$) of the argon configuration ($1s^2$ $2s^2$ $2p^6$ $3s^2$ $3p^6$). This shows how Na and Cl atoms get bonded when brought close because of their easy ionization to Na^+ and Cl^- ions. Each ion has the tendency of being surrounded by as many ions of the opposite type as possible. The coordination number and the nearest neighbour distance are governed jointly by the geometrical factors and the repulsive interaction between the like ions. But the tendency of one type of ions to

surround an ion of the opposite type binds a continuous network of ions to form the crystal instead of forming small discrete molecules.

Ionic bonds are neither saturable nor directional. Even then they are strong enough as is confirmed by the hardness, high melting point and low coefficient of expansion of ionic crystals.

5.2.4 Mixed Bonding:

The discussion of primary bonds remains inconclusive without commenting on the purity of bonds in crystals. The mixture of bonding types has been discussed by Pauling in great detail. In particular it is a matter of general observation that the bonding in the most of the covalent crystals has a small component of ionic bonding. Ash croft and Mermin elucidate it for GaAs by picturizing the electron density of Ga and As ions in the crystal. Phillips developed a semi-empirical theory of fractional covalent or ionic character bonds in a dielectric crystal. His calculations show that bonds in Ge and Si are purely covalent. But the bonding in most of the binary crystals turns out to be of mixed nature LiF, NaCl and RbF are formed to be almost purely ionic as the fraction of ionic character in them turns out to be 0.92, 0.94 and 0.96 respectively.

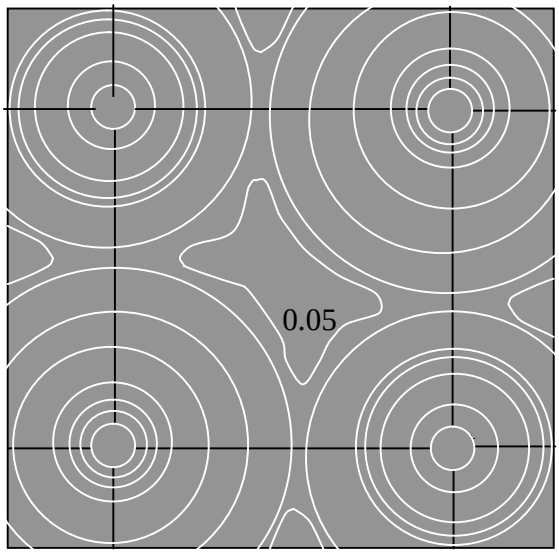


Fig:5.2: Electron density distribution in the basal plane of NaCl.
The number of contours expresses the relative electron density.

The ionic character of a solid may be determined by analyzing its X-ray scattering data. This is an effective tool as the scattering power depends on the number of electrons possessed by the constituent ions. For example in ionic bonded KCl each of the ions

$K^+$and Cl^- has 18 electrons. Hence the scattering power for K^+ and Cl^- is found to be equal. The electron density distribution in the basal plane of NaCl as derived by G. Schoknecht is shown in fig.5.2.

5.3 Secondary Bonds:

The limiting cases of secondary bonds are not easily separable. Van der Waal and hydrogen bonds will be discussed in this category. The electric dipole-dipole interaction forms the basis of these bonds.

5.3.1 The van der Waal's Bond:

One may be curious to know why neutral molecule or noble gases should undergo liquefaction and crystallization. Consider the noble gas atoms, which have closed-shell structure and are represented as spherical rigid charge distributions. Such electron distributions are reluctant to overlap when any two atoms are brought close to each other. But there must be some attractive forces between atoms bringing about the cohesion and finally the solidification. Of these forces one is Van der Waal force we are interested in. These are weak forces arising out of the attractive interactions between fluctuating electric dipoles. By chance for a fraction of a second there could be more number of electrons on one side of the nucleus than on the other. This destroys the spherical symmetry of the electron charge distribution and momentarily displaces the center of the negative charge (the electrons) from the positive charge (nucleus). Thus an atom becomes a tiny electric dipole, which is capable of inducing an electric dipole moment in the neighbouring atoms. The two dipoles attract each other through weakly resulting in Van der Waal binding. As the electron charge distribution keeps fluctuating the electric dipoles are called ***fluctuating dipoles***.

Suppose we have two atoms 1and 2 of a noble gas separated by a distance $\boldsymbol{r}$. As described above when atom 1 acquires an instantaneous electric dipole–moment $\mathbf{p_1}$, an instantaneous electric dipole-moment $\mathbf{p_2}$ may be induced in atoms 2 because of the polarization caused by the electric field **E** of the atomic dipole 1 at atom 2. It is now a trivial exercise in electrostatics to show that the interaction energy of two dipoles $\mathbf{p_1}$ and $\mathbf{p_2}$ ($= -\mathbf{p_2}.\mathbf{E}$) varies with separation as $1/\mathbf{r}^6$. These forces have been treated by London and Margenau on a quantum mechanical basis. An approximate expression for the interaction energy of two atoms or ions with filled shell electron configuration is

$\varepsilon(r) = -\frac{3}{2}\frac{\alpha_1\alpha_2}{r^6}\frac{I_1 I_2}{I_1 + I_2}$ where I_1 and I_2 refer to ionization energies of the particles involved and α_1, α_2 refer to polarizabilities. The nature of these forces is essentially a quantum effect, although the fact that they vary with the sixth power of the distance may easily be shown from classical considerations.

The interaction is expressed as $\mathbf{A/r^6}$ where **A** is a constant. It must be noticed that the Van der Waal bonds are neither saturable nor directional. Barring helium all the noble gases crystallize in the closed packed fcc structure. These weakly bound crystals are transparent insulators and characterized by low melting points.

5.3.2 The Hydrogen Bond:

A number of covalently bonded molecules behave as permanent electric dipoles. **HF** and **H_2O** are the two most talked about examples. The attraction between the positive end of one dipole and the negative end of the neighbouring dipole forms clusters or large aggregates of molecules on cooling when the crystallization may take place. The mechanism of bond formation involves the attraction of a hydrogen atom to two strongly electro-negative atoms. These bonds are called hydrogen bonds. These bonds have the directional property and are stronger than Van der Waal bonds. This is endorsed by the observation that super molecules like **H_2F_2, H_3F_3 . . . .** are formed even in the gaseous phase.

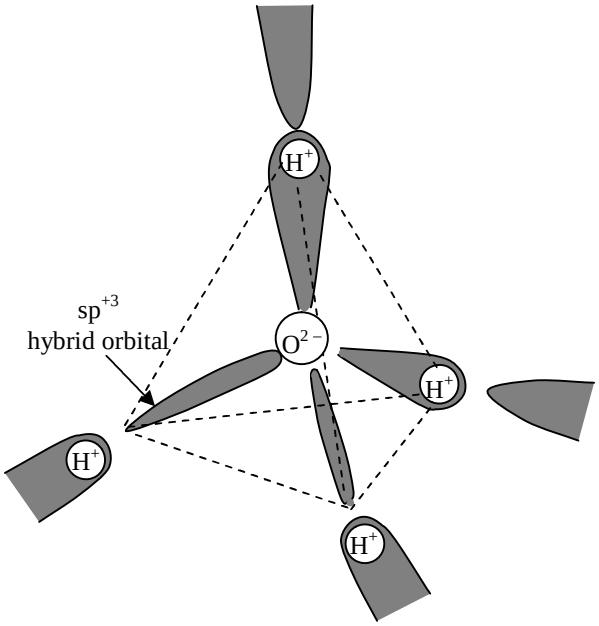


Fig 5.3: The tetra hedral covalent bonding in ice involving the sp^3 hybrid orbitals of H_2O.

In an effort to understand the mechanism of hydrogen bonding clearly we take up ice crystals, which exhibit some unusual behavior. The outer electron configuration of the oxygenation in H_2O is represented by four sp^3 hybrid orbitals. Fig 5.3 shows that two of these are involved in covalent bonds to the hydrogen atoms and the rest two are doubly occupied by oxygen electrons. Out of the six outer oxygen electrons, two are engaged in bonding and the other four find a place in space orbital. This bonding pattern gives a tetrahedral shape to bonded molecules in ice. Two vertices of a tetrahedral are occupied by hydrogen atoms, which are essentially the localized protons. The other two vertices exhibit relatively diffused negative charges.

The negligible small size of the proton is responsible for peculiar structure of ice crystals as it practically sits on the oxygen ions something impossible for any other positive ion to do. One of the many phases of ice crystals as shown in Pauling's book is pictured in fig5.4

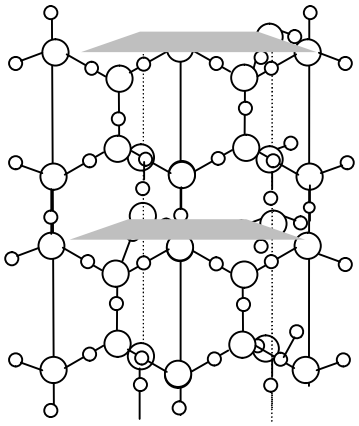


Fig5.4: The hydrogen bonding in ice. The large circles denote oxygen atoms and the small circles are protons.

Two oxygen atoms are bound by protons that is localized close to one oxygen atom along the line joining it to one of its neighbouring oxygen atoms. There are two protons close to each oxygen atom, giving a large number of a ways to attach a proton to either end of the bond. This is reflected in the irregular positions of protons which are very well accounted by the observed large residual entropy of ice crystals at low temperatures. Lastly, a word about the floating property of ice. The molecular clusters of water are smaller and less stable in liquid state than they are in ice because of which they are on the

average more closely packed in this state. This increases the density of water and ice floats when immersed in it.

5.4 The cohesive energy:

As stated earlier, the cohesive energy of a crystal is the measure of how strongly its constituent atoms or particles are bound. It may be defined as the energy required to break the crystal into its isolated neutral constituent parts. By constituent parts we precisely mean atoms or molecules. For example, it sounds more reasonable to define the cohesive energy to solid nitrogen as the energy required to disassemble it into isolated nitrogen molecules instead of atoms. The two definitions are, however, inter convertible. In the early development of the solids classification, the theory based it self heavily on the nature of cohesion. The role of spatial electronic arrangement was almost sidelined. But with the increasing need to study the non–equilibrium properties of the development of devices, the concern for the cohesive energy has ceased to be of any significant note.

We present a brief sketch of classical theory to discuss the cohesive energy of crystals at **0 K**. Calculations are made in the static lattice approximation which treats the atoms to be at rest at equilibrium sites. It neglects zero point motion, which relates to one of the basic tenets of quantum theory. The resulting error for most of the crystals turns out be one percent less. But there may be a question mark against this approximation while treating the lighter noble gases as they have the origin of their cohesion in the zero point motion. This is not to suggest the noble gases cannot be handled with a simplified theory like the one under consideration. Surprisingly, the theory which is reasonably simple in this case enjoys a high degree of success.

The over simplified theory discussed above is applied to ionic crystals with maximum ease as the dominating long–range attractive interaction between the oppositely charged ions renders other interactions of little concern. On the other hand, it is difficult to have a simple theory for calculating the cohesive energy of covalent and metallic crystals. The electronic configurations in crystals have far more distorted forms than what they are in isolated atoms or ions. This makes it imperative to calculate the energy treating the valence electrons in the field of the periodic crystal potential. It leads to the problem of band structure calculation complicating the whole procedure. The level of the present

book cannot accommodate the theory of covalent and metallic crystals. The description of simple methods of calculating the cohesive energy of ionic and noble gas crystals would be sufficient.

5.4.1 Ionic Crystals:-

The potential energy of an ionic crystal is considered to be composed of two components, one representing the electrostatic energy and the other belonging to the repulsive overlap, which has its origin in the Pauli exclusion principle. These crystals will be studied in detail in another lesson.

Properties of Ionic Crystals:

All ionic or electrovalent compounds are rigid and crystalline in nature. It has seen by X-ray diffraction that the constituents of the crystals are ions and not atoms. The ionic bond is fairly strong.

All ionic solids have high melting points and boiling points. This is because more energy is needed to make the ions mobile, there being a necessity to overcome the strong electrostatic force of attraction created due to the ionic bond in addition to the already existing interatomic force of gravitational attraction. This explains why ionic compound possesses high melting point and boiling point.

Pure and dry ionic compounds are insulators, because of the non availability of free electrons. However, in solution they conduct electricity because of the movement of the charged ions instead of free electrons. In a solution the ionic bond is weakened by the solvent molecules. Thus the ions become free to move about and thereby become conductors of electricity and the solutions are good electrolytes.

Ionic solids are easily soluble in polar solvent like water. This is because the molecules of the polar solvent interact strongly with the ions so as to reduce the attraction between the ions. Also, the polar solvents possess high dielectric constants, for example, water has a high dielectric constant of 81, i.e., it will reduce the electrostatic force of attraction between the ions to $\frac{1}{81}$ of the original value.

Ionic compounds are insoluble in non-polar solvents like benzene (C_6H_6), carbon tetrachloride (CCl_4), because their dielectric constants are very low.

Reaction between ionic compounds in solution state is always fast. This is because in a solution, ionic substances exist as ions and chemical reactions take place between the ions. The ionic crystals are transparent for all frequencies up to the value called the *fundamental absorption frequency*. At frequencies higher than this, they are opaque.

High hardness and low conductivity are typical properties of these solids. When subjected to stresses, ionic crystals tend to cleave (break) along certain planes of atoms rather than to deform in a ductile fashion as metals do.

5.4.2 Covalent Crystals:

When a covalent bond is formed we imagine that an electron from each atom is transferred to the region between the two atoms joined by the bond.

In an ionic bond it is a good approximation to think of the valence electrons as attached to definite atoms. The Pauli principle applied to ions with filled electronic shells ensures a low electron density in the region between the two ions where the charge shells make contact. Ions with filled shells do not generally form covalent bonds. In a covalent or homopolar bond the charge density between the two atoms may be rather high, and the valence electrons are to an appreciable extent shared between two atoms. The covalent bond is the normal electron-pair bond of chemistry, encountered particularly in organic chemistry. It is characterized by a high density of electrons between the ions and also by marked directional properties. The carbon bond is a good example of the directional properties of the covalent bond: carbon atoms often prefer to join onto each other or to other atoms by four bonds making tetrahedral angles with each other. That is, each carbon atom will be at the center of the tetrahedron formed by the nearest neighbor atoms. Diamond and methane, CH_4, are typical examples of the tetrahedral covalent bond. The diamond structure is loosely packed in a geometrical sense: the tetrahedral bond allows only four nearest neighbours, while a closest-packed structure would require twelve nearest neighbour atoms. The covalent bond is usually formed from two electrons, one from each atom participating in the bond. The spins of the two electrons in the bond are antiparallel. The carbon atom ($2s^2\ 2p^2$) tends, in a sense, to fill up the $2p^6$ electron shell by sharing electrons with four neighbors.

There is apparently a continuous range of crystals between the ionic and the covalent limits. It is often of importance to estimate the extent to which a given bond is ionic or covalent, but this may be difficult to do with any confidence. We think of NaF as an ionic crystal, and perhaps of InSb as largely covalent, but it is at present difficult to know what to say about the nature of the bonding of ZnS or PbS, for example. Pauling has formulated on a semi-empirical basis an electronegativity scale of some of the elements. Electronegativity is a chemical term meaning the power of an atom in a molecule to attract electrons to itself. The electronegativity is approximately proportional to the sum of the ionization energy and the electron affinity of the atom. A suggested empirical connection between the ionic character of a bond and the difference in electronegativity of the atoms being joined is shown in fig 5.5.

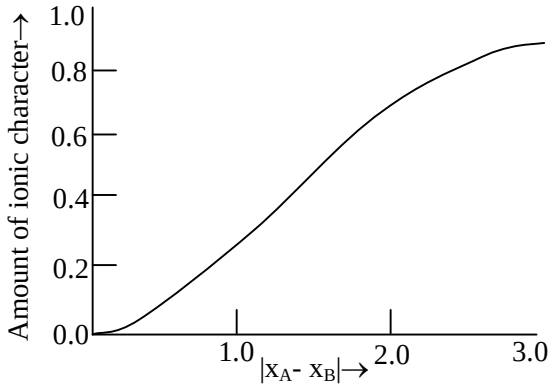


Fig 5.5: Curve relating the ionic character of a bond A–B to the difference in electro negativity $x_A - x_B$ of the atoms.

Atoms with nearly filled shells (Na, Cl) tend to be ionic, whereas atoms not close in the periodic table to the inert gases tend to be covalent (C, Ge, Si, Te).

5.4.3 Metal Crystals:

Metals are characterized by high electrical conductivity, and so a portion of the electrons in a metal must be free to move about. The electrons available to participate in the conductivity are called conduction electrons. In some metals such as the alkali metals the interaction of the ion cores with the conduction electrons is largely responsible for the binding energy. We may think of an alkali metal crystal as an array of positive ions embeded in a more-or-less uniform sea of negative charge. In some metals such as the transition metals it has been suggested that there may also be binding effects from

covalent-type bonds among the inner electron shells. Transition group elements have incomplete d-electron shells and are characterized by high binding energy.

The binding energy of an alkali metal crystal is seen from table 5.1 to be very considerably less than that of an alkali halide crystal, so the bond formed by a quasi-free conduction electron is not very strong. Part of the explanation is that the interatomic distances are relatively large in the alkali metals because the kinetic energy of the conduction electrons favors large interatomic distances, leading thus to weak binding. In the transition metals such as iron and tungsten the inner electronic shells make a substantial contribution to the binding. The binding energy of tungsten, for example, is 210 kcal/mole.

5.4.4 Molecular Crystals (or) Noble Gas Crystals:-

The discussion under Subsection 5.3.1 needs extension for obtaining a complete picture of interactions in noble gas crystals. The repulsive potential energy is conventionally expressed in the form of an empirical power law as B/r^{12}. The total potential energy of a pair of atoms at the separation r is generally represented by

$$U(r) = -\frac{A}{r^6} + \frac{B}{r^{12}} \quad \text{----------} \quad (5.5)$$

and this is known as the Lennard–Jones potential. The constants A and B are empirical parameters.

On the requirement that the repulsive force increases faster than the attractive force at short distances, the exponent in the repulsive term has to be more than 6. In addition, if the analytic simplicity is taken into account the choice falls on 12. Further, with this choice the theoretical estimates of several physical properties of noble gases, exclusive of He^3 and He^4, are found to be in excellent accord with the experiment. A satisfactory explanation for the disagreement in the case of helium is not possible at this stage to avoid digression. However, it should suffice to remark that helium is identified as a unique quantum matter in the solid state theory which has developed provisions for dealing with it especially. Finally, it must be emphasized that the van der Waals interaction is present in all the three states of every material. Being weak in nature it is not the main cause of cohesion in many crystals where other strong interactions are present and the crystal bindings are named after them.

Equation [5.5] is usually expressed as

$$U(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \quad \text{----------} \quad (5.6)$$

Where

$$\sigma = \left(\frac{B}{A}\right)^{1/6} \; and \; \varepsilon = \frac{A^2}{4B}$$

The Lennard-Jones potential in the form of relation (5.6) is shown in Fig 5.6. Suitable values of the parameters and σ have been obtained by fitting the theoretically calculated values of certain physical properties involving these parameters with those experimentally observed in the gaseous phase. In principle, these values cannot be used for solids where the interaction is not a sum of pair potentials on account of high densities. The values of parameters serve as a measure of the strength of attraction and the radius of the repulsive core.

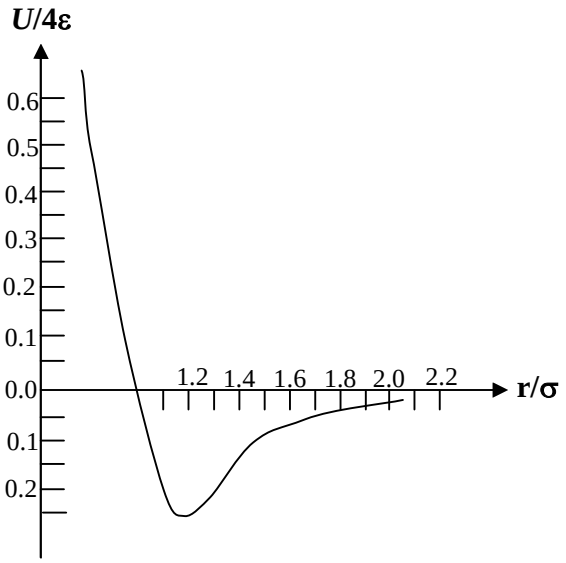


Fig 5.6 The Lennard–Jones potential as given by equation (5.6)

The hydrogen crystal are among the weakest bonded crystals with a cohesive energy of 0.01 eV per molecule which is one-tenth of that of methane's (0.1 eV per molecule). These weakly bound crystals are transparent insulators and characterized by low melting points.The cohesive energies of a number of crystals of different types are given in Table 5.1 for the purpose of a comparative study. For a thorough study the reader may consult the book by Ashcroft and Mermin which gives a more comprehensive view of the subject

and treasures large data on the related physical quantities of molecular and ionic crystals in particular.

It must now be abundantly clear that the calculations we have discussed in this section are closely linked with lattice constants (or lattice parameters) of crystals. Let us briefly comment on the involvement of lattice constants in a few other important physical properties. The pressure required to maintain a certain volume can be determined by calculating the rate of variation of cohesive energy with lattice constants. In effect, we succeed in reproducing the experimental value of the equilibrium lattice constant maintainable at zero pressure. Similarly, it is possible to study the change in volume caused by a change in pressure and calculate the compressibility. The significance of the lattice constant and compressibility must be emphasized in view of their utility in estimating the empirical parameters λ and ρ. As the lattice constants have a close relationship with the bond length and the atomic radii, we propose to give a short account of the same in the following section.

Table 5.1 Cohesive energy of crystals

Crystal	Binding type	Cohesive energy	Melting point (K)
LiF	Ionic	10.31 eV/ion	1143
NaCl	Ionic	7.96 eV/ion pair 3.28 eV/atom	1074
RbI	Ionic	6.17 eV/ion pair	915
Diamond	Covalent	7.37 eV/atom	>3773
Si	Covalent	4.63 eV/atom	1687
SiC	Covalent	12.04 eV/molecule	2873 (subl.)
Na	Metallic	1.11 eV/atom	371
Cu	Metallic	3.49 eV/atom	1358
Au	Metallic	3.81 eV/atom	1338
Ne	van der Waals	0.02 eV/atom	24.6
H_2	van der Waals	0.01 eV/molecule	14
CH_4	van der Waals	0.10 eV/molecule	90
Ice (H_2O)	Hydrogen	0.51 eV/molecule	273
HF	Hydrogen	0.30 eV/molecule	180.7

5.4.5 Hydrogen-Bonded Crystals:

As neutral hydrogen has only one electron, it should form a covalent bond with only one other atom. It is known, however, that under certain conditions an atom of hydrogen is attracted by rather strong forces to two atoms, thus forming what is called a hydrogen

bond between them, with a bond energy of about 5 kcal/mole. It is believed that the hydrogen bond is largely ionic in character, being formed only between the most electronegative atoms. The hydrogen atom loses its electron to one of the other atoms in the molecule; the proton forms the hydrogen bond. The small size of the proton permits only two nearest neighbor atoms because they are so close that more than two of them would get in each other's way; thus the hydrogen bond connects only two atoms.

The hydrogen bond is an important interaction between H_2O molecules and is responsible, together with the electrostatic attraction of the electric dipole moments, for the striking physical properties of water and ice. The hydrogen bond restrains protein molecules to their normal geometrical arrangements. It is also responsible for the polymerization of hydrogen fluoride and formic acid, for example. It is important in certain ferroelectric crystals, such as potassium dihydrogen phosphate.

5.5 Summary:

1. Bonds are broadly classified on the basis of strength as (i) Primary bonds (ii) Secondary bonds. These limiting cases of primary bonds are identified as (a) Covalent (b) Metallic and (c) ionic. (iii) Some primary bonds are mixed type. (iv) Van der Waals and hydrogen bonds fall in the category of secondary bonds.
2. Van der Waals interaction (induced dipole – dipole interaction) varies with interatomic separation as $1/ r^6$. It is a quantum effect. The interaction potential vanishes when the Planck constant h = 0. Example, inert gas solids.
3. The cause of repulsive interaction between atoms lies generally in the electrostatic repulsions of overlapping charge distributions and the Pauli exclusion principle that forces overlapping electrons of parallel spin occupy higher energy states.
4. The potential energy per ion in an ionic crystal is given by

$$U_{ion} = -\frac{\alpha q^2}{4\pi\varepsilon_0 r_0}\left(1-\frac{\rho}{r_0}\right)$$

where

α is the Madelung constant

q is the charge on an ion

ρ is an empirical parameter

r_0 is the separation in a pair of ions of opposite type at equilibrium.

5. The overlap of charge distributions of antiparallel electron spin gives rise to covalent bonding.

5.6 Key words:

Binding forces – Cohesion of atoms – Primary bonds – Secondary bonds – The Covalent bond – The metallic bond – The ionic bond – Mixed Bonding – The Van der Walls bond – The Hydrogen bond – The Cohesive energy – Lattice Constant.

5.7 Review questions:

1. Classify the different types of crystal bindings and their origin of bond formation?
2. What do you understand by cohesive energy and the repulsive overlap energy?
3. What is the inference of Noble gas crystals Explain the relation between atomic radic and lattice constants?
4. Explain different types of bonding in solids and their characteristics what is the nature of bond between molecules in ice. Explain the properties of diamond and graphite on the basis of their structure.

5.8 Text and Reference books:

1. Elements of Solid State Physics by J.P.Srivastava (PHI)
2. Solid State Physics by M.A. Wahab (Narosa)
3. Elements of Solid State Physics by A.Omar (Pearson education)
4. Solid State Physics by S.O. Pillai (New Age)
5. Solid State Physics by C.Kittel (Asia Publishing house)
6. A Text Book of Solid State Physics by S.L.Kakani and C.Hemrajani (S.Chand)
7. Fundamentals of Solid State Physics by Saxena Gupta Saxena (Pragati Prakasan)

UNIT – II

LESSON 6

LATTICE ENERGY OF IONIC CRYSTALS

Aim: To study the nature of forces acting in ionic crystals and their stability

Objectives:

- To study the results of Born's theory and the refinements suggested
- To study the relation between ionic radii of atoms involved and stability of various crystal structures.

Structure:

Introduction:

In lesson 5 we have learnt about the nature of forces that exist between atoms in crystals to some extent. Because of the importance of ionic crystals and in particular of alkali halide crystals we now study them in detail .

The potential energy of an ionic crystal is considered to be composed of two components, one representing the electrostatic energy and the other belonging to the repulsive overlap, which has its origin in the Pauli exclusion principle.

6.1 Born's theory:

Borns's theory of the lattice energy is based on the assumption that the crystals under consideration are built up of positive and negative ions. If we assume that the charge distribution in these ions is spherically symmetric, the force between two such ions depends only on their distance apart and is independent of direction. As an example, consider a lattice of NaCl structure, represented in Fig 6.1. We shall denote the shortest interionic distance by r and consider this quantity a variable for the moment. A given sodium ion is surrounded by 6 Cl^- ions at a distance r, 12 Na^+ ions at a distance $r\sqrt{2}$, 8 Cl^- ions at a distance $r\sqrt{3}$, etc. The Coulomb energy of this ion in the field of all other ions is therefore.

$$\in_e = -\frac{e^2}{r}\left(\frac{6}{\sqrt{1}} - \frac{12}{\sqrt{2}} + \frac{8}{\sqrt{3}} - \frac{6}{\sqrt{4}} + \frac{24}{\sqrt{5}} - \cdots\right) \quad \text{----------} \quad (6.1)$$

where e is the charge per ion. Note that because Coulomb forces decrease relatively slowly with distance, it is not sufficient to consider only a few shells of ions around the central ion.

Evidently, the coefficient of e^2/r is pure number, determined only by the crystal structure. Series of this type have been calculated by Madelung, Ewald, and Evjen. For the NaCl structure the result is

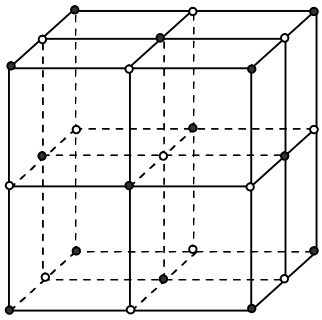


Fig 6.1 The sodium chloride structure

$\in_e = -Ae^2 / r$ with $A = 1.747558.....$ ---------- (6.2)

The constant *A* is called the Madelung constant. For other crystal structures composed of positive and negative ions of the same valancy, the Madelung constants are

Cesium chloride $A = 1.762670$

Zincbelende (ZnS) $A = 1.6381$

Wurtzite (ZnS) $A = 1.641$

Note that e in (6.2) represents in general the electronic charge times the valence of the ions under consideration. The minus sign in (6.2) indicates that the average influence of all other ions on the one under consideration is of an attractive nature. To prevent the lattice from collapsing, there must also be repulsive forces between the ions. These repulsive forces become noticeable when the electron shells of neighboring ions begin to overlap, and they increase strongly in this region with decreasing values of *r*. These forces, as other overlap forces, can best be discussed on the basis of wave mechanics, because they are of a non-classical nature. Born in his early work made the simple assumption that the repulsive energy between two ions as function of their separation could be expressed by a power law of the type B'/r^n, where B' and n are as yet undetermined constants characteristic of the ions in the solid under consideration. Focusing our attention again on one particular ion, we may thus write for the repulsive energy of this ion due to the presence of all other ions.

$\in_{rep} = B / r^n$ ---------- (6.3)

where *B* is related to B' by a numerical factor. In view of the fact, that, repulsive forces depend so strongly on the distance between the particles, the repulsive energy (6.3) is mainly determined by the nearest neighbors of the central ion. The total energy of one ion due to the presence of all others is then obtained by adding (6.2) and (6.3).

$\in = -Ae^2 / r + B / r^n$ ---------- (6.4)

Assuming that the two types of forces just discussed are the only ones we have to take into account and neglecting surface effects, we thus find for the total binding energy of a crystal containing *N* positive and *N* negative ions.

$$E(r) = N\left(-A\frac{e^2}{r} + \frac{B}{r^n}\right) = N \in (r) \quad \text{----------} \quad (6.5)$$

We multiply by N rather than by $2N$ because otherwise the energy between each pair of ions in the crystal would have been counted twice. The two contributions to $E(r)$ are represented schematically in Fig (6.2). If we consider the crystal at absolute zero, the equilibrium conditions require E to be a minimum, which will be the case for the equilibrium value $r = a_0$, where a_0 represents the smallest interionic distance in the crystal at $T = 0$. For this minimum

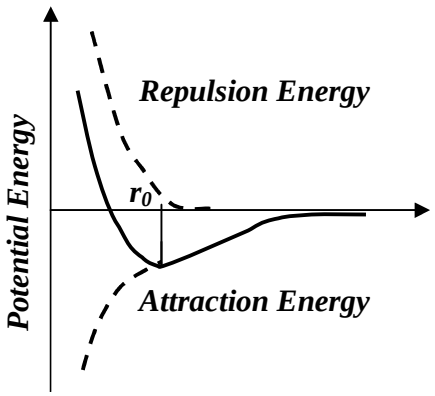


Fig 6.2 Schematic representation of the energy of attraction and repulsion as function of the lattice parameter. The resultant exhibits a minimum for a lattice constant a_0, corresponding to equilibrium.

$$\left(\frac{dE}{dr}\right)_{r=a_0} = 0 \quad \text{----------} \quad (6.6)$$

From the last two expressions one thus obtains the following relation between the two unknown parameters B and n:

$$B = \left(\frac{Ae^2}{n}\right) a_0^{n-1} \quad \text{----------} \quad (6.7)$$

Substitution into (6.5) yields for the lattice energy E_L,

$$E_L = E(a_0) = -NA\frac{e^2}{a_0}\left(1-\frac{1}{n}\right) = N \in_L \quad \text{----------} \quad (6.8)$$

where $\in_L = \in(a_0)$. The interionic distance can be obtained from X-ray diffraction data; the charge per ion is also known, and thus the lattice energy can be calculated if the repulsive exponent n is known. The information regarding n may be obtained is discussed in the next two sections.

6.2 Calculation of the repulsive exponent from compressibility data:

Born obtained the unknown repulsive exponent n from measurements of the compressibility of the crystals as follows: The compressibility K_0 at absolute zero is given by

$$\frac{1}{K_0 V_0} = \left(\frac{d^2E}{dV^2}\right)_{V=V_0} \quad \text{----------} \quad (6.9)$$

where V_0 is the volume of the crystal corresponding to an interionic distance a_0; V corresponds to the variable r. The relation between volume and interionic distance must of course be of the form

$V = cNr^3$ ---------- (6.10)

Where c is a constant determined only by the type of lattice. For NaCl, for example, c = 2. Hence

$$\frac{dE}{dV} = \frac{1}{3cNr^2}\cdot\frac{dE}{dr} \quad \text{and} \quad \frac{d^2E}{dV^2} = \frac{1}{9c^2N^2r^2}\cdot\frac{d}{dr}\left(\frac{1}{r^2}\cdot\frac{dE}{dr}\right) \quad \text{----------} \quad (6.11)$$

From (6.5) we thus obtain

$$\frac{1}{K_0 cNa_0^3} = \left(\frac{d^2E}{dV^2}\right)_{a_0} = \frac{1}{9c^2Na_0^2}\left[\frac{-4Ae^2}{a_0^5} + \frac{n(n+3)B}{a_0^{n+4}}\right] \quad \text{----------} \quad (6.12)$$

Substituting B from (6.7), we find

$n = 1 + 9ca_0^4 / K_0 e^2 A$ ---------- (6.13)

from which the parameter n can be calculated if K_0 is known. Some experimental values for alkali halides according to Slater, and obtained by extrapolation of compressibility measurements to T = 0, are given below

LiF	n = 5.9,	NaCl	n = 9.1
LiCl	n = 8.0,	NaBr	n = 9.5
LiBr	n = 8.7.		

We note that there is a marked variation from one crystal to another. However, even an appreciable error in n leads to a relatively small error in the lattice energy, which is proportional to $(1 - 1/n)$. If we change n by unity, E_L changes by only 1 or 2 per cent. According to (6.8) and in view of the relatively large values of n, most of the lattice energy is due to the Coulomb interaction, and the repulsion contributes only a relatively small fraction. On the other hand, the repulsive and attractive forces acting on any one ion just balance for $r = a_0$ and thus are equal in magnitude.

6.3 The repulsive exponent as function of electron configuration:

It will be obvious that the repulsive forces acting between two ions will depend on the distribution of the electronic charges in the ions and especially on the number of electrons in the outer shells. For example, we would expect *n* to be larger for NaCl than for LiCl, because the Na^+ ion has eight outer electrons and the Li^+ ion has only two. From an approximate treatment of the interaction between closed-shell electronic configurations, Pauling arrived at the following values of n as a function of the occupation of electronic shells.

Table 6.1 Repulsive Exponent as Function of Electron Configuration

Ion type	Electron configuration					
	K	L	M	N	O	n
He	2	...…..	……..	……..	…….	5
Ne	2	8	……..	……..	…….	7
Ar (Cr)	2	8	8(18)	……..	…….	9
Kr (Ag)	2	8	18	8(18)	…….	10
Xe (Au)	2	8	18	18	8(18)	12

This table should be used by taking the average value of *n* for the two ion types occurring in the crystal. For NaCl, for example, one takes the average of 7 and 9; for NaF the average of 7 and 7, etc. Note that this table is in qualitative agreement with the experimental values of Slater referred to above.

6.4 Calculated and experimental lattice energies:

The lattice energy $\in_L$ may now be calculated from (6.8) by substituting the proper values for the charge of the ions, the interatomic distance and the Born exponent *n*. Values for $\in_L$ so obtained are given in Table 6.2 for alkali halides and the alkaline earths oxides. The charge per ion in the latter group is assumed to be 2*e*; it is not quite certain that these oxides can be considered ionic compounds. It may be remarked that CsCl, CsBr and CsI crystallize in the cesium chloride structure (see fig 6.3), whereas all other compounds in the table have the NaCl structure. The expansion of the lattice, entering

through the interionic distance a_0, can usually be neglected; the coefficient of expansion of ionic crystals at room temperature is of the order of 10^{-4} per degree.

Table 6.2 Lattice Energies for Alkali Halides and Alkaline Earth Oxides.

Compound	a_0 in Angstroms	n	$\in_L$ in ev calc.	$\in_L$ in ev exp.
LiF	2.07	6.0	10.5	……
NaF	2.31	7.0	9.3	……
KF	2.66	8.0	8.3	……
RbF	2.82	8.5	7.9	……
CsF	3.00	9.5	7.5	……
LiCl	2.57	7.0	8.4	8.6
NaCl	2.81	8.0	8.0	7.9
KCl	3.14	9.0	7.1	7.1
RbCl	3.27	9.5	6.9	7.0
CsCl	3.56	10.5	6.5	6.7
LiBr	2.74	7.5	7.9	8.2
NaBr	2.97	8.5	7.5	7.5
KBr	3.29	9.5	6.8	6.8
RbBr	3.42	10.0	6.6	6.6
CsBr	3.71	11.0	6.2	6.4
LiI	3.03	8.5	7.4	7.8
NaI	3.23	9.5	7.0	7.2
KI	3.53	10.5	6.5	6.6
RbI	3.66	11.0	6.2	6.5
CsI	3.95	12.0	5.9	6.3
Mgo	2.10	7.0	41.0	……
CaO	2.40	8.0	36.5	……
SrO	2.57	8.5	34.5	……
BaO	2.75	9.5	32.5	……

6.5 Born Haber Cycle:

An experimental check on the calculated values of the lattice energies may be obtained from what is known as a Born-Haber cycle. Consider, for example, 1 gram atom of solid sodium reacting with $\frac{1}{2}$ gram molecule of Cl_2 gas. As a result of the reaction, solid NaCl is formed and a certain amount of heat Q (the “heat of formation”) is

given off. The change in energy due to such a reaction may be calculated by considering the following steps

$$\mathrm{Na}_{\mathrm{solid}} + S_{\mathrm{Na}} \rightarrow \mathrm{Na}_{\mathrm{vapor}}$$

$$\mathrm{Na}_{\mathrm{vapor}} + I_{\mathrm{Na}} \rightarrow \mathrm{Na}^{+} + \mathrm{electron}$$

$$\tfrac{1}{2}\mathrm{Cl}_2 + \tfrac{1}{2} D_{\mathrm{Cl}_2} \rightarrow \mathrm{Cl}$$

$$\mathrm{Cl} + \mathrm{electron} \rightarrow \mathrm{Cl} + E_{\mathrm{Cl}}$$

$$(\mathrm{Na}^{+}\mathrm{Cl}^{-})_{\mathrm{gas}} \rightarrow \mathrm{NaCl}_{\mathrm{solid}} + \in_L$$

$$\mathrm{Na}_{\mathrm{solid}} + \tfrac{1}{2}\mathrm{Cl}_2 + S_{\mathrm{Na}} + I_{\mathrm{Na}} + \tfrac{1}{2} D_{\mathrm{Cl}_2} \rightarrow \mathrm{NaCl}_{\mathrm{solid}} + E_{\mathrm{Cl}} + \in_L$$

The quantities introduced all refer to the formation of one ion pair of solid NaCl. Here S_{Na} represents the sublimation energy of sodium per atom. Sublimation energies in general can be determined experimentally by direct caloric measurements or from measurements of the vapor pressure as function of temperature.

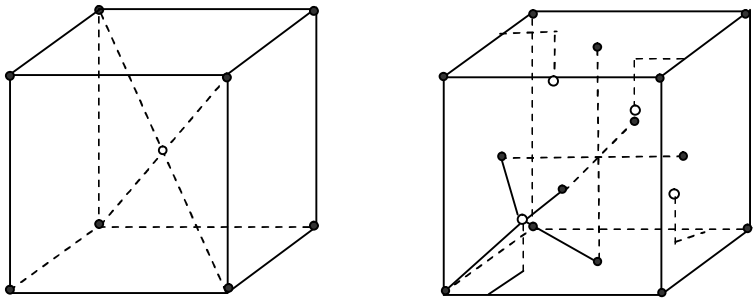


Fig 6.3 The CsCl and the ZnS structures. The open circles in the ZnS structure are located at points obtained by displacements of ¼ along three cube edges of the corresponding corner point. For one of the open circles we have indicated how it is surrounded by four black dots occupying the corner points of a regular tetrahedron, with the open circle at the center.

The ionization energy I_{Na} represents the energy required to take away the outer electron of the sodium atom, and can be obtained experimentally either from optical measurements or by bombardment of atoms with electrons and measuring the minimum energy of the latter required to produce ions. The dissociation energy D_{Cl_2} required to separate the two Cl atoms in a Cl_2 molecule can be obtained by determining the dissociation constant as function of temperature. The electron affinity E_{Cl} is the energy gained by combining an electron and a Cl-atom. Electron affinities can be determined by measuring the ionization energy of the negative ions, or by measuring the density of halide ions in alkali halide vapor. Now, we also know that

$$\mathrm{Na}_{\mathrm{solid}} + \tfrac{1}{2}\mathrm{Cl}_2 \rightarrow \mathrm{NaCl}_{\mathrm{solid}} + q$$

where q refers again to the heat of formation per "molecule" NaCl formed. Subtracting this equation from the one obtained above, we find for the lattice energy per ion pair,

$$\in_{L\,\mathrm{exp.}} = S_{\mathrm{Na}} + I_{\mathrm{Na}} + \tfrac{1}{2}D_{\mathrm{Cl}_2} - E_{\mathrm{Cl}} + q \quad \text{---------- (6.14)}$$

For NaCl, all quantities on the right-hand side are known from experiments and thus we are able to give an experimental value for $\in_L$, which may be compared with the one calculated with the Born theory. For NaCl we find, for example, from (6.14)

$$\in_{L\,\mathrm{exp.}} = 1.1 + 5.1 + 1.2 - 3.8 + 4.3 = 7.9\ \mathrm{eV}$$

whereas Born's theory yields 8.0 eV. The experimental values obtained in this way are listed in Table 6.2, and we see that theory and experiment agree within a few per cent, indicating that the relatively simple approach is essentially correct.

For the fluorides and oxides, the electron affinities are not known from experiment, and they are usually calculated by replacing $\in_{L\mathrm{exp.}}$ by $\in_{L\mathrm{exp.}}$ in (6.14). We note that for oxygen the electron affinity is negative, i.e., it requires energy to add 2 electrons to the atom. This is not surprising, because after the first electron has been added, we have a negative O^- ion and we would expect addition of a second electron to require appreciable energy. An experimental determination of the affinity of a neutral oxygen atom for the first electron added gives 2.2eV according to Lozier. Now the total electron affinity for the addition of 2 electrons is –7.3 eV when calculated from the lattice energy of oxides in the manner indicated above.

Table 6.3 Electron Affinities and Dissociation Energies

Atom	Electron affinity	Dissociation energy
F	4.25 eV	2.75 eV
Cl	2.50	2.50
Br	3.8	2.01
I	3.45	1.58
O	- 7.3	1.52
S	- 3.5	2.75
Se	- 4.2	2.50

Thus addition of the second electron requires about 9.5 eV. The usually accepted values of the electron affinities are given in Table 6.3 together with the dissociation energies of the diatomic molecules (in electron volts).

6.6 Stability of structure and ionic radii:

Ionic compounds of the composition A^+B^- occur in the sodium chloride structure, the cesium chloride structure, and the Zincblende structure (ZnS). The latter two are represented in Fig. 6.3. In the CsCl structure each ion is surrounded by 8 nearest neighbors of opposite sign; in NaCl by 6, and in Zincblende by 4. one may thus ask why a certain compound crystallizes in a particular structure.

The answer must obviously be sought in the fact that the energy should be a minimum, and the problem is thus reduced to explaining why for a given compound its natural structure has a lower energy than any other structure. We shall see that some insight into this problem may be obtained from considerations of the size of the ions.

For metals one defines the atomic radius as half the distance between nearest neighbors, although it is recognized that the meaning of the size of an atom is necessarily vague. For ionic crystals one could try a similar approach, but one is immediately faced with the difficulty that these compounds consist of at least 2 types of ions, so that the lattice constant provides information only about the sum of two radii. A little consideration of the interionic distances as given in the preceding section shows that to a fair approximation ionic radii are additive quantities. For example, if one calculates the difference ($r_{K^+} - r_{Na^+}$) from Table 6.2 for the halides of these metals, one finds from the fluorides,

$$r_{K^+} - r_{Na^+} = a_{KF} - a_{XaF} = 0.35 \text{ Angstrom}$$

and from the chlorides, bromides, and iodides in the same manner 0.33Å, 0.32Å, and 0.30Å, respectively. We see that the difference is roughly constant and that it has meaning to associate a rather definite radius with each ion. It is also obvious that a table of ionic radii can be obtained only if the radius of one ion is known.

Atomic and ionic radii were compiled by Pearson. An examination of this data shows that on ion formation the size of a positive ion is smaller and that of a negative ion larger than the size of the neutral atom. The best example illustrating this fact are the radii of

C($0.72\,\mathring{A}$), C^{4+}($0.15\,\mathring{A}$) and C^{4-}($2.60\,\mathring{A}$). This can be understood by taking simple examples of Na and Cl. The Na atom, whose radius is 1.86 $\mathring{A}$, loses one electron in becoming the Na^+ ion of radius of $0.90\,\mathring{A}$. On losing this outermost electron ($3s^1$) which is loosely bound, the atom assumes a more tightly bound configuration of Na^+ ion with the reduced size. This is one of the best examples to state, as the Na^+ ion turns out to be smaller than even the iso-electronic Ne atom of the preceding period. Similarly, the bigger size of the Cl^- ion makes sense on the ground that the Cl atom is readily willing to turn into a Cl^- ion whose electronic configuration is identical with the rigid spherical charge distribution of the nearest (the next in the same period) noble gas atom (Ar).

The extent of charge clouds of a constituent ion in any crystal is generally referred to as the ***crystal ionic radius***. It needs to be distinguished from the free ion radius. According to the quantum mechanical interpretation, the free ion radius may be defined as the radius at which the probability amplitude $|\psi|^2$ for the outermost electrons to be in the orbital represented by the wave function ψ is maximum. The wavefunction ψ is then a solution of the Schrodinger wave equation whose Hamiltonian includes the ordinary Coulomb potential. The boundary condition requires that the magnitude and the derivative of the wavefunction ψ vanish at infinitely large distances. In contrast, for an ion in the crystal the magnitude and derivative of the wavefunction must be zero at the boundary of the ion on the demand of the boundary condition. The potential in this case is not Coulomb but of the type shown in fig 6.2.

The above discussion indicates that the ionic radii may be calculated by solving one-electron Schrödinger wave equation. This is usually done by the Hartree-Fock self-consistent field method, treated in great detail by Slater. Some semi-empirical methods of calculating the ionic radii in crystals are also in practice. For example, the distance between two atoms shown as r_0 (the nearest neighbour distance in a crystal) is approximately equal to the sum of the radii of the two neighbouring atoms. This property is known as the additive rule. The diffraction method is the standard way at present to measure lattice constants. From the diffraction pattern it is possible to determine the radius ratio of any two type of atoms from which the radii of other atoms comprising the crystal can be determined. Goldschmidt in 1927 has tabulated ionic radii based on a

radius of the F^- ion of 1.33Å, a value which he decided upon on the basis of work by Wasastjerna on the relation between polarizability and ion size. Pauling, in the same year, independently published ionic radii based on theoretical calculations of the radii of some ions. The two sets are not equal, which is not surprising because of the inaccuracies involved. One commonly refers to the Goldschmidt and the Pauling radius of a given ion. In Table 6.4 the Goldschmidt radii (G) do not refer to the original set but include many recent X-ray diffraction data, especially those of Zachariasen. Contrary to the tables by Goldschmidt, the radius for O^{2-} is 1.45Å rather than 1.35Å. The radii according to Pauling are also given in Table 6.4.

Table 6.4 Goldschmidt (G) and Pauling (P) Ionic Radii in A

Ion	G	P	Ion	G	P	Ion	G	P
H^-	1.54	2.08	Be^{2+}	0.30	0.31	B^{3+}	0.2	0.20
F^-	1.33	1.36	Mg^{2+}	0.65	0.65	Al^{3+}	0.45	0.50
Cl	1.81	1.81	Ca^{2+}	0.94	0.99	Sc^{3+}	0.68	1.81
Br^-	1.96	1.95	Sr^{2+}	1.10	1.13	Y^{3+}	0.90	0.93
I^-	2.19	2.16	Ba^{2+}	1.29	1.35	La^{3+}	1.04	1.15
			Zn^{2+}	0.69	0.74	Ga^{3+}	0.60	0.62
O^{2-}	1.45	1.40	Cd^{2+}	0.92	0.97	In^{3+}	1.81	1.81
S^{2-}	1.90	1.84	Hg^{2+}	0.93	1.10	Tl^{3+}	1.91	0.95
Se^{2-}	2.02	1.98	Pb^{2+}	1.17	1.21			
Te^{2-}	2.22	2.21				Fe^{3+}	0.53	……
			Mn^{2+}	0.80	0.80	Cr^{3+}	0.55	……
Li^+	0.68	0.60	Fe^{2+}	0.76	0.75			
Na^+	0.98	0.95	Co^{2+}	0.70	0.72	C^{4+}	0.15	0.15
K^+	1.33	1.33	Ni^{2+}	0.68	0.69	Si^{4+}	0.38	0.41
Rb^+	1.48	1.48	Cu^{2+}	0.92	……	Ti^{4+}	0.60	0.68
Cs^+	1.67	1.69				Zr^{4+}	0.77	0.80
Cu^+	0.95	0.96				Ce^{4+}	0.87	1.01
Ag^+	1.13	1.26				Ge^{4+}	0.54	0.53
Au^+	……..	1.37				Sn^{4+}	0.71	0.71
Tl^+	1.51	1.44				Pb^{4+}	0.81	0.84

Using the additive rule of atomic radii it is possible to predict the bond lengths or the inter atomic separations in a crystal. This is successfully done for crystalline phases even before they are crystallized. But we must be aware that the charge distribution is not rigid and spherically symmetric in every atom. This may introduce an appreciable error in

the estimate. Therefore, it is a *must* to know accurately the average of charge clouds of the constituent atoms in a crystal for making any estimates.

At first instance the knowledge of atomic radii can be used to determine the coordinates of an atom and suggest the probable arrangement of atoms in the crystal. Though this point has more relevance, it is worth digressing to emphasize the role of atomic radii being discussed here. Beiser has excellently explained the threefold coordinates for a binary ionic crystal. It is shown that the coordination is stable only when the positive ion A^+ touches all the surrounding negative ions B^-. The mutual contact among the three B^- ions is not the necessary requirement for the coordination to crystallize. A simple geometrical calculation shows that the minimum ionic radius ratio r_A/r_B for this coordination to materialize is 0.155. We calculate below this ratio for the fourfold coordination in a similar ionic crystal.

In the fourfold coordination, the A^+ ion positioned at the center of a cube is surrounded by four B^- ions located at the alternate corners of the cube (Fig 6.4) whose edge is 'a'. For stable coordination, all the four B^- ions should touch the A^+ ion. Two B^- ions on every face are also in contact with each other, the distance between their centers being equal to the length of the face diagonal $\left(\sqrt{2}a\right)$. Let θ be the angle between the line joining a B^-ion to an A^+ ion and the line joining the two B^-ions. From fig 6.4, we have,

$$\cos\theta = \frac{\sqrt{2}a/2}{\sqrt{3}a/2} \qquad \text{(length of the body diagonal = } \sqrt{3}a\text{)}$$

$$= \sqrt{\frac{2}{3}}$$

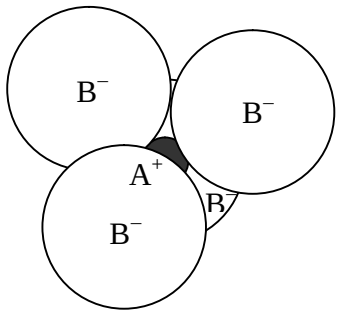


Fig 6.4: The fourfold coordination. The smaller ion A^-at the center of the cube is surrounding by four B^-ions located at the alternate corners of the cube. The minimize value of r_A/r_B for the coordination is 0.225.

Table 6.6 Data on the nearest neighbour separation and lattice constant, predicted from the radii of the constituent ions/atoms, for some crystals

Crystal	Radius of the Constituent ions/atoms (Å)	Predicted nearest neighbour separation, d (Å)	Relation between d and a (lattice constant)	Lattice constant a Å	
				Predicted	Observed
NaF	Na^+ (0.97)	2.33	a = 2d	4.66	4.62
	F^- (1.63)				
NaCl	Na^+ (0.97)	2.78	a = 2d	5.56	5.63
	Cl^- (1.81)				
Diamond	C (0.72)	1.54	$a = \frac{4d}{\sqrt{3}}$	3.56	3.56
Ne	Ne (1.58)	3.16	$a = \sqrt{2}d$	4.47	4.46

Also $$\cos\theta = \frac{r_B}{r_A + r_B} = \sqrt{\frac{2}{3}} = 0.816$$

This gives $\frac{r_A}{r_B}$ = 0.225 which is the minimum value for the fourfold coordination to be stable. These calculations show that for the formation of a stable threefold coordination, the ion radius ratio must lie between 0.155 and 0.225.

Finally we demonstrate how successfully the bond length and the lattice constants are reproduced with the knowledge of the self-consistent atomic/ionic sizes. The degree of success achieved can be perceived by the data in table 6.6.

Returning now to the question of stability, we would expect at first sight that the CsCl structure should always be more stable than the other structures, because it has the highest coordination number. Now, although it is true that a high coordination number will lead to strong binding and thus high stability, there is another requirement to be fulfilled, viz., that ions of opposite sign should be separated by as small a distance as possible. In other words, positive and negative ions should "touch," because any increase in their separation would give as higher energy (less binding) according to equation (6.8). It is at this point that a consideration of the relative radii of the ions can provide at least a guiding principle. To illustrate this, let us consider an ionic crystal of the type A^+B^+ with ionic radii r_1 and r_2, where we assume $r_1 < r_2$. Suppose we build a CsCl structure with

these ions, assuming that positive and negative ions touch each other. The cube edge, corresponding to the separation of ions of equal sign, is then $a = (2/\sqrt{3})(r_1 + r_2)$

Suppose the ion of radius r_1 is the ion in the center of the cube. If we now increase the radius r_2 gradually, leaving $r_1$constant, we reach a value of r_2 such that further increase makes it impossible for the central ion to touch the ones at the corners. This critical value is clearly reached when

$a = 2r_2 = (2/\sqrt{3})(r_1 + r_2)$ or $r_2 = 1.37r_1$

Thus r_2 >1.37 r_1 would lead to an increase in the distance between positive and negative ions and consequently to an increase in energy. The competition between coordination number on the one hand and separation between positive and negative ions on the other will thus set in as soon as the ratio of the radii becomes larger than 1.37 and a more favorable structure may result, viz., the **NaCl** structure. The stability limits of the latter may be investigated in the same way. In this case the critical ratio of the radii is determined by $2r_2 = (r_1 + r_2)\sqrt{2}$ or $r_2 = 2.44r_1$

Again, if the ratio becomes larger than 2.44, positive and negative ions cannot touch each other, leading to an increase of the energy and consequently to the formation of the more stable Zincblende structure (fig 6.3). For this structure, positive and negative ions cannot touch each other if r_2 > **4.55** r_1**.** The stability limits as derived from the above simplified billiard ball model for the ions are therefore

Cesium chloride ……$1 < r_2/r_1 < 1.37$

Sodium chloride ……$1.37 < r_2/r_1 < 2.44$

Zincblende ……$2.44 < r_2/r_1 < 4.55$

It must be emphasized that these result can be looked upon only as a rough rule. In general, however, one may say that the CsCl structure is found in those compounds for which the ionic radii are nearly equal, whereas the zincblende structure occurs only when the ratio of the radii is about two or more. This may be illustrated by a few examples in Table 6.5.

It is finally of interest to note that structure transformations have been observed under high pressures. A review of this subject may be found in a book by Bridgman.

Table 6.5 Ratio of Negative and Positive Ion Radii for Salts with the Cesium Chloride and Zincblende Structure

Cesium chloride structure	r_-/r_+	Zincblende structure	r_-/r_+
CsCl	1.1	ZnS	2.1
CsBr	1.2	ZnSe	2.3
CsI	1.3	BeS	5.1
TlCl	1.2	BeSe	5.6
TlBr	1.3	CuCl	1.9
TlI	1.5	CuBr	2.0
		CuI	2.3

6.7 Refinements of the Born theory:

The development of wave mechanics provided a better understanding of the chemical bond and interatomic forces in general. As a result, several refinements of the Born theory have been made, in particular by Born and Mayer and their collaborators. The essential refinements were the following:

1. Quantum mechanical calculations of the forces between ions indicate that a simple power law for the repulsive forces (6.3) cannot be rigorous. One therefore replaced this law by an exponential one of the form

$$\in_{\text{rep}}(r) = ce^{-r/\rho} \quad \text{----------} \quad (6.15)$$

where c and ρ are constants.

2. One added an attractive term to the lattice energy corresponding to the van der Waals forces which act between ions or atoms with a rare gas electron configuration.

3. One takes into account the "zero-point energy" of the crystal.

We shall not go through the calculation of the lattice energy which includes the modifications just mentioned, because the method is in principle the same as the one followed above. Also, the differences in the results obtained are slight. However, a few remarks about the modifications themselves, in particular about those mentioned under (2) and (3) may be in order.

The van der Waals forces are responsible for the cohesion in the liquid and solid states of rare gases as well as for most organic crystals. These forces have been treated by London and Margenau on a quantum mechanical basis. An approximate expression for the interaction energy of two atoms or ions with filled shell electron configuration is

$$\in (r) = -\frac{3}{2} \cdot \frac{\alpha_1 \alpha_2}{r^6} \cdot \frac{I_1 I_2}{I_1 + I_2} \qquad \text{---------- (6.16)}$$

where I_1 and I_2 refer to the ionization energies of the particles involved and α_1, α_2 refer to the polarizabilities. The nature of these forces is essentially a quantum effect, although the fact that they vary with the sixth power of the distance may easily be shown from classical considerations.

A homogeneous electric field E induces in an atom a dipole:

$\mu = qx = \alpha E$ ---------- (6.17)

where q and x are, respectively, the effective charge and displacement; α is the polarizability of the atom. The energy of the atom in the field is then

$$\in = -\int_0^x qE dx = -\int_0^E qE \frac{\alpha}{q} dE = -\tfrac{1}{2}\alpha E^2 \quad \text{---------- (6.18)}$$

For a field strength varying with time, one would have for the average energy,

$\in = -(\alpha/2)(E^2)$ ---------- (6.19)

Now, suppose the atom is under influence of another atom at a distance r. The latter may be considered a system of oscillating dipoles formed by the nucleus and the electrons. The electric field strength of a dipole varies as r^{-3} and hence, according to (6.18), the energy of one atom in the field of another may be written

$\in = -$ constant $/ r^6$ ---------- (6.20)

The mutual energy of two atoms would then be given by the sum of two terms of the type (6.20). From the classical point of view, therefore, these forces are a consequence of the dipole-dipole interaction between the atoms.

Actually, the energy corresponding to (6.16) is only part of the van der Waals energy and there is an infinite series of rapidly converging terms. The next one corresponds to dipole-quadrupole interaction and varies as r^{-8}.

For the alkali halides, the attractive energy corresponding to (6.16) is of the order of a few per cent of the total lattice energy. For the silver halides it is appreciably more; e.g., for AgBr it is about 14 per cent. This is a consequence of the relatively high polarizability of the silver ion. We should note that the van der waals energy sometimes plays an important role in the discussion of the stability of different lattice structures.

The zero-point energy of the crystal is also a consequence of quantum mechanics. The possible energy levels of a harmonic oscillator are given by

$$\in = (n + \tfrac{1}{2})hv \qquad \text{----------} \quad (6.21)$$

where n is an integer and v is the frequency. Thus, even at absolute zero an oscillator has a zero point energy of $hv/2$. Now, in the Debye theory of the specific heat of solids, a crystal is represented formally by a system of harmonic oscillators with a frequency spectrum given by

$$F(v)dv = 4\pi V\left(\frac{2}{c_t^3} + \frac{1}{c_1^3}\right)v^2 dv \qquad \text{----------} \quad (6.22)$$

where V is the volume of the crystal and c_t and c_l are, respectively, the velocities of propagation of transverse and of longitudinal elastic waves. Making use of the definition of the Debye frequency v_D, one may write

$$F(v)dv = \left(\frac{9N}{v_D^3}\right)v^2 dv \qquad \text{----------} \quad (6.23)$$

where N stands for the total number of atoms or ions in the crystal. Hence, at absolute zero, the contribution of the zero-point energy is

$$\tfrac{1}{2}\int_0^{v_D} F(v)hv\, dv = \tfrac{1}{2}Nhv_D \qquad \text{----------} \quad (6.24)$$

Per ion pair this corresponds to $9hv_D/4$. with a Debye frequency of the order of 10^{12}-10^{13} sec^{-1} this gives about 0.1 eV. As a correction to the lattice energy the zero point energy thus contributes about 1 per cent. Note that this correction reduces the values given in Table 6.2,

	LiF	CsI
Coulomb……	–12.4	– 6.4
Repulsive ……..	+ 1.9	+0.63
Dipole-dipole ……	– 0.17	– 0.48
Dipole-quadrupole …	– 0.03	– 0.04
Zero-point …….	+ 0.17	+ 0.3

whereas the van der Waals correction raises them. In general, the van der Waals correction is more important for heavy elements (large polarizabilities), and the zero-point energy for light elements (high Debye frequency). As an example, we give here the

various contributions to the lattice energy for the two extreme cases LiF and CsI (all energies in eV).

6.8 Born-Mayer revised theory of ionic crystals

This theory closely follows the earlier theory given by Born but for the form of repulsive energy term. The details of the theory are given below.

6.8.1 The electrostatic energy:

Consider an ionic crystal having N molecule, given a total of 2N ions. The electrostatic energy of the crystal can be written as

$$U_e = NU_i \quad \text{----------} \quad (6.25)$$

Where U_i is the average potential energy of a single ion in the field of the other remaining ions. Also,

$$U_i = \sum_{i=1}^{2N-1} U_{ij} \quad \text{----------} \quad (6.26)$$

with $U_{ij} = \pm \frac{1}{4\pi\varepsilon_0} \frac{q^2}{|\mathbf{r}_{ij}|}$; $\mathbf{r_{ij3} = r_i - r_j}$

Here U_{ij} represents the electrostatic interaction energy in SI units between two ions bearing an equal charge q and their positions being given by the vectors $\boldsymbol{r_i}$ and $\boldsymbol{r_j}$. Taking the origin of the coordinate system at the position of one of the positive ions, we have

$$\mathbf{r_{ij}} = \left(\hat{i}n_1 + \hat{j}n_2 + \hat{k}n_3\right)r$$

$$|\mathbf{r_{ij}}| = \left(n_1^2 + n_2^2 + n_3^2\right)^{\frac{1}{2}} r \quad \text{----------} \quad (6.27)$$

where $\mathbf{n_1, n_2, n_3}$ represent the number of units of the nearest neighbour distance $\boldsymbol{r}$ along the **x, y, z** axes of the crystal. Having a positive ion at the origin, we observe that $\mathbf{(n_1, n_2, n_3)}$ $\boldsymbol{r}$ represents the location of

a negative ion, if $\mathbf{(n_1, n_2, n_3)}$ is an odd integer

a positive ion, if $\mathbf{(n_1, n_2, n_3)}$ is an even integer.

The Coulomb energy between the ion at the origin and any other ion located at $\boldsymbol{r_j}$ will be

$$U_{ij} = \frac{(-1)^{n_1+n_2+n_3} q^2}{4\pi\varepsilon_0 (n_1^2 + n_2^2 + n_3^2)^{\frac{1}{2}} r} \quad \text{----------} \quad (6.28)$$

or

$$U_i = \frac{\alpha q^2}{4\pi\varepsilon_0 r} \quad \text{----------} \quad (6.29)$$

Where α is a constant known as the *Madelung constant*, expressed by the following relation:

$$\alpha = \sum_{n_1=0}^{2N-1} \sum_{n_2=0}^{2N-1} \sum_{n_3=0}^{2N-1} (-1)^{n_1+n_2+n_3} (n_1^2 + n_2^2 + n_3^2)^{-\frac{1}{2}} \quad \text{----------} \quad (6.30)$$

$(n_1, n_2, n_3) \neq (0, 0, 0)$

For NaCl crystal, the first term is $-6/\sqrt{1}$ that is contributed by the six nearest neighbours of the opposite type located at (±1, 0, 0), (0, ±1, 0) and (0, 0, ±1) in units of the nearest neighbour distance. The second term $12/\sqrt{2}$ comes from the 12 nearest neighbours of the same type, located at (±1, ±1, 0), (0, ±1, ±1) and (±1, 0, ±1) and at a distance of $\sqrt{2}r$ from the origin. The sum in (6.30) can be made to converge rapidly using some tricky mathematical method. The calculations yield the value of Madelung constant as –1.7476 for a NaCl crystal. For some other important ionic crystals such as CsCl, Zincblende (ZnS) and *würtzite* (ZnS), the values are –1.7627, –1.6381 and –1.641, respectively. A detailed account of the method of calculation has been provided by Born and Huang. respectively.

6.8.2 The Repulsive Overlap Energy:

An approximate analytical form of the potential energy for pair of ions as originally introduced by Born and Mayer is given by

$$U_{ij}^{(r)} = \lambda_{ij} \exp\left(\frac{-|r_{ij}|}{\rho}\right) \quad \text{----------} \quad (6.31)$$

where λ_{ij} and ρ are the empirical parameters which depend on the nature of the ions *i* and *j* but are independent of the distance between them.

The repulsive term is representative of the fact that the overlap between the electron configurations of neighbouring ions is resisted. The constants λ and ρ stand for the strength and range of interaction, respectively, and can be determined with the knowledge of the experimental values of the lattice constant and compressibility. The range ρ is defined as the values of $|r_{ij}|$ for which the interaction is reduced to 1/e of the interaction λ_{ij} when the two ions are in contact, treating them as ideal point charges.

Thus the total potential energy of the ions r_i and r_j can be written as the sum of the attractive potential energy $U_{ij}^{(a)}$ and the repulsive potential energy $U_{ij}^{(r)}$ such that

$$U_{ij}^{(t)} = U_{ij}^{(a)} + U_{ij}^{(r)} = -\frac{1}{4\pi\varepsilon_0}\frac{q^2}{|r_{ij}|} + \lambda_{ij}\exp\left(\frac{-|r_{ij}|}{\rho}\right) \quad \text{---------- (6.32)}$$

The total potential energy of a crystal of NaCl structure will be written as

$$U = N\left(-\frac{\alpha q^2}{4\pi\varepsilon_0 r} + 6\lambda_{+-}\exp\left(\frac{-\sqrt{1}r}{\rho}\right) + 12\lambda_{++}\exp\left(\frac{-\sqrt{2}r}{\rho}\right) + ...\right) \quad \text{---------- (6.33)}$$

The subscript of λ in (6.33) indicate that its value is different with reference to the interaction between the like ions and that between unlike ions.

If we consider the repulsion only between the nearest neighbours the relation (6.33) assumes the general from

$$U = N\left(-\frac{\alpha q^2}{4\pi\varepsilon_0 r} + z\lambda e^{-r/\rho}\right) \quad \text{---------- (6.34)}$$

Where ***z*** stands for the number of nearest neighbours (six in NaCl structure).

Since the cohesive energy is referred to as the minimum value of the potential energy, we can achieve the objective by expressing (6.34) in terms of the nearest neighbour distance, for which the minimum potential energy occurs .For *U* to be minimum,

$$\frac{dU}{dr} = 0$$

From (6.34), we get

$$\frac{dU}{dr} = N\left(\frac{\alpha q^2}{4\pi\varepsilon_0 r^2} - \frac{z\lambda}{\rho}e^{-r/\rho}\right)_{r=r_0} = 0$$

where r_0 is the nearest neighbour distance at equilibrium .This condition gives

$$z\lambda e^{-r_0/\rho} = \frac{\alpha q^2 \rho}{4\pi\varepsilon_0 r_0^2} \quad \text{---------- (6.35)}$$

From (6.34) and (6.35), we get

$$U_0 = \frac{-N\alpha q^2}{4\pi\varepsilon_0 r_0^2}\left(1 - \frac{\rho}{r_0}\right) \quad \text{---------- (6.36)}$$

As an exercise, we apply (6.36) to a NaCl crystal. The potential energy of a single ion in the crystal is given by

$$U_{ion} = \frac{\alpha q^2}{4\pi\varepsilon_0 r_0^2}\left(1-\frac{\rho}{r_0}\right) \quad \text{----------} \quad (6.37)$$

For NaCl, $\alpha = 1.748$, $r_0 = 2.81\,\overset{o}{A}$ and ρ is generally found to be $r_o/9$. (A small value of ρ shows the short range of repulsive interaction.)

Substituting these value in (6.37), we get

$U_{ion} = -7.97$ eV

This is not correct energy of a single ion as every ion is counted twice while taking the pair interaction. Therefore, the true potential energy of an ion is –7.97/2 eV = –3.99 eV, which is essentially the contribution of a single ion to the potential energy .With reference to ionic crystals the quantity, 7.97 eV, is referred to as the lattice energy which only the cohesive energy per ion pair .This is in excellent agreement with the measured value of 7.96eV.

The cohesive energy per atom may as well be calculated easily .The ionization potential of the Na is 5.14 eV and the electron affinity of Cl is –3.62 eV. In the formation of $Na^+ + Cl^-$ ion pair, the energy spent in the electron transfer is equal to (5.14-3.61) eV i.e. 1.53 eV, meaning thereby that each atom contributes 0.77 eV to the cohesive energy. Thus, the cohesive energy/atom of NaCl

= (–3.99 + 0.77) eV

= –3.22 eV

The success of the underlined theory is further emphasized by the closeness of the measured value (3.28 eV/atom).

6.9 Summary:

The potential energy of an ionic crystal is considered to be composed of two components, one representing the electrostatic energy and the other belonging to the repulsive overlap, which has its origin in the Pauli exclusion principle.

$\in = -Ae^2/r + B/r^n$

Born's theory of the lattice energy is based on the assumption that the crystals under consideration are built up of positive and negative ions.

For the NaCl structure the result is

$\in_e = -Ae^2 / r$ with $A = 1.747558.....$

The constant *A* is called the Madelung constant. For other crystal structures composed of positive and negative ions of the same valency, the Madelung constants are

Cesium chloride A = 1.762670

Zincbelende (ZnS) A = 1.6381

Wurtzite (ZnS) A = 1.641

Assuming that the two types of forces just discussed are the only ones we have to take into account and neglecting surface effects, we thus find for the total binding energy of a crystal containing *N* positive and *N* negative ions.

Born obtained the unknown repulsive exponent n from measurements of the compressibility K_0 of the crystals

The repulsive exponent $n = 1 + 9ca_0^4 / K_0 e^2 A$

It is obvious that the repulsive forces acting between two ions will depend on the distribution of the electronic charges in the ions and especially on the number of electrons in the outer shells. For example, we would expect *n* to be larger for NaCl than for LiCl, because the Na^+ ion has eight outer electrons and the Li^+ ion has only two. From Born Haber method we can obtain an expression for lattice energy as

$\in_{L\,\text{exp.}} = S_{\text{Na}} + I_{\text{Na}} + \frac{1}{2} D_{\text{Cl}_2} - E_{\text{Cl}} + q$

For NaCl, all quantities on the right-hand side are known from experiments and thus we are able to give an experimental value for $\in_L$, which may be compared with the one calculated with the Born theory. For NaCl we find

$\in_{L\,\text{exp.}} = 1.1 + 5.1 + 1.2 - 3.8 + 4.3 = 7.9\ \text{eV}$

whereas Born's theory yields 8.0 eV.

From a study on alkali halide crystals it is concluded that depending on the ratio of anion to cation radii of various structures result as

Cesium chloride ……$1 < r_2/r_1 < 1.37$

Sodium chloride ……$1.37 < r_2/r_1 < 2.44$

Zincblende ……$2.44 < r_2/r_1 < 4.55$

Born and Mayer suggested refinements to the original Born's theory particularly to the form of repulsive potential.

Quantum mechanical calculations of the forces between ions indicate that a simple power law for the repulsive forces cannot be rigorous. One therefore has to replace this law by an exponential one of the form

$$\in_{\text{rep}} (r) = ce^{-r/\rho}$$

where c and ρ are constants.

Other refinements include adding an attractive term has to be to the lattice energy corresponding to the van der Waal's forces which act between ions or atoms with a rare gas electron configuration and taking into account the "zero-point energy" of the crystal.

The van der Waal's forces are responsible for the cohesion in the liquid and solid states of rare gases as well as for most organic crystals.

The van der Waal's contribution to energy can be written as $\in$ = – constant / r^6

For the alkali halides, the attractive energy corresponding to van der Waal term is of the order of a few per cent of the total lattice energy. For the silver halides it is appreciably more; e.g., for AgBr it is about 14 per cent. This is a consequence of the relatively high polarizability of the silver ion. We should note that the van der Waal's energy sometimes plays an important role in the discussion of the stability of different lattice structures.

6.10 Key words:

Lattice energy – Born's theory – Medelung constant – Repulsive exponent – Heat of formation – Born Haber cycle – Ionic radii – Zero point energy – Polarizabilities – Debye frequency – Lattice constants.

6.11 Review questions:

1. Define Madelung constant and obtain its value for NaCl crystal. Discuss its physical significance.
2. How do you classify the crystal based on binding energy.

3. Discuss clearly giving examples the relationship between the ionic radii and stability of different types of structures for ionic crystals. Obtain the necessary limits for the different structures.
4. Obtain an expression for the binding energy of an ionic crystal.
5. How does the stability of structures depend on ionic radii?
6. (a) Classify solids based on the nature of inter atomic forces. (b) Derive an expression for the binding energy of NaCl crystal. (c) What is Madelung constant?
7. Discuss Born's theory for determining the binding energy of an ionic crystal.
8. (a) Obtain an expression for binding energy of an ionic crystal based on Born's theory. (b) How the stability of structures depends on ionic radii? (c) Comment on the refinement of the Born's theory.

6.12 Text and reference books:

1. Elements of Solid State Physics by J.P.Srivastava (PHI)
2. Solid State Physics by M.A. Wahab (Narosa)
3. Elements of Solid State Physics by A.Omar (Pearson education)
4. Solid State Physics by S.O. Pillai (New Age)
5. Solid State Physics by C.Kittel (Asia Publishing house)
6. A Text Book of Solid State Physics by S.L.Kakani and C.Hemrajani (S.Chand)
7. Fundamentals of Solid State Physics by Saxena Gupta Saxena (Pragati Prakasan)

UNIT – II **LESSON: 7**

Elastic constants of Crystals

Aim:

To know about elastic constants of crystal and also to learn about elastic waves in cubic crystal.

Objectives:

- To know the relation between crystal binding (energy) and elastic properties of crystals.
- To know definitions of elastic stress, elastic strain and the respective relationship in between them
- To know definitions of dilation, elastic compliance and stiffness constants and how these are applied to cubic crystals.
- To know about elastic waves in crystal and propagation of waves in the (100) direction and (110) directions.

Structure Of The Lesson:

7.1 Introduction:

The crystal binding discussed in the earlier lesson establishes how this binding controls the elastic behavior of solids .The nature of binding forces in a solid often reflected in its elastic response. The study of elastic properties is essential for the interpretation of several properties of solids. Elastic properties relate themselves to thermal properties like the Debye temperature. Now we will know stress-strain relationship in a crystal treating it as a homogeneous continuous elastic medium. The crystal's picture as a periodic array of atoms is replaced with homogeneous elastic medium.

7.1.1: Elastic stress:

Let us consider uniform deformation of an elementary cube of a crystal. Under the action of deforming force, an internal force develops within the crystal as a function of the applied force. The internal force acting on the unit area of a crystal is defined as the stress. For the present treatment, it is assumed that the applied force is not large and the Hookes' law is valid (stress is proportional to strain). The stress acting on the six faces of the cube is expressed by nine components $\sigma_{xx,}$ σ_{xy} , σ_{xz} , σ_{yx} ,σ_{yy} ,$\sigma_{yz,}$ σ_{zx} ,σ_{zy} ,$\sigma_{zz,.}$ The first subscript denotes the direction of the applied force and the second subscript denotes the direction of the normal to the force on which the force is applied. The stress is a tensor of second rank and denoted by a (3x3) matrix namely

$$\left[\sigma_{\alpha\beta}\right] = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{bmatrix} \quad \text{---------- (7.1)}$$

with α, β = x, y, z.

The components σ_{xx}, $\sigma_{yy,}$ σ_{zz} express the normal stress components acting on the yz , zx, and xy faces respectively. The remaining six components represent the tangential stress (two components of each of the three pairs of the faces). If the cube is in the state of static equilibrium and it does not rotate under the influence of tangential stress components $\sigma_{\alpha\beta}$ and $\sigma_{\beta\alpha}$ would produce equal and opposite rotations. Hence $\sigma_{\alpha\beta} = \sigma_{\beta\alpha}$ and the nine stress components reduce to six independent components. Figure 7.1 shows a normal component σ_{yy} and a tangential component σ_{yz} acting on the respective faces of

a cube. The stress components have the dimension of force per unit area or energy per unit volume.

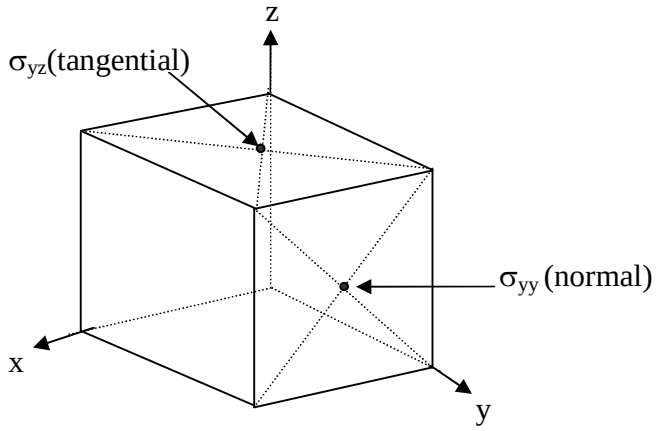


Fig 7.1:The normal stress component σ_{yy}, produced by a force applied in the +y direction on the xz face (with the normal along the +y direction) of a cube. The tangential stress component σ_{yz} produced by a force applied in the +y direction on the xy face (with normal along the +z direction) of a cube.

7.1.2 Elastic Strain:

Let the three orthogonal unit vectors $\hat{i}, \hat{j}, \hat{k}$ are embeded in an unstrained crystal. Suppose on straining the crystal by a small deforming force these vectors are transformed into non-orthogonal vector $\bar{l}, \bar{m}, \bar{n}$ respectively, with individual magnitudes differing from unity. We treat the sets of vectors as the new and old coordinate axes. The new coordinate axes may be expressed as

$$\left.\begin{aligned} \bar{l} &= (1+\varepsilon_{xx})\hat{i} + \varepsilon_{xy}\hat{j} + \varepsilon_{xz}\hat{k} \\ \bar{m} &= \varepsilon_{yx}\hat{i} + (1+\varepsilon_{yy})\hat{j} + \varepsilon_{yz}\hat{k} \\ \bar{n} &= \varepsilon_{zx}\hat{i} + \varepsilon_{xzy}\hat{j} + (1+\varepsilon_{zz})\hat{k} \end{aligned}\right\} \quad \text{---------- (7.2)}$$

where $\varepsilon_{\alpha\beta}$ define the deformation. These coefficients are dimensionless and very small (<<1) for a small strain.

By taking the dot products $\bar{l}\cdot\bar{l}, \bar{m}\cdot\bar{m}, \bar{n}\cdot\bar{n}$, we can easily show that the magnitude of the each of the three new vectors is different from unity. Also the dot products $\bar{l}\cdot\bar{m}, \bar{m}\cdot\bar{n}, \bar{n}\cdot\bar{l}$ do not vanish indicating that $\bar{l}, \bar{m}$ and $\bar{n}$ are not orthogonal vectors.

The strain components $e_{\alpha\beta}$ are defined in terms of $\varepsilon_{\alpha\beta}$ as

$e_{xx}= \varepsilon_{xx}$, $e_{yy} = \varepsilon_{yy}$, $e_{zz} = \varepsilon_{zz}$ ---------- (7.3(a))

and

$$\left.\begin{aligned} e_{xy} &= \bar{l}.\bar{m} \cong \varepsilon_{yx} + \varepsilon_{xy} \\ e_{yz} &= \bar{m}.\bar{n} \cong \varepsilon_{zy} + \varepsilon_{yz} \\ e_{zx} &= \bar{n}.\bar{l} \cong \varepsilon_{xz} + \varepsilon_{zx} \end{aligned}\right\} \quad \text{---------- (7.3(b))}$$

In relation (7.3b) the sign ≅ may be replaced by the sign =, if the terms of order ε^2 are neglected ($\varepsilon_{\alpha\beta}<<1$) . It is to be noted that the strain components in (7.3b) are defined in terms of the changes in angle between the axes. Therefore for a rigid rotation in which angles do not change $e_{xy} = e_{yz} = e_{zx} = 0$ and $\varepsilon_{xy} = -\varepsilon_{yx}$, $\varepsilon_{yz} = -\varepsilon_{zy}$ and $\varepsilon_{zx}= -\varepsilon_{xz}$. If we do not consider pure rotation since they are not deformed, we may always take

$\varepsilon_{xy} = \varepsilon_{yx}$, $\varepsilon_{yz} = \varepsilon_{zy}$, and $\varepsilon_{zx} = \varepsilon_{xz}$ ---------- (7.4)

using (7.4) in (7.3b) we have

$$\left.\begin{aligned} \varepsilon_{xy} &= \varepsilon_{yx} = \frac{1}{2} e_{xy} \\ \varepsilon_{yz} &= \varepsilon_{zy} = \frac{1}{2} e_{yz} \\ \varepsilon_{zx} &= \varepsilon_{xz} = \frac{1}{2} e_{zx} \end{aligned}\right\} \quad \text{---------- (7.5)}$$

Now consider an atom at position '$\bar{r}$' in the unstrained crystal. The position vector is expressed as

$$\bar{r} = x\hat{i} + y\hat{j} + z\hat{k} \quad \text{---------- (7.6)}$$

Let the position of the atom in the strained crystal be given by

$$\bar{r}' = x\bar{l} + y\bar{m} + z\bar{n} \quad \text{---------- (7.7)}$$

The displacement of the atom under the action of the deforming force is

$\Delta\bar{r} = \bar{r}' - \bar{r}$ ---------- (7.8)

Making use of (7.2) & (7.5) to (7.7) the above relation can be expressed as

$$\Delta\bar{r} = \left[xe_{xx} + \frac{1}{2} ye_{yx} + \frac{1}{2} ze_{zx} \right]\hat{i} + \left[\frac{1}{2} xe_{xy} + ye_{yy} + \frac{1}{2} ze_{zy} \right]\hat{j} + \left[\frac{1}{2} xe_{xz} + \frac{1}{2} ye_{yz} + ze_{zz} \right]\hat{k}$$

---------- (7.9)

We rewrite it as

$$\Delta(\bar{r}) = u_1\hat{i} + u_2\hat{j} + u_3\hat{k} \quad \text{---------- (7.10)}$$

where

$$\left.\begin{aligned} u_1 &= xe_{xx} + \frac{1}{2}ye_{yx} + \frac{1}{2}ze_{zx} \\ u_2 &= \frac{1}{2}xe_{xy} + ye_{yy} + \frac{1}{2}ze_{zy} \\ u_3 &= \frac{1}{2}xe_{xz} + \frac{1}{2}ye_{yz} + ze_{zz} \end{aligned}\right\} \text{ ---------- (7.11)}$$

Where u_1, u_2, u_3 are the displacement components along the coordinate axes of the unstrained crystal.

All the six dimensionless strain components can now be defined as follows by taking partial derivatives of u_1, u_2 and u_3

$$\left.\begin{aligned} e_{xx} &= \frac{\partial u_1}{\partial x}; & e_{xy} &= \frac{\partial u_2}{\partial x} + \frac{\partial u_1}{\partial y} \\ e_{yy} &= \frac{\partial u_2}{\partial y}; & e_{yz} &= \frac{\partial u_3}{\partial y} + \frac{\partial u_2}{\partial z} \\ e_{zz} &= \frac{\partial u_3}{\partial z}; & e_{zx} &= \frac{\partial u_1}{\partial z} + \frac{\partial u_3}{\partial x} \end{aligned}\right\} \text{ ---------- (7.12)}$$

Relation in (7.12) give alternative definition of strain components. Relation (7.2) and (7.3a) indicate that e_{xx}, e_{yy} and e_{zz} represent linear strain i.e., changes in length per unit length along the three axes. Each of the other type of three components interpret a combination of two simple shears. Take for example,

$$e_{yz} = \frac{\partial u_3}{\partial y} + \frac{\partial u_2}{\partial z}$$

It describes two shears one in which the planes normal to y-axis slide in the z-direction and other in which the planes normal to z-axis along the y-direction. Like stress, strain is also a tensor of second rank. In general it is described by nine components with the matrix representation as

$$\left[e_{\alpha\beta}\right] = \begin{bmatrix} e_{xx} & e_{xy} & e_{xz} \\ e_{yx} & e_{yy} & e_{yz} \\ e_{zx} & e_{zy} & e_{zz} \end{bmatrix} \text{ ---------- (7.13)}$$

With α,β=x,y,z.

7.1.3 Dilation: The fractional increase in volume created by deformation is called dilation. It is useful in determining some elastic constants such as the bulk modulus.

Volume of a unit cube after deformation is

$V' = \bar{l}.\bar{m} \times \bar{n}$ ---------- (7.14)

Substituting for $\bar{l}, \bar{m}, \bar{n}$ from (2.2) into (7.14) neglecting the product of 2 strain components ($e_{\alpha\beta}$<<1), we get

$V' = 1 + e_{xx} + e_{yy} + e_{zz}$ ---------- (7.15)

Therefore the dilation $\delta = \frac{V' - V}{V} = e_{xx} + e_{yy} + e_{zz}$ ---------- (7.16)

7.2 ELASTC COMPLIANCE AND STIFFNESS CONSTANTS:

According to Hooke's law the strain is directly proportional to the stress for sufficiently small deformations. Therefore for appreciable small elastic deformations of a crystal the stress tensor components and the strain tensor components are related as

$$\begin{bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{zx} \\ \sigma_{xy} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{21} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{31} & C_{32} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{41} & C_{42} & C_{43} & C_{44} & C_{45} & C_{46} \\ C_{51} & C_{52} & C_{53} & C_{54} & C_{55} & C_{56} \\ C_{61} & C_{62} & C_{63} & C_{64} & C_{65} & C_{66} \end{bmatrix} \begin{bmatrix} e_{xx} \\ e_{yy} \\ e_{zz} \\ e_{yz} \\ e_{zx} \\ e_{xy} \end{bmatrix} \text{ ---------- (7.17)}$$

Conversely, the strain components can be expressed as the linear function of the stress components.

$$\begin{bmatrix} e_{xx} \\ e_{yy} \\ e_{zz} \\ e_{yz} \\ e_{zx} \\ e_{xy} \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & S_{13} & S_{14} & S_{15} & S_{16} \\ S_{21} & S_{22} & S_{23} & S_{24} & S_{25} & S_{26} \\ S_{31} & S_{32} & S_{33} & S_{34} & S_{35} & S_{36} \\ S_{41} & S_{42} & S_{43} & S_{44} & S_{45} & S_{46} \\ S_{51} & S_{52} & S_{53} & S_{54} & S_{55} & S_{56} \\ S_{61} & S_{62} & S_{63} & S_{64} & S_{65} & S_{66} \end{bmatrix} \begin{bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{zx} \\ \sigma_{xy} \end{bmatrix} \text{ ---------- (7.18)}$$

The coefficients C_{12}, C_{11} etc are called ***elastic stiffness constants*** and represent moduli of elasticity with dimensions of $\frac{\text{force}}{\text{area}}$ or $\frac{\text{energy}}{\text{volume}}$. The other coefficients S_{11}, S_{12} are called elastic compliance constants and have dimensions of $\frac{\text{area}}{\text{force}}$ or $\frac{\text{volume}}{\text{energy}}$.

7.2.1 ELASTIC ENERGY DENSITY:

By analogy with the expression for the energy of a stretched spring, the elastic energy density φ is a quadratic function of strains in the approximation of Hooke's law

$$\phi = \frac{1}{2}\sum_{\mu=1}^{6}\sum_{V=1}^{6}\overline{C}_{\mu v}e_{\mu}e_{v} \text{ ---------- (7.19)}$$

where the indices 1 to 6 should read as

1 ≡ xx, 2 ≡ yy, 3 ≡ zz, 4 ≡ yz, 5 ≡ zx, 6 ≡ xy ---------- (7.20)

The coefficients $\overline{C}_s$ are found to be related to C_s of (7.17) as we will see below. We will exploit the definition (7.19) to show that 36 coefficients in (7.17) or (7.18) can be reduced in number. The very definition of potential energy allows us to obtain the stress components from the derivative of ϕ with respect to the associated strain component. For example, when the stress σ_{xx} acts on one face of a unit cube and the opposite face is held at rest, we have

$$\sigma_{xx} = \frac{\partial\phi}{\partial e_{xx}} = \frac{\partial\phi}{\partial e_1} = \overline{C}_{11}e_1 + \frac{1}{2}\sum_{\beta=2}^{6}(\overline{C}_{1\beta} + \overline{C}_{\beta 1})e_{\beta} \text{ ---------- (7.21)}$$

An inspection of relation (7.21) reveals that only the combination $\frac{1}{2}(\overline{C}_{\alpha\beta} + \overline{C}_{\beta\alpha})$ enters the stress strain relations ,implying that the elastic stiffness constants are symmetrical. Thus we have

$C_{\alpha\beta} = \frac{1}{2}(\overline{C}_{\alpha\beta} + \overline{C}_{\beta\alpha}) = C_{\beta\alpha}$ ---------- (7.22)

The above symmetrical property reduces the number of constants from 36 to 21

7.2.2 Application to cubic crystals:

In accordance with the Neumann's principle, the number of independent elastic stiffness constants decreases as the symmetry of a crystal increases. This number 21, 13, 5 and 3 respectively for triclinic, monoclinic, hexagonal, and cubic systems. The cubic crystals being the most symmetric have the least number of independent elastic stiff constants. We now derive this result.

We pronounce that the relations for the elastic energy density of a cubic crystal has the form

$$\phi = \frac{1}{2}C_{11}(e_{xx}^2 + e_{yy}^2 + e_{zz}^2) + \frac{1}{2}C_{44}(e_{yz}^2 + e_{zx}^2 + e_{xy}^2) + C_{12}(e_{yy}e_{zz} + e_{zz}e_{xx} + e_{xx}e_{yy}) \text{ ------ (7.23)}$$

The above relation does not have the quadratic terms

$$(e_{xx}e_{xy} + ...); (e_{yz}e_{zx} + ...); (e_{xx}e_{yz} + ...); \text{ ---------- (7.24)}$$

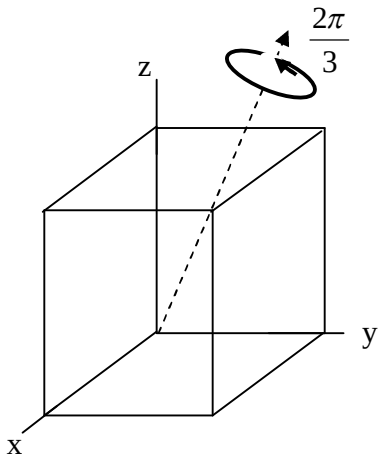


Fig 7.2 One of the four threefold Rotation axes (the direction of the body diagonal) in a cubic crystal . The [111] direction and its equivalent directions represent other threefold rotation axes (i.e. the direction of the other three body diagonals)

The correctness of (7.23) can be confirmed by showing that ϕ is invariant under all symmetry operations permitted in a cubic crystal. The minimum symmetry requirement for cubic symmetry is the presence of four three fold rotation along the directions of the four body diagonals of the cube (fig 7.2).

Counter clock wise rotations by $2\pi/3$ about the (111)direction and three other equivalent directions interchange the x, y,z axes according to the following four schemes.

$$\left.\begin{array}{ll} x \to y \to z \to x; & -x \to z \to -y \to -x \\ x \to z \to -y \to x; & -x \to y \to z \to -x \end{array}\right\} \text{----------(7.25)}$$

It is straight forward to check that the relation for ϕ remains unchanged when x, y, z are interchanged in (7.23) according to any of the four schemes (7.25). But every term appearing in (7.24) is odd in one or more indices. One of the schemes in (7.25) is surely such that its application to (7.23) would change the sign of the term in (7.24) .This confirms that the terms included in (7.24) have rightly been excluded from the relation (7.23) for ϕ. We may now easily derive the stress components from (7.23) Thus

$$\sigma_{xx} = \frac{\partial \phi}{\partial e_{xx}} = C_{11} e_{xx} + C_{12}(e_{zz} + e_{yy}) \text{ ---------- (7.26)}$$

The corresponding relation is given by (7.17) is

$$\sigma_{xx} = C_{11} e_{xx} + C_{12} e_{yy} + C_{13} e_{zz} + C_{14} e_{yz} + C_{15} e_{zx} + C_{16} e_{xy} \text{ ---------- (7.27)}$$

On comparing (7.26) and (7.27) we get

$C_{12} = C_{13}$; $C_{14} = C_{15} = C_{16} = 0$ ----------(6.28)

Further, from (7.23)

$$\sigma_{xy} = \frac{\partial \phi}{\partial e_{xy}} = C_{44} e_{xy} \text{ ---------- (7.29)}$$

The corresponding relation from (7.17) is

$$\sigma_{xy} = C_{61} e_{xx} + C_{62} e_{yy} + C_{63} e_{zz} + C_{64} e_{yz} + C_{65} e_{zx} + C_{66} e_{xy} \text{ ---------- (7.30)}$$

Comparison of (7.29) with (7.30) gives

$C_{66} = C_{44}$; $C_{61} = C_{62} = C_{63} = C_{64} = C_{65}0$ ---------- (7.31)

Proceeding this way for other stress components, we find that the array of values of the elastic stiffness constants of a cubic crystal may be expressed in the following matrix representation.

$$\left[C_{\alpha\beta}\right] = \begin{bmatrix} C_{11} & C_{12} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{12} & C_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{44} \end{bmatrix} \text{ ---------- (7.32)}$$

Where $\alpha, \beta = 1, 2, 3, 4, 5, 6$.

Matrix (7.32) shows that a cubic crystal has only three independent stiffness constants by calculating the inverse matrix to (7.32), we obtain the following relationship between the stiffness and compliance constants for cubic crystals.

$C_{44} = 1/S_{44}$; $C_{11} - C_{12} = (S_{11} - S_{12})^{-1}$, $C_{11} + 2C_{12} = (S_{11} + 2S_{12})^{-1}$ ---------- (7.33)

7.2.3 Bulk Modulus and Compressibility:

Let us consider a strained crystal which is uniformly dilated. This refers to the mathematical condition

$$e_{xx} = e_{yy} = e_{zz} = \frac{1}{3}\delta \quad \text{---------- (7.34)}$$

From (7.23) we obtain the following relation for the elastic charge density of a cubic crystal.

$$\phi = \frac{1}{6}(C_{11} + 2C_{12})\delta^2 \quad \text{---------- (7.35)}$$

The bulk modulus B is usually defined by the relation

$$\phi = \frac{1}{2}B\delta^2 \quad \text{---------- (7.36)}$$

Comparing (7.35) with (7.36), we express the bulk modulus of a cubic crystal as

$$B = \frac{1}{3}(C_{11} + 2C_{12}) \quad \text{---------- (7.37)}$$

The inverse of B has been interpreted as another useful elastic property called the compressibility K. The compressibility of a cubic crystal is accordingly given by

$$K = \frac{3}{C_{11} + 2C_{12}} \quad \text{---------- (7.38)}$$

7.3 Elastic waves in a cubic crystal:

Consider of an elementary cube of edges $\Delta x = \Delta y = \Delta z$ with in the volume of a cubic crystal. The cube edges are in the direction of x-, y-, and z- axes. When the cube is strained, let the stress $\sigma_{xx}(x)$ act on the face at x. Assuming that the variation of σ_{xx} is uniform along the x-direction, the stress on the face parallel to that at x can be expressed as $[\sigma_{xx}(x) + (\partial\sigma_{xx}/\partial x)\Delta x]$ (see fig 7.3). The net force on the cube due to a σ_{xx} component is equal to $[(\partial\sigma_{xx}/\partial x)\Delta x]\ \Delta y \cdot \Delta z$. The other forces in the x-direction arise from the variation of σ_{xy} and σ_{xz} across the cube. Therefore the net force on the cube along the x-direction is

$$F_x = \left(\frac{\partial\sigma_{xx}}{\partial x} + \frac{\partial\sigma_{xy}}{\partial y} + \frac{\partial\sigma_{xz}}{\partial z}\right)\Delta x \cdot \Delta y \cdot \Delta z \quad \text{---------- (7.39)}$$

The above force is actually the restoring force that tends to bring the cube to its unstrained state. As a result, the particles in the crystal are thrown to a motion described by the relevant equation of motion. If ρ is the density of the crystal, the force per unit volume on the crystal along the x-direction is $\rho(\partial^2 u_1/\partial t^2)$ and the equations of motion in the x-direction becomes

$$\rho \frac{\partial^2 u_1}{\partial t^2} = \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z} + \frac{\partial \sigma_{xy}}{\partial y} \quad \text{---------- (7.40)}$$

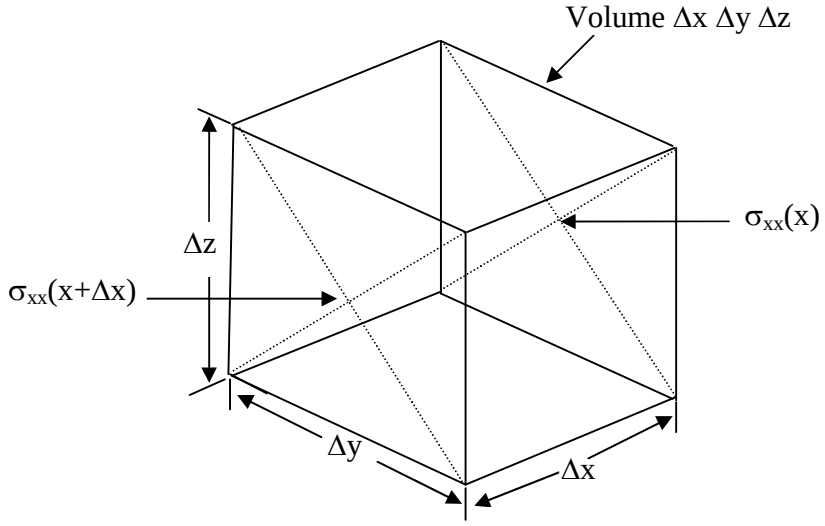


Fig (7.3) Variation of normal stress component (σ_{xx}) with in a cube along x edge

Using (7.17) and (7.32) the above equation is reduced to the following form for a cubic crystal.

$$\rho \frac{\partial^2 u_1}{\partial t^2} = C_{11} \frac{\partial e_{xx}}{\partial x} + C_{12}\left(\frac{\partial e_{yy}}{\partial x} + \frac{\partial e_{zz}}{\partial x} \right) + C_{44}\left(\frac{\partial e_{xy}}{\partial y} + \frac{\partial e_{zx}}{\partial z} \right) \quad \text{---------- (7.41)}$$

On eliminating the strain components with the help of (7.12) above equations assumes the form

$$\rho \frac{\partial^2 u_1}{\partial t^2} = C_{11} \frac{\partial^2 u_1}{\partial x^2} + C_{44}\left(\frac{\partial^2 u_1}{\partial y^2} + \frac{\partial^2 u_1}{\partial z^2} \right) + (C_{12} + C_{44})\left(\frac{\partial^2 u_2}{\partial x \partial y} + \frac{\partial^2 u_3}{\partial x \partial z} \right) \quad \text{---------- (7.42)}$$

Here u_1, u_2, u_3 are the components of the atomic displacement $\Delta(\bar{r})$ along the axes of the unstrained crystal [see equation (7.10)]. The solution to (7.42) turns out to have a waveform, indicating that waves propagate within the crystal when it is strained in such a way as there exists a non-zero stress on the crystal. These waves are called ***elastic waves*** because they are produced here in an elastic continuum by elastic deformation.

Similar to (7.42) there are equations of motion in the y- and z-directions. We write them by symmetry on the basis of (7.42).

$$\rho \frac{\partial^2 u_2}{\partial t^2} = C_{11} \frac{\partial^2 u_2}{\partial y^2} + C_{44}\left(\frac{\partial^2 u_2}{\partial x^2} + \frac{\partial^2 u_2}{\partial z^2} \right) + (C_{12} + C_{44})\left(\frac{\partial^2 u_1}{\partial x \partial y} + \frac{\partial^2 u_3}{\partial x \partial z} \right) \quad \text{---------- (7.43)}$$

$$\rho \frac{\partial^2 u_3}{\partial t^2} = C_{11} \frac{\partial^2 u_3}{\partial z^2} + C_{44}\left(\frac{\partial^2 u_3}{\partial x^2} + \frac{\partial^2 u_3}{\partial y^2} \right) + (C_{12} + C_{44})\left(\frac{\partial^2 u_1}{\partial x \partial z} + \frac{\partial^2 u_2}{\partial z \partial x} \right) \quad \text{---------- (7.44)}$$

Now we proceed to study the solutions of these equations in some common directions in cubic crystals.

7.3.1 Propagation of waves in the [100] direction:

First, consider a longitudinal or compressional wave that propagates along the x-cube edge. Its propagation constant or wavevector *k* is parallel to the particle displacement u_1, given by

$u_1 = (u_1)_0 \exp [i(kx - \omega t)]$ ---------- (7.45)

where $k = 2\pi / \lambda$, and ω is the angular frequency. When (7.45) is used as a trial solution of (7.42) and placed in it, we obtain the dispersion relation.

$\omega^2 \rho = C_{11} k^2$ ---------- (7.46)

This gives the velocity of the longitudinal wave in the [100] direction in a cubic crystal as

$$v_l = \frac{\omega}{k} = \left(\frac{C_{11}}{\rho} \right)^{1/2} \quad \text{---------- (7.47)}$$

Next, we consider a transverse or shear wave with its wavevector *k* along the x-cube edge and the particle displacement u_2 in the y-direction. Thus,

$u_2 = (u_2)_0 \exp [i(kx - \omega t)]$ ---------- (7.48)

The substitution of (7.48) into (7.43) gives

$\omega^2 \rho = C_{44} k^2$ ---------- (7.49)

Therefore, the velocity of a transverse wave in the [100] direction in a cubic crystal is

$$v_s = \frac{\omega}{k} = \left(\frac{C_{44}}{\rho} \right)^{1/2} \quad \text{---------- (7.50)}$$

It can be shown that a transverse wave with wavevector along the x-cube edge and the particle displacement u_3 in the z-direction moves with the identical velocity. This result asserts that two independent shear waves whose wavevectors point along the [100] direction propagate in a cubic crystal with equal velocities. For a general direction of the wavevector this result is not applicable. The geometry of the longitudinal and transverse (wave) propagation in the [100] direction in a cubic crystal is shown in fig 7.4(a)

7.3.2 Propagation of Waves in the [110] Direction:

The study of elastic waves propagation in the [110] direction (the direction of a face diagonal) is especially gainful because the three elastic constants can be obtained simply from the three measured propagation velocities.

Let us first consider a shear wave with wave vector $\bar{k} = k_x\hat{i} + k_y\hat{j}$ propagating in the xy plane and causing a particle displacement u_3 in the z-direction. Thus

$u_3 = (u_3)_0 \exp [i(k_x x + k_y y - \omega t)]$ ---------- (7.51)

Substituting (7.51) into (7.44) we have

$$\omega^2 = \left(\frac{C_{44}}{\rho}\right)k^2 \quad \text{---------- (7.52)}$$

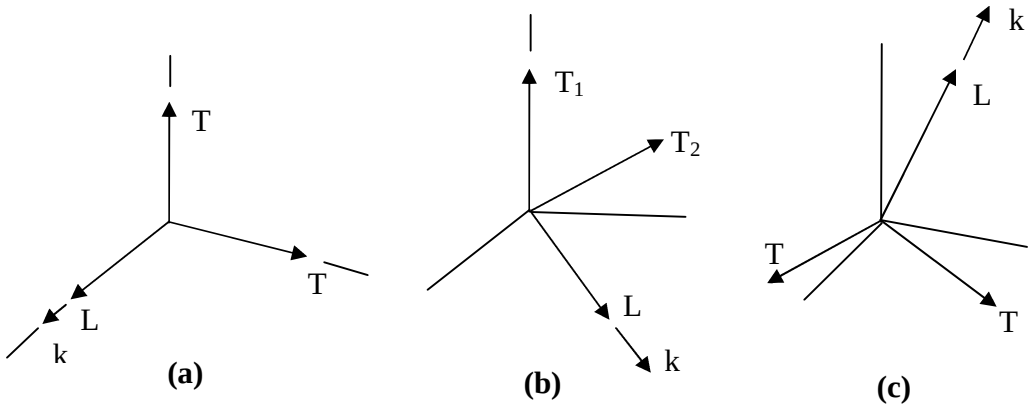


Figure 7.4: The propagations of elastic waves in a cubic crystal L denotes longitudinal waves and T denotes a shear wave.

(a). Waves in the [100] direction (k parallel to i) one longitudinal wave with velocity characterized by C_{11}, two degenerate shear waves with velocity characterized by C_{44}.

(b). Waves in the [110] direction (k parallel to (i+j)); one longitudinal wave with velocity characterized by ½ ($C_{11}+C_{12}+2C_{44}$), two shear waves with velocities characterized respectively by C_{44} and ½ ($C_{11}-C_{12}$) .

(c) Waves in the[111] direction [k parallel to (i+j+k)] .

One longitudinal wave with velocity characterized by 1/3($C_{11}+2C_{12}+C_{44}$), two degenerate shear waves with velocity characterized by 1/3($C_{11}-C_{12}+C_{44}$) i,j,k denotes the unit vectors along the cube edges defining the X-,Y-,Z- coordinate axes.

Next, consider the other two waves propagating in the xy plane with particle velocity in the xy plane. Let these be represented as

$$\left.\begin{aligned} u_1 &= (u_1)_0 \exp\ i(k_x x + k_y y - \omega t) \\ u_2 &= (u_2)_0 \exp\ i(k_x x + k_y y - \omega t) \end{aligned}\right\} \quad \text{---------- (7.53)}$$

Placing these solutions in (7.42) and (7.43), we get the following pair of equations.

$$\left.\begin{aligned} \omega^2 \rho u_1 &= (C_{11}k_x^2 + C_{44}k_y^2)u_1 + (C_{12} + C_{44})k_x k_y u_2 \\ \omega^2 \rho u_2 &= (C_{11}k_y^2 + C_{44}k_x^2)u_2 - (C_{12} + C_{44})k_x k_y u_1 \end{aligned}\right\} \quad \text{---------- (7.54)}$$

For a wave in [110] direction for which $k_x = k_y = k/\sqrt{2}$, equation (7.54) have a characteristically simple solution. The solution exists only if the determinant of the coefficients u_1 and u_2 in (7.54) vanishes. Thus

$$\begin{vmatrix} \frac{1}{2}(C_{11} + C_{44})k^2 - \omega^2\rho & \frac{1}{2}(C_{12} + C_{44})k^2 \\ \frac{1}{2}(C_{12} + C_{44})k^2 & \frac{1}{2}(C_{11} + C_{44})k^2 - \omega^2\rho \end{vmatrix} = 0 \quad \text{---------- (7.55)}$$

On solving equation (7.55) we get two dispersion relations

$$\omega^2 = \left[\frac{C_{11} + C_{12} + 2C_{44}}{2\rho}\right]k^2 ; \omega^2 = \left[\frac{C_{11} - C_{12}}{2\rho}\right]k^2 \quad \text{---------- (7.56)}$$

These two roots refer to two different types of waves. We now determine the nature of the waves by finding the direction of the particle displacement caused by the respective waves. When we substitute the first root in the first equation of (7.54) we get $u_1 = u_2$. Since the particle displacement occurs in the xy plane, $\Delta(\bar{r}) = u_1(\hat{i} + \hat{j})$. This shows that the displacement takes place in the [110] direction that represents the direction of $(\hat{i} + \hat{j})$ and happens to the direction of propagation of the wave (see fig 7.4(b)). Thus we infer that the first root in (7.56) belongs to a longitudinal wave.

Similarly, on substituting the second root in the first equation of (7.54), we obtain $u_1 = -u_2$ implying that, $\Delta r = u_1\ (\hat{i} + \hat{j})$. The direction of the vector $(\hat{i} - \hat{j})$ is indicated as $[1\bar{1}0]$ which is perpendicular to the[110] direction, the direction of propagation of the wave. Hence the second root in (7.56) must refer to a shear wave [fig 7.4 (b)]. The treatment of waves in the [111] direction is relatively lengthy though not complicated as for other general directions. The main features of propagation may be found in 7.4(c).

The subject matter of the last section is closely linked to the analysis of normal modes of vibration of crystals. Generally the direction of particle displacement or the polarization of normal modes may not be exactly parallel or perpendicular to the direction of the wave vector k. The analysis becomes easier when $\bar{k}$ is along any symmetry axis of the crystal.

7.4 Experimental Determination Of Elastic Constants:

In the experimental determination of elastic constants, a pulse of sound, generated by a piezoelectric quartz transducer, travels through the crystal. After reflection from the rear surface of the crystal, it travels back through the crystal and on emerging out of the crystal is detected by an electronic device as shown in fig 7.5. The elastic wave velocity can be determined by dividing the round trip distance by the time taken, and hence, the elastic constants of a crystal can be determined from the velocities of these waves through the relationship as mentioned above. In table 7.1 the measured elastic constants for some materials are given

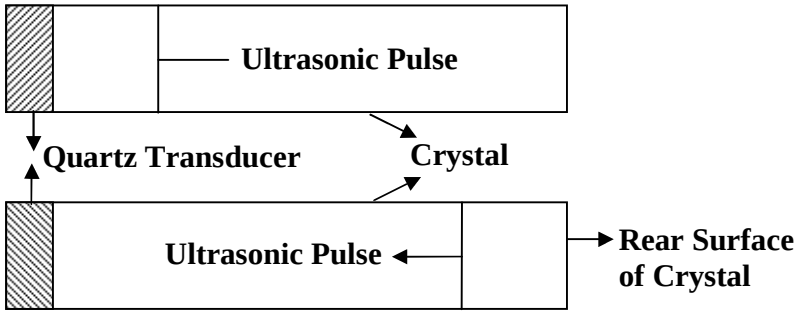


Fig 7.5 Ultrasonic pulse method for determination of elastic constants

Table 7.1 Elastic constants (in units of Gpa) of some cubic materials

Material	C_{11}	C_{44}	C_{12}
Diamond	1040	550	170
Si	165	79.2	64
GaAs	118	59.4	53.5
NaCl	49.1	12.8	12.8
ZnS	101	44.3	64.4
$BaTiO_3$	222	112	139
$SrTiO_3$	316	123	102

7.5 Summary:

The nature of binding forces in a solid often reflected in its elastic response. The study of elastic properties is essential for the interpretation of several properties of solids. Elastic properties relate themselves to thermal properties like the Debye temperature.

Under the action of deforming force, an internal force develops within the crystal as a function of the applied force. The internal force acting on the unit area of a crystal is defined as the stress. For the present treatment, it is assumed that the applied force is not large and the Hookes' law is valid (stress and strain). The stress acting on the six faces of the cube is expressed by nine components $\sigma_{xx,}\ \sigma_{xy}\ ,\sigma_{xz}\ ,\sigma_{yx}\ ,\sigma_{yy}\ ,\sigma_{yz,}\ \sigma_{zx}\ ,\sigma_{zy}\ ,\sigma_{zz,.}$

The components σ_{xx}, $\sigma_{yy,}$ σ_{zz} express the normal stress components acting on the yz, zx, and xy faces respectively. The remaining six components represent the tangential stress components (two components of each of he three pairs of the faces).

The fractional increase in volume created by deformation is called dilation. It is useful in determining some elastic constants such as the bulk modulus.

The coefficients C_{12}, C_{11} are called elastic stiffness constants and represent moduli of elasticity with dimensions of force volume. The other coefficients S_{11}, S_{12} are called elastic compliance constants and have dimensions of area or volume.

The expression for the energy of a stretched spring, the elastic energy density φ is a quadratic function of strains in the approximation of Hooke's law

$\phi = \frac{1}{2}\sum_{\mu=1}^{6}\sum_{V=1}^{6}\overline{C}_{\mu v}e_{\mu}e_{v}$, where the indices 1 to 6 should read as

1 ≡ xx, 2 ≡ yy, 3 ≡ zz, 4 ≡ yz, 5 ≡ zx, 6 ≡ xy

In accordance with the Neumann's principle, the number of independent elastic stiffness constants decreases as the symmetry of a crystal increases. This number is 21, 13, 5 and 3, respectively for triclinic, monoclinic, hexagonal, and cubic systems. The cubic crystals being the most symmetric have the least number of independent elastic stiff constants. The three independent elastic stiffness constants for a cubic crystal are C_{11}, C_{12}, C_{44}.

The velocity of a transverse wave in the [100] direction in a cubic crystal is

$v_s = \frac{\omega}{k} = \left(\frac{C_{44}}{\rho}\right)^{1/2}$. For a wave in [110] direction for which $k_x = k_y = k/\sqrt{2}$,

we get two dispersion relations

$\omega^2 = \left[\frac{C_{11}+C_{12}+2C_{44}}{2\rho}\right]k^2; \omega^2 = \left[\frac{C_{11}-C_{12}}{2\rho}\right]k^2$ These two roots refer to two different types of waves. The first root belongs to a longitudinal wave and the second root refers to a shear wave

7.6 Keywords:

Elastic constant – Elastic stress – Elastic strain – Dilation – Elastic Compliance – Stiffness Constants – Elastic energy density – Bulk modulus and Compressibility – Elastic waves.

7.7 Review questions:

1. Define elastic constants of a crystal and obtain respective relationship in between them.
2. Define dilation, elastic compliance and stiffness constants and elastic energy density, how these are applicable to cubic crystals.
3. Define bulk modulus and compressibility. What do you understand by elastic waves in cubic crystals?. How do these waves propagate along [100] and [110] directions?
4. Show that the longitudinal and shear wave velocities in the [111] direction in a cube are respectively, given by

$$v_1 = \left[\frac{C_{11}+2C_{12}+4C_{44}}{3\rho}\right]^{1/2}; \quad v_2 = \left[\frac{C_{11}-2C_{12}+4C_{44}}{3\rho}\right]^{1/2}$$

5. Show that in a cubic crystal, the effective elastic constant for a shear across the (110) plane in the $[\bar{1}11]$ direction is equal to $\frac{(C_{11}-C_{12})}{2}$
6. Discuss the solution for a longitudinal wave in a [110] direction and longitudinal shear wave in a [110] direction.
7. A cubic crystal is subjected to tension in the [100] direction. Find the expression for Young's modulus and Poisson's ratio in terms of elastic compliances or stiffnesses.
8. Show that in a cubic crystal the condition for a longitudinal wave in the [111] direction to have the same velocity as a longitudinal velocity in the [110] direction is that $C_{11}- C_{12} = 2C_{44}$.

9. Give symmetry arguments to show that there are three independent elastic constants for a cubic system.

10. For a cubic crystal, obtain an expression for velocity of elastic waves in (100) and (110) directions. Express Bulk Modulus and compressibility in terms of elastic constants.

11. What are the stiffness and compliance constant so that the stress and strain are symmetric tensors of the second rank. Describe an experiment to determine the elastic constants of a cubic crystal.

12. Set of the equation of motion for propagation of elastic waves in cubic crystals. Derive the relation between the velocity of propagation and elastic constants for (100) and (110) directions.

13. Discuss the analysis of elastic strain and stress in crystals. How are the elastic coefficient determined experimentally.

7.8 Text and reference books:

1. Elements of Solid State Physics by J.P. Srivastava (PHI)
2. Solid State Physics by M.A. Wahab (Narosa)
3. Elements of Solid State Physics by A. Omar (Pearson education)
4. Solid State Physics by S.O. Pillai (New Age)
5. Solid State Physics by C.Kittel (Asia Publishing house)
6. A Text Book of Solid State Physics by S.L. Kakani and C. Hemrajani (S.Chand)
7. Fundamentals of Solid State Physics by Saxena Gupta Saxena (Pragati Prakashan)

UNIT – II **LESSON: 8**

Lattice vibrations

Aim: To know about the theory of lattice vibrations.

Objectives of the lesson:

- To know about the introductory remarks about lattice vibrations.
- To know the balls and springs model of harmonic crystals.
- To look into one-dimensional atomic chain under normal model.
- To determine the periodic values of K through periodic boundary condition.
- Extension of normal model to one-dimensional diatomic chain.
- To look into the salient features of dispersion curves.
- To know about the restrahlen frequency band.
- To highlight the general theory of harmonic approximation.
- To know about normal model of real crystals.
- Quantised concept of lattice vibrations.
- Phonon dispersion by inelastic neutron scattering.

Structure of the lesson

.

8.1. Introduction:

The basis of crystal structure is often described in terms of ions for the interpretation of the properties of solids. The valence electrons are considered to have been placed in the force field of the lattice of ions. The roles of ionic and electronic motions are crucial in the determination of the properties of solids. While some properties depend heavily on the electronic motion, several others are closely linked to the ionic dynamics lattice heat capacity, thermal expansion and hardness are some examples of properties that belong to the latter class. In this lesson small vibrations of crystalline solids in terms of normal modes (independent motions of characteristic frequency) of motion of constituent ions. In normal modes all the ions mere with well-defined amplitude and phase. A normal mode has the same amplitude in each cell but varies from one unit cell to the other across the crystal like a wave with a certain wave vector. Such a wave is called lattice wave and the vibration with which it is associated is commonly know as *lattice vibration*.

The analysis of lattice vibrations is faced with the major difficulty of finding a way to treat the motion of ions that are heavy, separately from that of light electron,. This is accomplished by working in the so called adiabatic approximation, which in this context of molecules is famous as the 'Born–Oppenheimer approximation'. The electron velocity at the Fermi level $V_f \approx 10^6$ m s^{-1} where as typical ionic velocities are at the most of the order of 10^3 m s^{-1}. This simplifies the calculation of potential energy of ions.

8.2 The balls and springs model of a harmonic crystal:

Previously we observed that the equilibrium interionic separation in a crystal is determined by the balance between the attractive forces (large separations) and the repulsive force (small separation). When thermal agitations displace ions from their equilibrium points each of them experiences a net force in the form of restoring force. This force tends to bring the ions back to their equilibrium positions and accounts for the

elastic property of solids. In a simple mechanical, model we visualize a crystal as a three-dimensional periodic array of balls each of which is connected to its neighbours by massless springs. The balls and springs are respectively the representation for the ion and the coupling bonds of the real crystal. The force on an ion in the crystal is calculated by drawing an analogy between the crystal and its mechanical model. For small contractions and extensions of springs the Hooke's law is applicable to the mechanical model.

Each ion inside a crystal moves in a potential well of force field of its neighboring ions. For small displacements or small deviations from the equilibrium value of separation of ions, the potential energy curve is parabolic since this form of potential energy is a well-known attribute of simple harmonic motion, the ionic motion involving small displacements may be treated as simple harmonic in nature.

Some of the rough estimates made below demonstrate that the harmonic approximation is generally acceptable in solids. When the separation of ions changes by δa from the equilibrium value, the change in potential energy of the ion is written as

$$\delta U = \frac{1}{2} f(\delta a)^2$$

where f is called force constant or the spring constant associated with the massless spring joining the two ions viewed as the harmonic oscillator.

For δa=1Å, it gives the value of δU as 1 eV, yielding f = 30 N m^{-1}. Using this value of for a simple cubic crystal, we can readily deduce the macroscopic elastic modulus E for extension along one edge of the crystal. Thus, we have

$$E = \frac{f(\delta a)/a^2}{(\delta a)/a} = \frac{f}{a} = 10^{11} \text{ N m}^{-2} \qquad (\text{for } a = 3 \text{ Å})$$

This value of E is typically of the order of Young's modulus identified alternatively as a stiffness constant. By ignoring the fact that an ion is coupled to several other ions, let us calculate the frequency of vibration of an ion attached to a spring whose other end is rigidly fixed. We find that the order of magnitude of frequency for f = 30 N m^{-1} refers to the IR region which indeed represents the region of vibration frequencies of solids.

Another important aspect is the propagation of long waves in solids. To these waves such as sound waves, a crystal behaves as an elastic continuum. It enables us to estimate the velocity v_s of these acoustic waves as

$$v_s = \sqrt{\frac{E}{\rho}} = a\left(\frac{f}{M}\right)^{1/2}$$

Where ρ is the density M/a^3, of a simple cubic crystal with M as the ionic mass.

The above relation gives $v_s = 10^4$ ms^{-1}, which is a correct order of magnitude; most of the solids have v_s in the range 1–10 Km s^{-1}. This calculation, though rough, also underlines the importance of stiffness constants in the continuum approach to wave propagation in solids. In a commonly employed technique for measuring stiffness constants, an ultrasonic pulse is sent through the crystal. The ultrasonic waves whose wavelength (~μm) measures several thousand times the atomic spacing treat the crystal as a continuum.

8.3 Normal modes of a one-dimensional Monoatomic chain:

Let us consider the simplest crystal which can be a linear chain of N identical atoms (Fig 8.1). This is equivalent to a linear chain of N primitive cells. With one atom in each of them. If we wish to describe the vibrations of the chain we are confronted with the problem of accounting approximately the motion of the ions in the middle and the motion of two ions at the ends. The broad feature of the motion may still be obtained by considering only the nearest neighbour interactions and ignoring the ends. The results of such a calculation are most acceptable when the number of atoms is large. If we denote the displacement of an ion at any moment from its site l in the static lattice by s_l, the effective potential energy of a chain of interionic separation, a, in the harmonic approximation is

$$U^{\text{eff}} = \sum_l \frac{1}{2} f (s_l - s_{l+a})^2 \quad \text{----------} \quad (8.1)$$

where only the nearest neighbour interactions are included.

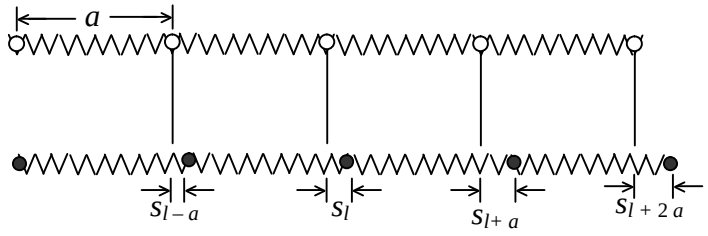


Fig 8.1 The displacement of ions in a linear chain of identical ions connected by springs.

The restoring force on the ion situated in the static lattice at the site *l* can be expressed as

$M\ddot{\mathbf{s}}_l = -f[(\mathbf{s}_l - \mathbf{s}_{l+a}) + (\mathbf{s}_l - \mathbf{s}_{l-a})]$

$= -f(2\mathbf{s}_l - \mathbf{s}_{l+a} - \mathbf{s}_{l-a})$ ---------- (8.2)

The preceding equation represents the equation of motion of a single ion in the chain. We expect from symmetry a normal mode solution in the form

$\mathbf{s}_l = \mathbf{s}_0 \exp[i(\mathbf{k} \cdot \mathbf{1} - \omega t)]$ ---------- (8.3)

which is a wave with wavevector **k** varying as exp(***i*k .x**) along the line of the ions (the x direction) with amplitude s_0.

Substituting $\mathbf{s}_l$ from (8.3) in (8.2) we get

$M\omega^2 = f[2 - \exp(ika) - \exp(-ika)]$

or

$$\omega^2 = \frac{2f}{M}(1-\cos ka) = \frac{f}{M}4\sin^2\left(\frac{1}{2}ka\right)$$

Which given the angular frequency of oscillation as

$$\omega(k) = 2\sqrt{\frac{f}{M}}\left|\sin\left(\frac{1}{2}ka\right)\right|$$

$$= \omega_m\left|\sin\left(\frac{1}{2}ka\right)\right| \quad \text{---------- (8.4)}$$

if the positive root is chosen.

The ω_m is the frequency maximum $[2\sqrt{(f/m)}]$ observed at $k = \pm \pi/a$. It must be noted that *l* does not figure in (8.4), indicating that the equation of motion of every ion gives the same algebraic relation between ω and *k*. This shows that the trial function **s** is indeed a solution of (8.2). Relation (8.4) is the required dispersion relation and the ω versus k plot derived from this is known as the *dispersion curve*.

Another important observation is that we started with equations of *N* coupled harmonic oscillators (8.2), implying that if one ion starts vibrating, it does not continue with constant amplitude, but transfers energy to others in a complicated way. Thus the vibrations of individual ions are not simple harmonic on account of this energy exchange. Our solutions, on the other hand, are uncoupled oscillations called normal modes. Each normal mode has a characteristic *k*-value with a definite ω and, therefore, the oscillations of ions in different normal modes are independent of each other.

8.3.1 The Periodic Boundary Condition:

The dispersion relation (8.4) shows that $\omega = 0$ at $k = 0$ and $\omega = \omega_m$ at $k = \pm \pi/a$. This indicates that all the normal mode frequencies lie in the range of k values from $-\pi/a$ to $+\pi/a$ defining the extent of the first Brillouin zone of a linear lattice. The complete spectrum of the normal modes is usually determined by involving the periodic boundary condition originally given by Born and Karman for the electron gas. The condition requires the two ends of the crystal to be joined. It is appreciated most easily in the case of the electronic conduction. The electrical conductivity depends on the details of electronic motion within the crystal. It implies as if an electron enters the crystal at one end as soon as it leaves the other end. The fact is equivalent to starting that the two ends of the crystal behave as being in contact. In the present case the ions at the ends of the linear chain are imagined to have been joined by an additional spring which is identical with those considered to couple the successive ions in the chain. The linear chain is thus transformed into a ring as shown in fig 8.2.

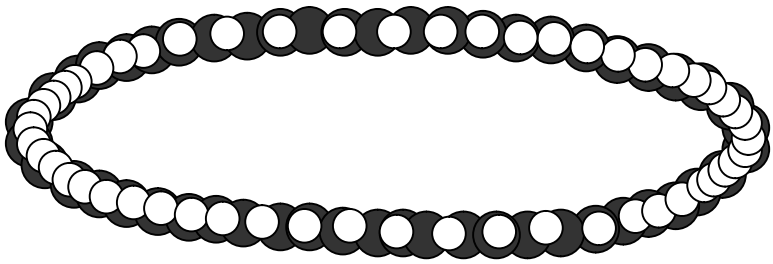


Fig 8.2: The Born – Karman periodic boundary condition for a linear chain of N identical ions.

When we move away from a certain ion along the ring and take N steps, each of length a (the interatomic spacing), we are back to the same ion. Then we require that

$\mathbf{s}_l = \mathbf{s}_{l+Na}$ ---------- (8.5)

Where $l = na$, n being an integer

Na = the length of the chain L.

Using the solution (8.3) in the condition given by (8.5), we have

$\exp(ik \,.\, Na) = 1 = \exp(i2\pi n)$ ---------- (8.6)

or $$k = \frac{2\pi}{a} \cdot \frac{n}{N} = n \cdot \frac{2\pi}{L}$$ ---------- (8.7)

Where n is an integer.

So, the allowed values of k are given by

$$k = 0, \pm \frac{2\pi}{L}, \pm 2\frac{2\pi}{L}, \ldots \quad \text{----------(8.8)}$$

From (8.3) it is evident that if *k* changes by $2\pi/a$ and or by its multiples, the solution remains unchanged. But such changes ($2\pi n/a$) in *k* denotes the reciprocal lattice vector of the one dimensional lattice under discussion, as such a root of frequency ω corresponding to a certain *k*-value repeats whenever the *k*-value is changed by a reciprocal lattice vector. This result forms one of the founding principles of the theory of solid state. The range of the first Brillouin zone is from $-\pi/a$ to $+\pi/a$ totaling $2\pi/a$ width. Therefore for every *k*-value in the second, third or any other zone, there is a *k*-value in the first zone such that the vibration frequency for the two *k*-values is common. It happens so because the points of *k*-values in the set are connected by the reciprocal lattice vectors (Fig 8.3).

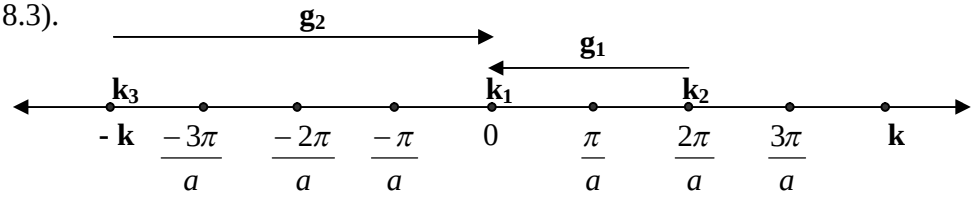


Fig 8.3 Brillouin zones of a one-dimensional crystal of lattice constant *a*. Wave vectors k_2 (extending to the second zone) and k_3 (extending to the third zone) one connected to wave vector k_1 (with in the first zone) by reciprocal lattice vectors g_1 and g_2, respectively. The form of the solution to the equation of motion is such as to yield the same value of vibration frequency for k_1, k_2 and k_3. Thus all important *k*-values (known as unique values) lie within the first zone.

This leads to a conclusion of great significance that all unique values of *k* that satisfy the solutions (8.3) lie within the first zone and we need not search for the *k*-values in other zones as they yield no new roots of ω.

Now, we are in a position to fix the last term in the set of the allowed *k*-values given by (8.8) and rewrite the set as

$$k = 0, \pm\frac{2\pi}{L}, \pm 2\frac{2\pi}{L}, \ldots\ldots, \frac{N}{2}\frac{2\pi}{L}\left(=\frac{\pi}{a}\right) \quad \text{----------} \quad (8.9)$$

The last term coincides with the boundary π/a of the first Brillouin zone. The negative sign is dropped since $-\pi/a$ is not independent on account of being connected to $+\pi/a$ by the shortest reciprocal lattice vector of magnitude $2\pi/a$. The total number of unique *k*-values in the allowed set is *N* and so will be the total number of normal modes.

8.3.2 Salient Features Of The Dispersion Curve:

The dispersion relation (8.4) and the dispersion curve displayed in fig 8.4 show only the nearest neighbour interactions. In this spirit a monatomic linear chain may be considered to act as a low pass filter. The features of the dispersion curve are further revealing in the extreme limit of the wave vector **k**. In the limit of small *k*-values, more often referred to as the long wavelength limit, sin (*ka*/2) in relation (8.4) may be replaced with *ka*/2 in the first approximation. The dispersion relation would then read as

$$\omega(k) = \sqrt{\frac{f}{M}}ka \quad \text{----------} \quad (8.10)$$

From (8.10) we infer that the frequency varies linearly with the wave vector for small vectors. This behaviour is clearly evident in fig 8.4.

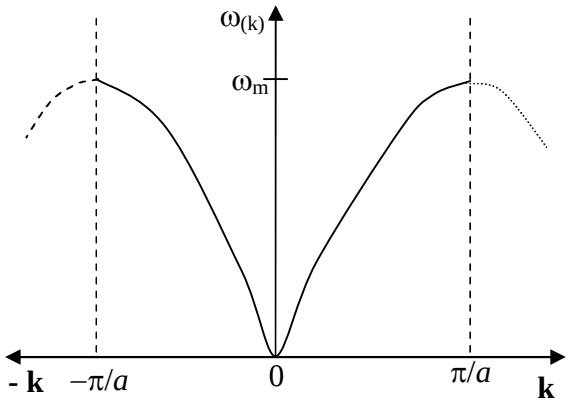


Fig 8.4 The dispersion curve for a one-dimensional monatomic chain .

The behaviour of elastic waves in continuum is of exactly similar nature. In the long wavelength limit of the waves where the wavelength is much greater than the interatomic spacing, the medium behaves as an elastic continuum since these waves pass smoothly through the medium. The chain of atoms under this condition acts like a heavy elastic string.

In the limit, the group velocity $d\omega/d\mathbf{k}$ and the phase velocity $\omega/\mathbf{k}$ of the elastic waves (or sound waves) are equal and both become independent of frequency. But as ***k*** changes to larger values, the discreteness of the medium begins to show up and at the zone boundary ($k = \pm \pi/a$), the tangent to the dispersion curve is horizontal showing there by

that the group velocity is zero here. This refers to the total dispersion and no waves propagate through the crystal indicating that it acts as a discrete medium in this situation. This sounds perfectly logical as at $k = \pm \pi/a$, the wavelength is twice the inter atomic spacing ($\lambda = 2a$) for which the crystal cannot be treated as continuous medium. The zero value of the group velocity also shows that the motions of the neighbouring atoms are out of phase and the elastic waves suffer Bragg reflection at this point in ***k***-space. This behaviour is consistent with the condition for Bragg reflection in a one-dimensional crystal. The elastic waves are no more traveling waves defined by (8.3) and get transformed into standing waves at the zone boundary.

8.4 Normal modes of one-dimensional Diatomic Chain:

This chain is different from the earlier one in the sense that there are two different types of atoms whose positions alternate along its length (Fig 8.5).

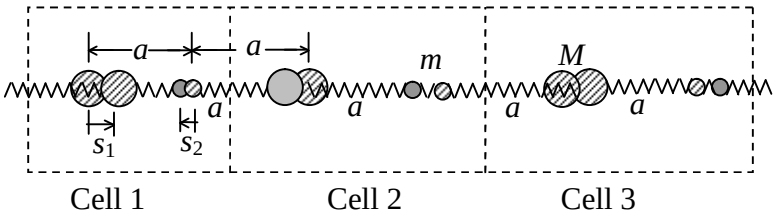


Fig 8.5: The displacement of ions in three consecutive unit cells of a one – dimensional diatomic crystal.

The chain may be viewed as a one-dimensional crystal whose primitive cell contains two atoms of different masses *M* and *m*. The relevance of treating one-dimensional atomic chains as against the reality of crystals being three-dimensional may be questioned. But these calculations bring out a few broad features of the vibrations of real crystals, as seen in the earlier section. In other words, it suffices to acknowledge that these exercises introduce us to the basics of lattice dynamics using simple mathematics.

From fig.8.5 we see that the springs are identical. Therefore, if we consider only the nearest neighbour interactions, a single force constant (say, *f*) will be involved in the equations of motion. Supposing that the heavy ion (*M*) occupies the site 1 and the light one (*m*) is at the site 2 in each primitive cell, we obtain the following equations of motion for these ions in that order:

$$M\ddot{\mathbf{s}}_{1n} = -f(2\mathbf{s}_{1n} - \mathbf{s}_{2n} - \mathbf{s}_{2,n-1}) \quad \text{----------} \quad (8.11)$$

$$m\ddot{\mathbf{s}}_{2n} = -f(2\mathbf{s}_{2n} - \mathbf{s}_{1n} - \mathbf{s}_{1,n+1}) \quad \text{----------} \quad (8.12)$$

where $\boldsymbol{s}_{1n}$ stands for the displacement of the ion at site 1 in the primitive cell n.

We seek the following solutions to the above equations:

$$\mathbf{s}_{1n} = \mathbf{u}\exp[i(kna - \omega t)]$$

and $\mathbf{s}_{2n} = \mathbf{v}\exp[i(kna - \omega t)]$ ---------- (8.13)

Substituting these solutions in (8.11) and (8.12), we have

$$(2f - M\omega^2)u - f[1 + \exp(-ika)]v = 0$$

$$-f[1 + \exp(ika)]u + (2f - m\omega^2)v = 0 \quad \text{----------} \quad (8.14)$$

where u and v (the amplitudes) are unknowns in the above homogeneous equations. The equations have solutions only if the determinant of the coefficients of u and v in them vanishes. That is,

$$\begin{vmatrix} (2f - M\omega^2) & -f[1 + \exp(-ika)] \\ -f[1 + \exp(ika)] & (2f - m\omega^2) \end{vmatrix} = 0 \quad \text{----------} \quad (8.15)$$

or

$$Mm\omega^4 - 2f(M + m)\omega^2 + 2f^2(1 - \cos ka) = 0 \quad \text{----------} \quad (8.16)$$

Based on our experience with the monatomic lattice it is advisable to solve (8.16) for small k (i.e. long wavelength limit) and for the largest k, i.e. at the first zone boundary. These describe the distinct features of the dispersion curves.

For small values of k, we have

$$\cos ka = 1 - \frac{1}{2!}(ka)^2 + ...$$

Retaining the first two terms of the series and putting this value of cos ka in (8.16), we get the following two roots:

$$\omega^2 \cong 2f\left(\frac{1}{M} + \frac{1}{m}\right) \quad \text{----------} \quad (8.17)$$

and

$$\omega^2 = \frac{\frac{1}{2}f}{M + m}k^2a^2 \quad \text{----------} \quad (8.18)$$

The dispersion curve obtained from (8.17) is called ***optical branch*** while the one from (8.18) is known as ***acoustical branch.***

For the maximum value of k, i.e. at $k = \pm\pi/a$, the roots are

$$\omega^2 = \frac{2f}{m} \qquad (\textit{optical branch}) \qquad \text{---------- (8.19)}$$

and $$\omega^2 = \frac{2f}{M} \qquad (\textit{acoustical branch}) \qquad \text{---------- (8.20)}$$

8.4.1 Salient Features Of The Dispersion Curves:

The most prominent feature of the dispersion curves shown in Fig. 8.6 is the manifestation of a frequency gap between the acoustical and optical branches. This brings out the fact that a diatomic linear chain acts as a band pass filter. Other features of the observed two branches representing two different types of normal modes are discussed below.

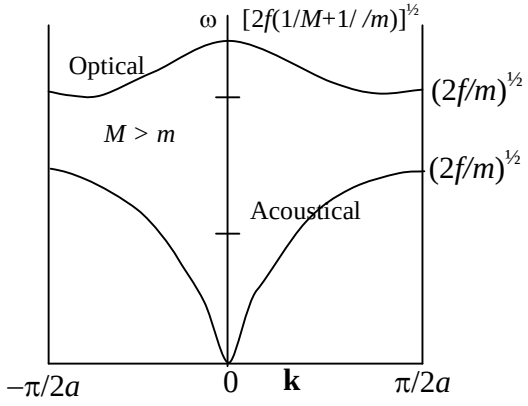


Fig 8.6 Optical and acoustical branches of dispersion curves for a one-dimensional diatomic crystal of lattice constant a. The frequency extreme for two branches are given.

The frequency of the optical branch is nearly constant in the limit $k \to 0$ as made out by the approximate nature of (8.17). But it decreases slowly as k increases, dropping to the value $2f/m$ at the zone boundary.

The acoustic branch corresponds to the single branch, obtained for the linear chain of monatomic atoms. The linear behaviour of ω with k in the limit of small k (or long wavelength) is in the limit of sound waves which are longitudinal and treat the crystal as elastic continuum.

We may, further, exploit the above treatment to derive the state of ionic motions in the two branches, again for the same two limiting cases.

Substituting ω^2 from (8.17) in (8.14), at $k = 0$, we have for the optical branch,

$$\frac{u}{v} = -\frac{m}{M}$$

or $$Mu = -mv \qquad \text{----------} \quad (8.21)$$

Relation (8.21) shows that the movements of the heavy and light ions are out of phase, i.e., they move towards each other or away from each other such that their centre of mass

remains at rest. This refers to the situation in ionic crystals where such a motion of positive and negative ions may displace the centre of positive charges with the centre of negative charges creating an electric dipole. The oscillating electric dipole on interaction with the electromagnetic radiation may absorb the radiation. From a rough calculation we can find that the frequency of vibration of ionic crystals (e.g. NaCl, KBr, LiF etc.) lies in the infrared region. That is why ionic crystals are known to absorb infrared light. This forms the basis for giving the name optical branch to the branch under discussion. The relevance of treating a diatomic linear chain is thus amply justified. The results can be only instructive as we have allowed the linear chain to produce only longitudinal waves. In real crystals, there can be two transverse waves for each longitudinal wave. Generally, the frequencies of all the modes are different with the exception that the two transverse modes along directions of high symmetry in the crystal are degenerate (Fig 8.10). The longitudinal and transverse vibrational modes of a crystal can be clearly separated only in certain symmetry directions of crystals. The modes for any arbitrary direction are mixed in character.

For acoustical branch at $k = 0$, we obtain

$$\frac{u}{v} = \frac{m}{M} \quad \text{----------} \quad (8.22)$$

which shows that the two ions move in phase.

The state of motions of the ions in the optical and acoustical branches in illustrated for a transverse wave in fig 8.7. The example of a transverse wave is chosen since the difference between the motions of ions in the two branches is more striking in appearance for the transverse motion. Though this motion is not allowed in the linear chain, it is present in all real crystal.

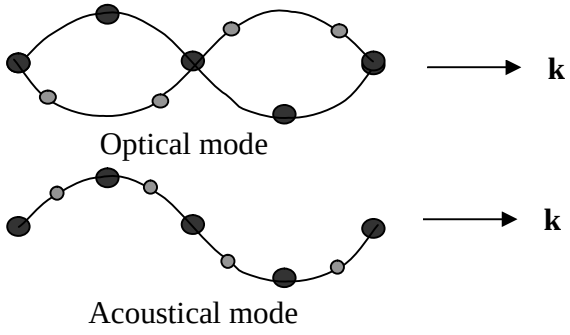


Fig 8.7 Illustration of the movement of ions in the transverse optical and acoustical modes at equal wavelength in a diatomic linear crystal.

The variation of the amplitude ratio u/v with the wavevector can be seen in Fig 8.8. The state of motion at the zone end ($k = \pm \pi/a$, the maximum k) can be interpreted with the help of this figure. The light ions are at rest in the acoustical mode while the heavy ions are at rest in the optical mode.

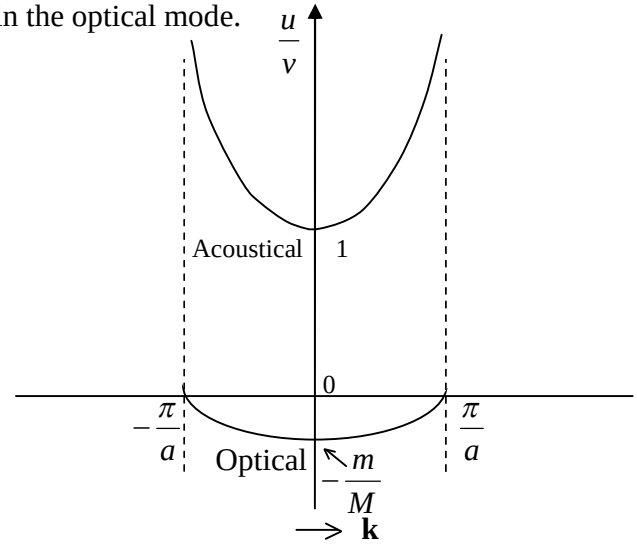


Fig 8.8 The amplitude ratio u/v as a function of the wave vector k for the acoustical branch (upper curve) and the optical branch (lower curve), u belongs to *M* and *v* to *m* (*M* > *m*).

Another significant feature of the dispersion curves is that the gap at the zone end decreases with the decrease in the mass of heavy ion and approaches zero as $M \rightarrow m$. But it would surprise us to know that we have still two different branches though $M = m$ which should give only the acoustical branch. For the present chain, if we have monatomic basis, we need to correct the size of the first zone which will now be double in size owing to the interionic spacing or the lattice constant being halved $a/2$. This ensures that the acoustical branch is continuous over whole region of the modified zone ($-2\pi/a$ to $2\pi/a$). The illusory optical branch of $M = m$ in the original zone gets effectively reflected on to the regions added to the original zone.

8.5 The Reststrahlen Band:

In the previous section to treat an extremely interesting phenomenon observed in the infrared absorption spectra of ionic and partly ionic crystals. These crystals show very intense reflection of infrared radiation over a small range of frequencies. The range of frequencies over which this reflection occurs is termed ‘reststrahlen band’. ‘Reststrahlen’

is a German word which means residual rays. This underlines the fact that these crystals show strong reflection to a radiation of certain frequency, for which they also act as strong absorbers. NaCl and GaAs are good examples of such crystals.

It was mentioned a little earlier how an optical mode in an ionic crystal is excited when an electromagnetic radiation is incident on its surface. The transverse electric field of the radiation does the trick by exerting force on cations and anions in opposite directions. The strong absorption (resonance) takes place when the frequency of radiation matches the frequency of a transverse optical mode of the crystal. A proper explanation to the reststrahlen phenomenon follows from the theory of optical constants. According to this theory, when

$\omega_0 < \omega < \varepsilon^{1/2}\omega_0$ ---------- (8.23)

no wave can propagate through the crystals.

Here ω_0 is the frequency of a transverse optical mode, ω the frequency of the incident radiation and ε is the dielectric constant of the crystal. In the above condition we have an evanescent wave that decays exponentially with the increasing distance in the crystal therefore, in the specified range of frequencies the radiation incident on the crystal from outside suffers total external reflection. This is what we know as the reststrahlen phenomenon. The range of frequencies over which this occurs is called the *reststrahlen band*.

Now we make an estimate of ω_0, the frequency at which the strong absorption occurs using equations of motion set up in earlier for a diatomic lattice. This also gives the measure of the reststrahlen frequency which is the same as ω_0. For NaCl crystal, the lattice constant a is equal to 5.63 Å. With $\lambda = 2a$, we get the maximum cut-off wavevector $k_m = 2\pi/\lambda \approx 10^{10}$ m^{-1}. The wavevector associated with a typical infrared radiation of wavelength 10,000Å is about 10^6m^{-1}. Therefore, the vibration with these relatively small wavevectors in the optical branch can be determined in the limit $k \to 0$. In this limit, relations (8.14) reduce to

$-M\omega^2 u = 2f(v - u)$

$-m\omega^2 v = -2f(v - u)$ ---------- (8.24)

Relations (8.24) essentially give the force on masses M and m, respectively. When an electromagnetic radiation with the transverse electric field $E = E_0 \exp(i\omega t)$ is incident, these force terms get corrected by $\mp eE_0$. Then relations (8.24) reduce to

$$-M\omega^2 u = 2f(v-u) - eE_0$$

$$-m\omega^2 v = -2f(v-u) + eE_0 \quad \text{----------} \quad (8.25)$$

Relations (8.25) give the two amplitudes of vibration as

$$u = \frac{-\frac{eE_0}{M}}{\omega_0^2 - \omega^2} \quad \text{----------} \quad (8.26)$$

$$v = \frac{\frac{eE_0}{m}}{\omega_0^2 - \omega^2}$$

with

$$\omega_0^2 = 2f\left(\frac{1}{M} + \frac{1}{m}\right) \quad \text{----------} \quad (8.27)$$

Though relations (8.26) indicate a single frequency at which the maximum absorption should occur, experiments show that there is a small range of frequencies of the electromagnetic radiation showing the reststrahlen phenomenon. This small range of frequencies forms the reststrahlen band. The frequency ω_0 is referred to as the ***reststrahlen frequency***.

8.6 General Theory of Harmonic Approximation:

Let the position of a lattice site in the static crystal be represented by **l.** Assume that the average position of the ion at the site in question is still given by **l** when the ion vibrates. Thus the symmetry of the Bravais lattice is supposed to remain unchanged in a vibrating crystal. Further, it is assumed that the displacement or deviation of the ion from its equilibrium position is much small in comparison with the interatomic distance. We denote the displacement of the ion by s_l and its displaced position by r_l (fig 8.9) having the relation

$$\mathbf{r}_l = \mathbf{l} + \mathbf{s}_l \quad \text{----------} \quad (8.28)$$

Small values of the atomic displacements allow us to expand the total potential energy about its equilibrium value in a Taylor's series. This makes the exercise of obtaining the equation of motion quite straightforward, which is otherwise extremely tedious.

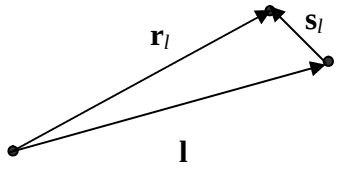


Fig 8.9 An ion at site l of the static lattice being displaced to position $\mathbf{r}_l$ during a vibration. The displacement $\mathbf{s}_l$ from site l is very small.

The kinetic energy T of the crystal can be written as

$$T = \sum_n \frac{1}{2} M_l \left|\dot{\mathbf{s}}_{\ln}\right|^2 \quad \text{----------} \quad (8.29)$$

where M_l is the mass of the ion at site l and n stands for the unit cell running through all the unit cells.

And the potential energy U is defined by its Taylor's expansion as

$$U = U_0 + \sum_{l\,nj} \mathbf{s}_{l\,nj} \left(\frac{\partial U}{\partial \mathbf{s}_{l\,nj}}\right)_0 + \frac{1}{2} \sum_{ll',nn',jj'} \mathbf{s}_{l\,nj}\, \mathbf{s}_{l'n'j'} \left[\frac{\partial^2 U}{\partial \mathbf{s}_{l\,nj} \cdot \partial \mathbf{s}_{l'n'j'}}\right]_0 + \ldots\ldots \quad \text{----------} \quad (8.30)$$

where index j denotes the three Cartesian x, y, z components.

The first term U_0 refers to the potential energy at equilibrium which is a constant. It can be dropped from the expression as it is the changes in potential energy from U_0 that determine the vibration spectrum of a crystal. The coefficient of $\mathbf{s}_{lnj}$ in the linear term is equal to the magnitude of the net force on the ion at site l exerted by all the other ions when each of the ions is at its equilibrium position. At equilibrium, this force vanishes and so should the linear term. The terms of higher orders beyond the third term are not considered in the harmonic approximation. Therefore, the lone non-vanishing third term in the expansion expresses the effective potential energy of the crystal in the harmonic approximation. That is,

$$U^{eff} = \frac{1}{2} \sum_{ll',nn',jj'} \mathbf{s}_{l\,nj}\, \mathbf{s}_{l'n'j'} \left[\frac{\partial^2 U}{\partial \mathbf{s}_{l\,nj} \cdot \partial \mathbf{s}_{l'n'j'}}\right]_0 \quad \text{----------} \quad (8.31)$$

We combine (8.29) and (8.31) to write the usual Lagrange's equation from the equations of motion for the Cartesian components obtained in the form

$$M_l\, \ddot{\mathbf{s}}_{l\,nj} = \sum_{l',n',j'} \left[\frac{\partial^2 U}{\partial \mathbf{s}_{l\,nj} \cdot \partial \mathbf{s}_{l'n'j'}}\right] \mathbf{s}_{l'n'j'} \quad \text{----------} \quad (8.32)$$

Relation (8.32) represents the *j*th component (which could be the x-, y- or z-component) of the net force on the ion at site *l* in the unit cell *n* owing to the rest of the atoms in the crystal. This equation marks the beginning of all theories of lattice dynamics.

In view of the elastic nature of crystals, the coupling between any two atoms is customarily represented in the material form of a spring. Each spring in a crystal is assigned a constant, whose value is characteristic of the nature of the ions being coupled. In classical mechanics, this constant is called by the name *spring constant* or *force constant.*

A comparison of (8.32) with the equation of motion of a particle executing simple harmonic motion shows that the derivatives in (8.32) serve as force constants in these equations of motion. For convenience, we will use from now onwards a shorter notation for these derivatives given by

$$\left(\frac{\partial^2 U}{\partial \mathbf{s}_{l\,nj} \cdot \partial \mathbf{s}_{l'n'j'}}\right) = U_{l\,nj}^{l'n'j'} \quad \text{----------} \quad (8.33)$$

The equation of motion (8.32) then reads as

$$M_l \ddot{\mathbf{s}}_{l\,nj} = -\sum_{l'n'j'} U_{l\,nj}^{l'n'j'} \mathbf{s}_{l'n'j'} \quad \text{----------} \quad (8.34)$$

Equation (8.34) is the general equation of motion for ions in crystals. It is to be observed that the equations of motion for the linear chains of monatomic basis (8.2) and diatomic basis [(8.11) and (8.12)] are only the special cases of this equation with V_{lnj} having the dimensions of force constants. They denote the generalized force constants of a system with many degrees of freedom. The isotropy of space, the translational invariance and the point group symmetry require the coupling constants to satisfy certain conditions.

It must be noted that each term within the summation in (8.32) denotes a force that depends on the relative position of unit cells *n* and *n′* and not on their absolute position. It is the consequence of the translation invariance which effectively requires that

$$U_{l\,nj}^{l'n'j'} = U_{loj}^{l'(n'-n)j'} \quad \text{----------} \quad (8.35)$$

We have, now, come to grips with the basic workable technique of solving the equations of motion in crystals.

8.7 Normal Modes Of Real Crystals:

This exercise concerns real crystals which are three-dimensional and in general may have a polyatomic basis having atoms of different elements. The problem is most general and, therefore, the task of finding its solution is bound to be tedious. But, in principle, the solution is tractable. We give below the procedure for calculating the frequencies of normal modes in the framework of a theory based on the harmonic approximation.

The equation of motion of an ion, whose site in the unit cell n is denoted by the position vector l_n, is given by (8.34). If the crystal is composed of N unit cells with p atoms in a unit cell, we get in total $3pN$ equation analogous to (8.34). It is proper to write the solutions or the displacement s_{lnj} in terms of a plane wave with respect to the cell coordinates. That is,

$$\mathbf{s}_{l\,nj} = \frac{1}{\sqrt{M_1}}\mathbf{u}_j(\mathbf{k})\exp[i(\mathbf{k}\cdot\mathbf{l}_n - \omega t] \quad \text{----------} \quad (8.36)$$

This plane wave is defined only at the lattice sites, unlike a normal plane wave.

Using this solution for (8.34), we have

$$-\omega^2\mathbf{u}_j(\mathbf{k}) + \sum_{l'j'}\sum_{n'}\frac{1}{\sqrt{M_l M_{l'}}}U_{l\,nj}^{l'n'j'}\exp[i\mathbf{k}\cdot(\mathbf{l'}_{n'}-\mathbf{l}_n)]\cdot\mathbf{u}_{l'j'}(\mathbf{k}) = 0 \quad \text{----------} \quad (8.37)$$

Let us write (8.37) as

$$-\omega^2\mathbf{u}_j(\mathbf{k}) + \sum_{l'j'} D_{lj}^{l'j'}(\mathbf{k})\,\mathbf{u}_{l'j'}(\mathbf{k}) = 0 \quad \text{----------} \quad (8.38)$$

Where

$$D_{lj}^{l'j'}(\mathbf{k}) = \sum_{n'}\frac{1}{\sqrt{M_l M_{l'}}}U_{l\,nj}^{l'n'j'}\exp[i\mathbf{k}\cdot(\mathbf{l'}_{n'}-\mathbf{l}_n)] \quad \text{----------} \quad (8.39)$$

According to (8.35), the terms in the sum of the above equations depend only on the difference ($n' - n$) and not on the absolute values of n and n'. The quantity $D_{lj}^{l'j'}(\mathbf{k})$, which is obtained by summing over n', is independent of n. It is instructive to note that it couples amplitudes with each other without having to depend on n. This also explains why the amplitudes in (8.13) appear without index n. A very common term in lattice dynamics known as the dynamical matrix is formed by the quantities defined by (8.39). The set of equations (8.38) belongs to a set of linear homogeneous set of order $3p$. The set of linear homogeneous equations has a solution only if the,

Determinant: $\{ D_{lj}^{l'j'}(\mathbf{k}) - \omega^2\}$ vanishes. ---------- (8.40)

The above equation gives $3p$ different solutions, i.e., for one value of **k** there are $3p$ values of ω, each of which lies on a separate branch. A branch is characterized by its dispersion relation. Thus, there are in total $3p$ branches out of which three are acoustical – one longitudinal (LA) and two transverse (TA). The number of acoustical branches is three in all crystals as it does not depend on the number of atoms. The rest $(3p - 3)$ belong to the optical branches; $(p - 1)$ longitudinal (LO) and $2(p - 1)$ transverse (TO). Since there are N (the number of unit cells) unique values of **k** for each branch, the total number of vibrational modes comes to $3pN$.

A complete description of the normal modes of a crystal requires the knowledge of directions along which the longitudinal and transverse waves move for a certain value of **k**. These directions are known as the directions of polarization. In an isotropic crystal, the three solutions can be manipulated such that the longitudinal mode is polarized along the direction of propagation (parallel to **k**) and the two transverse modes along directions perpendicular to **k**. The picture in anisotropic crystals is not so straightforward. We can describe the dispersion curves of these crystals in terms of longitudinal and transverse branches, only if **k** lies along any n-fold axis of rotation permitted by the crystal symmetry. In such a case the longitudinal mode is polarized along **k**, and transverse modes, that are degenerate, along directions perpendicular to **k**.

It can now well be imagined that the picture of dispersion curves for crystals with polyatomic basis will be fairly complex. We take a relatively simple example of diamond whose primitive cell has two identical atoms. The dispersion curves plotted from the experimental data are shown in fig 8.10 and refer to the propagation of waves along two important directions [100] and [111]. The wavevector is expressed as a dimensionless quantity, $q = k(2\pi/a)$, called the reduced wavevector as measured from the centre of the Brillouin zone. In principle, there should be six branches in total. But we observe only four, as the transverse modes along [100] and [111] directions are degenerate in the acoustical and the optical branches, separately. The other important point exhibited by the curves is that the LA and LO modes at the zone boundary are degenerate. This confirms the zero gap between the acoustical and the optical branches at this point for the linear chain of diatomic lattice composed of identical atoms.

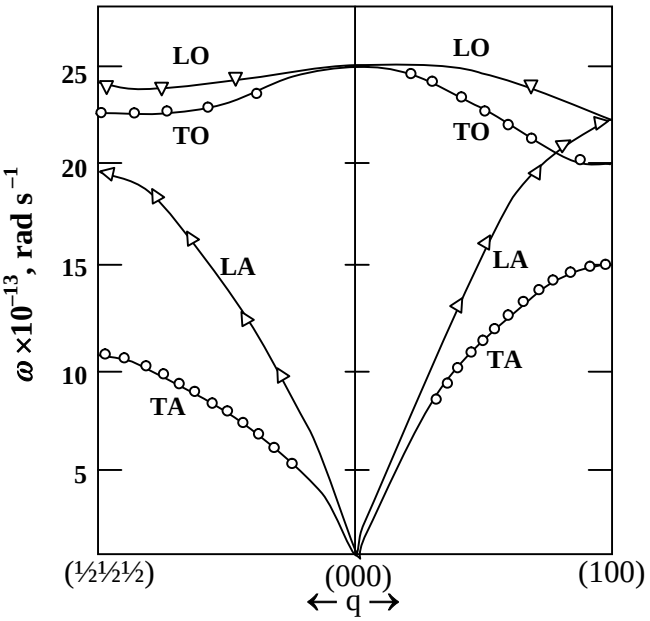


Fig 8.10 Dispersion curve for diamond crystal

8.8 Quantization of Lattice Vibrations:

The energy of the lattice waves is quantized the same way as that of the electromagnetic waves. In a harmonic crystal, the atomic oscillators are treated analogous to the Plank's radiation oscillators. The energy of a vibrational mode of angular frequency ω with wavevector **k** in the branch s is expressed as $\left(n_{ks}+\frac{1}{2}\right)\hbar\omega_s(\mathbf{k})$, where n_{ks} is an integer denoting the excitation number or the order of the excitation state of the classical normal mode. The fact of the normal mode being in its n_{ks} excited state is expressed in the language of quantum theory by saying that there are n_{ks} phonons of wavevector **k** in branch *s*. The usage of the term phonon is analogous to the term photon for the electromagnetic radiation. Phonon is the corpuscular representation for a quantum of vibration of energy $\hbar\omega_s(\mathbf{k})$ carried by a sound wave in the same way as photon represents a quantum of radiation in an electromagnetic wave. The n_{ks} is also defined as the phonon occupancy expressed by the Plank's distribution function which is a function of $\omega_s(\mathbf{k})$ and **k**.

The thermal energy of a harmonic crystal is given by

$$E=\sum_{\mathbf{k}\,s}\left(n_{ks}+\frac{1}{2}\right)\hbar\omega_s(\mathbf{k}) \qquad \text{---------- (8.41)}$$

Relation (8.41) shows that the energy of an oscillator is not zero even in the lowest vibration state (n = 0) and has a value $\frac{1}{2}\hbar\omega_s(\mathbf{k})$. It indicates that even the lowest state is not vibrationless. This is known as zero point motion and finds its interpretation only with the use of quantum mechanics. The quantity $\frac{1}{2}\hbar\omega_s(\mathbf{k})$ is called the zero point energy.

8.9 Measurement of Phonon Dispersion by Inelastic Neutron Scattering:

The experiment performed to obtain the phonon dispersion curves are based on the exchange of energy between lattice vibrations and a probe. In principle, X-rays or thermal neutrons can be used as the probe in these experiments. The energy of thermal neutrons is ~0.025 eV, which is of the order of their average thermal energy (k_BT) at room temperature. Because of small energy, neutrons show an appreciable change in their energy on being scattered while exchanging the energy of a vibrational quantum. This change in energy is easily measurable on the respective energy scale. On the other hand, the energy of X-rays is much high (~keV). The energy change for such an energetic photon is only 1 part in 10^6 which is difficult to measure on the scale of the photon energy. On account of this reason, thermal neutrons, also known as slow neutrons, are preferred for the measurement of phonon spectra. The assembly of apparatus used for this purpose is called a triple-axis spectrometer. Its schematic diagram is shown in Fig.8.11. The experiment is generally performed around a nuclear reactor which acts as the source of thermal neutrons. The term slow neutrons used for these neutrons is only relative on the energy scale as the typical order of their velocity is 10^5 cms^{-1}, which is large compared to the motion of material objects in our daily life.

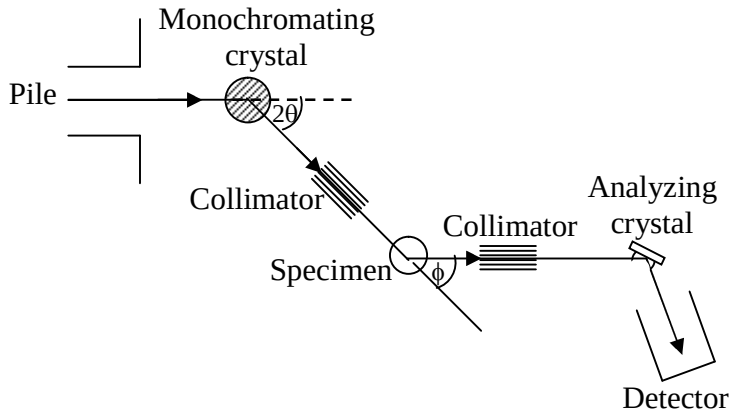


Fig 8.11 Outlines of the experimental set-up for measuring phonon spectra using a triple-axis neutron spectrometer.

The information from the inelastically scattered neutrons is extracted mainly by exploiting the consequences of the momentum and the energy conservation laws. If a neutron of a wavevector **k** is scattered to a state of **k′** and a phonon of wavevector $\mathbf{k}_s$ is created or destroyed in the process, the momentum conservation requires that

$\mathbf{k} - \mathbf{k}' = \pm\mathbf{k}_s + \mathbf{g}$ ---------- (8.42)

Where (+) and (–) signs refer respectively to the creation and destruction of a phonon of energy $\hbar\omega_s(\mathbf{k})$. The symbol *s* stands for the branch.

The reciprocal lattice vector **g** appears for the reason that phonons are plane waves that are modified for a crystalline medium on the ground of its periodic character. The value of $\mathbf{k}_s$ in (8.42) lies within the first Brillouin zone.

According to the principle of energy conservation,

$$\frac{\hbar^2\mathbf{k}^2}{2m_n} = \frac{\hbar^2\mathbf{k}'^2}{2m_n} + \hbar\omega_s(\mathbf{k}) \quad \text{----------} \quad (8.43)$$

Where m_n is the neutron mass.

The wavevector of the scattered neutron **k′** is measured experimentally. For a given incident neutron **k,** the wavevector of phonon $\mathbf{k}_s$ absorbed or emitted and its frequency $\omega_s(\mathbf{k})$ can be determined. This is based on finding the energy gain or loss of the scattered neutrons as a function of the scattering direction (**k – k′**).

8.10 Summary:

- Ionic velocities in crystals are at least a thousand times smaller than the electron velocity at the Fermi surface. Therefore, the electrons can be assumed to remain in their ground state for an ionic configuration.
- For small displacements from the equilibrium positions, the ions in crystals may be treated as harmonic oscillators.
- All lattice vibrations (elastic waves) can be described by wavevectors that lie within the first Brillouin zone.
- A one-dimensional monatomic chain acts as a low-pass filter when only the nearest neighbour interactions are considered.
- A one-dimensional diatomic chain acts as a band-pass filter by virtue of a frequency gap occurring between the acoustical and optical branches.

- The reststrahlen frequency for an ionic crystal containing ions of masses *M* and *m* is $\omega_0 = 2f\left(\frac{1}{M}+\frac{1}{m}\right)$ where *f* denotes the force constant.
- For a crystal with *p* atoms in the primitive cell:
  Number of acoustical branches = 3 (independent of *p*) = 1(LA) +2(TA)
  Number of optical branches = 3(*p* – 1) = (*p* – 1) ….(LO) + 2(*p* – 1)…. (TO)
- A quantized crystal vibration or the quantum unit of a crystal vibration is called a phonon. Its energy is given by hω, where ω is the angular frequency of a crystal vibration.
- The thermal energy of a harmonic crystal is given by

  $$E = \sum_{\mathbf{k}\,s}\left(n_{ks}+\frac{1}{2}\right)\hbar\omega_s(\mathbf{k})$$

  where n_{ks} denotes the number of phonons with frequency ω and wavevector **k** in the branch *s*; and $\frac{1}{2}\,\hbar\omega_s(\mathbf{k})$ is the zero point energy.
- When in a crystal a neutron of wavevector **k** is inelastically scattered to a state of wavevector **k′** and a phonon of wavevector $\mathbf{k_s}$ is created or destroyed in the process, the momentum conservation requires that

  $$\mathbf{k} - \mathbf{k}' = \pm\mathbf{k}_s + \mathbf{g}$$

  where (+) and (–) signs refer respectively to the creation and destruction of a phonon of energy $\hbar\omega_s(\mathbf{k})$; and **g** is a reciprocal lattice vector.

8.11 key words:

Lattice vibrations – Ionic motion – Electronic motion – Balls and springs model – Force constant – Lattice constant – Normal modes dispersion curve – Optical branch – Acoustical branch – Restrahlen frequency.

8.12 Review questions:

1. Describe the balls and springs model of a harmonic crystal write the significance normal modes of one dimentional monatomic chain.
2. What do you infer from periodic boundary condition. Sketch and illustrate dispersion curve for a one dimensional monatiomic chain.

3. What is explained in normal modes of one-dimensional diatomic chain and illustrate the salient features of the dispersion curve.
4. Explain the resthlen band / frequency and what is to be inferred from general theory of harmonic approximations. Illustrate the dispersion curve for diamond crystal.
5. Write briefly about quantization of lattice vibration and also measurement of phonon dispersion by inelastic neutron scattering.

8.13 Text and reference books:

1. Elements of Solid State Physics by J.P.Srivastava (PHI)
2. Solid State Physics by M.A. Wahab (Narosa)
3. Elements of solid State Physics by A. Omar (Pearson education)
4. Solid State Physics by S.O. Pillai (New Age)
5. Solid State Physics by C.Kittel (Asia Publishing house)
6. A Text Book of Solid State Physics by S.L.Kakani and C.Hemrajani (S.Chand)
7. Fundamentals of Solid State Physics by Saxena Gupta Saxena (Pragati Prakashan)

UNIT – III **LESSON: 9**

THERMAL PROPERTIES OF SOLIDS - I

Aim: To learn about the thermal properties of Solids.

Objectives of the lesson:

- To know about the outstanding properties of solids namely high electrical conductivity and thermal conductivity, metals obeying Ohms law, Wiedemann-Frantz ratio etc.
- To know about the free electron model
- To derive an expression for the density of electronic states.
- To understand the effect of temperature on the parameters of free electrons gas.

Structure:

9.1 Introduction:

The outstanding properties of metals are

(1) **High electrical and thermal conductivities**,

(2) **Obeying Ohm's law:** Metallic conductors obey Ohm's law, which states that current density ($\bar{j}$) in the steady state is proportional to electric field strength ($\overline{E}$) i.e., $\bar{j} = \sigma \overline{E}$ where σ the electrical conductivity is independent of $\bar{j}$ or $\overline{E}$.

(3) **Temperature dependence of conductivity:** Resistivity of metals (ρ) above Debye temperature is proportional to the absolute temperature i.e. $\rho \propto T$ and at low temperatures the resistivity of metals is proportional to the fifth power of absolute temperature i.e. $\rho \propto T^5$.

(4) **Wiedemann - Frantz ratio:** The ratio of thermal to the electrical conductivity $\frac{k}{\sigma}$ has the same value for nearly all metals at a given temperature and $\left(\frac{1}{T}\ \frac{k}{\sigma}\right)$ has the same value at any temperature.

The idea that the large electrical and thermal conductivity of metallic substances might be explained by the presence of large concentration of mobile free electrons in the metals was first proposed by Drude in 1900. The implications of this hypothesis were exhaustively investigated subsequently by Lorentz. Drude and Lorentz attempted an explanation of these properties on the basis of the assumption that these free electrons are capable of moving through the lattice there by suffering collisions with the atoms. As such the free electrons in metals as an ideal gas of free particles, which when in thermal equilibrium would obey Maxwell-Boltzmann statistics. One of the greatest achievements of the theory was that it led to semi quantitative agreement with the Wiedemann-Frantz law.

The free electron theory in its simplest form, however, led to a prediction of the electronic component of specific heat, which was in serious disagreement with experimental results. Another difficulty encountered in classical theory pertains to the magnetic properties of free electrons. Each electron has a magnetic moment associated with its spin and classically should therefore give rise to a paramagnetic susceptibility

inversely proportional to the temperature. Experimental results on the other hand show that paramagnetism of metals is nearly independent of temperature.

Further another difficulty is: Using the classical theory we cannot understand the occurance of long electronic mean free paths. From many types of experiments it is abundantly clear that a conduction electron in metal can move freely in a straight path over many atomic distances undeflected by collisions with other conduction electrons or by collisions with the atom cores. In a very pure specimen at low temperatures the mean free path may be as long as 10^8 or 10^9 interatomic spacings (more than 1 cm), vastly longer than we would expect from the known sizes of atoms. We must ask why condensed metallic matter is so transparent to conduction electrons. The conduction electrons act as a gas of non-interacting particles. To explain this we have to invoke quantum concepts and classical statistics is not valid.

There are two parts to the answer to our question.

(a) A conduction electron is not deflected by ion cores arranged on a periodic lattice because matter waves propagate freely in a periodic structure.

(b) A conduction electron is scattered only infrequently by other conduction electrons. This property is a consequence of the Pauli's exclusion principle. These difficulties were resolved by Sommerfield using Fermi-Dirac statistics rather than classical Boltzmann statistics. The Fermi – Dirac free electron picture serves as a very simple and conceptually quite direct way of discussing and visualizing transport effects in metals.

9.2 Free Electron Model:

In the free electron model, the conduction electrons are assumed to be free, except for a potential at the surface (fig 9.1), which has effect of confining the electrons to the interior of the specimen. According to this model, the conduction electrons move about inside the specimen without collisions except for an occasional reflection from the surface much like the molecules in ideal gas. The forces between the conduction electrons and ion cores are neglected in the free electron approximation., all calculations proceed as if the conduction electrons were free to move everywhere within the specimen. The total energy in all like kinetic energy, the potential energy is neglected.

However the free electron gas in metals differs from ordinary gas in some important respects. First the free electron gas is charged (in ordinary gases the molecules are mostly neutral). Free electron gas is thus actually similar to a plasma to a certain extent. Secondly the concentration of electrons in metals is large: $N \cong 10^{29}$ electrons /m^3.

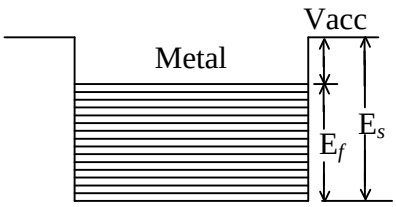


Fig 9.1: The Sommerfeld model. E_s is the energy difference between an electron at rest inside the metal and one at rest in vacuum. At T = 0 all energy levels up to E_f are filled; all higher ones are empty. Work function $\varphi = E_s - E_f$.

In contrast the ordinary gas has about 10^{25} molecules /m^3. We may thus think free electron gas in a metal as dense plasma. Our model for metals in short, is a box full of electrons and nothing else. Our problem is how to treat a box of full of electrons quantum mechanically. First we have to apply Pauli's exclusion principle-no two electrons can have all the same quantum numbers, but we do not know the quantum numbers of particles in a box. The electrons in box are free, the only constraint to which they are subjected is that they are in a box. We must find the allowed states in a box.

9.3 Free Electron Gas In A One–Dimensional Box:

Consider an electron of mass m confined to a length L by infinite barriers Fig 9.2. The wave function $\psi_n(x)$ of the electron is a solution of the Schrodinger equation.

$$H\psi = E\psi \quad \text{----------} \quad (9.1)$$

where Hamiltonian H is equal to the sum of the Kinetic energy and Potential energy and E is a set of allowed energies or eigen values of the electron in the orbital. Since the total energy is Kinetic (Potential energy is neglected).

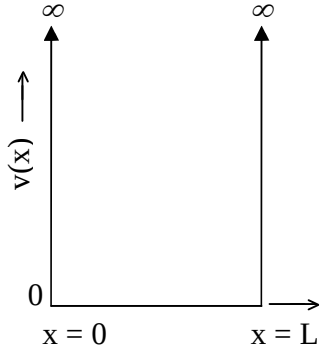


Fig 9.2: An electron of mass m confined to a length L by infinite barriers.

We have $H = \frac{P^2}{2m}$ where P is the momentum in quantum mechanics. P the momentum is represented by $-i\hbar \frac{d}{dx}$, so that

$$H\psi_n = -\frac{\hbar^2}{2m}\frac{d^2\psi_n}{dx^2} = E_n\psi_n \quad \text{----------} \quad (9.2)$$

The boundary conditions are $\psi_n(0) = 0$ and $\psi_n(L) = 0$ as imposed by the infinite potential energy barrier.

The general solution of equation (9.2) is

$$\psi_n(x) = Ae^{ikx} + Be^{-ikx} \quad \text{----------} \quad (9.3)$$

$$\text{where } k^2 = \left(\frac{2m}{\hbar^2}\right)E_n \quad \text{----------} \quad (9.4)$$

and A and B are arbitrary constants

The first boundary condition $\psi_n(o)$ yields $A = -B$ leaving the wave function to be sine like. Applying the 2nd boundary condition one singles out those solutions for which

$$\text{Sin } (kL) = 0 \text{ or } k_n = \frac{n\pi}{L} \text{ with } n = 1,2,3\ldots \quad \text{----------} \quad (9.5)$$

and the allowed energy eigen value E_n are given by

$$E_n = \frac{\hbar^2}{2m}\cdot\left(\frac{n\pi}{L}\right)^2 \quad \text{----------} \quad (9.6)$$

and the corresponding wave functions are

$$\psi_n = A \sin\left(\frac{n\pi x}{L}\right) \quad \text{----------} \quad (9.7)$$

Thus only for integral values of n, there are allowed wave functions $\psi_n(x)$ and the corresponding energy values E_n. The number 'n' is called quantum number. The plot of energy E_n versus n is given in fig 9.3.

The energy spectrum evidently consists of discrete energy levels with their separation depending on $\frac{n^2}{L^2}$. If L is large, the energy levels are thus spaced very close together.

The energy levels are spaced very closely together. For example if L = 1 cm, $E_n - E_{n+1} = 3.5 \times 10^{-19}$ eV. It is only when L is of atomic dimensions that the separation between the levels become appreciable.

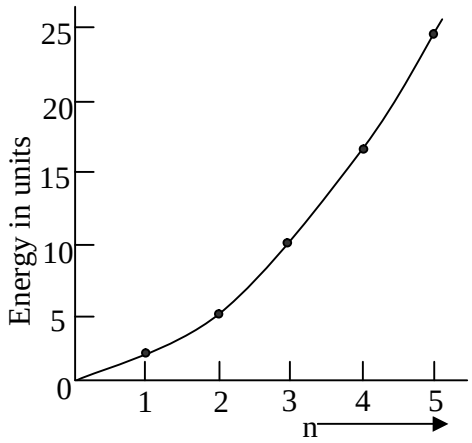


Fig 9.3: The energy is a quadratic function of the quantum number n, for a free electron confined to a length L in one dimension

The value of the constant *A* in equation 9.6 is so chosen that there is unit probability of finding the electron somewhere on the line. Because $\psi_n^*(x)\psi_n(x)$ is the probability that the electrons in the line segment *dx* at *x*, we require that

$$\int_0^L \psi_n^*(x)\,\psi_n(x)dx = 1$$

or $$\int_0^L A^2 \sin^2 \frac{n\pi x}{L} dx = 1$$

or $$A^2 = \frac{2}{L} \; i.e. \; A = \sqrt{\frac{2}{L}}$$

So the normalized wave functions are

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L} \text{ ---------- (9.8)}$$

The first few lower energy state wave functions are represented in fig 9.4.

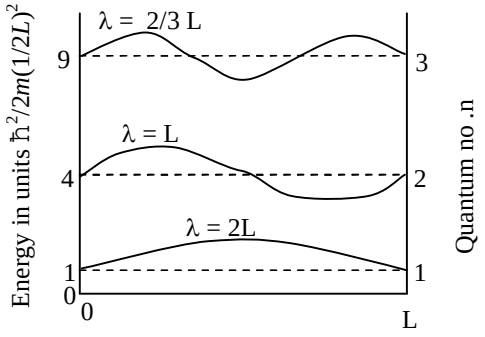


Fig 9.4: First three energy levels and wave functions of a free electron of mass m confined to a line of length L, the energy levels are labeled according to the wave functions. The wavelengths are indicated on the wave functions. The energy E_n of the level of quantum no. n is equal to $(\hbar^2 2m)(n/2L)^2$.

Now consider that we are interested in to accommodate *N*-electrons on this line. This can be done with the help of Pauli's exclusion principle. The elementary statement of the Pauli's exclusion principle is that no two electrons can have all their quantum numbers identical. That is each orbital can be occupied by at most one electron. This applies to electrons in atoms, molecules or solids. In a solid the quantum numbers of an electron in

a conduction electron orbital are n and m_s, while n is any positive integer in eq (9.8) and $m_s = \pm\frac{1}{2}$, according to the spin orientation. Each energy level with quantum number n can accommodate two electrons. One with spin up and other with spin down. If there are eight electrons, then in the ground state of the system the filled orbitals are those given in the table:

Let n_f denote the top most filled energy level, where we start filling the level from bottom (n = 1) and continue filling the higher levels with electrons until all N-electrons are accommodated. Considering N to be even we have

$2n_f = N$ ---------- (9.9)

Table:

n	m_s	Electron occupancy	n	m_s	Electron occupancy
1	↑	1	4	↑	1
1	↓	1	4	↓	1
2	↑	1	5	↑	0
2	↓	1	5	↓	0
3	↑	1	6	↑	0
3	↓	1	6	↓	0

determines the value of n for the upper most filled level. The Fermi energy E_f is defined as the energy of the top-most filled level at 0^0K. From equation (9.7) with $n = n_f$ we have

$$E_f = \frac{\hbar^2}{2m}\left(\frac{n_f\pi}{L}\right)^2 = \frac{h^2}{2m}\left(\frac{N/2}{2L}\right)^2 \quad \text{---------- (9.10)}$$

We see that the energy of the top electron depends on the number of electrons in the box and on the size of the box just as in an atom the ground state depends on the number of electrons (z) and on the size of the central coulomb field strength *i.e.*, the number of protons in the nucleus.

For example if N/L = 0.8 electrons/angstrom = 8×10^7 from equation (9.10) we have,

$$E_f = \frac{h^2}{4\pi^2 2m}\cdot\frac{N^2\pi^2}{4L^2} = \frac{(6.6\times10^{-27})^2}{32\times9\times10^{-28}}\times64\times10^{14}$$
$$= 9.68\times10^{-12}\ \text{ergs} = 6\text{eV}.$$

Consider the case of N electrons in the lowest energy state of the entire system, the total energy E_0 can be obtained by summing the individual energies E_n between n = 1 and $n_f = \frac{1}{2}N$; where the factor 2 is introduced due to the spin degeneracy. Thus

$$E_0 = 2\sum_{n=1}^{N/2} E_n = \frac{2\hbar^2}{2m}\left(\frac{\pi}{L}\right)^2 \sum_{n=1}^{N/2} n^2$$

But

$$\sum_{n=1}^{N/2} n^2 = 1^2 + 2^2 + 3^2 + \cdots\cdots + S^2 \quad \text{where} \quad S = N/2$$

$$= \frac{1}{6}S(2S^2 + 3S + 1) \cong \frac{1}{3}S^3 \quad \text{for } S \gg 1$$

$$\cong \tfrac{1}{3}(N/2)^3$$

Thus

$$E_0 = \frac{2\hbar^2}{2m}\left(\frac{\pi}{L}\right)^2 \frac{1}{3}\left(\frac{N}{2}\right)^3 = \frac{1}{3}\frac{h^2}{2m}\left(\frac{N/2}{2L}\right)^2 N$$

$$= \tfrac{1}{3}NE_f \quad \text{---------- (9.11)}$$

Hence for one dimensional problem, the average K.E in the ground state is one third of Fermi energy.

The Density Of States:

We now find number of orbital per unit energy, often called the density of states, equation 9.7 gives

$$E_n = \frac{\hbar^2}{2m}\left(\frac{n\pi}{L}\right)^2$$

Differentiating we have

$$dE_n = \frac{\hbar^2}{2m}\left(\frac{\pi}{L}\right)^2 2n\,dn$$

Now $\frac{dn}{dE_n}$ denoted the energy levels per unit energy. Corresponding the to spin states, there are two quantum states and hence the density of states of free electron gas is given by

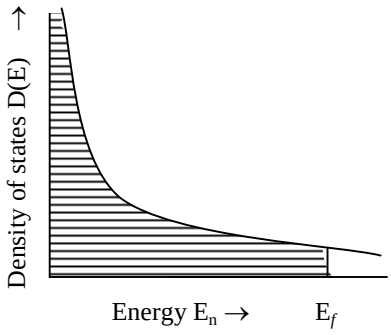


Fig 9.5 Density of electronic states as a function of energy, for one dimensional line. At absolute zero, all the states are filled upto Fermi energy E_f.

$$D(E) = \frac{2dn}{dE_n} = 2 \cdot \frac{2m}{\hbar^2}\left(\frac{L}{\pi}\right)^2 \frac{1}{2n} = \frac{4L}{h}\left(\frac{m}{2E_n}\right)^{1/2} \quad \text{---------- (9.12)}$$

This result plotted in fig.9.5

9.4 Free Electron Gas In Three Dimensions:

The free electron Schrodinger equation in three dimensions is

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)\psi_k(r) = E_k\psi_k(r) \quad \text{---------- (9.13)}$$

Where E_k is the total energy (in this case kinetic energy) of electron in K-state.

If the electron is confined to a cube of edge L, the analogue to wave function (9.8) is

$$\psi_k(r) = \sqrt{\left(\frac{8}{L^3}\right)}\sin\left(\frac{\pi n_x x}{L}\right)\sin\left(\frac{\pi n_y y}{L}\right)\sin\left(\frac{\pi n_z z}{L}\right) \quad \text{---------- (9.14)}$$

where n_x, n_y and n_z are positive integers and $\sqrt{\left(\frac{8}{L^3}\right)}$ is normalizing constant. This is standing wave.

The solution of the wave equation is a consequence of a particular boundary conditions imposed. It is better to introduce wave functions, which satisfy periodic boundary conditions as we did for phonons. We now require the wave functions to be periodic in *x, y* and *z* with period *L*. The appropriate boundary conditions are

$$\left.\begin{aligned}\psi(x,y,z) &= \psi(x+L,y,z)\\ \psi(x,y,z) &= \psi(x,y+L,z)\\ \psi(x,y,z) &= \psi(x,y,z+L)\end{aligned}\right] \quad \text{---------- (9.15)}$$

The wave functions satisfying the free particle Schrodinger equation (9.13), the normalization condition over the volume L^3 and the periodic boundary (9.15) are of the traveling plane wave form

$$\psi(x, y, z) = Ae^{(ik\cdot r)} = Ae^{i(k_x x + k_y y + k_z z)} \quad \text{---------- (9.16)}$$

Provided that the components of the wave vector k (k_x, k_y, k_z) satisfy

$$k_x = 0, \pm\frac{2\pi}{L}, \pm\frac{4\pi}{L} + \cdots\cdots \quad \text{---------- (9.17)}$$

and similar k_y and k_z. That is any component of k is of the form $\frac{2n\pi}{L}$ where n is positive or negative integer. The components of *k* are the quantum numbers of the problem; along

with quantum number m_s for the spin direction that is the state of electron is specified completely when we are given the values of k_x, k_y, k_z and m_s.

On substituting (9.16) in (9.13) we have

$$E_k = \frac{\hbar^2}{2m}(k_x^2 + k_y^2 + k_z^2) = \frac{\hbar^2}{2m}k^2 \quad \text{----------} \quad (9.18)$$

for the energy E_k of the orbital with the wave vector $\boldsymbol{k}$. The magnitude of the wave vector is related to the wave length λ by

$$k = \frac{2\pi}{\lambda} \quad \text{----------} \quad (9.19)$$

Evidential the energy spectrum consists of discrete energy levels which be usually very close together (~10^{-15}eV apart) and thus for most purposes may be taken as continuous; such energy levels are said to be quasi continuous. In fact the separation between energy levels depends on the size of the box. The above separation (≅10^{-15}eV) is for the box having laboratory dimensions, on the other hand if the dimensions are atomic, the levels are discrete as well as widely spaced (~3.6 eV for box of 5 *A* side)

Using the normalizing condition $\int \psi^*(r)\psi(r)d\tau = 1$. Constant *A* can be determined as

$$\int_0^L \int_0^L \int_0^L A^2 e^{-ik\cdot r} \cdot e^{ik\cdot r} dxdydz = 1$$

or $$A = \left(\frac{1}{V}\right)^{1/2}$$

Hence the normalized wave function is

$$\psi_k(r) = \left(\frac{1}{V}\right)^{1/2} e^{ik\cdot r} \quad \text{----------} \quad (9.20)$$

Suppose there are *N* non-interacting electrons contained in the box at absolute zero temperature. Using the effect of Pauli's exclusion principle on the energy distribution of conduction electrons – as we did in the case of one dimensional box, it can be shown that at absolute zero all the levels below a certain level will be filled with electrons and all the levels above it will be empty. The level which divides the filled and vacant-levels is called the Fermi level at absolute zero and is denoted by $E_{f(0)}$. We note that according to quantum mechanics even at absolute zero not all the electrons are condensed into the state of zero energy as happens in the classical theory. They rather occupy states between

the energy values 0 and $E_{f(0)}$. This feature modifies the parameters of the free electron gas in a significant manner.

Let us now discuss the effect of temperature. It is clear at non zero temperatures, the distribution of filled states would not end abruptly rather some of the states below $E_{f(0)}$ would be empty and some above it would be filled. Enrico Fermi has shown that the probability that a particular quantum state at energy E is filled is given by the so called Fermi- function.

$$f(E) = \frac{1}{\exp\{(E - E_f)/k_B T\} + 1} \quad \text{---------- (9.21)}$$

Where E_f is the Fermi-energy. Fig (9.6) shows the plot of this function for several value of temperature including zero.

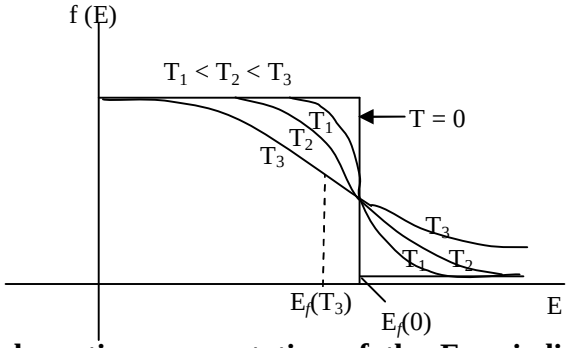


Fig 9.6: Schematic representation of the Fermi distribution function for four different temperatures. Note the variation of the Fermi energy with temperature. The temperature dependence of the Fermi energy depicted here is, typical of a three-dimensional free-electron gas but the actual variation in any particular systems will depend critically upon the density of states functions (or level degeneracies) for that system.

From(9.21) it is clear that at absolute zero

$$\left.\begin{aligned} f(E) &= 1 \quad \text{for} \quad E < E_{f(0)} \\ &= 0 \quad \text{for} \quad E > E_{f(0)} \end{aligned}\right\} \quad \text{---------- (9.21a)}$$

i.e, the Fermi distribution function becomes simply the step function implying that all the energy states below $E_{f(0)}$ are filled and all those above it are empty. As the temperature increases the edges of the step are rounded off, and the distribution function varies rapidly from nearly unity to nearly zero over an energy range of a few times $k_B T$ around the value $E = E_{f(0)}$. At the same time the value E_f (energy at the highest filled level) itself changes as shown in fig (9.6). At very high temperatures, the distribution function loses

its step like character and varies much more slowly with energy. From (9.21) it is clear that the value of $f(E)$ at $E = E_f$ is first $\frac{1}{2}$,That is $f(E_f) = \frac{1}{2}$.

Hence a quantum state at the Fermi level has a probability of occupations of $\frac{1}{2}$. This in turn means that E_f represents a level lying half way between the filled and empty states. But in case of metals since the allowed energy levels are very closely spaced (~ 10^{-19}eV) it is usually taken as reference to the highest filled (or partially filled) energy level.

It may therefore be concluded that if the electrons are heated up from absolute zero to a temperature T each electron would not increase its energy (log k_BT) as is expected classically, but those which are already with an energy range k_BT below the Fermi level can do it of course. This is because absorption of energy by an electron requires its excitation into a higher energy state; for electrons lying deep into the energy there are no empty states with in few k_BT into which they can be excited but for electrons lying below the Fermi level the empty states are available. Thus according to quantum mechanics very small fraction, that is

$$\frac{k_BT}{E_f} \cong \frac{1}{100} \quad \text{for } E_f = 0.3\,\text{eV}$$

of electrons can absorb energy from an external heat source, Moreover, ordinarily it is very unlikely that an electron in the distribution will be excited at room temperature, if it lies more than 0.1eV below Fermi level because this requires that the electron should absorb about 4 k_BT (0.1 eV ≅ 4 k_BT) from external surrounding, this is not available. These modifications in the behaviour of free electron gas have an important bearing on the problem of the heat capacity of metals.

9.5 Density Of Available Electronic States $D(E)$:

By $D(E)$ we mean that total number of available electronic states (number of orbitals) per unit energy range at E. This quantity is useful in the description of the behaviour of the free electron gas. To find the expression for D(E) we consider the linear momentum which in quantum mechanics is represented by operator

$P = -i\hbar\nabla$

Whence for the energy state (orbital) 9.16

$$P\psi_k(r) = -i\hbar\nabla\psi_k(r) = \hbar k\psi_k(r) \quad \text{----------} \quad (9.22)$$

So that the plane wave ψ_k in an eigen function of the linear momentum with eigen value $\hbar k$. The particle velocity is given by

$$v = \frac{\hbar k}{m} \quad \text{----------} \quad (9.23)$$

In the ground state of a system of N free electrons the occupied orbitals may be represented as points inside a sphere in K-space. The energy at the surface of the sphere is the Fermi energy, the wave vectors at the Fermi surface have a magnitude k_f in the fig (9.7) such that

$$E_{f(0)} = \frac{\hbar^2 k_f^2}{2m} \quad \text{----------} \quad (9.24)$$

From equation (9.17) and fig (9.7) we see that there is one allowed vector (one distinct triplet of quantum numbers k_x, k_y, k_z) for the volume element $\left(\frac{2\pi}{L}\right)^3$ of k space. Thus in the sphere is called Fermi sphere of volume $\frac{4\pi}{3}k_f^3$ the total number of orbitals is

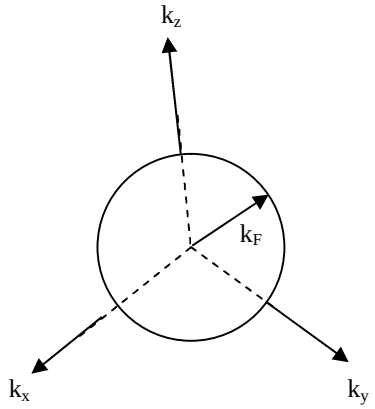


Fig 9.7: **In the ground state of a system of N electrons the occupied orbitals of the system fill a sphere of radius k, where $E_f = \frac{\hbar^2 k_f^2}{2m}$ is the energy of an electron having wave vector k_f on the surface of a sphere**

$$\frac{2(4\pi k_f^3/3)}{(2\pi/L)^3} = \frac{V}{3\pi^2}k_f^3 = N \quad \text{----------} \quad (9.25)$$

where the factor 2 on the left comes from the two allowed values of m_s, the spin quantum number of each allowed value of $\boldsymbol{k}$. We have put the number of orbitals equal to N, the number of electrons from (9.25) we have

$$k_f = \left(\frac{3\pi^2 N}{V}\right)^{1/3} \quad \text{----------} \quad (9.26)$$

This depends only on the particle concentration and not on the mass using (9.24) we have

$$E_{f(0)} = \frac{\hbar^2}{2m}\left(\frac{3\pi^2 N}{V}\right)^{2/3} \quad \text{----------} \quad (9.27)$$

This relates the Fermi energy to the electron concentration $\frac{N}{V}$ and mass m. The electron velocity v_f at Fermi surface is

$$v_f = \frac{\hbar k_f}{m} = \frac{\hbar}{m}\left(\frac{3\pi^2 N}{V}\right)^{1/3} \quad \text{---------- (9.28)}$$

The density of states function $D(E)$ defined from the fact that all the energy states below $E_{f(0)}$ are occupied and this is equal to the total number of electrons i.e

$$\int_0^{E_{f(0)}} D(E)dE = N \quad \text{---------- (9.29)}$$

Substituting the value of N from (9.27) we have

$$\int_0^{E_{f(0)}} D(E)dE = \frac{V}{3\pi^2}\left(\frac{2mE_{f(0)}}{\hbar^2}\right)^{3/2}$$

Expressing the integrals as an indefinite integral we have

$$\int D(E)dE = \frac{V}{3\pi^2}\left(\frac{2mE}{\hbar^2}\right)^{3/2}$$

or $$D(E) = \frac{V}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} E^{1/2} = CE^{1/2} \quad \text{---------- (9.30)}$$

where $$C = \frac{V}{2\pi^2}\cdot\left(\frac{2m}{\hbar^2}\right)^{3/2}$$

The result may be obtained and expressed most simply by writing (9.27) as

$$\log N = \frac{3}{2}\log E_{f(0)} + \text{constant}$$

or $$\frac{dN}{N} = \frac{3}{2}\frac{dE_{f(0)}}{E_{f(0)}}$$

and we have

$$D(E_{f(0)}) = \frac{dN}{dE_{f(0)}} = \frac{3}{2}\frac{N}{E_{f(0)}} \quad \text{---------- (9.31)}$$

With in a factor of the order of unity, the number of orbitals per unit energy range at the Fermi-energy is just the total number of conduction electrons divided by the Fermi energy.

These results apply to free electrons with energy (E) proportional to k^2. We can obtain a result for a general relation $E(k)$ by direct analogy as

$$D(E) = \frac{2V}{(2\pi)^3}\int \frac{ds_E}{(grad^E k)} \quad \text{---------- (9.32)}$$

Where the factor 2 arises from the two spin orientations. *V* is the volume of the specimen and *ds* is the element of area in *k* space of the surface of constant energy *E*.

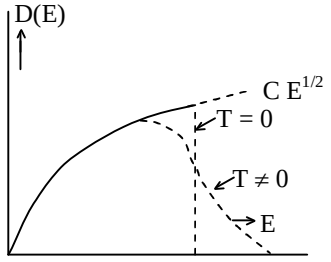


Fig 9.8: The curve $CE^{1/2}$ represents D(E) in accordance with (9.30). The energy distribution is obtained by multiplying D(E) with F(E).

It is seen from expression (9.30) that *D*(*E*) is a parabolic function of energy. This is allowed in fig (9.8).

Also it increases with increase in the crystal volume; this is in order to accommodate the total number of electrons present, which also increases with the size of the crystal.

9.6. Effect of Temperature on the parameters of the free Electron Gas:

Hitherto we have assumed that the electron system is at 0^0K and in accordance with the Pauli's principle. We fill the levels until we have used up all electrons at the Fermi level of energy $E_{f(0)}$. Now the kinetic energy of the electron gas increases as the temperature is increased and some energy levels are occupied which were vacant at 0^oKand some levels are vacant which were preciously occupied at 0^oK.

As an example of Fermi function at T > 0 is given in fig. (9.6), For energies below E_f such that $E_f - E \gg k_B T$ (degenerate Fermi gas) the value of $f(E)$ is still practically unity i.e. energy distribution in that region is the same as that for $T = 0$. It is only in the vicinity of E_f minus a few $k_B T$ the $f(E)$ begins to drop below the vaule at $T = 0$. For energies above E_f such that $E - E_f \gg k_B T$ one may neglect the term unity in the denominator of distribution function and obtain

$$f(E) = e^{-(E-E_f)/k_B T}$$

the Fermi distribution becomes identical with Boltzmann distribution.

Here it is of the prime importance to take into account the effect of temperature on the Fermi energy and average kinetic energy. In order to do this we shall have to evaluate the integral

$$I = \int_0^\infty f(E)\frac{\partial \phi(E)}{\partial E} dE \quad \text{---------- (9.33)}$$

where $f(E)$ is the Fermi function and $\phi(E)$ is a function of energy(E) which has the property that

$\phi(0) = 0$ ---------- (9.34)

Integrating (9.33) by parts, we have

$$I = \left[f(E)\phi(E) \right]_0^\infty - \int_0^\infty \phi(E) \frac{\partial f(E)}{\partial E} dE$$

$$= - \int_0^\infty \phi(E) \frac{\partial f(E)}{\partial E} dE \quad \text{---------- (9.35)}$$

Since the product $f(E)\ \phi(E)$ vanishes at both the limits. Expanding $\phi(E)$ as a Taylor's series about the point $E = E_f$ we have

$$\phi(E) = \phi(E_f) - (E - E_f)\left(\frac{\partial \phi}{\partial E} \right)_{E=E_f} + a_2 \left(\frac{\partial^2 \phi}{\partial E^2} \right)_{E=E_f} + \cdots\cdots \quad \text{---------- (9.36)}$$

Where by (9.35) can be expressed as

$$I = a_0 \phi(E_f) + a_1 \left(\frac{\partial \phi}{\partial E} \right)_{E=E_f} + a_2 \left(\frac{\partial^2 \phi}{\partial E^2} \right)_{E=E_f} + \cdots\cdots \quad \text{---------- (9.37)}$$

where $$a_n = -\frac{1}{n!} \int_0^\infty (E - E_f)^n \frac{\partial f(E)}{\partial E} dE \quad \text{---------- (9.38)}$$

But

$$\frac{\partial f(E)}{\partial E} = - \frac{e^{(E-E_f)/k_B T}}{\left\{ 1 + e^{(E-E_f)/k_B T} \right\}^2} \cdot \frac{1}{k_B T} = - \frac{1}{k_B T} f_E (1 - f_E) \quad \text{---------- (9.39)}$$

The plot of $\frac{\partial f(E)}{\partial E}$ versus E is given in fig 9.9.

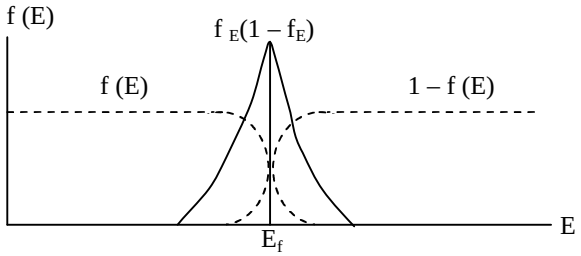


Fig 9.9: The function f(E), 1–f(E) and f_E . (1 – f_E) for a Fermi system, plotted as a function of energy,, The vertical scale factor for the f_E .(1– f_E) curve is not the same as for the other curves.

Now if $k_B T << E_f$, $\frac{\partial f(E)}{\partial E}$ is negligible for negative values of E and the limits of integral (9.38) may be extended to $+\infty$ to $-\infty$. If this is done then, we have

$$a_0 = -\int_{-\infty}^{+\infty} \frac{\partial f(E)}{\partial E}\, dE = -\left[f(E)\right]_{-\infty}^{+\infty} = 1 \quad \text{---------- (9.40)}$$

$$a_1 = -\int_{-\infty}^{+\infty} (E - E_f) \frac{df}{\partial E}\, dE = 0 \quad \text{---------- (9.41)}$$

The integral vanishes because $\frac{\partial f(E)}{\partial E}$ is an even function of $(E\text{-}E_f)$ and integrand from $-\infty$ to $+\infty$ must yield zero

$$a_2 = -\tfrac{1}{2}\int_{-\infty}^{+\infty} (E - E_f) \frac{df(E)}{\partial E}\, dE$$

and putting $x = \frac{E - E_f}{k_B T}$, we have

$$^1a_2 = \frac{(k_B T)^2}{2} \int_{-\infty}^{+\infty} \frac{x^2 e^x}{(1+e^x)^2} dx = \frac{\pi^2}{6}(k_B T)^2 \quad \text{---------- (9.42)}$$

Substituting these values in (9.37), we have

$$I = \int_0^{\infty} f(E) \frac{\partial \phi(E)}{\partial E} dE \; = \phi(E_f) + \frac{\pi^2}{6}(k_B T)^2 \left\{ \frac{\partial^2 \phi(E)}{\partial E^2} \right\}_{E=E_f} + \cdots\cdots \text{---------- (9.43)}$$

This formula is convenient for working out approximate values of Fermi integrals .It is restricted by the condition (9.34) and by the condition that $k_B T << E_f$. The room temperature quantities the latter condition. Suppose now that we choose

$$\phi(E) = \int_0^E D(E) dE$$

whence $\frac{\partial \phi(E)}{\partial E} = D(E) and \frac{\partial^2 \phi(E)}{\partial E^2} = \frac{\partial D(E)}{\partial E}$

with $D(E)$ as given by (9.30) for the free electrons system, we have from (9.43)

$$N = \int_0^{\infty} f(E) D(E) dE$$

$$= \int_0^{E_f} D(E) dE + \frac{\pi^2}{6}(k_B T)^2 \left\{ \frac{\partial D(E)}{\partial E} \right\}_{E=E_f} \quad \text{---------- (9.44)}$$

But from (9.29), we have

$$\int_0^{E_{f(0)}} D(E) dE = N$$

Substituting the values N we have in (9.44) we have

$$\int_0^{E_{f(0)}} D(E)dE = \int_0^{E_f} D(E)dE + \frac{\pi^2}{6}(k_B T)^2 \left\{\frac{\partial D(E)}{\partial E}\right\}_{E=E_f}$$

or
$$O = \int_{E_{f(0)}}^{E_f} D(E)dE + \frac{\pi^2}{6}(k_B T)^2 \left\{\frac{\partial D(E)}{\partial E}\right\}_{E=E_f}$$

$$\simeq D(E_f)\{E_f - E_{f(0)}\} + \frac{\pi^2}{6}(k_B T)^2 \left\{\frac{\partial D(E)}{\partial E}\right\}_{E=E_f} \quad \text{---------- (9.45)}$$

In the above equation we have assumed that $D(E)$ does not vary much in the interval from $E_{f(0)}$ to E_f which for $k_B T << E_f$ will be only a small fraction of $E_{f(0)}$. Since $D(E)$ is of the from $CE^{1/2}$ (9.30) with C, a constant, we have

$$\left\{\frac{\partial D(E)}{\partial E}\right\}_{E=E_f} = \tfrac{1}{2} C E_f^{-1/2} = \tfrac{1}{2}\frac{C E_f^{1/2}}{E_f} = \tfrac{1}{2}\frac{D(E_f)}{E_f}$$

and equation (9.45) can be written as

$$O = D(E_f)\{E_f - E_{f(0)}\} + \frac{\pi^2}{12}(k_B T)^2 \frac{D(E_f)}{E_f}$$

or
$$E_f = E_{f(0)} - \frac{\pi^2}{12}\frac{(k_B T)^2}{E_f} \quad \text{---------- (9.46)}$$

For $k_B T << E_f$, the second term will be a small correction to be subtracted from relative by large quantity

$$E_{f(0)}, \left\{ e.g., \; for E_{f(0)} \simeq 5eV, \frac{\pi^2}{12}\frac{(k_B T)^2}{E_f} \simeq 10^{-4} eV \right\}$$

and the difference between E_f and $E_{f(0)}$ will be small compared to $E_{f(0)}$. Under these circumstances not much error will be made if E_f in the connection term (9.46) is replaced with $E_{f(0)}$, giving finally

$$E_f = E_{f(0)}\left[1 - \frac{\pi^2}{12}\left(\frac{k_B T}{E_{f(0)}}\right)^2\right] \quad \text{---------- (9.47)}$$

Thus the Fermi energy is not constant but decreases slowly as the temperature rises, the slowness of the variations following immediately from the occurrence of the factor

$\frac{\pi^2}{12}\left(\frac{k_B T}{E_{f(0)}}\right)^2$ which is very small. The result (9.47) is quite satisfactory but while using it is applicable only at temperature such that $k_B T << E_f$, which incidentally includes the whole range of temperature for which the metals are solids (500^oc to 3000^oc).

9.7 Summary:

1. When atoms are packed together to form a metal, valance electrons detach from their own atoms and more through out the crystal without any collision and one treated like molecules of an ideal gas. This is called fee electron gas.

2. The energy spectrum of one dimension free electron gas is discrete with the energy level separation depending on $\frac{n^2}{L^2}$. For laboratory dimensions the levels are very closely spaced.

3. The energy spectrum of three dimensional free electron gas is also discrete and at absolute zero temperature all levels below a certain level will be filled and all the above it will be empty. The levels dividing filled and vacant levels is called the Fermi level ($E_{f(0)}$).

4. Fermi energy is determined by the concentration of electrons. Its value is $\frac{\hbar^2}{2m}\left(\frac{3\pi^2 N}{V}\right)^{2/3}$ and the density of states is given by $\frac{V}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} E^{1/2} = CE^{1/2}$. This, at absolute zero, can be expressed as $D(E_{f(0)}) = \frac{3}{2} N / E_{f(0)}$.

9.8. Keywords:

Thermal conductivity – Electrical conductivity – Wiedemannn–frantz ratio – Fee electron gas – Density of states – Fermi level – Fermi sphere.

9.9.Review Questions:

1. Describe the free electron model by explaining wave function and eigen function.
2. Explain in detail the free electron model in three dimensions, explain the Fermi function and also sketch Fermi distribution function for different temperatures.
3. How do you understand the density of electronic states and also effect of temperature on the parameters of free electron gas.

4. Explain Sommerfield's treatment of free electron model and work out expressions for ` Fermi Energy ' and `density of states.'

   Using the free electron model calculate the number of orbitals between 1.0 $*10^{-8}$ cm^{-1} and 1.2 $10^{-8}cm^{-}1$ for the magnitudes of wave vector for a cubic sample of dimensions 1mm * 1mm * 1mm.

5. Discuss the successes and failures of free electron theory of metals.

9.10 Text and Reference Books:

1. Elements of Solid State Physics by J.P. Srivatsava (PHI)
2. Solid State physics by M.A. Wahab (Narosa)
3. Elements of Solid State Physics by A. Omar (Pearson education)
4. Solid State Physics by S.O. Pillai (New Age)
5. Solid State Physics by C. Kittel (Asia Publishing House)
6. Solid State Physics by S.L. Kakani and C.Hemrajani (S.Chand)
7. Solid State Physics by Saxena Gupta Saxena (Pragati Prakashan).
8. Solid State Physics by C.J. Dekker (Macmillan)

UNIT – III **LESSON: 10**

THERMAL PROPERTIES OF SOLIDS - II

Aim: To know about thermal properties of solids

Objectives of the lesson:

- To know about thermal capacity and its dependence on temperature
- To discuss the classical and quantum theories of heat capacity
- To discuss about the Einstein model and its deficiencies.
- To know about Debye model the criticism on Debye model

STRUCTURE OF LESSON:

10.1 Thermal capacity of the free-electron System:

The classic electron theory proposed by Drude-Lorentz, in which the free electrons are assumed to obey Maxwell-Boltznamm satistics leads to the conclusion that (assuming one free electron per atom) the heat capacity of free electrons should be $\frac{3}{2} Nk_B$ .The total heat capacity would be then the sum of the lattice contribution given by Debye theory and electron contribution $\frac{3}{2} Nk_B$. At temperature large compared to Debye temperature, the total specific heat of the metal can be expressed as

$C_v = C_v$ (lattice)+ C_v (electronic)

$$=3Nk_B+\tfrac{3}{2}Nk_B \quad = \tfrac{9}{2}\ Nk_B \qquad \text{----------} \quad (10.1)$$

Dulong and Petit observed that the specific heat above the characteristic temperature is the same for all metals and is equal to $3Nk_B$, as predicted by Debye theory for the lattice contribution alone. Since the large specific heat predicted by (10.1) is not observed in metals, it appears that the heat capacity of free electron is much smaller than the value predicted by Drude–Lorentz–Boltzmann model. This was one of the major shortcomings of the original Drude –Lorentz model .

With the advent of the Quantum mechanics it became apparent that this shortcoming is due to incorrect assessment of electrons that absorb thermal energy. When Fermi–Dirac statistics is used to describe the free electron energy distribution, the calculated heat capacity is much smaller than the $\frac{3}{2}Nk_B$ and in fact is negligible compared to the lattice contribution at moderate and high temperatures. The Fermi free-electron model is thus much better accord with the experimental observations, than the original Drude –Lorentz theory .The explanation of the electronic specific heat of metallic substances was one of the original triumphs of quantum statistics.

The reason for much smaller heat capacity of the Fermi-Dirac free electron system is that only electrons with in a few k_BT of the surface of the Fermi sphere can receive energy from an external heat source. Ordinarily a particle of external heat bath at temperature T has an energy of only few times k_BT to give an electron of the Fermi free electron system. Electrons deep inside the Fermi sphere are incapable of interacting with such external exciting bodies because there are no occupied states with in a few times k_BT in energy into which they can be excited. As such these electrons do not contribute towards specific heat. Only electrons near the Fermi surface where there are unoccupied states available can participate in interactions with an external heat source and contribute towards the specific heat. The quantum mechanics therefore modifies the thermal behavior of the free electrons in a simple and satisfying manner.

Contribution of conduction electrons to specific heat of solids:

In order to make an accurate calculation of heat capacity, it is necessary to find the average energy for the Fermi free electron gas. The average energy U is given by

$$U = \int_0^\infty E\ \mathrm{D(E)}\ \mathrm{f(E)}\ \mathrm{dE} \qquad \text{----------} \quad (10.2)$$

In order to solve the integral, we choose

$$\phi(E) = \int_0^E E\ \mathrm{D(E)}\,\mathrm{dE}$$

Where $\dfrac{\partial\phi(E)}{\partial E} = E\ \mathrm{D(E)}$

and $\dfrac{\partial^2\phi(E)}{\partial E^2} = \dfrac{\partial}{\partial E}\{E\ \mathrm{D(E)}\}$

Substituting in (10.2), we have

$$U = \int_0^\infty E\ \mathrm{D(E)}\,\mathrm{f(E)}\,\mathrm{dE} = \int_0^\infty f(E)\frac{\partial\phi(E)}{\partial E}dE$$

and $$U = \int_0^{E_f} E\ \mathrm{D(E)}\,\mathrm{dE} = \frac{\pi^2}{6}(\mathrm{k_B}T)^2\left[\frac{\partial}{\partial E}E\ \mathrm{D(E)}\right]_{E=E_f} \quad \text{---------- (10.3)}$$

In lesson VIII we derived an expression for the density of states as

$$D(E) = CE^{1/2} \quad \text{(or)} \quad E\,D(E) = CE^{3/2}$$

and from this $$\frac{\partial}{\partial E}\{E\,D(E)dE\}_{E=E_f} = \frac{3}{2}CE^{1/2} = \frac{3}{2}D(E_f) \quad \text{---------- (10.4)}$$

Substituting in (10.3), we have

$$U = \int_0^{E_{f(0)}} E\ \mathrm{D(E)}\,\mathrm{dE} = \int_{E_{f(0)}}^{E_f} E\ \mathrm{D(E)}\,\mathrm{dE} + \frac{\pi^2}{6}(k_BT)^2\frac{3}{2}D(E_f)$$

$$\cong U_0 + [E_f - E_{f(0)}]E_{f(0)}\ D(E_{f(0)}) + \frac{\pi^2}{4}(k_BT)^2 D(E_{f(0)}) \quad \text{---------- (10.5)}$$

Here U_0 is the absolute zero value of the internal energy as represented by the integral of $E\,D(E)$ from zero to $E_{f(0)}$ and the integral of this quantity over the small range $E_{f(0)}$ to E_f has been approximated by the same procedure, which was used in connection with equation (9.45) of lesson nine. In the last term of eqn.10.5, $D(E)$ is evaluated at $E_{f(0)}$ rather than at E_f; which introduce a small error . Substituting the value given by eqn.9.47 for

Ef - $E_{f(0)}$ into (10.5), we have

$$U = U_0 - \frac{\pi^2}{12}(k_BT)^2 D(E_{f(0)}) + \frac{\pi^2}{4}(k_BT)^2 D(E_{f(0)})$$

$$= U_0 + \frac{\pi^2}{6}(k_BT)^2 D(E_{f(0)}) \quad \text{---------- (10.6)}$$

Equation 9.31 gives

$$D(E_{f(0)}) = \frac{3}{2}\frac{N}{E_{f(0)}} = \frac{3}{2}\frac{N}{k_BT_F}$$

where T_F is Fermi temperature given by $E_{f(0)}$ =$k_B T_F$ and T_F =30,000^{0}K, a typical value for simple metals (alkali and noble) calculated from equation (9.27). Substituting value of $D(E_{f(0)})$ in (10.6), we have

$$U = U_0 + \frac{\pi^2}{6}(k_B T)^2 \cdot \frac{3}{2}\frac{N}{k_B T_F} = U_0 + \frac{\pi^2}{4}\frac{N k_B T^2}{T_F} \quad \text{----------} \quad (10.7)$$

The heat capacity is obtained in the usual way as

$$C_v = \frac{\partial U}{\partial T} = \frac{\pi^2}{2}\frac{N k_B T}{T_F} \quad \text{----------} \quad (10.8)$$

The electronic component of specific heat as calculated using Fermi statistics thus amounts to only small fraction of Drude–Lorentz value $\frac{3}{2}Nk_B$ and except at very low temperatures is small compared to the lattice contribution. At temperature small compared to the Debye temperature, the lattice contribution becomes quite small approaching zero like T^3 as $T \rightarrow 0$. In this range the electronic contribution is often a significant factor and in certain temperature may even be the dominant effect.

Comparison between theory and experiment:

At temperatures much below the Debye temperatures and very much below Fermi temperature, the heat capacities of metals may be written as the sum of electronic and lattice contributions**:**

$$C_v = \gamma T + AT^3 \quad \text{----------} \quad (10.9)$$

Where γ and A are constants characteristic of the material and are given by (10.8) ?. The electronic term is linear in *T* and is dominant at sufficiently low temperature. The measurements made in this limit determines primarily the electronic contributions and the direct comparison of the above theory with experiment is possible.

Kok and Keesom determined the value of γ for few metals as a test of the above theory. By applying these considerations, they, in fact represented the experimental result by plotting a curve C_v/T versus T^2 and from the slope of the straight line graph thus obtained, determined value of γ. Such a plot for Potassium is shown in fig 10.1 and corresponding value of γ is obtained. The observed values of the co-efficient γ are of the expected magnitude, but often do not agree very closely with the value calculated for free electrons of the mass m . By the use of relation (10.8).

$$\gamma = \frac{\pi^2 N k_B^{\ 2} Z}{2E_f} \quad \text{----------} \quad (10.10)$$

for a mole of material, where *N* is Avagadro's number *Z* is the valancy of the element.

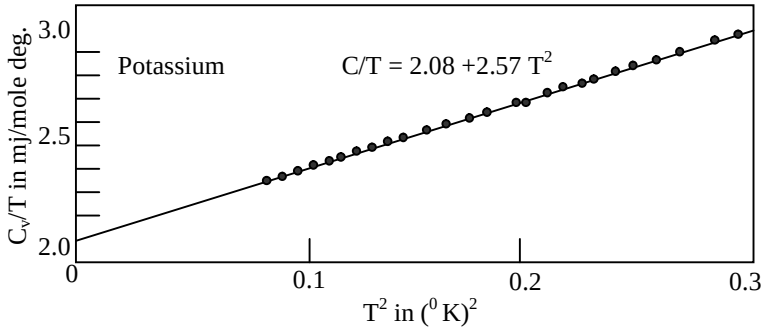


Fig 10.1: Experimental heat capacity values for Potassium, plotted C_v/T versus T^2. The solid points were determined with an adiabatic demagnetization of crystal

This is the difficulty which one encounters also for other metals and is a consequence of the over simplifying assumptions made in the free electron model .The departure involves three separate effects**:**

(i) The interaction of the conduction electron with periodic potential of the rigid crystal lattice.

(ii) The interaction of the conduction electron with phonons .An electron tends to polarize or distort lattice in its neighborhood so that, the moving electron tries to drag nearby ions along, thereby increasing effective mass of the electron .In ionic crystal, the effect has a name – the polarization effect.

(iii) The interaction of conduction electrons with themselves. A moving electron causes an inertial reaction in the surrounding electron gas, there by increasing the effective mass of the electron.

The effect of electron –electron interaction are usually described with in the framework of what is called London theory of Fermi liquid.

10.2 Heat capacity: Classical Theory:

The classical theory is based on the following assumptions

(1) Solids consist of large number of atomic particles executing simple harmonic motion about their equilibrium position. They are called atomic oscillatros.

(2) The atomic oscillators vibrate with the same frequency, but, their energies are different, because, they vibrate with different amplitudes

(3) The internal energy is largely due to vibrational energy of all atomic oscillators.

(4) There is no restriction on the energy values of the atomic oscillator and it can take up continuum of energy values right from 0 to ∞ though, the probability of occupation of each level is different

(5) The total energy of a solid at a given temperature is N times the average energy ($\overline{\in}$) of the oscillator.

For a single three-dimensional isotropic harmonic oscillator the total energy is

$$E = \frac{P^2}{2m} + V(r) \quad \text{---------- (10.11)}$$

with $V(r) = \frac{1}{2} m\omega_0{}^2 (x^2+y^2+z^2) = \frac{1}{2} m\omega_0{}^2 r^2 = f(P, q)$ ----------(10.12)

where m is the mass and ω_0 the natural frequency of the oscillator. Assuming Maxwell-Boltzmann distribution law for energy, we have

$$f(\in) = Ae^{-\varepsilon/k_BT} = Ae^{-P^2/2mk_BT} \cdot e^{-m\omega_0^2 r^2/2k_BT} \quad \text{---------- (10.13)}$$

Where A is constant. The average energy is then given by

$$\overline{E} = \frac{\int E\ f(E)\, dp\, dq}{\int f(E)\, dp\, dq} \quad \text{---------- (10.14)}$$

If we choose spherical coordinate(r, θ, ϕ) in co-ordinate space, and (p, θ_p, ϕ_p) in momentum space, we have

$$\overline{E} = \frac{\int_p \int_r \left(\frac{p^2}{2m} + \frac{1}{2} m\omega_0^2 r^2 \right) e^{-P^2/2mk_BT} \cdot e^{-m^2\omega_0^2 r^2/2k_BT} \cdot p^2 \sin\theta_p dp\, d\theta_p d\varphi_p\, r^2 \sin\theta \cdot d\theta\, dr\, d\varphi}{\int_p \int_r e^{-P^2/2mk_BT} \cdot e^{-m^2\omega_0^2 r^2/2k_BT} \cdot p^2 \sin\theta_p dp\, d\theta_p d\varphi_p \cdot r^2 \sin\theta \cdot d\theta \cdot dr\, d\varphi}$$

---------- (10.15)

Since the integrand has no dependence on θ, ϕ, θ_p and ϕ_p, the integration over these variables give simply 4π both in numerator and denominator leaving

$$\overline{E} = \frac{\int_0^\infty \frac{p^4}{2m} e^{-P^2/2mk_BT} dp \cdot \int_0^\infty r^2 e^{-m\omega_0^2 r^2/2k_BT} \cdot dr}{\int_0^\infty e^{-P^2/2mk_BT} dp \cdot \int_0^\infty r^2 e^{-m\omega_0^2 r^2/2k_BT} dr}$$

$$+ \frac{\int_0^\infty \frac{1}{2} m\omega_0^2 r^4 e^{-m\omega_0 r^2/2k_BT} dr \int_0^\infty p^2 e^{-P^2/2mk_BT} dp}{\int_0^\infty r^2 e^{-m\omega_0^2 r^2/2k_BT} dr \int_0^\infty p^2 e^{-P^2/2mk_BT} dp}$$

$$= \frac{1}{2m}\frac{\int_0^\infty p^4 e^{-P^2/2mk_BT}\,dp}{\int_0^\infty p^2 e^{-P^2/2mk_BT}\,dp} + \tfrac{1}{2}m\omega_0^2\frac{\int_0^\infty r^4 e^{-m\omega_0^2 r^2/2k_BT}\,dr}{\int_0^\infty r^2 e^{-m\omega_0^2 r^2/2k_BT}\,dr} = \overline{E}_k + \overline{E}_p \quad \text{---------- (10.16)}$$

It is apparent that the first term above represents average kinetic energy and the second, the average potential energy. Evaluating integrals with the help of Maxwell–Boltzmann integrals, we have

$\overline{E}_k = \overline{E}_P = \frac{3}{2}\mathrm{k_B}\,T$ ---------- (10.17)

For an assembly of *N* independent oscillators, the total internal energy is

$U = N\overline{E} = 3Nk_B\,T$ ---------- (10.18)

and the heat capacity C_v is by definition

$$C_v = \left(\frac{\partial U}{\partial T}\right)_v = 3Nk_B = 3R \quad \text{---------- (10.19)}$$

The heat capacity is found to be a constant, independent of temperature. This is Dulong and Pettit law. This result is certainly in agreement with experimental values at high temperature but the actual value of C_v decreases as *T* decreases, and in fact, vanishes entirely as $T \to 0^0K$. This discrepancy between the theory and experiment was one of the outstanding paradoxes in physics until 1905, when it was resolved by Einstein involving quantum mechanics.

10.3 The Einstein model:

The physical model employed by Einstein was simplified but his results definitely indicate that quantum theory contained the answer to the difficulty encountered in the classical theory. The salient features of Einstein's model are**:**

1. All the atoms vibrate independently of each other
2. The solid containing *N* atoms is considered to be equivalent to 3*N* harmonic oscillators.
3. All the atomic oscillators have the same frequency (*v*) because of their assumed identical surroundings.
4. The energy spectrum of oscillators is not continuous but discrete. The possible energy levels of a harmonic oscillator are given by

$$\in_n = (n + \tfrac{1}{2})\mathrm{h}v \quad \text{---------- (10.20)}$$

5. Any number of oscillators may be in the same quantum state of the system.

6. Since these atomic oscillators form an assembly of systems which are distinguishable or identifiable by virtue of their location at separate and discrete lattice sites, hence the oscillators are regarded as having Maxwell-Boltzmann distribution of energies.

Equation (10.20) refers to an isolated oscillator but the atomic oscillations in solid are not isolated. They are continuously exchanging energy with the ambient thermal bath surrounding the solid .The energy of the oscillators is therefore continually changing but its average value at thermal equilibrium is given by

$$\overline{\in}_n = \frac{\sum_{n=0}^{\infty} \in_n e^{-\in_n / k_B T}}{\sum_{n=0}^{\infty} e^{-\in_n / k_B T}} = \frac{\sum_{n=0}^{\infty} (n+\frac{1}{2}) h\nu\, e^{-(n+\frac{1}{2})h\nu / k_B T}}{\sum_{n=0}^{\infty} e^{-(n+\frac{1}{2})h\nu / k_B T}} = \frac{h\nu \sum_{n=0}^{\infty} (n+\frac{1}{2})\, e^{(n+\frac{1}{2})x}}{\sum_{n=0}^{\infty} e^{(n+\frac{1}{2})x}} \quad \text{---------- (10.21)}$$

where $x = -\dfrac{h\nu}{k_B T}$ ---------- (10.22)

Equation (10.21) can be written as

$$\overline{\in} = h\nu \frac{\frac{1}{2}e^{x/2} + \frac{3}{2}e^{3x/2} + \frac{5}{2}e^{5x/2} + ...}{e^{x/2} + e^{3x/2} + e^{5x/2} + ...}$$

$$= h\nu \frac{d}{dx} \log[e^{x/2} + e^{3x/2} + e^{5x/2} + ...]$$

$$= h\nu \frac{d}{dx} \log[e^{x/2}(1 + e^{x} + e^{2x} + e^{3x} + ...)]$$

$$= h\nu \frac{d}{dx} [\tfrac{1}{2}x - \log(1 - e^{x})]$$

$$= h\nu \left[\tfrac{1}{2} + \frac{e^{x}}{(1-e^{x})}\right] = h\nu \left[\tfrac{1}{2} + \frac{1}{(e^{-x}-1)}\right]$$

$$= h\nu \left[\tfrac{1}{2} + \frac{1}{(e^{h\nu / k_B T} - 1)}\right] \quad \text{---------- (10.23)}$$

The internal energy is obtained by multiplying the average energy per oscillator by the number of oscillators ($3N$). Hence

$$U = 3N\,\overline{\in} = \frac{3Nh\nu}{2} + \frac{3Nh\nu}{e^{h\nu / k_B T} - 1} \quad \text{---------- (10.24)}$$

and $$C_v = \left(\frac{\partial U}{\partial T}\right)_v = 3Nk_B \left(\frac{h\nu}{k_B T}\right)^2 \frac{e^{h\nu / k_B T}}{(e^{h\nu / k_B T} - 1)^2} \quad \text{---------- (10.25)}$$

putting $h\upsilon = \theta_E k_B$ ---------- (10.26)

Here θ_E is called “Einstein’s temperature”. Substituting (10.26) in (10.25), we have

$$C_v = 3Nk_B\left(\frac{\theta_E}{T}\right)^2 \frac{e^{\theta_E/T}}{(e^{\theta_E/T}-1)^2} \quad \text{---------- (10.27)}$$

If we plot C_v verses T using Equation (10.27), we obtain a curve of general shape as in Fig (10.2), which indicates that the theory is now in agreement with experiment at least qualitatively; over the entire temperature range.

Note in particular that $C_v \approx 0$ as $T \to 0^0K$ a new and important feature, which was lacking in the classical theory. The temperature θ_E is an adjustable parameter chosen to produce the best fit to the measured values over the whole temperature range. Fig (10.2) illustrates the procedure for copper where θ_E is found to be 240 *K*.

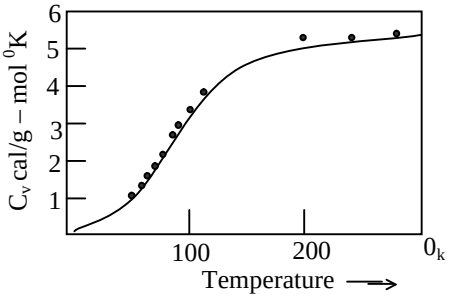


Fig 10.2: Specific heat of copper versus temperature. The dots represent experimental values and the curve is given by the Einstein's expression.

We can calculate Einstein's frequency $\omega_E = 2\pi\nu_E$, once we have determined the temperature θ_E. Thus for $\theta_E = 240K$, the $\omega_E = \frac{k_B\theta_E}{\hbar} = 2.5\times10^{+13}\,\text{sec}^{-1}$ which is in the infrared region.

Let us now examine the behaviour of C_v as given by (10.27) in the extreme temperature limits. If $k_BT >> h\nu(T >> \theta_E)$, i.e., high temperature limit, we can write

$$e^{h\nu/k_BT} = 1 + \frac{h\nu}{k_BT}$$

$$U = \frac{3Nh\nu}{k_BT} + 3Nk_BT$$

and $C_v = \left(\frac{\partial U}{\partial T}\right)_\infty = 3Nk_B = 3R.$

which is Dulog and Petit law. Thus the model is satisfactory high at the high temperature limit because the quantum aspects become irrelevant.

At low temperatures $T << \theta_E$ and

$$e^{h\nu/k_BT} >> 1$$

and $$C_v = 3Nk_B\left(\frac{\theta_E}{T}\right)^2 e^{-\theta_E/T} \quad \text{---------- (10.28)}$$

The specific heat approaches zero exponentially as the temperature approaches zero.

Physically, the reason that specific heat becomes small at low temperatures can be understood by assuming that the crystal is placed in contact with external heat bath consisting of an ideal mono-atomic gas at some given temperature; and allowed to absorb energy of the ideal gas. The average energy of gas atoms is $\frac{3}{2} k_B T$ and if the temperature of the system is high enough so that $k_B T$ is of the order of, or greater than, the energy $h\nu$ required to excite one of the vibrating atoms of the crystal to a higher energy state, then such excitations will occur frequently when the gas atoms from the heat bath collide with the crystal. These collisions will then be elastic in mechanical sense and energy initially belonging to the gas atoms will readily transferred to the crystal lattice as vibrational energy.

At low temperatures the mean thermal energy $k_B T$ is very much less than the quantum of energy $h\nu$ and is insufficient to excite the thermal vibrations of oscillators-most of them are in the state of zero point energy or frozen energy state and execute zero point oscillations and not the thermal oscillations. Under these circumstances an occasional gas atom having much higher energy than average will be capable of effecting such an excitation and transferring heat to the crystal lattice resulting in excitation of a small fraction of oscillators. Hence the mean thermal energy of the oscillator will be much less than $k_B T$. Thus at low temperatures the specific heat of quantum oscillator will be less than the specific heat of the oscillator.

In most respects Einstein's model has been a remarkable success; its results are in good agreement with experiment over most of the temperature range. However its drawbacks are;

(1) It explains the specific heat curves up to a certain point (upto $0.2\theta_E$) but fails completely at extremely low temperatures (below $0.2\theta_E$) where T^3 law holds true.

(2) ν and θ_E are obtained empirically and can not be verified from any other independent physical data.

(3) Einstein himself recognized that the assumption of the monochromatic vibrations did not correspond to actual facts. The vibrations of a particular atom must be very complex indeed because it is under the field of force of thousands of other vibrating atoms.

Nevertheless Einstein's model was a great step forward in the right direction.

10.4. DEBYE MODEL:

The main difference between Einstein's model and Debye model is, that, Debye considers the vibrational modes of a crystal as a whole ($\omega=v_0k$) the dispersion relation for a homogeneous solid where as Einstein starting point was to consider the vibration of a single atom assuming that the atomic vibrations to be independent of each other.

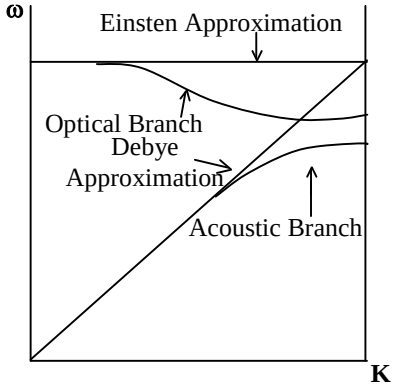


Fig 10.3: Comparision of Einsteins and Debye approximation for dispersion relation of lattice waves, referred to a linear diatomic lattice.

The Debye model is based on the following assumptions:

1) A solid is treated as an isotropic elastic continuum, so that the the longitudinal and transverse wave velocities are independent of propagation direction relative to the crystal axes.

2) The Debye approximation considers the thermal energy in acoustical mode where as Einstein approximation considers the thermal energy in optical mode fig (10.3).

3) The atoms of the solid have $3N$ frequencies ranging from 0 to a maximum ω_m. The restriction to the maximum frequency is a consequence of the fact that the waves of the half wave length shorter than interatomic distances ($\lambda/2 < a$) can not be propagated through the crystal.

4) With each frequency Debye associated a distribution function $D(\omega)$ such that $D(\omega)\ d\omega$ represents the number of modes of vibrations lying between the frequency interval ω and $\omega + d\omega$.

5) These modes of vibration are distributed in energy according to the Maxwell-Boltzmann distribution law.

6) Since any vibrational motion of the system can be thought as a superposition of independent normal-mode vibrations, those normal modes are regarded as independent harmonic oscillators whose allowed energy levels are given by $\in_n = (n + \frac{1}{2} v)\ hv$ and whose average energy $\in (\omega)$ is given by $hv\left[\frac{1}{2} + \frac{1}{(e^{hv/kT} - 1)}\right]$.

7) The contribution to the internal vibrational energy of the crystal from these modes of vibration is given by

$$dU = \overline{\in}(\omega)\, D(\omega) d\omega \quad \text{----------} \quad (10.29)$$

The total energy of the crystal is now obtained by integrating (10.29). Thus

$$U = \int_0^{\omega_m} \overline{\in}(\omega)\, D(\omega) d\omega \quad \text{----------} \quad (10.30)$$

where from, the heat capacity can be obtained using

$$C = \frac{dQ}{dT} = T\left(\frac{\partial S}{\partial T}\right) = \left(\frac{\partial U}{\partial T}\right) \quad \text{----------} \quad (10.31)$$

The procedure was first given by Born and von Karman and according to it the problem of calculating heat capacity is essentially of calculating the quantity $D(\omega)$. We therefore proceed to calculate this quantity.

Density of Modes in three Dimensions:

Let us consider a crystal with dimensions (Lx, Ly, Lz). The Boundary conditions are

$$\left.\begin{aligned} S(Lx, y, z, t) &= S(0, y, z, t) \\ S(x, Ly, z, t) &= S(x, 0, z, t) \\ S(x, y, Lz, t) &= S(x, y, 0, t) \end{aligned}\right\} \quad \text{----------} \quad (10.32(a))$$

The equations of motion for the mechanical vibrational amplitude S(x, y, z, t) then lead to vibrational solution of the form

$$S(x, y, z, t) = Ae^{i(\mathbf{k}.r+\omega t)} = Ae^{i\{(\mathbf{k}_x x+\mathbf{k}_y y+\mathbf{k}_z z)+\omega t\}}$$

where $\mathbf{k}^2 = \mathbf{k}_x^2 + \mathbf{k}_y^2 + \mathbf{k}_z^2$

using the boundary conditions we have

$$\mathbf{k}_x = \frac{2\pi n_x}{L_x}; \qquad \mathbf{k}_y = \frac{2\pi n_y}{L_y}; \qquad \mathbf{k}_z = \frac{2\pi n_z}{L_z} \quad \text{----------} \quad (10.32b)$$

where n_x, n_y, n_z are integers (positive or negative). A set of allowed values of $\mathbf{k}_x$, $\mathbf{k}_y$, $\mathbf{k}_z$ corresponds to a single normal mode of vibration satisfying the periodic boundary conditions. The various allowed normal modes can thus be represented as points in an orthogonal **k**-space. Therefore there is one allowed value of k per volume $\frac{8\pi^2}{L_x L_y L_z}$ in **k**-space; or allowed values of **k** per unit volume of **k**-space, for each polarization are

$$\frac{L_x L_y L_z}{8\pi^2} = \frac{V}{8\pi^2}$$

Here L_x, L_y, L_z = V the volume of specimen.

The constant **k** value surface

$$\mathbf{k}^2 = \mathbf{k}_x^2 + \mathbf{k}_y^2 + \mathbf{k}_z^2 = \frac{8\pi^2 n_x n_y n_z}{V}$$

in **k**-space are spheres and hence one may calculate D(**k**)d**k** the allowed number of normal modes in an interval d**k** about the wave number **k** by evaluating the volume of **k**-space between the spherical surface of radius **k** and **k**+d**k**. The required volume is $4\pi\mathbf{k}^2$d**k**. Hence the number of allowed modes are

$$\mathrm{D}(\mathbf{k})\mathrm{d}\mathbf{k} = 4\pi\mathbf{k}^2\mathrm{d}\mathbf{k}\frac{V}{8\pi^2} = \frac{V\mathbf{k}^2}{2\pi^2}d\mathbf{k}$$

For each allowed value of **k** (allowed point in **k**-space) there are three independent normal modes, one corresponding to longitudinal vaibration and two others corresponding to two mutually orthogonal transverse vibrations. To take them into account the right-hand side of the above equation should be increased by a factor of 3 giving

$$\mathrm{D}(\mathbf{k})\mathrm{d}\mathbf{k} = \frac{3V\mathbf{k}^2}{2\pi^2}d\mathbf{k}$$

In order to find D(ω) knowing the relation for D(**k**) requires the knowledge of dispersion relation. Unfortunately such a relation is not known in any convenient form. Further, the mathematical form of equation (10.30) is such that if the actual variation is used even some how or other, the problem of its integration stands up. Debye simplified the problem by assuming the continuum approximation in which phase velocity (v_0) is taken to be constant which he assumed to be equal to the velocity of long wave-length elastic vibrations propagating through the crystal i.e.

$\omega(\mathbf{k}) = v_0\mathbf{k}$

Substituting this in the above equation we have

$$\mathrm{D}(\omega)\mathrm{d}\omega = \frac{3V\omega^2}{2\pi^2 v_0^3}d\omega \quad \text{---------- (10.32c)}$$

Debye formula for specific heat:

Substituting 10.23 and 10.32c in 10.30 we obtain

$$U = \int_0^{\omega_m}\left[\frac{1}{2}\hbar\omega + \frac{\hbar\omega}{e^{\hbar\omega/k_BT}-1}\right]\frac{3V\omega^2}{2\pi^2 v_0^3}d\omega$$

$$= \frac{3\hbar V}{2\pi^2 v_0^3}\int_0^{\omega_m}\left[\frac{1}{2}\omega^2 + \frac{\omega^2}{e^{\hbar\omega/k_BT}-1}\right]d\omega$$

$$= \frac{3\hbar V\omega_m^4}{16\pi^2 v_0^3} + \frac{3\hbar V}{2\pi^2 v_0^3}\int_0^{\omega_m} \frac{\omega^2}{e^{\hbar\omega/k_BT}-1}\,d\omega \quad \text{---------- (10.32d)}$$

The first term (10.32d) is zero point contribution to the internal energy which, however, is independent of temperature and contributes nothing to the specific heat. Let us now put

$$x = \frac{\hbar\omega}{k_BT}$$

and $$x_m = \frac{\hbar\omega_m}{k_BT} = \frac{\hbar v_0}{k_BT}\left(\frac{6N\pi^2}{V}\right)^{1/3} \quad \text{---------- (10.33)}$$

and $$\theta_D = \frac{\hbar\omega_m}{k_B} = \frac{\hbar v_0}{k_B}\left(\frac{6N\pi^2}{V}\right)^{1/3} \quad \text{---------- (10.34)}$$

The parameter θ_D usually refers to as the Debye temperature, is seen to have dimensions of temperature and plays the role of characteristic temperature in much the same way as does the Einstein's temperature θ_E in Einstein theory. It is of course independent of temperature except for slight temperature variation introduced by the variation of v and v_0 with temperature. Substituting (10.34) (10.32d) along with (10.31) gives

$$U = \frac{9}{8}Nk_B\theta_D + 9Nk_BT\left(\frac{T}{\theta_D}\right)^3\int_0^{x_m}\frac{x^3}{e^x-1}dx \quad \text{---------- (10.35)}$$

Differentiating (10.32d) *w.r.t* temperature, the heat capacity can be found. The result is

$$C_v = \frac{3\hbar^2 V}{2\pi^2 v_0^3 k_B T^2}\int_0^{\omega_m}\frac{\omega^4 e^{\hbar\omega/k_BT}}{(e^{\hbar\omega/k_BT}-1)^2}d\omega$$

$$= 9Nk_B\left(\frac{T}{\theta_D}\right)^3\int_0^{\theta_D/T}\frac{x^4e^x}{(e^x-1)^2}dx \quad \text{---------- (10.36)}$$

Tables have been calculated on the Debye theory for U, C_v and other quantities and are given in the Landolt-Bornstein tables and also in the Jahnke-Emde-Losch tables. The heat capacity is plotted in fig (10.4). At $T >> \theta_D$, the heat capacity approaches the classical value $3Nk_B$

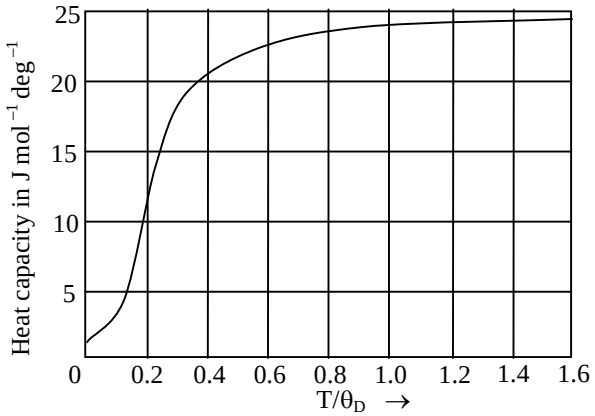


Fig 10.4: Heat capacity C_v of a solid according to Debye approximation. The region of T^3 law is below $0{\cdot}1\theta_D$. The asymptotic values of T/θ_D is $3Nk_B$ (~25 Jmol^{-1}deg^{-1}).

Let us try to evaluate the integral in equation (10.35) at the extreme value of temperature i.e., $T\to 0$ and $T\to\infty$. As $T\to\infty$, $x\to 0$ imply $T >> \theta_D$. Here we can write

$$e^x = 1+x$$

$$\therefore \quad \int_0^{x_m} \frac{x^3}{e^{\propto}-1}dx \to \int_0^{x_m} x^2\,dx = \frac{x^3_m}{3} \quad \text{----------} \quad (10.37)$$

$$\therefore \quad U = \frac{9}{8}Nk_B\theta_D + 9Nk_BT\left(\frac{T}{\theta_D}\right)^2 \cdot \frac{1}{3}\left(\frac{\theta_D}{T}\right)^3 = \frac{9}{8}Nk_B\theta_D + 3Nk_BT \quad \text{----------} \quad (10.38)$$

and $C_v = 3Nk_B = 3R$ ---------- (10.39)

The Dulong and Petit law in agreement with both classical and Einstein approximations.

Physically this means that at high temperatures ($T >> \theta_D$) there is sufficient thermal energy to excite all $3N$ modes regardless of their frequency so that they are contributing to C_v. It is thus equivalent to the Einstein approximation because the frequency is unimportant and to the classical approximation because the quantum discontinuities are negligible.

At low temperatures $T\to 0$ implies that $x\to\infty$ so that we can let the upper limit of the integral can be solved by contour integration method and comes out to be equal to $\frac{\pi^4}{15}$ i.e.

$\int_0^{\infty} \frac{x^3}{e^{\propto}-1}dx = \frac{\pi^4}{15}$, giving

$$U = \frac{9}{8}Nk_B\theta_D + 9Nk_BT\left(\frac{T}{\theta_D}\right)^3\frac{\pi^4}{15}$$

$$= \frac{9}{8}Nk_B\theta_D + \frac{3}{5}\pi^4 Nk_BT\left(\frac{T}{\theta_D}\right)^3 \quad \text{----------} \quad (10.40)$$

and $C_v = \frac{12}{5}\pi^4 N k_B \left(\frac{T}{\theta_D}\right)^3$ ---------- (10.41)

This exhibits the Debye T^3 law. Its predictions for specific heat in the low temperature region are in good agreement with experimental data for many substances. The cubic dependence may also be appreciated from the following qualitative argument.

At low temperatures only a few of these modes are excited. These are the modes whose quantum energy $\hbar\omega$ is less than k_BT. The number of these modes may be estimated by drawing a sphere in the k-space whose frequency $\omega = k_BT/\hbar$, and counting the number of points inside as shown in fig (10.5). This sphere may be called as thermal sphere. The number of modes inside the thermal sphere is proportional to $K^3 \sim \omega^3 \sim T^3$. Each mode is fully excited and has an average energy equal to k_BT. Therefore the total energy of excitation is proportional to T^3.

The value of the Debye temperature θ_D may be calculated from (10.33). Methods for determination of a suitable average sound velocity to be used in calculating θ_D have been given by Blackmann. The value of θ_D can also be obtained experimentally by choosing that value of θ_D, which leads to the best fit between the experimental data and the theoretical expression (10.35). For the most common metallic elements the Debye temperature generally lies in the range 150–450*K*.

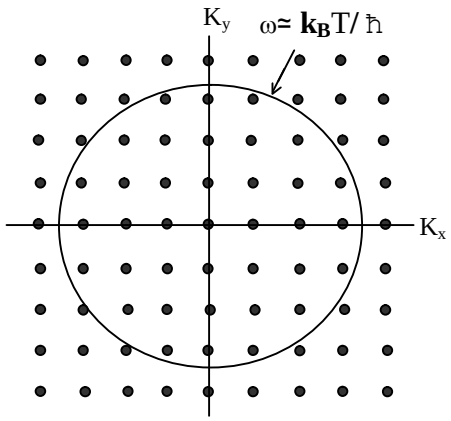


Fig 10.5: The thermal sphere which is the frequency contour $\omega = k_BT/\hbar$.

For actual lattice, the temperature at which the T^3 approximation holds true are quite low. It may be below $T = \theta_D/50$ to get reasonably pure T^3 behaviour. The heat capacity is, however, relatively insensitive to changes in the density of modes.

Criticism of Debye model:

1) The continuum model is valid for long wave-lengths i.e., the model assumes that only low frequencies are active in the solid.

2) Debye assumed that the number of modes of vibrations are $3N$ but this is contradictory to the assumption that solid is an elastic continuum, which is supplied to possess infinite frequencies i.e., there is little justification for the equation

$$\int_0^{\omega_m} D(\omega)d\omega = 3N$$

3) ω_m is taken to be the same for longitudinal and transverse waves. But this assumption can not be justified on physical grounds because of the difference in nature of the two waves. Born however proposed to cut off the spectrum in such a manner that the longitudinal and transverse modes have a common minimum wavelength (λ_{min}). So that one obtains two Debye frequencies one for longitudinal modes ($\nu_t = c_t/\lambda_{min}$) and other for transverse modes ($\nu_t = c_t/\lambda_{min}$). This is actually more sound theoretically speaking and in line with the theory of lattice vibrations developed by Born and von Karmann.

4) If the Debye model were strictly valid, the values of θ_D varies with temperature (table 10.1). The deviation reaching as much as 10% or even more in some cases.

5) Actually crystalline nature of the solid has not been taken into account.

TABLE 10.1: Deviation from the T^3 law

NaCl			KCl			Li		
T	θ_D	$10^4 C_v/T^3$	**T**	θ_D	$10^4 C_v/T^3$	**T**	θ_D	$10^4 C_v/T^3$
20^0K	288^0K	0.388	14^0K	213^0K	0.960	30^0K	288^0K	0.101
15^0K	297^0K	0.356	8^0K	222^0K	0.832	20^0K	297^0K	0.118
10^0K	308^0K	0.334	4^0K	236^0K	0.708	15^0K	308^0K	0.131
			3^0K	227^0K	0.798			

10.5 Summary:

- Experiments show that the heat capacity of conduction electrons is much smaller than predicted by classical statistical mechanics. This is explained on the basis of exclusion principle. All the energy levels up to the Fermi level are occupied and when the system is heated, only those electrons near the Fermi levels are excited. The electron heat capacity $\frac{\pi^2}{2}\frac{Nk_BT}{T_F}$.
- The Fermi energy $E_{f(0)}$ is function of temperature and varies as

$$E_f = E_{f(0)}\left[1-\frac{\pi^2}{12}\left(\frac{k_B T}{E_{f(0)}}\right)^2\right].$$

- The atoms in the lattice are arranged as a set of harmonic oscillators and the thermal energy is the average energy of these oscillators.
- According to classical theory, the average energy for one dimensional oscillator is $\bar{\varepsilon} = KT$ and the total thermal energy per mole is $U = 3Nk_BT$ and the molar specific heat $C_v = \frac{\partial U}{\partial T} = 3Nk_B M = 3R.$
- This result, known as Dulong and Petit law, asserts that C_v is a constant independent of temperature. This is found to be valid only at high temperatures and at low temperatures the specific heat decreases and then vanishes at $T = 0K$.
- Einstein rectified this discrepancy by treating the oscillator quantum mechanically. The average thermal energy for the oscillator is then given by $\bar{\varepsilon} = \frac{h\omega}{e^{h\omega/k_BT}-1}$ which approaches the classical value k_BT only at high temperatures. At low temperatures, the quantum mechanical energy decreases very rapidly because of the freezing of the motion.
- Treating the atoms as independent oscillators, vibrating with common frequency, Einstein found that the specific heat is $C_v = 3R\left(\frac{\theta_E}{T}\right)^2 \frac{e^{\theta_E/T}}{e^{\theta_E/T}-1}$.
- Specific heat approaches the classical value $3R$ at high temperatures and vanishes at $T = 0K$. Both these facts are in accord with experiment.
- Careful measurements show that the decrease in C_v near absolute zero is slower than predicted by Einstein. Debye explained this by treating the atoms not as independent oscillators but as coupled oscillators vibrating collectively as sound waves.
- Making the long-wavelength approximation, Debye found that the specific heat is given by $C_v = 9R\left(\frac{T}{\theta_D}\right)^3 \int_0^{\theta_D/T} \frac{x^4 e^x}{(e^x-1)^2} dx$.
- This expression for C_v approaches the classical value $3R$ at high temperatures but at low temperatures $C_v \sim T^3$. This latter result is known as Debye T^3 law and is in agreement with observation.

- The agreement of Debye model with experiment is good through out the entire range of temperature. Better agreement can be achieved only by removing the long wave-length approximation and treating the crystal as discrete lattice.

10.6 Key words:

Drude – Lorentz model – The Einstein model – The Debye model – Thermal sphere.

10.7 Review Questions:

1. Explain in detail the thermal capacity of free electron system.Draw the plot of T^2 versus C_v and explain the comparison between theory and experiment.
2. Distinguish between classical and quantum mechanical aspects of heat capacity.
3. Explain Einstein model, draw the merits and demerits of Einstein model of specific heat of solids.
4. Explain Debye model of heat capacity, draw the merits of Debye model over Einstein model and also explain the criticism of Debye model.
5. Derive an expression for the electronic contribution to the specific heat of a metal. How is it verified experimentally?
6. Explain phonon dispersion curves. How are they studied experimentally?
7. Derive an expression for the specific heat of a solid on the Einstein model and show that at low temperatures it drops exponentially with decreasing temperature.
8. What are Einstein and Debye models for the specific heat of solids?. Derive T^3 law for the specific heat of solids at low temperatures.
9. What are optical and acoustc phonons?. Discuss with theory any experiment to determine the properties of individual phonons.
10. Discuss the specific heat of a 2-D square lattice on the basis of Debye approximations. Compare the Debye, Einstein and lattice mode theories of specific heats of solids. Comment on Reshtrhlen frequency.

10.8 Text and Reference Books:

1) Elements of Solid State Physics by J P Srivatsava (PHI)
2) Solid State Physics by M A Wahab (Narosa)
3) Elements of Solid State Physics by A. Omar (Pearson education)
4) Solid State Physics by S O Pillai (New age)
5) Solid State Physics by C Kittel (Asian students edn.)
6) A Text Book of Solid State Physics by S.L.Kakani & Hemrajani (S.Chand)
7) Solid State Physics by Saxena Gupta Saxena (Pragati Prakashan)

UNIT – III
LESSION: 11

Thermal properties of solids - III

Aim: To explain the thermal properties of solids on the basis of lattice vibrations in the harmonic approximation and to consider the anharmonic effects.

Objective of the lesson

- To distinguish between classical and quantum mechanical concept of heat capacity.
- To consider anharmonic effects of lattice vibration to look into thermal expansion, finite value of thermal conductivity, temperature dependence of elastic constants. To know about the phonon collision process, phonon thermal conductivity, Size effect.

Structure of the lesson:

Introduction:

Thermal properties of solids can be classified into two groups.

1. Properties whose broad features can be explained in the harmonic approximation.
2. Properties whose broad features have to be explained considering anharmonic effects also

11.1 Classical heat capacity:

The major contribution to the heat capacity of solids comes from the lattice vibrations. In non-magnetic insulators, lattice vibrations are the only contributors. Other contributions come from conduction electrons in metals and ordering in magnetic materials. Here we deal with lattice vibrations only.

Classically, the energy of solid, can be expressed as

$$\varepsilon = \varepsilon_s + \varepsilon_v \quad \text{----------} \quad (11.1)$$

where

ε_s is the energy of the static lattice

and ε_v is the vibration energy of the solid.

The classical equipartition energy of vibration of atom is k_BT at temperature T, k_B being the Boltzmann constant. This gives the total vibration energy $3Nk_BT$. If there are N atoms in the solid,.substituting this into (11.1) We have

$$\varepsilon = \varepsilon_s + 3Nk_BT \quad \text{----------} \quad (11.2)$$

At $T = 0$, $\varepsilon = \varepsilon_s$ which does not include the zero point energy and in classical theory, there is no concept of zero point motion. The heat capacity is defined as (the value at constant volume).

$$C_v = \left(\frac{\partial \varepsilon}{\partial T}\right)_V \quad \text{----------} \quad (11.3)$$

Using (11.2), we can have

$$C_v = 3Nk_B \quad \text{----------} \quad (11.4)$$

The experimentally measured value of heat capacity is referred as the heat capacity at constant pressure C_P. But for a harmonic solid $C_V = C_P$, since the difference between the two is known to depend on the square of the temperature coefficient of linear expansion which is zero for harmonic crystal. Later it will be known that the expansion of solids occurs because of the anharmonic motion of atoms. The equation of heat capacity given

by (11.4) was originally derived by Dulong and Petit in 1869. This is known as the Dulong–Petit law. It states that the heat capacity is constant for a solid and independent of temperature. The measured value of heat capacity approaches this value at high temperatures and remains constant thereafter in the solid state (Fig11.1). Dulong and Petit explained the behavior of solids at high temperatures. This law, however, fails to explain the variation of heat capacity with temperature. Most significantly, the heat capacity drops to zero at 0^0K. The failure of Dulong–Petit law's to hold at all temperatures is later explained only with the use of quantum theory

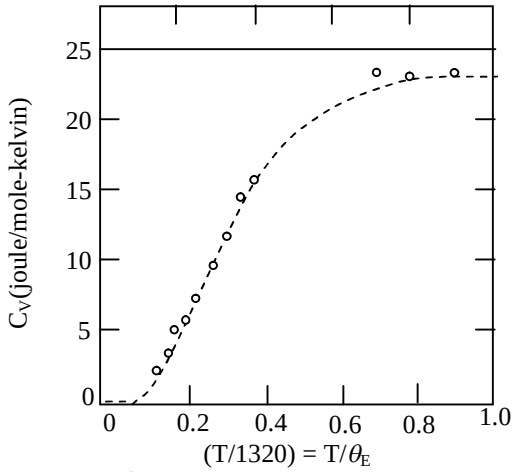


Fig 11.1 A comparison of the experimental curve (represented by the circles) with Einstein's theoretical curve (dotted) for molar heat capacity of diamond with Q_E = 1320K. The horizontal line gives the classical Dulong-Petit value.

11.2 Quantum Theory of Heat Capacity:

Here we build up the structure for calculating the heat capacity on the principles of quantum theory. Since lattice vibrations are pictured as quantized in this theory, the term phonon heat capacity is more appropriate than the conventional term. Phonon heat capacity is mainly concerned with the evaluation of the average thermal energy, which in turn requires the knowledge of the vibration spectrum of solid under study. Our first attempt would be to calculate the average thermal energy of harmonic oscillator.

11.2.1: Average Thermal Energy of a Harmonic Oscillator:

The energy levels of a quantum harmonic oscillator are given by

$$\varepsilon_n = \left(n + \frac{1}{2}\right)\hbar\omega \quad \text{----------} \quad (11.5)$$

where ω is the angular frequency of the oscillator.

The average thermal energy of the oscillator ε in thermal equilibrium at temperature T is

$$\varepsilon = \sum_n P_n \varepsilon_n \quad \text{----------(11.6)}$$

Where P_n represents the probability of finding the oscillator in the energy level ε_n. It is given by

$$P_n \propto \exp\left(-\frac{\varepsilon_n}{k_B T}\right) \text{ (the Boltzmann distribution) ---------- (11.7)}$$

The proportionality constant ensures that the oscillator is in one of the allowed levels. Therefore

$$\sum_{n=0}^{\infty} \mathbf{P}_n = 1 \text{----------} \quad (11.8)$$

or $$\sum_{n=0}^{\infty} \mathbf{P}_0 \exp\left[-\frac{\left(n+\frac{1}{2}\right)\hbar\omega}{k_B T}\right] = 1$$

where P_0 is the constant of proportionality. Therefore,

$$\mathbf{P}_0 = \exp(-\hbar\omega / 2k_B T) \sum_{n=0}^{\infty} \exp(-\hbar\omega\, n / k_B T) = 1$$

$$\text{or} \qquad \mathbf{P}_0 = \frac{\exp(\hbar\omega / 2k_B T)}{1 - \exp(-\hbar\omega / k_B T)} \quad \text{----------} \quad (11.9)$$

Therefore,

$$\mathbf{P}_n = \exp(-n\hbar\omega / k_B T)(1 - \exp(-\hbar\omega / k_B T)) \quad \text{----------} \quad (11.10)$$

Substituting P_n from (11.10) in (11.6), we get

$$\varepsilon(\omega T) = [1 - \exp(-\hbar\omega / k_B T)]\hbar\omega \sum_{n=0}^{\infty}\left(n + \frac{1}{2}\right)\exp(-\hbar\omega\, n / k_B T)$$

Putting exp $(-\hbar\omega / k_B T)$ = x, we have

$$\varepsilon(\omega, T) = (1-x)\hbar\omega\left[\sum_{n=0}^{\infty}(nx^n) + \frac{1}{2(1-x)}\right] = \frac{1}{2}\hbar\omega + (1-x)\hbar\omega \sum_{n=0}^{\infty} nx^n \quad \text{----------} \quad (11.11)$$

Since

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \quad \text{----------} \quad (11.12)$$

on differentiating (11.12) and then multiplying both sides by x, we get

$$\sum_{n=0}^{\infty} nx^n = \frac{1}{(1-x)^2} \quad \text{----------} \quad (11.13)$$

From (11.13) and (11.11), we have

$$\varepsilon(\omega,T)=\left(\frac{1}{2}+\frac{1}{\exp(\hbar\omega/k_BT)-1}\right)\hbar\omega \quad \text{---------- (11.14)}$$

or

$$\varepsilon(\omega,T)=\left(\langle n\rangle+\frac{1}{2}\right)\hbar\omega \quad \text{---------- (11.15)}$$

where

$$\langle n\rangle=\frac{1}{\exp(\hbar\omega/k_BT)-1} \quad \text{---------- (11.16)}$$

<n> must denote the phonon occupancy. This is also known Planck's distribution function obtained by Planck for the radiation oscillators. The above exercise is a simple demonstration of the fact that lattice vibration one quantized the same way as the radiation oscillators. The <n> stands for the occupancy of phonons in the energy level ε_n which is also interpreted as the expected number of photons in energy state ε_n of an oscillator in thermal equilibrium at temperature *T*.

The calculation of total thermal energy of solid is, not a straightforward one. Practically it is done by making use of dispersion curves which provide the valuable information on the number of modes per unit frequency interval. The term giving this information is known as the density of states. This model proposed by Einstein, serves as a landmark in the progress of the theory of phonon heat capacity.

10.3 Anharmonic effects:

While dealing with lattice vibrations, the harmonic vibration is the basis for discussion. Keeping aside the exceptions like solid helium in which harmonic approximation is not applicable, we come across a number interesting physical phenomena that cannot be explained without relaxing the approximation. Some of these are

(1) Thermal expansion
(2) Thermal conductivity has a finite value
(3) The temperature dependence of elastic constants
(4) The failure of Debye heat capacity to attain the classical value at high temperatures ($T>>\theta_D$)
(5) The line broadening or measurable width of one phonon peaks in the neutron inelastic scattering pattern.

The cause of failure of the harmonic approximation in explaining the above phenomena rests in questionable relevance of two basic assumptions.

1. Vibrations of atoms may be treated as small oscillations in which the displacement of atoms from their respective equilibrium positions are small.

2. Only the leading non-vanishing term in the expansion of potential energy about its equilibrium is retained.

The validity of the first assumption is always questionable at high temperatures where the amplitude of vibration is large. And when the displacements become large, we cannot afford to ignore the higher order terms in displacement beyond the quadratic term in potential energy expansion. In this situation both the assumptions become irrelevant. This is indicated by the shape of the potential energy curve at large interatomic separations (fig11.2).

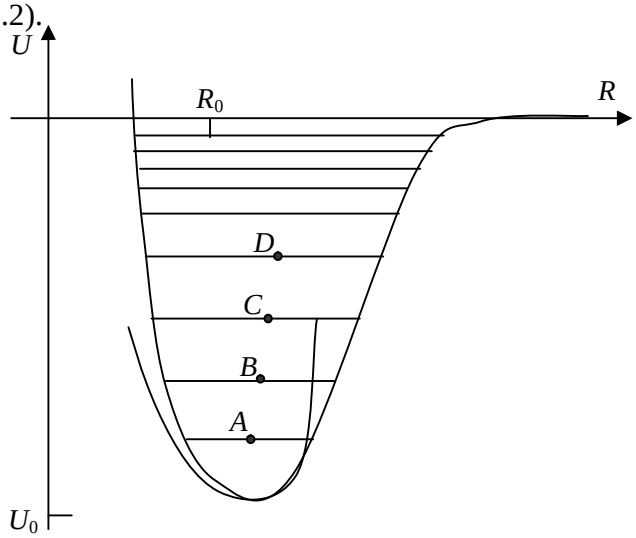


Fig 11.2: An exaggerated potential energy curve based on Lennard-Jones potential for real crystals. The asymmetrical nature of the curve at higher energies does not allow the motion to be interpreted in parabolic approximation. The increase in the interatomic separation R at high energies (excited at high temperatures) produces thermal expansion. The R_0 is the interatomic distance in the ground state. Because of anharmonicity, the vibration quantum (the spacing between vibrational levels) does not remain constant. It decreases with increase in energy

Where it is more asymmetric showing thereby deviation from the harmonic nature. This is understandable as atoms can never oscillate like independent harmonic oscillators in a real crystal because the motion of the adjacent atoms is always correlated. In general we must accept the presence of anharmonicity to a certain degree in the vibrations of a solid.

The physical phenomena stated earlier are explained successfully when connections on the requirement of anharmonicity are included in the calculations.

The anharmonic effects are most reasonably accounted when the potential energy expansion is truncated not before the quartic term. Dropping the constant equilibrium term, we write potential energy as

$$U(x) = fx^2 - gx^3 - hx^4 \quad \text{----------} \quad (11.17)$$

Where x is the deviation from the equilibrium separation at absolute zero and all the coefficients f, g and h are positive.

The first term in (11.17) is the usual harmonic component and the other two refer to the anharmonic effects. The cubic term stands for the asymmetry of the mutual repulsion of the atoms and the quartic term represents the softening of the vibration at large amplitudes.

As the applications of anharmonic potential we will deal with first two of the phenomena cited earlier as examples. While the thermal expansion refers to an equilibrium property the thermal conductivity is a well-known transport phenomenon.

11.3.1 Thermal Expansion:

It is an established fact that solids expand on heating. In the harmonic approximation, all atoms vibrate about their equilibrium positions within a perfectly symmetric well (parabola) of interatomic interaction. But the potential energy curve (fig.11.2) as derived from a Lennard–Jones potential type for real solids, matches this behaviour only in the range low thermal excitations that correspond to low temperatures.

The potential well gets asymmetrical for larger interatomic separations (fig11.2) occurring at higher energies. The mean interatomic separations at a few vibrational energies taken in increasing order are denoted by points A, B, C, D respectively. The values represented by these points are in increasing order as shown by the trend of shift forward larger values relative to the mean equilibrium separation R_0 in the ground state. It is then imperative that at higher temperatures when higher vibrational states are sufficiently populated, solids would show expansion. The range of vibration frequencies of solids is such as can be excited by the thermal energy. These arguments explain the thermal expansion of solids. Since this property follows from the anharmonic nature of

the potential energy curve, the thermal expansion is attributed to the presence of anharmonicity in atomic vibrations.

Now we use (11.17) to calculate the thermal expansion in terms of the mean displacement $\langle x \rangle$ of the atoms in a solid. It is nonzero because of the cubic term. A crude way of determining $\langle x \rangle$ is to use the condition,

$$\frac{\partial U}{\partial x} = 0$$

and ignore the term in h. This gives

$$\langle x \rangle = \frac{3g\langle x^2 \rangle}{2f} \quad \text{----------} \quad (11.18)$$

If the mean square displacement is calculated classically in the harmonic approximation, we get

$$\langle x^2 \rangle = \frac{\int_{-\infty}^{+\infty} x^2 \exp(-fx^2 / k_B T)dx}{\int_{-\infty}^{+\infty} \exp(-fx^2 / k_B T)dx} = \frac{\frac{k_B T}{4f}\sqrt{\frac{\pi}{f / k_B T}}}{\frac{1}{2}\sqrt{\frac{\pi}{f / k_B T}}}$$

or

$$\langle x^2 \rangle = \frac{k_B T}{2f} \quad \text{----------} \quad (11.19)$$

Substituting (11.19) into (11.18), we express the mean displacement as

$$\langle x \rangle = \frac{3gk_B T}{4f^2} \quad \text{----------} \quad (11.20)$$

The relation (11.20) shows an increase in the interatomic separations and therefore, a thermal expansion of solid.

We may define the coefficient of linear expansion α as

$$\alpha = \frac{\partial \langle x \rangle}{\partial T} \quad \text{----------} \quad (11.21)$$

Using (11.20), we have

$$\alpha = \frac{3gk_B}{4f^2} \quad \text{----------} \quad (11.22)$$

Though changes in size because of thermal expansion are small in solids, the knowledge of their expansion coefficients in none the less of great practical value in industry where even simple jobs like making a permanent joint between two materials have their tactical importance

Another striking feature that is associated with the thermal expansion, has reference to the change of characteristic frequency in a solid. The levels of vibrational energy come closer at high-energy values (fig.11.2). In the harmonic approximation the levels are expected to remain evenly spaced and unaffected by the change of temperature. The spacing can alter only of the energy quantum $h\omega$ changes. By taking anharmonic considerations into account with the use of (11.17), it can be shown that the characteristic frequency ω decreases with the increase in temperature. This is consistent with observation of closer energy levels at high energies to which the thermal excitation can be made by raising the temperature one must appreciate that the anharmonic contributions at these values of energy and temperature are at their maximum.

The above frequency effect may as well be interpreted as the change in characteristic frequencies in a solid because of the change in its volume. It is conventionally put in the form of an assumption that allows both the characteristic frequencies and the volume V to suffer the same relative change. That is,

$$\frac{\Delta\omega}{\omega} \propto \frac{\Delta V}{V}$$

or

$$\frac{\Delta\omega}{\omega} = \gamma \frac{\Delta V}{V} \quad \text{---------- (10.23)}$$

Where the proportionality constant γ is called Grümeisen constant. It is a measure of the anharmonic coupling. This anharmonic effect, though small, may be studied by neutron scattering measurements that provide the change in characteristic frequencies with temperature.

An exact treatment of anharmonic effects is not so easy. We have to take recourse to some approximate method like perturbation method of quantum mechanics. As a first approximation to the true solution, we start with solutions of harmonic potential (the phonons). Obviously, the phonons are not the exact eigensolutions to the equation of motion. The description of a state of motion in an anharmoinic solid in terms of a phonon or a plane wave changes with time and the accuracy of the description drops progressively. For an accurate description with time, it is required to recognize the status of a spectrum of some other phonons. It means that the anharmonic phonons, unlike harmonic ones, have only a limited lifetime after which they merge or decay to produce

new phonons. In the next section we discuss these phonon processes as they enable us to explain the basis for thermal resistance and related features in the next section.

10.3.2 Phonon Collision Process:

An exact quantum mechanical treatment based on the first order perturbation theory shows that:

a) The cubic term in the potential energy accounts for the following processes.

i) One phonon decay into two

ii) Phonons merge into one

b) The quantum term in the potential energy accounts for the following processes:

i) One phonon decays into three

ii) Two phonons gets converted into two others

iii) Three phonons merge into one.

We observe that the cubic term is related to the three-phonon processes and the quartic to the four-phonon processes. It is now necessary to examine the probable collision processes, especially in the context of thermal resistance which is in our plan of discussion. Let us consider the collision involving three phonons. The use of the term collision should not be objectionable as phonons are treated as particles. If two phonons have wave vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ collide or merge to produce a new phonon with a wave vector $\mathbf{k}_3$, the momentum conservation requires that

$$\mathbf{k}_1 + \mathbf{k}_2 = \mathbf{k}_3 \quad \text{----------(11.24)}$$

The total energy of the initial phonons is completely held by the final phonon. The net phonon momentum is written as

$$\mathbf{j} = \sum_{\mathbf{k}s} n_{\mathbf{k}s} \hbar \mathbf{k}_s \quad \text{----------(11.25)}$$

where n_{ks} is the number of phonon with wave vector **k** in the branch s.

The net phonon momentum **j** constitutes the phonon thermal current in a solid between whose two ends a difference of temperature is maintained. In view of (11.24) the net phonon moment is a constant and, therefore, the phonon current **j** remains unchanged. The distribution of phonons (or the distribution function) at any stage when $\mathbf{j} \neq 0$ tends to appear its equilibrium value (condition of zero temperature). But in the present case it is not permitted as **j** has to remain unaffected. Any distribution of phonon

flows down the solid unhindered. As if the phonon mean free path were infinite. Consequently the thermal resistance, a measure of the rate at which equilibrium distribution is approached, becomes zero. This gives an infinite value of thermal conductivity an unacceptable result. Therefore the collision processes described by (11.24) do not give rise to thermal resistance. These processes are called normal processes or simply N-processes. Figure 11.3(a) shows a normal process of collision in a two dimensional square lattice.

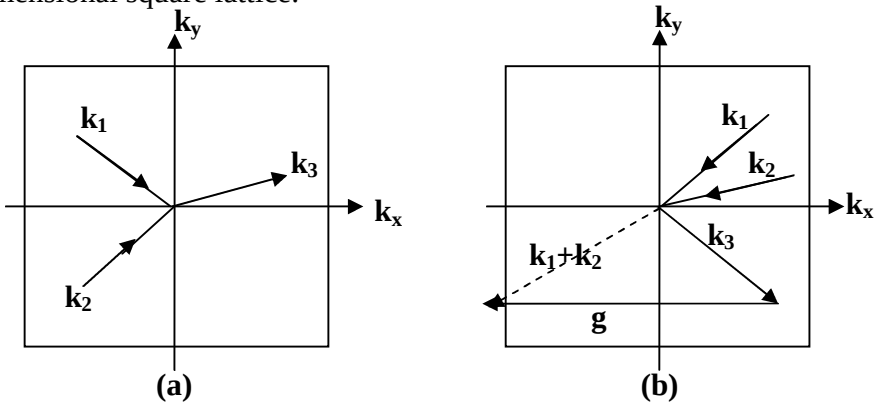


Fig 10.3: Phonon scattering processes in k-space for a two-dimensional square lattice: (a) normal process. (b) Umklapp process – two phonons with wavevector k_1 and k_2 merge into k_3 with the help of a reciprocal lattice vector g (= $2\pi/a$ in this case). Note that the direction of the x-component of the phonon flux has been reversed.

It should be, however surprising to note that N-processes explain the conduction of heat in gases where we deliberately impose the conditions that no mass transport of particles takes place. In this picture, the hot particles while moving along one direction lose their energy to the cold particles moving in the opposite direction [fig 11.4(a)].

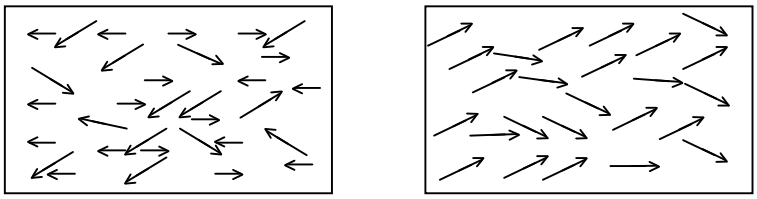


Fig 11.4: (a) Conduction of heat in a gas: hot electrons moving in one direction lose their energy to cold electrons moving in the opposite direction. During the conduction process the number of electrons is conserved. (b) Conduction of heat by phonons in a solid: There is no net particle conservation. Phonons carrying heat from the hot end get destroyed at the cold end.

But these ideas are not tenable in the case of phonon gas where there is no net particle conservation in the true picture. The phonons that carry heat from the hot end one destroyed at the cold end [fig 11.4 (b)].

The phonons could also suffer collisions with the immobile imperfections and the crystal bounding surface. But these collisions are known to be elastic in nature, meaning there by that the collisions do not change the frequency of the individual phonons. This makes impossible for an equilibrium distribution of phonons to be established locally.

The puzzle of thermal resistance is solved by a different kind of collisions expressed as

$$\mathbf{k}_1 + \mathbf{k}_2 = \mathbf{k}_3 + \mathbf{g} \quad \text{---------- (11.26)}$$

Where $\mathbf{g}$ is a reciprocal lattice vector.

Since in a periodic lattice the energy of a phonon with a wavevector ($\mathbf{k}_3 + \mathbf{g}$) is the same as that for a phonon with a wavevector $\mathbf{k}_3$, the phonon $\mathbf{k}_3$ in (11.26) must be carrying total energy. The specific feature of (11.26) is that it destroys the momentum $\hbar\mathbf{g}$ and changes the direction of energy flow as show in fig 11.4(b). The *x*-component of $\mathbf{k}_1$ and $\mathbf{k}_2$ are directed opposite to the *x*-component of $\mathbf{k}_3$. Peierls called such collisions by the name 'Umklapp processes' or 'U-processes'. Umklapp is the German term for 'folding over'. Because of the non-conservation of the net momentum, the equilibrium distribution of phonons can be restored by these processes at a certain rate that determines the thermal resistance. This confirms that the U-processes of collisions give rise to thermal resistance, limiting the value of thermal conductivity to a finite value.

For the condition (11.26) to be satisfied, the sum of $\mathbf{k}_1$ and $\mathbf{k}_2$ must extend beyond the first Brillouin zone and the individual values should not be less than ½ $\mathbf{g}$. The sum of ($\mathbf{k}_1 + \mathbf{k}_2$) can be translated to the first zone by the reciprocal lattice vector – $\mathbf{g}$ [fig 10(b)]. The $\mathbf{k}_3$ is called the reduced value of ($\mathbf{k}_3 + \mathbf{g}$) and belongs to a set of unique $\mathbf{k}$-values. U-processes can be taken up while discussing the effect of temperature on thermal conductivity. In the refined theoretical picture, a phonon collision is referred to as the scattering of phonons because of phonon–phonon interaction.

11.3.3 Phonon Thermal Conductivity:

Metals are good conductors of heat and electricity. The energy is transported mainly by the electrons. But it must not be constructed that thermal conductivity depends mainly on

the number of available free electrons. Some of the insulators for example the crystalline sapphire (Al_2O_3) and quartz (SiO_2) possess high thermal conductivity than copper at low temperature. The maximum value of thermal conductivity of sapphire is 200 $Wcm^{-1}k^{-1}$ as against corresponding value 100 $Wcm^{-1}k^{-1}$ of copper. The observations indicate to the involvement of some other carriers in addition to the free electrons for the heat transport. These carriers are phonons as confirmed by the theory explaining the experimental data. Thus the phonons are unambiguous by the main carriers of heat in insulators and in some of the insulators the passage of phonons is smooth enough so as to result in large values of thermal conductivity.

An exact treatment of phonon thermal conductivity is a problem of great mathematical complexity, we present an elementary classical based on kinetic theory of gases.

Consider an insulating solid cylinder (fig11.5) whose two ends are maintained at different temperatures T_1 and T_2 (T_2>T_1). Let the temperature gradient down its length be defined as $-\frac{\partial T}{\partial x}$. At steady state there is a steady flow of heat from the hot and to the cold end maintaining the temperature difference (T_2-T_1) at a constant value.

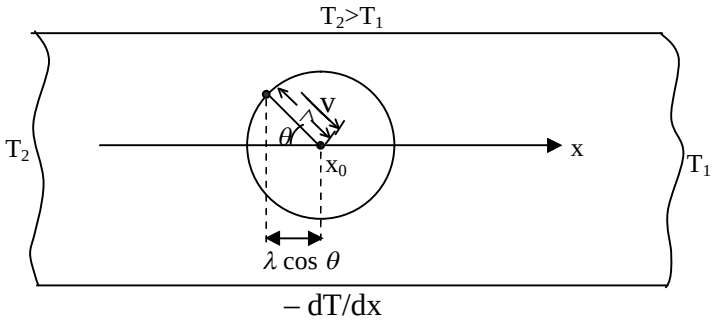


Fig 11.5: Conduction of heat by phonons in an insulating solid cylinder along its length when a temperature gradient is maintained between the two ends of the cylinder. Phonons approaching the point x_0 at an angle θ with the cylinder's axis (x-direction) and moving with velocity v make a collision at the point. The temperature at the point is T.

It is observed that the energy flowing down the cylinder per unit of its circular cross section per unit time *j* is proportional to the temperature gradient $-\frac{\partial T}{\partial x}$. That is

$$\mathbf{j} \propto -\frac{\partial T}{\partial x} \quad \text{----------} \quad (11.27)$$

or

$$\mathbf{j} = -K_{ph}\left(\frac{\partial T}{\partial x}\right) \quad \text{----------} \quad (11.28)$$

where $\boldsymbol{K}_{ph}$ is called thermal conductivity and $|\mathbf{j}|$ is the thermal flux per unit time.

The flow of heat cannot be linear as in that case **j** would have been proportional to T. The heat energy differs from the hot end to the cold end through a random process. As we consider here only the phonons as heat carriers, the random process involve mainly a phonon-phonon collisions. The extra phonons created at hot end are destroyed at the cold end this picture of phonon gas better suited to be treated by the kinetic theory of molecular gases.

Frequently collisions among phonons limit the phonon mean free path. In a rigorous theory the mean freepath is treated as a function of the temperature and the phonon frequencies. But we take an average value (Λ) over the length of the solid cylinder. Let the non-equilibrium energy density contributed by phonons coming out collisions at the point x be denoted by $u(x)$. If T be the temperature at x, $u(x)$ is supposed to be proportional to the equilibrium energy density $u[T(x)]$. Then, the contribution to thermal current density j from a single phonon is $v_x \cdot u(x)$, where v_x is the x-component of the phonon velocity v.

Fig 10.5 shows that the collision at x_0 involves the phonons approaching it at an angle θ with the x-axis. The immediate last collision of these phonons on the average should have occurred at a point whose x-coordinate is shorter by $\Lambda\cos\theta$, i.e. at $x = (x_0 - \Lambda\cos\theta)$. The net thermal current j can be calculated by using the fact that j is proportional to the product $v_x \cdot u(x_0 - \Lambda\cos\theta)$ average over all solid angles. Accordingly,

$$j = \langle v_x \cdot u(x_0 - \Lambda\cos\theta)\rangle$$

$$= \int_0^{\pi} v\cos\theta\, u(x_0 - \Lambda\cos\theta)\frac{2\pi\sin\theta}{4\pi}d\theta = \frac{1}{2}\int_{-1}^{1} vtu(x_0 - \Lambda t)dt \quad \text{----------} \quad (11.29)$$

Expanding $u(x_0 - t)$ absent u(x_0) and retaining only the first order terms in the derivative, we have

$$j = \frac{1}{2}\int_{-1}^{1} vt\left[u(x_0) - \Lambda t\frac{\partial u}{\partial x}\bigg|_{x_0}\right]dt = 0 - \frac{1}{2}v\Lambda\frac{\partial u}{\partial x}\left[\frac{t^3}{3}\right]_{-1}^{1} = -\frac{1}{3}v\Lambda\frac{\partial u}{\partial x}$$

or

$$j = \frac{1}{3}v\Lambda\frac{\partial u}{\partial T}\left(-\frac{\partial T}{\partial x}\right) \quad \text{----------} \quad (11.30)$$

Comparing (11.30) with (11.28), the contribution of phonons to the thermal conductivity may be expressed as

$$K_{\text{ph}} = \frac{1}{3} v \Lambda \frac{\partial u}{\partial T}$$

or

$$K_{\text{ph}} = \frac{1}{3} C_V \Lambda v = \frac{1}{3} C_V v^2 \tau \quad \text{----------} \quad (11.31)$$

Where C_V denotes the phonon heat capacity and τ is the relaxation time ($v\tau = \Lambda$).

We can define the phonon collision frequency as τ^{-1}. The variation of thermal conductivity with temperature in terms of the parameters Λ and C_V is explained below.

At high temperatures (T>> θ_D)

In its temperature range, the thermal equilibrium phonon occupation numbers given by

$$\langle n_k \rangle = \frac{1}{\exp(\hbar\omega(\mathbf{k}) k_B T) - 1}$$

reduce to

$$\sim \frac{k_B T}{\hbar\omega(\mathbf{k})} \quad \text{----------} \quad (11.32)$$

The total number of phonons is proportional to the temperature. Therefore, the higher the temperature, the more will be the collision frequency, resulting in smaller mean free paths. Since C_V approaches the constant Dulong-Petit value, the change in the thermal conductivity is predominantly controlled by the change in the mean free path. Thus, the thermal conductivity decreases with increase in temperature. Though this trend is confirmed by experiments, the observed rate of fall is given by

$$K_{\text{ph}} \propto \frac{1}{T^m} \text{----------} \quad (11.33)$$

where *m* lies between 1 and 2.

At low temperatures (T<<θ_D)

The phonon occupying in thermal equilibrium may be taken as ~ exp ($-\hbar\omega(\boldsymbol{k})k_B T$). The probability of occurrence of U-process [as expressed by (11.26)] is written as

$$n_{\mathbf{k}_1} \cdot n_{\mathbf{k}_2} = \exp(-\hbar\omega_{\mathbf{k}_1} / k_B T) \cdot \exp(-\hbar\omega_{\mathbf{k}_2} / k_B T)$$

$$= \exp[-\hbar(\omega_{\mathbf{k}_1} + \omega_{\mathbf{k}_2}) / k_B T = \exp(-\hbar\omega_{\mathbf{k}_3} / k_B T)$$

Which shows that U-processes are almost frozen at low temperatures and the thermal conductivity exponentially approaches infinity.

But, in practice, all solids have finite conductivity even in the ranges of lowest temperatures. Since no perfect crystal exists in nature, there may be collisions of phonons with imperfections, impurities or even with the bounding surfaces. At certain temperature the mean free path on account of these collisions becomes so large that it is comparable with the size of the sample. This being the maximum realistic value, the mean free path assumes the *temperature independent behaviour* below this temperature. This is known as the *size effect*. Then (11.31) is replaced by

$$K_{\text{ph}} = \frac{1}{3} C_V v D \quad \text{----------} \quad (11.34)$$

Where D stands for the size (diameter in the case of rod).

The behaviour of thermal conductivity in this temperature region is mainly determined by the Debye heat capacity which drops as T^3. In fact, a competition between the exponential nd the T^3 terms sets in at the temperature below which the mean *free path is temperature independent*. The drop of T^3 being faster than the exponential increase the thermal conductivity is limited to a finite value.

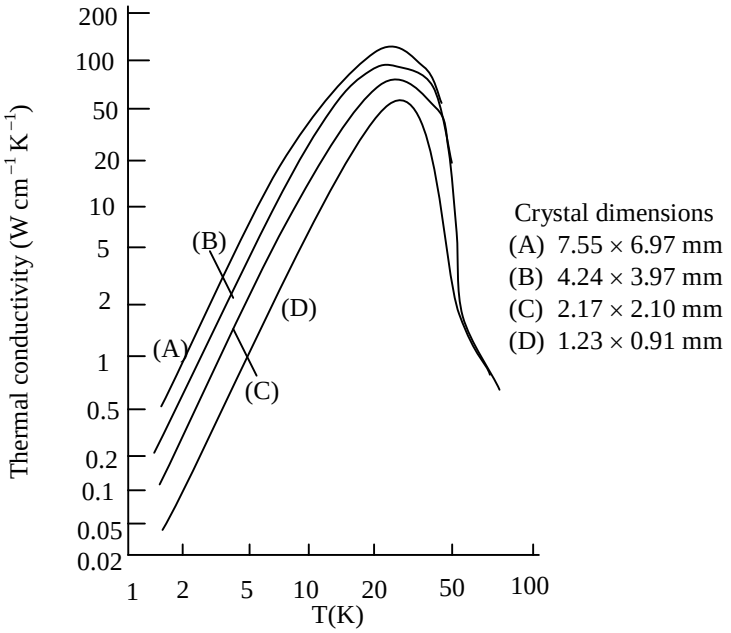


Fig 11.6: Thermal conductivity of isotopically pure LiF crystals as a function of temperature. The curve are the beautiful demonstration of size effect on thermal conductivity. The large the cross-sectional area of the sample, the higher the thermal conductivity.

The scattering of phonon from crystal walls was suggested by Casimir. The nature of the curves of Fig11.6 below 10 K provides a sound proof to the ideas of Casimir. The larger the cross-sectional area of the sample, the higher the conductivity. From below 10 K as the temperature rises, the U-processes become more frequent and the conductivity attains maximum value when the mean free path because of the phonon-phonon scattering becomes comparable with that owing to the scattering from the crystal's surface. With further rise in temperature the conductivity falls because of the increasing phonon-phonon scattering and assumes the expected exponential drop at higher temperatures where the heat capacity tends to be on level with the constant Dulong-Petit value.

The presence of rival phonon scattering mechanism is found to soften the 'sharp conductivity maximum' in most of the samples, irrespective of the material. The scattering from isotropic impurities is well identified among these. Its effect is demonstrated in fig 11.7 by the experimental curves of two samples of LiF crystals, one of which contains a mixture of ^{6}Li and ^{7}Li isotopes. On cooling, the rise of conductive is less steep in the sample having both isotopes than in the sample which is almost free from ^{6}Li isotope. In addition it has a lower and flatter maximum. The sharp maximum in the pure sample is almost totally contributed by U-processes in absence of the rival isotropic scattering.

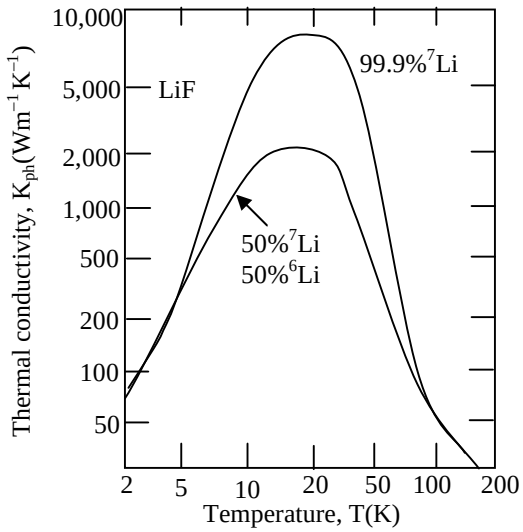


Fig11.7: The influence of isotope scattering on the maximum of thermal conductivity. The phonon scattering from isotopes lowers the thermal conductivity as is observed for the sample composed of ^{7}Li and ^{6}Li isotopes. The other sample almost free of ^{6}Li isotope shows a sharper and higher maximum. The presence of the rival phonon scattering mechanisms softens the sharp conductivity maximum.

Insulators characterized by large values of thermal conductivity have found wide applications in the field low temperature physics.

11.4 Summary:

1. For a harmonic crystal, heat capacity at constant volume, C_V is equal to heat capacity at constant pressure *Cp* because the difference between the two depends on the square of the temperature coefficient of linear expansions which is zero for a harmonic solid.

2. The achievements of Einstein model of phonon heat capacity are:

(i). It explains behaviour of heat capacity at room temperature and above. In the high temperature limit it gives the classical value ($3NK_B$) which the experimental value approaches at these temperatures.

(ii). It shows that the heat capacity approaches zero as $T \rightarrow 0$ K, at a fact that agrees with the experiment.

3. The limitation of Einstein model of phonon heat capacity is that the heat capacity at low temperatures decreses as $\exp(-\hbar\omega/k_BT)$ is contrast to the observed T^3 dependence.

The Debye T^3-law based on a continuum model explains very well at low temperatures at low temperature. The Debye model enjoys success over the entire useful temperature range.

4. The anharmonic effects are most reasonably accounted when the potential energy expansion is truncated not before the quadratic term. That is,

$U(x) = fx^2 - gx^3 - hx^4$

5. Two phonons of wave vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ may scatter against each other (or collide) and merge to produce a new phonon by either of the two types of scattering processes

Normal process: $\mathbf{k}_1 + \mathbf{k}_2 = \mathbf{k}_3$
(N-process)

Umklapp Process: $\mathbf{k}_1 + \mathbf{k}_2 = \mathbf{k}_3 + \mathbf{g}$
(U-process)

6. At high temperatures ($T>>\theta_D$), the phonon thermal conductivity (K_{ph}) changes according to

$K_{ph} \propto \frac{1}{T^m}$ where $1 < m < 2$

At low temperatures ($T<<\theta_D$), the U-process are almost frozen and K_{ph} for ideal crystals exponentially approaches infinity. Real crystals have, however, finite K_{ph} because of collisions of phonons with imperfections, impurities and bounding surfaces of specimens. At very low temperatures, the phonon mean free path is comparable with the size of specimen (or a constant), so that

$$K_{ph} = \frac{1}{3} C_V v D$$

where D is the diameter in the case of a rod. or

$K_{ph} \propto T^3$ (since v and D are constant and $C_V \propto T^3$)

11.5 Keywords:

Harmonic approximation – Anharmonic Effect – Vibrational energy – Heat capacity – Dulong-Petit law – Quantum theory of Heat capacity – Plank's law – Phonon heat capacity – Potential energy curve – Thermal expansion – Lennard – Jones Umklapp Process – Size effect.

11.6 Review Questions:

1. Distinguish between classical and Quantum theory of heat capacity of solids.
2. How thermal expansion of solids play an important role in the Heat capacity of solids.
3. How anharmonic effects explain heat capacity more effectively than harmonic approximation.
4. Describe the variety of phonon collision processes in the theory of heat capacity of solids.
5. How do you account the phonon thermal conductivity at low and high temperature region.

11.7 Text and Reference Books:

1. Elements of Solid State Physics by J.P. Srivatsava (PHI)
2. Solid State Physics by M.A.Wahab (Narosa)
3. Elements of Solid State Physics by A. Omar (Pearson education)
4. Solid State Physics by S.O. Pillai (New Age)
5. Solid State Physics by C. Kittel (Asia Publishing House)
6. Solid State Physics by S.L. Kakani and C. Hemrajani (S.Chand)
7. Solid State Physics by Saxena Gupta Saxena (Pragati Prakashan).
8. Solid State Physics by C.J.Dekker (Macmillan)

UNIT – III **LESSON: 12**

Energy band theory

Aim: To know about the Electron Energy Bands.

Objectives of the lesson:

- To know about the failure electron theory in explaining conductivity of materials.
- To know about the Bloch theorem and its proof, the consequences of periodicity.
- To know about wave mechanical interpretation of energy bands through Kronig Penney model.
- To know about the velocity of Bloch electrons and the dynamical effective mass m^*.
- To know about momentum, crystal momentum and physical origin of effective mass.
- To understand the concept of Hole.
- To know about the limiting cases of periodic potential.

Structure of the lesson:

12.1 Failure of free electron theory:

In lesson 11 we learnt how the free electron model in great detail gave invaluable information in accounting for the observed metallic properties. Nevertheless this model is an approximation and has its limitations and even the quantized free electron theory is unable to explain some property. Let us consider some of the limitations.

1. The method is unable to explain why some materials are metals why others are good insulators. Their electrical conductivity may differ by a magnitude of order 10^{30} which is unusually high for any physical property.
2. The model suggests that other things being equal, electrical conductivity is proportional to electron concentration. But it is surprising that the divalent metals (Be, Cd, Zn etc) and even trivalent metals (Al, In) are consistently low conductive than the monovalent metals (Cu, Ag and Au) despite the fact that the former have higher concentration of electrons.
3. A far more damaging testimony against the model is the fact that some metals exhibit positive Hall Coefficient, for example Be, Zn, Cd. The free electron model always predicts a negative Hall coefficient.
4. Complex transport phenomena in the presence of a magnetic field to, do not find satisfactory explanation.
5. The continuous parabolic energy curve renders us helpless in finding a clue to the sharp resonance like structures observed in the optical spectra of solids.
6. Measurements of the Fermi surface indicate that it is non spherical in shape. This contradicts the model, which predicts spherical Fermi surface.

These difficulties and other, which were not mentioned here, can be resolved by a more sophisticated theory, which takes into account the interaction of the electron with lattice.

12.2 Consequences of periodicity –The Bloch Theorem:

To understand the difference between insulators and conductors, one must extend the free electron model to take account of the periodic potential with periodicity of the lattice. In one electron model the periodic potential may be thought of arising from periodic charge distribution associated with ion cores situated on the lattice sites plus the

constant (average) contribution due to all other free electrons of the crystals. The later contribution to the potential accounts in an average sense, the interaction effects of the single electron with all others. A one dimensional representation of a periodic crystal of a cubic crystal of lattice constant 'a' is shown in Fig 12.1. So far as the one-electron quantum-mechanical picture is concerned, the crystal periodicity will usually be assumed to extend to infinity in all directions, but at surface of any actual crystal the periodicity will be interrupted, the potential function then behaving somewhat as shown at the left-hand edge of the Fig 12.1. The lattice spacing will be quite uniform near such a surface, but practically the periodicity will be usually found to be almost perfect after a few atomic spacings with in the crystal.

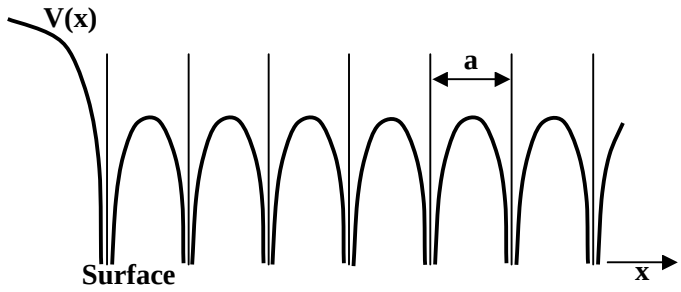


Fig 12.1 Schematic representation of the potential within a perfectly periodic crystal lattice. The surface potential barrier is shown at the left.

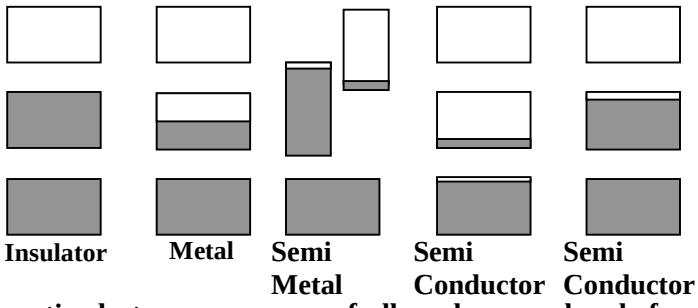


Fig 12.2 Schematic electron occupancy of allowed energy bands for an insulator, metal, semimetal and semiconductor. The vertical extent of the boxes indicates the allowed energy regions; the shaded areas indicate the regions filled with electrons.

We can see on the basis of above description that electrons in crystal are arranged in energy bands (fig 12.2); separated by regions in energy for which no wave like electron orbitals exists. Such forbidden regions are called energy gaps or band gaps; and will be shown to result, from the interaction of conduction electron waves with the ion cores of the crystal. The important question for electrical conductivity is how the electrons

respond to an applied electric field. The crystal will behave as an insulator if the number of electrons is such that the allowed energy bands are either filled or empty, for then no electron can move in an electric field. The crystal will act as a metal if one or more bands are partly filled. The crystal is a semi conductor or semimetal if all bands are entirely filled, except for one or two bands slightly filled or empty.

The possibility of band gap is the most important new property that emerges when we extend the free electron model to take into account of the periodicity of the lattice of the solid. We shall encounter other remarkable properties of electrons in crystals. For example, they respond to applied electric or magnetic field as if the electrons were endowed with an effective mass m^*, which may be larger or smaller than the free electron mass, or even may be negative.

12.3 The Bloch Theorem:

The Bloch Theorem is a mathematical statement regarding the form of the one-electron wave functions for a perfectly periodic potential.

Let the potential energy of an electron satisfy the equation

$$V(x) = V(x+a)$$

where 'a' is the period. The Sehrodinger wave equation is than

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{2m}{\hbar^2}[E - V(x)]\psi = 0 \quad \text{----------} \quad (12.1)$$

With reference to the solution of the equation, there is an important theorem known as Bloch theorem ; which states that the solutions are the plane waves modulated by the functions $U_k(x)$ that has the same periodicity as the lattice. In the theory of differential equations it is known as Floquet's theorem. Mathematically the solutions are of the form

$$\psi(x) = e^{\pm i\mathbf{k}x} U_{\mathbf{k}}(x) \quad \text{----------} \quad (12.2)$$

where $U_{\mathbf{k}}(x) = U_{\mathbf{k}}(x+a)$

The wave functions of the type (12.2) are called Bloch functions. We note that Bloch functions (12.2) have the property

$$\psi(x+a) = e^{\pm i\mathbf{k}(x+a)} U_{\mathbf{k}}(x+a) = e^{\pm i\mathbf{k}a}\psi(x)$$

Thus the Bloch functions have the property that

$$\psi(x+a) = \lambda\psi(x) \quad \text{----------} \quad (12.3)$$

where constant λ is

$$\lambda = e^{\pm i\mathbf{k}a}$$

If we imagine the crystal to be a form of a ring, N steps of the above displacement or translation will bring the electron back to the same atom at x from where it started.

It means that $\Psi(x + Na) = \Psi(x + L) = \Psi(x)$

But the net phase shift in the wave function requires that $\Psi(x + L) = C^N\Psi(x)$.

So C^N must be equal to 1 or $e^{i2\pi n}$ or $C = e^{i2\pi n/N}$ with n = 0, ±1, ±2, ……± N/2. It implies

$k = n\frac{2\pi}{L}$; for n $= \frac{\pm N}{2}$; $k = \pm\frac{\pi}{a}$.

It is clear that if we can show that the Schrodinger equation (12.1) has solutions with the property (12.3), the solutions can be written as Bloch functions and the theorem is proved.

Since (12.1) is a linear second-order differential equation, there are two independent solutions $g(x)$ and $h(x)$ such that

$$\psi(x) = Ag(x) + Bh(x) \quad \text{----------} \quad (12.4)$$

represents the most general solution of (12.1) with A and B arbitrary constants. Since $\psi(x) = \psi(x + a)$, not only g (x) and h (x) but also $g(x + a)$ and $h(x + a)$ also satisfy (12.1). Any solution of (12.1) is expressible as a linear combination of the independent solutions g (x) and h (x). Hence we must have the relations

$$\left.\begin{aligned} g(x+a) &= \alpha_1 g(x) + \alpha_2 h(x) \\ h(x+a) &= \beta_1 g(x) + \beta_2 h(x) \end{aligned}\right\} \quad \text{----------} \quad (12.5)$$

where α_2 and β_2, are constants and real functions of energy E. Then

$\psi(x + a) = Ag(x + a) + Bh(x + a) = (\alpha_1 A + \beta_1 B)\, g(x) + (\alpha_2 A + \beta_2 B)\, h(x)$ ------- (12.6)

Now $\psi(x + a)$ can always be expressed in the form

$$\psi(x + a) = \lambda\psi(x) = \lambda Ag(x) + \lambda Bh(x) \quad \text{----------} \quad (12.7)$$

Comparing co-efficients of $g(x)$ and $h(x)$ in (12.6) and (12.7) we have

$$\left.\begin{aligned} (\alpha_1 - \lambda)A + \beta_1 B &= 0 \\ \alpha_2 A + (\beta_2 - \lambda)B &= 0 \end{aligned}\right\} \quad \text{----------} \quad (12.8)$$

This system of homogenous equations in A and B has solutions other than $A = B = 0$ only if

$$\begin{vmatrix} \alpha_1 - \lambda & \beta_1 \\ \alpha_2 & \beta_2 - \lambda \end{vmatrix} = 0$$

or $$\lambda^2 - (\alpha_2 + \beta_2)\lambda + (\alpha_1\beta_2 - \alpha_2\beta_1) = 0$$

Now we can easily prove that $(\alpha_1\beta_2 - \alpha_2\beta_1) = 1$. Thus we have

$$\lambda^2 - (\alpha_1 + \beta_2)\lambda + 1 = 0 \quad \text{----------} \quad (12.9)$$

The solution of this quadratic equation in λ serves to determine the two possible values of λ for which Equation (12.7) is true. Hence there are two functions $\psi_1(x)$ and $\psi_2(x)$ which exhibit the property (12.7). Hence we consider two cases.

(1) For energy ranges in which $(\alpha_1 + \beta_2)^2 < 4$

The two roots λ_1, λ_2 of equation (12.9) will be complex and since $\lambda_1\lambda_2 = 1$, they will be conjugates. In those regions of energy we may write

and
$$\left.\begin{aligned}\lambda_1 &= e^{i\mathbf{k}a}\\ \lambda_2 &= e^{-i\mathbf{k}a}\end{aligned}\right\} \quad \text{----------} \quad (12.10)$$

The corresponding functions $\psi_1(x)$ and $\psi_2(x)$ then have the property

$\psi_1(x+a) = e^{i\mathbf{k}a}\psi_1(x)$ and $\psi_2(x+a) = e^{-i\mathbf{k}a}\psi_2(x)$ and thus are Bloch functions. We can also prove this starting with $g(x)$, $h(x)$ as given below;

Since $g(x)$ and $h(x)$ are real solutions of (12.1) we have

$$\frac{d^2g(x)}{dx^2} + \frac{2m}{\hbar^2}\{E - V(x)\}g(x) = 0$$
$$\frac{d^2h(x)}{dx^2} + \frac{2m}{\hbar^2}\{E - V(x)\}h(x) = 0$$

multiplying former by $h(x)$ and latter by $g(x)$ and subtracting we have

$$h(x)\frac{d^2g(x)}{dx^2} - g(x)\frac{d^2h(x)}{dx^2} = 0$$

or
$$W(x) = h(x)\frac{dg(x)}{dx} - g(x)\frac{dh(x)}{dx} = \text{constant.}$$

This is known as Green's identity. R.H.S. of this equation is called Wronskian, $W(x)$ of the two solutions and is constant in this case. Further, from equation (12.5) we have

$$h(x+a)\frac{d\,g(x+a)}{dx} - g(x+a)\frac{d\,h(x+a)}{dx} = (\alpha_1\beta_2 - \alpha_2\beta_1)\left\{h(x)\frac{dg(x)}{dx} - g(x)\frac{dh(x)}{dx}\right\}.$$

or
$$W(x+a) = (\alpha_1\beta_2 - \alpha_2\beta)W(x)$$

or
$$\alpha_1\beta_2 - \alpha_2\beta_1 = 1$$

and
$$\left.\begin{aligned}\psi_1(x+a) &= e^{i\mathbf{k}a}\psi_1(x)\\ \psi_2(x+a) &= e^{-i\mathbf{k}a}\psi_2(x)\end{aligned}\right\} \quad \text{----------} \quad (12.11)$$

Such functions are called Bloch functions. The functions are written in the form

$$\psi_k(x) = e^{\pm i\mathbf{k}x} U_{\mathbf{k}}(x) \quad \text{----------} \quad (12.11a)$$

where $U_k(x)$ is a periodic function of period 'a'. Thus the solutions of Schrodinger wave equation in this region can be expressed in the form of Bloch function

(2) For energy range in which $(\alpha_1 + \beta_2)^2 > 4$

In this case the two roots λ_1 and λ_2 are real. The solutions of Schrodinger equation are of the type

$$\left.\begin{aligned} &\psi_1(x) = e^{\mu x} U(x) \\ \text{and} \quad &\psi_2(x) = e^{-\mu x} U(x) \end{aligned}\right\} \quad \text{----------} \quad (12.12)$$

where μ is a real quantity. Although such solutions are mathematically sound, they can not in general be accepted as wave functions describing electrons; since they are not bounded. Thus there are no electronic states in the energy region corresponding to real roots λ_1 and λ_2. The above discussion thus leads also to the notion that the energy spectrum of an electron in a periodic potential consists of allowed and forbidden energy regions or bands.

In three dimensions the Bloch theorem becomes

$$\psi_k(r) = e^{\mathbf{ik}\cdot\mathbf{r}}\, U_{\mathbf{k}}(r) \quad \text{----------} \quad (12.13)$$

The wave vector, **k** used to label the Bloch functions has the properties

(a) Under a crystal translation which carries **r** to **r + T** we have

$$\psi_k(\mathbf{r}+\mathbf{T}) = e^{\mathbf{ik}\cdot\mathbf{T}}\, e^{\mathbf{ik}\cdot\mathbf{r}}\, U_{\mathbf{k}}(\mathbf{r}+\mathbf{T}) = e^{\mathbf{ik}\cdot\mathbf{T}} \psi_{\mathbf{k}}(\mathbf{r}) \quad \text{----------} \quad (12.14)$$

because $U_{\mathbf{k}}(\mathbf{r}+\mathbf{T}) = U_{\mathbf{k}}(\mathbf{r})$. Thus $e^{\mathbf{ik}\cdot\mathbf{T}}$ is the phase factor by which a Bloch function is multiplied when we make a crystal lattice translation **T**.

(b) If the lattice potential vanishes $U_{\mathbf{k}}(\mathbf{r})$ is constant. We must have

$$\psi_k(r) = e^{\mathbf{ik}\cdot\mathbf{r}}$$

as for a free electron.

(c) The value of k enters into conservation laws for collision processes of electrons in crystals. For this reason $\hbar k$ is called the crystal momentum of the electron. When an electron **k** collides with phonon of wave vector. $\mathbf{k}_f$, the selection rule is

$$\mathbf{k} + \mathbf{k}_f = \mathbf{k}' + \mathbf{G}$$

If the phonon is absorbed in the collision the electron has been scattered from a state **k** to a state **k′** and **G** is any reciprocal lattice vector.

12.4 Kronig-Penney model: Wave Mechanical Interpretation of Energy Bands:

In view of the discrete energy levels scheme of isolated atoms, it is unlikely possibilities that the energy bands would be infinitely continuous. This is really the case. In general, there is a region of forbidden energies between the two successive bands. However the bands overlap in some cases. Energy band is almost centered around its parent level, the N fold degenerate level in an isolated atom of a crystal composed of N atoms. Kronig and Penney (1930) demonstrated that regions of forbidden energies intervene the regions of allowed energies. They accomplished this task using a one dimensional square well crystal potential depicted in fig 12.3.

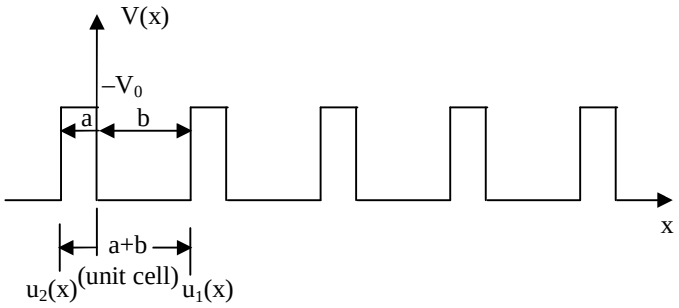


Fig 12.3 Ideal periodic square well potential well used by Kronig and Penney to illustrate the general characteristics of the quantum behaviour of electrons in periodic lattices.

The period of potential is $(a + b)$ in region such as $o < x < a$, the potential energy is assumed to be zero and in regions such as $-b < x < a$ the potential energy is V_0. For such an array of periodic square potentials the exact solution of Schrodinger equations is possible. Although the potential represented by fig12.3 is some what idealized periodic potential which is only a crude approximation to that found in the actual crystal, it is very useful because it serves to illustrate in a most explicit way many of other important characteristic features of the quantum behaviour of electrons in periodic lattice.

The wave function associated with this model may be calculated in the one-electron approximation by solving Schrodinger equation,

$$\frac{\partial^2\psi(x)}{\partial x^2} + \frac{2m}{\hbar^2}\{E - V(x)\}\psi(x) = 0 \quad \text{----------} \quad (12.15)$$

Since the potential is periodic, the wave function must have the Bloch form and we have

$$\psi(x) = e^{ikx}\, U(x) \quad \text{---------- (12.16)}$$

Substituting (12.16) in (12.15) and putting $\alpha = \left\{\frac{2mE}{\hbar^2}\right\}^{1/2}$, It is found that the function $U(x)$ must satisfy

$$\frac{d^2U}{dx^2} + 2ik\frac{dU}{dx} - \left\{k^2 - \alpha^2 + \frac{2mV(x)}{\hbar^2}\right\}U(x) = 0 \quad \text{---------- (12.17)}$$

If $U_1(x)$ and $U_2(x)$ represent the values of $U(x)$ in the two regions of potential function *i.e.* in the intervals $(0 < x < a)$ and $(-b < x < 0)$ respectively, we have

$$\frac{d^2U_1}{dx^2} + 2ik\frac{dU_1}{dx} - (k^2 - \alpha^2)U_1 = 0 \quad (0 < x < a) \text{---------- (12.18)}$$

$$\frac{d^2U_1}{dx^2} + 2ik\frac{dU_2}{dx} - (k^2 - \beta^2)U_2 = 0 \quad (-b < x < 0) \text{---------- (12.19)}$$

where $\beta = \left\{\frac{2m(E - V_0)}{\hbar^2}\right\}^{1/2}$ and is purely imaginary quantity for $0 < E < V_0$. The solutions of the differential equations (12.18) and (12.19) are

$$U_1 = Ae^{i(\alpha - k)x} + Be^{-i(\alpha + k)x} \quad (0 < x < a) \quad \text{---------- (12.20)}$$

$$U_2 = Ce^{i(\beta - k)x} + De^{-i(\beta + k)x} \quad (-b < x < 0) \quad \text{---------- (12.21)}$$

where A, B, C, D are constants. The continuity of the wave function ψ and its derivative at $x = a$ and $x = -b$ and remembering that $U(x)$ has the periodicity of the lattice demands that the wave function $U(x)$ satisfies the boundary condition

$$\left.\begin{array}{c} U_1(0) = U_2(0) \qquad U_1(a) = U_2(-b) \\ \left(\frac{dU_1}{dx}\right)_{x=0} = \left(\frac{dU_2}{dx}\right)_{x=0}, \quad \left(\frac{dU_1}{dx}\right)_{x=a} = \left(\frac{dU_2}{dx}\right)_{x=-b} \end{array}\right\} \quad \text{---------- (12.22)}$$

Using boundary conditions (12.22) we find that

$$\left.\begin{array}{l} A + B = C + D \\ (\alpha - k)A - (\alpha + k)B = (\beta - k)C - (\beta + k)D \\ Ae^{i(\alpha - k)a} + Be^{-i(\alpha + k)a} = Ce^{-i(\beta - k)b} + De^{i(\beta + k)b} \\ (\alpha - k)\, Ae^{i(\alpha - k)a} - (\alpha + k)\, Be^{-i(\alpha + k)a} \\ = (\beta - k)\, Ce^{-i(\beta - k)b} - i(\beta + k)De^{i(\beta + k)b} \end{array}\right\} \text{---------- (12.23)}$$

The coefficients A, B, C and D can thus be determined and the wave function can be calculated. For our purpose we are more interested in determining the values of energy

for which satisfactory solutions are obtained. The equations (12.23) have non-vanishing solutions if and only the determinant of coefficients vanishes. This requires that

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ \alpha - k & -(\alpha + k) & \beta - k & -(\beta + k) \\ e^{i(\alpha-k)a} & e^{-i(\alpha+k)a} & e^{i(\beta-k)b} & e^{-i(\beta+k)b} \\ (\alpha - k)e^{i(\alpha-k)a} & -(\alpha + k)e^{-i(\alpha+k)a} & (\beta - k)e^{i(\beta-k)b} & -(\beta + k)e^{-i(\beta+k)b} \end{vmatrix} = 0$$

Expanding the determinant, it can be shown that this leads to the following conditions.

$$-\frac{\alpha^2 + \beta^2}{2\alpha\beta} \sin \alpha a \sin \beta b + \cos \alpha a \cos \beta b = \cos k(a + b) \quad \text{----------} \quad (12.24)$$

Since in the range $0 < E < V_0$, β as defined earlier is imaginary, for these values of energy it is most convenient to express $\beta = i\gamma$

in this region, and noting that $\cos ix = \cosh x$ and $\sin ix = i \sinh x$, equation (12.24) can be written as

$$\frac{\gamma^2 - \alpha^2}{2\alpha\gamma} \sinh \gamma b \sin \alpha a + \cosh \gamma b \cos \alpha a = \cos k (a + b) \quad \text{----------} \quad (12.25)$$

where γ is a real quantity in the interval $0 < E < V_0$ just as β is in the interval $V_0 < E < \infty$. Thus we may use (12.24) most conveniently when $V_0 < E < \infty$ and (12.25) when $0 < E < V_0$.

The wave function (12.16) must, like all wave functions, be a well behaved function as x approaches $\pm\infty$. Since $U(x)$ is a periodic function whose values are the same in each unit cell, no difficulty arises on its account provided that the factor e^{ikx} in (12.16) is well behaved under these conditions. But e^{ikx} is well behaved at both $+\infty$ and $-\infty$ only if k is real, whereby e^{ikx} is oscillatory. If k were imaginary e^{ikx} would diverge to infinity either at $+\infty$ or $-\infty$ and resulting expression for $\psi(x)$ would not behave properly as a wave function. We must, therefore accept only functions of the form (12.16) with real values of k. The expressions (12.24) and (12.25) have on left hand side a function of the form $k_1 \sin \alpha a + k_2 \cos \alpha a$ which must the equal to $\cos (a + b)k$. If for a given value of energy, the function on the left hand side of these equations should have a value in the range between +1 and –1, there the required value for $\cos (a + b)k$ is obtained with a real value of $k(a + b)$. On the other hand, if the value of the function on the left hand side of (12.24) and (12.25) should be outside that range, it would mean that, $\cos (a + b)$ k would have to

be greater than +1 or less than –1 which would in turn, requires that, argument k(a+b) be complex number with imaginary part other than zero. Under these circumstances the solution (12.16) would not behave properly at infinity and would not satisfy the physical requirement for wave functions of the system. The energies associated with such values of k would simply be forbidden to the electron.

The equations (12.24) and (12.25) can be written as

$$\left[1+\frac{(\alpha^2-\beta^2)}{4\alpha^2\beta^2}\sin^2\beta\, b\right]^{1/2}\cos(\alpha\, a-\delta)=\cos k(a+b) \quad \text{----------} \quad (12.26)$$

where $\tan\ \delta=-\dfrac{\alpha^2+\beta^2}{2\alpha\beta}\tan\beta\, b \qquad (V_0<E<\infty)$

and

$$\left[1+\frac{(\alpha^2+\gamma^2)}{4\alpha^2\gamma^2}\sin h^2\gamma\, b\right]^{1/2}\cos(\alpha\, a-\delta)=\cos k(a+b) \text{----------} \quad (12.27)$$

where $\tan\ \delta=-\dfrac{\alpha^2+\gamma^2}{2\alpha\gamma}\tan h\gamma b \qquad (0<E<V_0)$

From these expressions it is clear that in both cases the left hand side has the form of cosine function times a modulating factor whose amplitude is invariably greater than unity. The value of this modulation factor is actually maximum for α = 0 (hence for E = 0) and approaches unity in the limit of large energies ; where $\alpha \cong \beta$.

When left hand side of (12.26) and (12.27) are plotted as a function of energy keeping in mind that $\alpha^2 - \beta^2 = \alpha^2 + \gamma^2 = 2mV_0/\hbar^2$ = constant, the results are as illustrated in fig (12.4). When the ordinate of the curve lies between +1 and –1, there exists a real value of k corresponding to physically possible wave functions.

Out side of these limits however k must be complex with a non zero imaginary part. Such values of k can never lead to physically well-behaved wave functions; the corresponding ranges of energy are forbidden and are shown in fig12.4, as shaded regions. There are thus formed alternate regions of allowed energy eigen values and forbidden regions. These regions are usually referred to as allowed and forbidden energy bands, and the grouping of the permitted energy values into these bands is one of the most important characteristic feature of the behaviour of electrons in periodic lattice. It can be shown that energy bands having the same qualitative aspects as those shown in fig

12.4 are formed no matter what the form of the potential is, so long as it is periodic. From the fig 12.4 the following conclusions may be drawn.

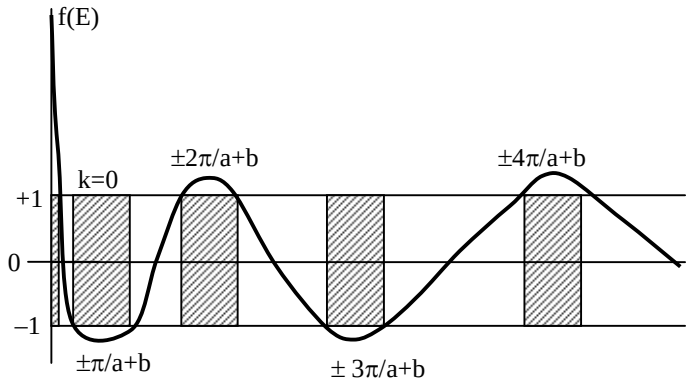


Fig 12.4 A plot of the functions on the left hand side of (12.26) and (12.27) versus energy. The shaded regions show forbidden energy bands where the value of k is complex, the unshared regions allowed energy bands corresponding to real value of k.

1. The energy spectrum of electrons consists of a number of allowed energy bands, separated by forbidden regions.
2. The width of the allowed energy bands increases with increasing values of energy.
3. The width of a particular allowed band decrease with increasing binding energy V_0. In extreme case with $V_0 \to \infty$ the allowed regions become infinity narrow and the energy spectrum becomes a line spectrum.

These conclusions can be appreciated much more easily by rewriting the equation 12.25 with the condition that even though as $V_0 \to \infty$ and b approaches zero the product V_0b remains finite. Under these circumstances (12.25) reduces to

$$(mV_0 b/\hbar^2\alpha)\sin\alpha a + \cos xa = \cos ka$$

Let us now define the quantity

$$P = mV_0 ba/\hbar^2$$

which is evidently a measure for the "area" V_0b of the potential barrier. In other words, increasing P has the physical meaning of binding a given electron more strongly to a particular potential well. From the last two equations we find that solutions for the wave functions exist only if

$$P\frac{\sin\alpha a}{\alpha a} + \cos\alpha a = \cos ka \quad \text{---------12.25a}$$

the above conclusions are summarized in fig 12.4a.

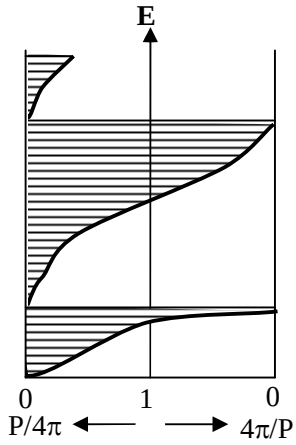


Fig 12.4a Allowed and forbidden energy ranges (shaded and open respectively) as function of P. The extreme left corresponds to P = 0(free electrons). The extreme right to P = ∝

where the energy spectrum is given as function of P. For P = 0, we simply have the free electron model and the energy spectrum is (quasi) continuous; for P = ∞, a line spectrum results as discussed under (3) above. For a given value of P the position and width of the allowed and forbidden bands are obtained by erecting a vertical line; the shaded areas correspond to allowed bands.

Using (12.26) and (12.27) and the value of $\alpha = (2mE/\hbar^2)^{1/2}$ and $\beta = \left\{\frac{2m(E-V_0)}{\hbar^2}\right\}^{1/2}$ it is possible to plot a curve showing the energy *E* as a function of k. The result is shown in fig12.5. The dotted curve is for large energies for which the function *E* (k) approaches the free electron relation

$$E = \frac{\hbar^2 k^2}{2m}$$

4. The discontinuities in *E* versus k curve occur for

$$k = \pm\frac{n\pi}{a+b},\ n = 1, 2, 3, 4\ldots\ldots$$

There *k*- values define the boundaries of 1st, 2nd, 3rd etc. Brillouin zones. The first zone extends from $\frac{-\pi}{a+b}$ to $+\frac{\pi}{a+b}$. Similarly the second zone consists of two parts, one extending from $\frac{\pi}{a+b}$ to $\frac{2\pi}{a+b}$ and the other part extending between $\frac{-\pi}{a+b}$ to $\frac{-2\pi}{a+b}$.

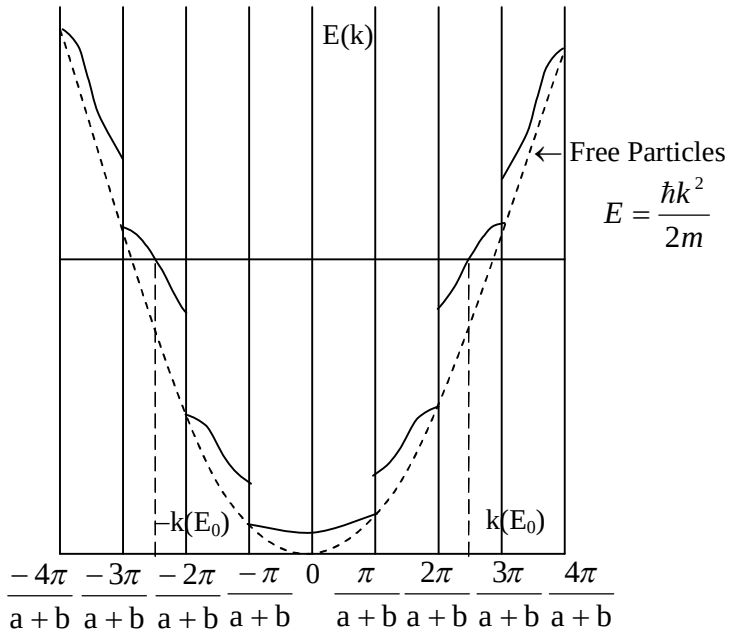


Fig 12.5 The energy *E* plotted as a function of k according to 12.26 and 12.27.

5. With in a given energy band the energy is a periodic function of k. For example, if one replaces k by k+$\frac{2\pi n}{a+b}$, where *n* is an integer, the right hand side of equation (12.26) and (12.27) remains the same. In other words k is not uniquely determined. It is therefore frequently convenient to introduce the "*reduced wave vector*" which is limited to the region

$$-\frac{\pi}{a+b} \le k \le \frac{\pi}{a+b}$$

The energy versus reduced wave vector is represented in fig 12.6

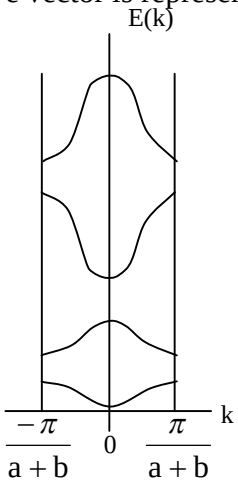


Fig 12.6 Schematic representation of the E versus a plot of figure 12.5, transformed to the reduced zone representation.

6. The number of possible wave functions per band can easily be shown to be equal to the number of unit cells N. Now as a result of spin of the electrons and the Pauli's exclusion principle each wave function can be occupied by at most two electrons. Thus each energy band provides for a maximum number of electrons equals to twice the number of unit cells. In other words, if there are 2N electrons in a band, the band is completely filled.

This conclusion as we shall see has far reaching consequences for the distinction between metals, insulators and semi conductors.

Properties of Bloch functions:

It has the form of a traveling plane wave, as represented by the factor $e^{i\mathbf{k}.\mathbf{r}}$, which implies that the electron propagates through the crystal like a free particle. The effect of the function $u_{\mathbf{k}}(\mathbf{r})$ is to modulate this wave so that the amplitude oscillates periodically from one cell to the next, as but this does not affect the basic character of the state function, which is that of a traveling wave.

Because the electron behaves as a wave of vector k, it has a deBroglie wavelength $\lambda = 2\pi/k$, and hence a momentum $\mathbf{p} = \hbar\mathbf{k}$, according to the deBroglie relation. We shall call the vector the crystal momentum of the electron.

The Bloch function of $\psi_{\mathbf{k}}$ is a crystal orbital, as it is delocalized throughout the solid, and not localized around any particular atom. Thus the electron is shared by the whole crystal.

The free electron and Bloch function electrons are not identical in their behaviour. The Bloch function of electron exhibits in many intriguing properties not shared by a free electron, properties which result from the interaction of the electron with the lattice.

12.5 Velocity of the Bloch electrons and the dynamical effective mass:

For convenience in visualizing the motion of Bloch electrons in solids we shall have to localize the wave function by superposing the solutions having different values of k. If this is done, then the group velocity associated with electron wave packet is

$$v_g = \frac{d\omega}{dk} = \frac{1}{\hbar}\frac{dE}{dk} \quad \text{----------} \quad (12.28)$$

Where of course E and ω are connected by Planck relation $E = \hbar\omega$. This in itself shows the importance of E versus k curves. In the case of free electrons $E = \frac{\hbar^2 k^2}{2m}$ and (12.28)

simply leads to the identity $v = \frac{\hbar k}{m} = \frac{P}{m}$. In the band theory, however E is in general not proportional to k^2 as may be seen from fig (12.6). Employing an $E(k)$ curve such as represented in fig 12.7(a) one obtains according to (12.28) for the velocity as a function of k, a curve of the type illustrated in fig 12.7(b). For the free electron v is proportional to k and is represented by dashed line in 12.7(b). At the top and bottom of the energy band $v = 0$, because from periodicity of the $E(k)$ curves it follows that $dE/dk = 0$, the absolute value of the velocity reaches a maximum for $k = k_0$ where k_0 corresponds to the inflection point of the $E(k)$ curve. It is of importance to note that beyond this point the velocity decreases with increasing energy; a feature which altogether different from the behaviour of free electrons.

Suppose that the electron is acted upon by an external electric field E_0 acquiring an increase in velocity dv_0, over a distance dx in time dt, it has gained an energy

$$dE = \frac{dE}{dk} \cdot dk = -eE_0 dx = -eE_0 v_0 dt = -\frac{eE_0}{\hbar}\frac{dE}{dk}dt$$

where by

$$dk = -\frac{eE_0}{\hbar}dt$$

or
$$\hbar\frac{dk}{dt} = \frac{dp}{dt} = -eE_0 = F \quad \text{----------} \quad (12.29)$$

Where we now use the symbol p to denote the crystal momentum. The equation (12.29) shows that the ratio of change of k is proportional to and lies in the same direction as the force $\boldsymbol{F}$. This relation is a very important one in the dynamics of Bloch electrons and is known as acceleration theorem.

Let us know consider consequence of the acceleration theorem. Considering the one-dimensional case, equation (12.29) can be written in the form

$$\frac{dk}{dt} = \frac{F}{\hbar}$$

showing that the wave vector k increases uniformly with time. Thus as t increases, the electron traverses the k-space at uniform rate as shown in fig 12.8. If we choose the repeated zone scheme, the electron starting from k = 0 for example moves up the band until it reaches the top (point A) and then starts to descend along the path BC.

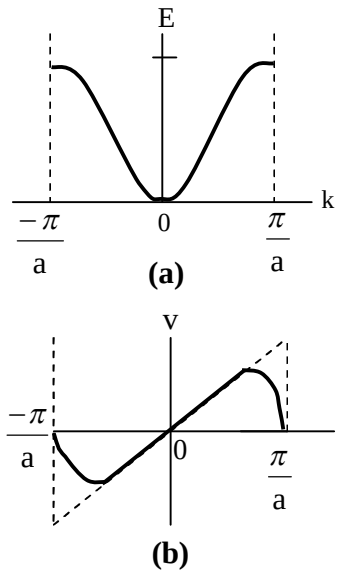


Fig 12.7 (a) The band structures, and (b) the corresponding electron velocity in a one dimensional lattice. The dashed line in (b) represents the free electron velocity.

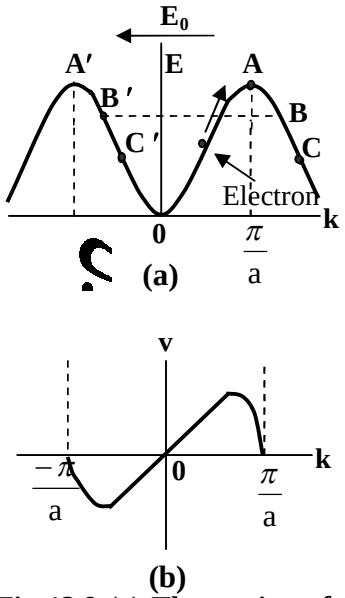


Fig 12.8 (a) The motion of an electron in k-space in the presence of an electric field (directed to the left) (b) The corresponding velocity.

If we use the reduced zone scheme then once the electron passes the zone edge at *A*, it immediately reappears at the equivalent point *A′*, then continues to descend along the path A′ B′ C′. Recalling the translational symmetry property the points B′, C′ are respectively equivalent to the points B, C so that we may use either of the schemes.

In the presence of electric field, the electron is in constant motion in *k*-space and never it is at rest. Fig (12.8) shows the velocity of electron as it traverses the k-axis. Starting at $k = 0$, as time passes the velocity increases, reaches maximum, decreases and then vanishes at the zone edge. The electron then turns round and acquires a negative velocity and so forth. The velocity we are discussing is the velocity in real space *i.e.* the usual physical velocity. It follows that a Bloch electron, in the presence of static electric field, executes an oscillatory periodic motion in real space, very much unlike a free electron. This is one of the surprising conclusions of electron dynamics in crystal.

If one differentiates (12.28) with respect to time, the result is

$$\frac{dv_g}{dt} = \frac{1}{\hbar}\frac{d}{dt}\left(\frac{dE}{dk}\right) = \frac{1}{\hbar}\frac{d^2E}{dk^2}\frac{dk}{dt}$$

By using (12.29) this can be written as

$$\frac{dv_0}{dt} = \frac{eE_0}{\hbar^2}\frac{d^2E}{dk^2} = \frac{F}{m^*} \quad \text{---------- (12.30)}$$

Where the effective mass m* is given by

$$m^* = \frac{\hbar^2}{\frac{d^2E}{dk^2}} \quad \text{---------- (12.31)}$$

Thus, in so far as the motion in an electric field is concerned, the Bloch electron behaves like a free electron whose effective mass is given by (12.31). The mass m* is inversely proportional to the curvature of the band, where the curvature is large that is d^2E/dk^2 is large-the mass is small, a small curvature implies a large mass. Thus if the energy is quadratic in k

$$E = \alpha k^2 \quad \text{---------- (12.32)}$$

Where α is a constant, then equation (12.32) yields

$$m^* = \frac{\hbar^2}{2\alpha}$$

Which is equivalent to $E = \frac{\hbar^2 k^2}{2m^*}$.

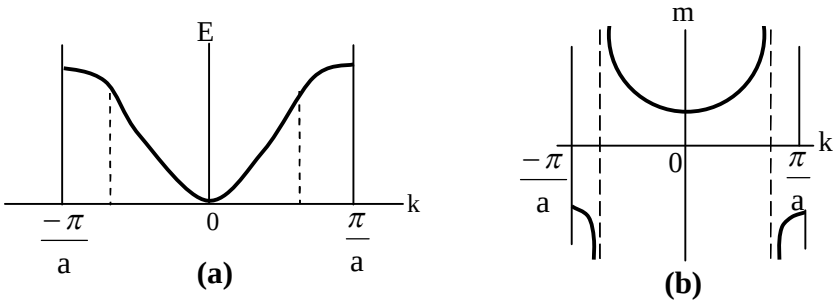


Fig 12.9 (a) The band structure and (b) the effective mass m* versus k.

Figure 12.9(a) and (b) shows respectively the band structure and the effective mass *m**, the latter calculated according to (12.31). Near the bottom of the band the effective mass *m** has a constant value which is positive, because quadratic relation (12.32) is satisfied near the bottom of the band. But as k increases, *m** is no longer a strict constant being now a function of k, because the quadratic relation is no longer valid.

Beyond the inflection k_0 the mass *m** becomes negative since the region is now close to the top of the band, and a negative mass is to be expected. The negative mass can be seen dynamically by noting that the velocity decreases for $k > k_0$. Thus acceleration is negative *i.e.* opposite to the applied force, implying a negative mass. This means that in this region of k-space the lattice

exerts such a large retarding force on the electron that it overcomes the applied force and produces a negative acceleration.

12.6 Crystal Momentum and Origin of the Effective Mass:

What has been said is that Bloch electron in the state ψ_k behaves as if it had a momentum $\hbar k$. Basically, there are three different reasons to support this statement.

(1). The Bloch function has the form (12.13).

$$\psi_k = e^{ik.r} U_k(r)$$

which, since U_k is periodic, appears essentially as a plane wave with k a constant of motion and that $\hbar k$ will have the dimensions of momentum. As the energy of electron increases (The particle becomes there by nearly free) the value of k in general approximates that of free-particle momentum divided by $\hbar$.

(2). When an electric field is applied, the wave vector varies with time according to equation (12.29)

$$\frac{d(\hbar k)}{dt} = F_{ext}.$$

again indicating that $\hbar k$ acts like a momentum. Here F_{ext}. refers to the external force applied to the crystal.

(3). In collision processes involving a Bloch electron the electron contributes a momentum equal to $\hbar k$.

It can easily be seen that these conclusions must hold no matter what particular form periodic potential takes. It is therefore customary to refer $\hbar k$ as the crystal momentum and is not equal to the actual momentum of Bloch electrons. To make the distinction between the actual momentum of electron (p) and the crystal momentum $p_c = \hbar k$ more clear, we evaluate the average momentum using quantum mechanical method.

$$p = \langle \psi_k | -i\hbar\nabla | \psi_k \rangle \quad \text{---------- (12.33)}$$

Where $-i\hbar\nabla$ is the momentum operator and ψ_k is the Bloch function. On solving this integral using the properties of the wave function ψ_k one finds that

$$p = m_0 \boldsymbol{v}$$

Where m_0 is the mass of the free electron and v is the velocity given by (12.28). Thus time average momentum of electron is equal to the time mass m_0 times the actual velocity v. If $p_c = \hbar k$ were the true momentum, then the force appearing on the right of (12.29) should have been

the total force and not just the external force, because there is a force exerted by the lattice, yet this force does not influence p_c.

The above idea may now be assembled to give a physical interpretation of the true momentum, one may write

$$\frac{m_0 dv}{dt} = F_{tot} = F_{ext} + F_L \quad \text{----------} \quad (12.34).$$

Where F_{tot} and F_L are, respectively, the total force and the lattice force acting on the electron. By lattice force, we mean the force exerted by the lattice on the electron as a result of its interaction with the crystal potential. Equation (12.34) can be expressed in terms of the effective mass as

$$m_0 \frac{dv}{dt} = m_0 \frac{F_{ext}}{m^*},$$

where m* is given by

$$m^* = \frac{m_0 F_{ext}}{F_{ext} + F_L} \quad \text{----------} \quad (12.35)$$

Now we see the reason why m^* is different from m_0, the free mass. If $\boldsymbol{F_L}$ were to vanish the effective mass would be come equal to true mass.

The effective mass may be smaller or larger from than m_0, or even negative depending on the lattice force. Suppose that the electron is "*Piled up"* primarily near the top of the crystal potential as shown in figure 12.10a. When an external force is applied it causes the electron to "*roll down hill*" along the potential curve.

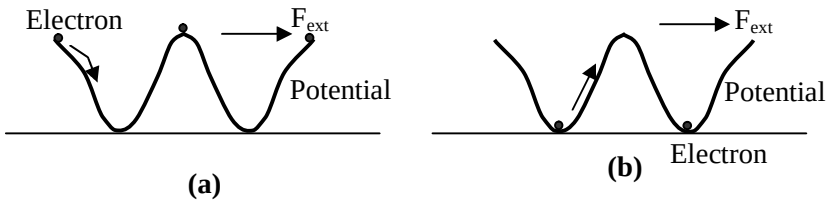


Fig 12.10 (a) Electron spatial distribution leading to an effective mass *m smaller than m_0 (b) A distribution leading $m^* > m_0$**

As a result, a positive force becomes operative and hence, according to (12.35) $m^* < m_0$. This is what happens in alkali metals for instance in the conduction band in semi conductors. Here m^* is less than m_0 because the lattice force assists the external force.

On the other hand, when electron is filled mainly near the bottom of the potential curve figure (12.10b) then clearly the lattice force tends to oppose the external force; resulting $m^* > m_0$. This is the situation in alkali halides. If the potential curve is sufficiently steep, then F_L becomes larger than F_{ext} and m^* becomes negative.

However, the crystal momentum $P_c = \hbar k$ is still a very useful quantity. In problems of electron dynamics in external field, the crystal momentum is much more useful than true momentum, since it is easier to follow motion in k-space than in real space. Therefore we shall continue to use P_c and refer to it as the momentum, when there is no ambiguity and even drop the subscript *c*.

In other words, the effective mass *m** and the crystal momentum $\hbar k$ are artifices which allow us - formally at least- to ignore the lattice force and concentrate on the external force only. This is very useful, because lattice force is not known a priori, nor is it easily found and manipulated as is the external force.

12.7 Concept of Hole:

In a one dimensional crystal some of the allowed bands will be entirely filled, some will be entirely empty and one may be partially filled, because at absolute zero the electrons of the system will occupy the states, one per state as required by the Pauli's exclusion principle from the lowest state up to a given energy determined by the number of available states, their distribution in energy, and number of electrons in the crystal. This energy is, of course, the Fermi energy of the crystal at zero temperature. If an electric field is applied, it is obvious that no current can be contributed from filled bands. This can be understood by nothing that the current density arises from a given band will be

$$I = -ne\bar{v}$$

where $\bar{v}$ is the average velocity and *n* is the number of electrons per unit volume belonging to that band. But $\bar{v}$ can be expressed as

$$\bar{v} = \frac{1}{nV}\sum_i v_i$$

where the summation is taken over all the velocities associated with individual electrons with in the volume *V* of the material. Hence current density *I* can be expressed as

$$I = -\frac{e}{V}\sum_i v_i \quad \text{----------} \quad (12.36)$$

This, however, must yield zero when summed over full band, because due to the symmetry of the curves of figure (12.6) about the *E*(k) axis, for every state of positive velocity $\hbar^{-1}\,\partial E/\partial k$ corresponding to a +ve slope, there is a state corresponding to a negative velocity of equal magnitude (negative slope) at k′ = k. We must conclude that only bands which are partially filled can contribute to current flow. This can be understood on physical grounds by observing

that in a partially filled band there are always electrons which can be excited gradually to unoccupied states of higher energy and momentum, where as in a filled band, on a account of Pauli's exclusion principle, their gradual field excitation can never occur since all the states are occupied.

For a band containing a relatively small number of electrons, the current obtained when a small voltage is impressed is most simply given by (12.36) summed over all the electrons in the band. If the band is nearly full, however, and there are only a few states (which in equilibrium state will be concentrated near the top of the band) the current equation (12.36) is best expressed by writing in the form

$$I = -\frac{e}{V}\sum_i v_i = -\frac{e}{V}\left[\sum_j v_j - \sum_k v_k\right] = +\frac{e}{V}\sum_k v_k \quad \text{----------} \quad (12.37)$$

Here the sum over *i* represents the sum over all velocity states occupied by electrons, the sum over *j* represent the sum over all velocity states in the band and sum over *k* represents and the sum over un occupied velocity states. As we have seen before the sum over *j*, taken over all states in band, must vanish. The remaining sum over the unoccupied states correspond to a current which could be produced by a corresponding number of positive charge carriers. It is possible (and advantageous also) to express the current from almost full band as a current derived from the motion of a comparatively small number of empty electronic *states* or *holes* which behave like positive particles, rather than a very large number of electrons. The velocity associated with a hole is that which an electron would have if it were to occupy the empty energy state, which is ordinarily near the top of the energy band. But since the *E* versus k relation there is concave downward, d^2E/dk^2 is negative, giving a negative electron effective mass from (12.31). A particle with negative effective mass experiences acceleration in a direction opposite to that of applied force. A negative charge particle with negative effective mass would thus be accelerated in the same direction as applied field and would thus exhibit the same dynamical behaviour as a positive particle of positive mass. We may, thus, regard the situation in a nearly filled band as one involving a relatively small number of positive particles of positive mass, which we shall refer to as holes, whose velocities and momenta are those corresponding to the un occupied electronic states in the band. We shall see later that in certain materials the physical nature and the dynamical behaviour of holes is very easily visualized in terms of defects in the electronic valence bonds which connect nearest neighbour atoms and provide the cohesive forces which hold the crystal together. We have not discussed in detail the behaviour of electrons and

holes under the influence of magnetic force, but it can be shown that they move as would negative or positive particles of effective mass m^* under the influence of usual magnetic force $\frac{q}{c}(v \times \mathbf{H})$.

12.8 Easy Limiting Cases of the Periodic Potential:

The periodic potential model led to the formation of allowed and forbidden energy bands resulting in the energy (E) versus k relation of the form shown in fig 12.5. The resulting electronic behavior was similar to free electron picture except that an effective mass had to be viewed as positive holes. We have emphasized these quantitative conclusions, which are independent of potential function as long as it is periodic.

The potential (fig 12.3) taken was only illustrative and is poorly related to the actual form of the potential, and consequently these qualitative conclusions are only illustrative. An exact description of the electronic behavior in a crystal must take into account the actual potential experienced by electron due to the ion cores and all other electrons in the crystals. This is a very complicated problem and an exact solution of the problem, even in one electron approximation is impossible. However two easy limiting approximations:-

(1) Nearly free electron approximation

(2) The light binding approximation

are of considerable interest. When the periodic potential is very weak, we can treat it as a perturbation and we have the "*nearly free electron approximation*". When the periodic potential is very strong each electron is almost bound to a minimum in the potential and so the rest of the lattice may be regarded as a perturbation on the above minimum. This is known as "*tight binding approximation*".

It may be remarked that in some solids the former approximation is quite good, while in others the latter one is more nearly correct. There are also solids where neither is adequate and the situation being intermediate between these two extreme cases. In this intermediate range we must use much more complex methods such as, orthogonalised plane wave (*OPW*) or an augmented plane wave (*APW*) method.

12.9 Summary:

1. The solutions of the wave equation in the periodic lattice are of the Bloch form $\psi_k(r) = e^{ik.r} U_k(r)$, where $U_k(r)$ is in variant under a crystal lattice translation.
2. The energy spectrum of the electron is comprised of a set of continuous bands separated by regions of forbidden region are complex.
3. The number of orbitals in a band is $2N$, Where N is the number of unit cells in the specimen.
4. Regarded as a function of k, the energy E(k) satisfies several symmetry properties. First it has transitional symmetry properties.

$$E(k + G) = E(k)$$

which enables us to restrict our consideration to the first Brillouin Zone only. The energy function $E(k)$ also has inversion symmetry, $E(-k) = E(k)$ and rotational symmetry in k-space.

5. An electron in Bloch state ψ_k moves through the crystal with a velocity

$$v = \frac{1}{\hbar}\Delta_m E(k)$$

This velocity remains constant so long as the lattice remains perfectly periodic.

6. In the presence of an electronic field, an electron moves in k space according to the relation

$$k = -eE_0/h$$

The motion is uniform and its rate proportional to the field. One obtains this relation if one regards the electron as having momentum $\hbar k$.

7. The effective mass of Bloch electron is given by

$$m* = \hbar^2/(d^2E/dk^2)$$

The mass is positive near the bottom of the band, where the curvature is positive. But near the top, where the band curvature is negative, the effective mass is also negative. The fact that the effective mass is different from free mass is due to the effect of lattice force on the electron.

8. A hole exists in a band, which is completely full, with one vacant site. The hole acts as a positive particle of positive charge +e.

12.10 Key Words:

Positive and negative hall coefficient – The Bloch Theorem – Periodic crystal potential – Energy gaps or band gaps – Kronig – Penney model – Dynamical effective mass m* – k-space.

12.11 Review Questions:

1. Investigate the motion of electrons in a periodic field discuss the electronic energy levels in a crystal
2. Prove that effective mass of an electrons in energy band is given by

$$\left(\frac{1}{m^*}\right)_{ij} = \frac{1}{\hbar^2}\frac{\partial^2 E}{\partial k_i \partial k_j}$$

3. Distinguish between extended zone, reduced zone and periodic zone scheme of plotting energy bands. Derive an expression for the effective mass m^* of the electron in a crystal and explain the physical basis of m^*
4. Discuss the essential features of the electron energy band structure on the basis of Kronig-Penney model.
5. (a) Prove the Bloch theorem and explain the reduced zone scheme.
   (b) Explain the significance of the effective mass of the electrons.

12.12 Text and Reference Books:

1. Elements of Solid State Physics by J.P .Srivatsava(PHI)
2. Solid State Physics by M.A. Wahab (Narosa)
3. Elements of Solid State Physics by A. Omar (Pearson education)
4. Solid State Physics by S.O. Pillai (New Age)
5. Solid State Physics by C. Kittel (Asia Publishing House)
6. Solid State Physics by S.L. Kakani and C. Hemrajani (S.Chand)
7. Solid State Physics by Saxena Gupta Saxena (Pragati Prakashan).
8. Solid State Physics by C.J. Dekker (Macmillan)

UNIT – IV **LESSON: 13**

BAND THEORY OF SEMICONDUCTORS

Aim: To know about energy bands, free electrons model and zone schemes for energy bands, free electrons model and zone schemes for energy band.

Objective of the lesson:

- To know about the nearly free electrons model and to show that the band gap energy is equal to twice the magnitude of the leading Fourier coefficient of the crystal potential is to show $\Delta E = 2V$.
- To identify the three different zone schemes for energy bands i.e. the extended zone scheme, the reduced zone scheme and the periodic zone scheme.
- To know the energy bands in a general periodic potential and to find solution near the boundary.
- To distinguish in between insulators, semi conductors and metals.
- To consider the tight binding approximation and to show $m^* = \frac{\hbar^2}{2ra^2}$, where m^* is the effective mass of the electron, a is the lattice constant, *r* is the overlap of atomic orbitals.
- To consider the Wigner – Seitz cellular method and to show $\varepsilon_k = \varepsilon_0 + \frac{\hbar^2 k^2}{2m}$ and also to show about estimation of cohesive energy.
- To depict qualitative method of band structure of calculation.

Structure Of The Lesson:

Introduction:

In lesson 12, we considered the consequences of the periodicity of crystalline lattice. There we observed that the potential energy function is a periodic function satisfying

$$V(\mathbf{r}) = V(\mathbf{r} + \mathbf{t}_n) \quad \text{----------} \quad (13.1a)$$

Where $\mathbf{t_n}$ is an arbitrary translation vector in the direct crystal lattice.

Being periodic in the crystal lattice, the potential V(r) may be expanded in a Fourier series and expressed as

$$V(r) = \sum_g V_g e^{ig.r} \qquad \text{----------} \quad (13.1b)$$

where g denotes a reciprocal lattice vector.

To find an appropriate solution to the following one electron time independent Schrodinger equation.

$$H\Psi_{(\mathbf{r})} = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r}) \right]\Psi_{(\mathbf{r})} == \varepsilon\Psi_{(\mathbf{r})} \quad \text{----------} \quad (13.1c)$$

we can try the plane wave expansion

$$\Psi_{(\mathbf{r})} = \Psi = \sum_{\mathbf{k}} C_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}}$$

solution and obtained

$$\left(\frac{\hbar^2 k^2}{2m} - \varepsilon \right) C_{\mathbf{k}} + \sum_{\mathbf{g}} V_{\mathbf{g}} C_{\mathbf{k}-g} = 0 \quad \text{----------} \quad (13.1d)$$

when subject to Born-von Karman periodic boundary condition, the number of allowed k values for the electrons waves will be equal to the number of unit cells, say N

the wave function with eigen value ε_k may be written as

$$\Psi_k(\mathbf{r}) = \sum_{\mathbf{g}} C_{\mathbf{k}-\mathbf{g}} e^{i(\mathbf{k}-\mathbf{g})\cdot\mathbf{r}} = u_k(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}} \quad \text{---------- (13.1e)}$$

with $u_k(\mathbf{r}) = \sum_{\mathbf{g}} C_{\mathbf{k}-\mathbf{g}} e^{-i\mathbf{g}\cdot\mathbf{r}}$ ---------- (13.1f)

From (13.1e) we infer that $u_k(\mathbf{r})$ modulates the plane wave $e^{i\mathbf{k}.\mathbf{r}}$ to a form that serves as a solution to the wave equation (13.1c) for single electron state **k** it is essentially the statement of Bloch function $\Psi_k(\mathbf{r})$ which is written as

$\Psi_{\mathbf{k}}(\mathbf{r}) = u_{\mathbf{k}}(\mathbf{r})\, e^{i\mathbf{k}.\mathbf{r}}$ ---------- (13.1g)

we note that $u_{\mathbf{k}}(\mathbf{r})$ is periodic in the direct crystal lattice. That is,

$u_{\mathbf{k}}(\mathbf{r} + \mathbf{t}_n) = u_{\mathbf{k}}(\mathbf{r})$ ---------- (13.1h)

In lesson 12 we applied Kronig penney model to obtained band structure in one dimension solids. Which was rather a crude approximation. The practical situation in this lesson we try to remove some approximation by using more appropriate models.

13.1 Nearly free electron model:

The energy bands of solids have generally complex structures. Therefore, it would be appropriate to study the influence of a crystal potential on electron in a certain limiting case and then extend the ideas to a general situation. Let us assume that the crystal potential grows slowly from zero value. In the limit of a vanishingly small potential, the Fourier coefficients V_g may be equated to zero in the first approximation. The descriptions of free electron states as given by the energy parabola (fig 13.1) is bound to alter in view of equation (13.1i).

$$\varepsilon_{k+g} = \varepsilon_k \quad \text{---------- (13.1(i))}$$

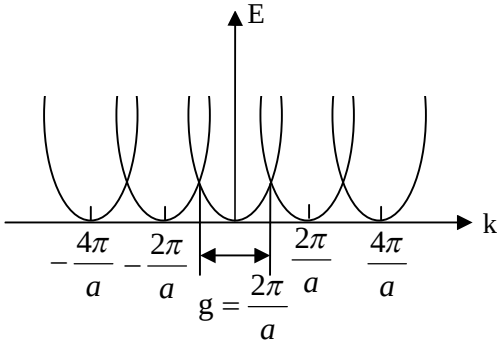


Fig 13.1: Periodic occurrence of the parabolic energy curves of a free electron in one-dimensional reciprocal space. The periodicity in the real space is 'a'. The electron is supposed to be placed in a periodic lattice with a vanishing potential.

The allowed energy states are no more confined to a single parabola in the *k*-space. The states are represented by other parabolae as well, that are displaced by any reciprocal lattice vector **g**.

$$\varepsilon_k = \varepsilon_{k+g} = \frac{\hbar^2}{2m}\left|k+g\right|^2 \quad \text{---------- (13.2)}$$

On the ground of simplicity, the parabolae for one-dimensional crystal are drawn in fig (13.1). The periodicity in real space is 'a' and in the *k*-space it is given by the reciprocal lattice vector $\mathbf{g} = \left[n \cdot \frac{2\pi}{\mathbf{a}} \right]$; n being an integer. At zone boundaries the energy values are degenerate as two parabolae intersect here. The first zone boundaries occur at $k = \pm\frac{1}{2}$ g = $\pm\frac{\pi}{a}$. Therefore the electron wave function with these *k*-values must be represented by a superposition of at least two plane waves which for a small potential may be taken as

$$\exp\left(i\mathbf{g} \cdot \frac{\mathbf{x}}{2} \right);\ \text{and}\ \exp\left[i\left(\frac{1}{2}\mathbf{g} - \mathbf{g} \right) \cdot \mathbf{x} \right] = \exp\left(-i\mathbf{g} \cdot \frac{\mathbf{x}}{2} \right) \quad \text{---------- (13.3)}$$

The waves move in opposite directions.

But here, the reciprocal lattice vectors larger than $\frac{2\pi}{a}$ should also be considered. The value of $C_{\mathbf{k}}$ as determined from (13.1d) is appreciably large when $\varepsilon_{\mathbf{k}}$ and $\varepsilon_{\mathbf{k+g}}$ both approach the value $\left(\frac{\hbar^2 k^2}{2m} \right)$. In this, the absolute magnitude of $C_{\mathbf{k}-\mathbf{g}}$ is approximately equal to that of $C_{\mathbf{k}}$. The two plane waves (13.3) at the first zone boundaries ideally correspond to this condition. Hence other reciprocal vectors can be ignored in the approximation for the construction of wave functions at the zone boundary. The wave functions may be expressed as

$$\Psi_{(+)} \sim e^{i\mathbf{g}\cdot\mathbf{x}/2} + e^{-i\mathbf{g}\cdot\mathbf{x}/2} \sim \cos\frac{\pi x}{a} \quad \text{---------- (13.4)}$$

$$\Psi_{(-)} \sim e^{i\mathbf{g}\cdot\mathbf{x}/2} - e^{-i\mathbf{g}\cdot\mathbf{x}/2} \sim \sin\frac{\pi x}{a} \quad \text{---------- (13.5)}$$

These standing waves appear as a result of Bragg reflection occurring

at k = ± $g/2$ = ± π/a. The electron plane waves when Bragg reflected superpose the waves moving towards the same zone edge where the former suffered Bragg reflection. The probability densities of the two sets of standing waves are

$$\Psi^*_{(+)}\Psi_{(+)} = |\Psi_{(+)}|^2 \sim \cos^2 \frac{\pi x}{a} \quad \text{---------- (13.6)}$$

$$\Psi^*_{(-)}\Psi_{(-)} = |\Psi_{(-)}|^2 \sim \sin^2 \frac{\pi x}{a} \quad \text{---------- (13.7)}$$

The electron potential energy in one-dimensional crystal is shown in fig 13.2

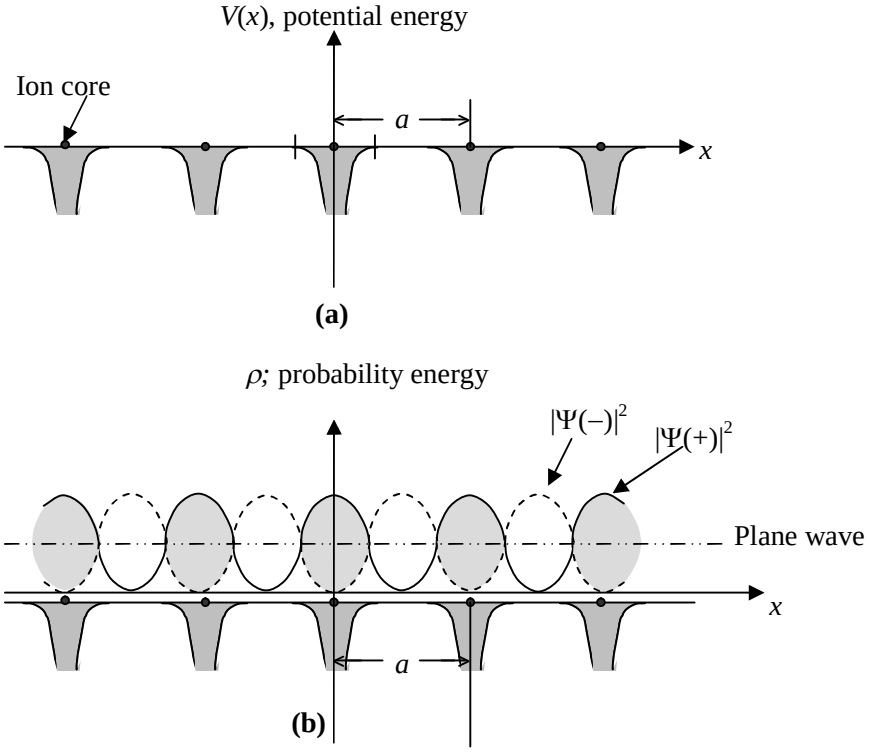


Fig 13.2 (a) Qualitative form of the electron potential energy V(x) in a one-dimensional crystal lattice. Dark circles denote the positions of the positive ion cores. (b) Distribution of probability density for the standing waves Ψ(±) and the plane wave inside the crystal lattice.

The potential field belongs to the positive ion cores whose valance electrons move in this field. The figure also depicts the probability distributions of the standing waves and the simple plane waves. The plane waves exp ($i\mathbf{k}\cdot\mathbf{x}$) have the same probability density at all points since exp (-$i\mathbf{k}\cdot\mathbf{x}$). exp ($i\mathbf{k}\cdot\mathbf{x}$) = 1. The distribution for $\Psi_{(+)}$ favours the piling of

electronic charge exactly above the ion cores. On the other hand the $\Psi_{(-)}$ waves push the electronic charge away from the iron cores. The eigenvalues of $\Psi_{(+)}$ and $\Psi_{(-)}$ differ, though both correspond to the same *k*-value $\left(\frac{\pi}{a} or -\frac{\pi}{a}\right)$. The energy dispersion curve which is continuous throughout the zone shows a gap with two unequal roots at the zone boundaries (fig 13.3). This explains the origin of the band gaps observed in the energy band structure.

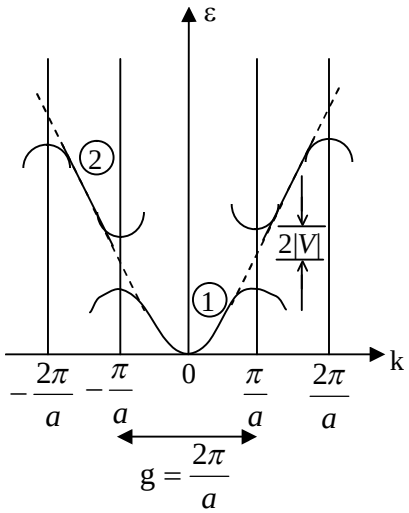


Fig 13.3 Discontinuity in the energy parabola of the free electron at the edges (k = ±π/*a*) of the first Brillouin zone in a one-dimensional lattice.

The Eigenvalue of $\Psi_{(+)}$ is lower in energy since the maxima of its probability density occurs at the points of minimum potential energy. The plane wave energy at the zone edges is centred between the two eigenvalues $\varepsilon_{(+)}$ and $\varepsilon_{(-)}$ corresponding $\Psi_{(+)}$ and $\Psi_{(-)}$ in that order. Now, we are in a position to make an approximate measure of the band gap. The Fourier expansion of a one-dimensional crystal potential has the form

$$V(x) = \sum_{\mathbf{g}} V_{\mathbf{g}} \exp(i\mathbf{g}.\mathbf{x}) \text{ ---------- (13.8)}$$

Appreciating that the potential energy function is real, one can rewrite (13.8) considering only the shortest reciprocal lattice vector as

$$V(x) = 2V \cos\frac{2\pi x}{a} \text{ ---------- (13.9)}$$

with $V = |V_g| = |V_{-g}|$.

Since the magnitude of V_g is known to decrease as g increases, for an approximate calculation one can ignore the contribution from the larger reciprocal vectors.

Using the first order perturbation theory, the band gap is written by

$$\Delta E = \varepsilon_{(+)} - \varepsilon_{(-)} = 2V\int \cos\frac{2\pi x}{a}\left[\Psi^*_{(+)}\Psi_{(+)} - \Psi^*_{(-)}\Psi_{(-)}\right]dx \quad \text{----------} \quad (13.10)$$

Normalizing the wave function at $k = \pm\dfrac{\pi}{a}$ over the crystal's length L, one can have

$$\left.\begin{aligned}\Psi_{(+)} &\sim \left(\frac{2}{L}\right)^{1/2}\cos\frac{\pi x}{a}\\ \Psi_{(-)} &\sim \left(\frac{2}{L}\right)^{1/2}\sin\frac{\pi x}{a}\end{aligned}\right\} \quad \text{----------} \quad (13.11)$$

Using(13.11) one can have from (13.10)

$$\Delta E = \frac{2V}{L}\int_0^L\left(1+\cos\frac{4\pi}{a}x\right)dx$$

or

$$\Delta E = 2V \quad \text{----------} \quad (13.12)$$

Thus, the band gap is equal to twice the magnitude of the leading Fourier coefficient of the crystal potential. The range of allowed energy states covered by the dispersion curve Fig (13.3) in the first Brillowin zone constitutes the first energy band. Similarly, the higher bands are identified with respective dispersion curves in other zones.

13.2 Zone Schemes For Energy Bands:

There are 3 schemes in which the energy bands are drawn.

1. The extended zone scheme. **2.** The reduced zone scheme. **3.**The periodic zone scheme

The Extended Zone Scheme

In this scheme different energy bands are drawn in different zones in the **k** space. The first band is shown in the first zone ($-\pi/a \le k \le \pi/a$) and the next higher in the second zone ($\pi/a \le k \le 2\pi/a$ and $-2\pi/a \le k \le -\pi/a$), and so on.

The Reduced Zone Scheme

All energy bands are shown in the first Brillouin zone in this scheme. As an example, the free electron parabola is shown in this scheme by fig13.4. The curves in the two

segments of the second zone are translated to the first zone by reciprocal vectors $2\pi/a$ and $-2\pi/a$, separately. Similarly the energy pictures in other zones are translated to the first zone by appropriate reciprocal lattice vectors.

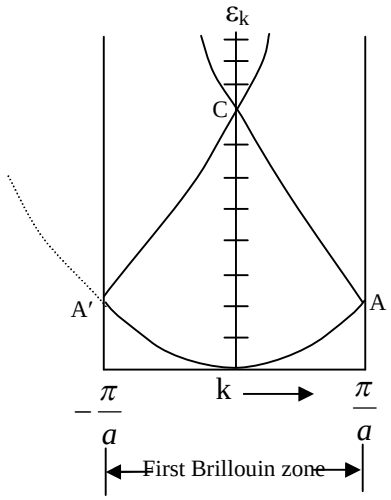


Fig 13.4 The free electron energy parabola plotted in the reduced zone scheme for a one-dimensional lattice. The curve AC, when translated by $-2\pi/a$, reproduces the usual free electron curve for the negative k-values represented by the dashed curve. Similarly, the translation of A′C by $2\pi/a$ will give the curve for the positive k-values. This often gives a useful description of the band structure of a crystal.

The Periodic Zone Scheme:

Every band is drawn in every zone in two schemes. The first three energy bands of a linear crystal as drawn in the three schemes are shown figure 13.5 for the purpose comparison.

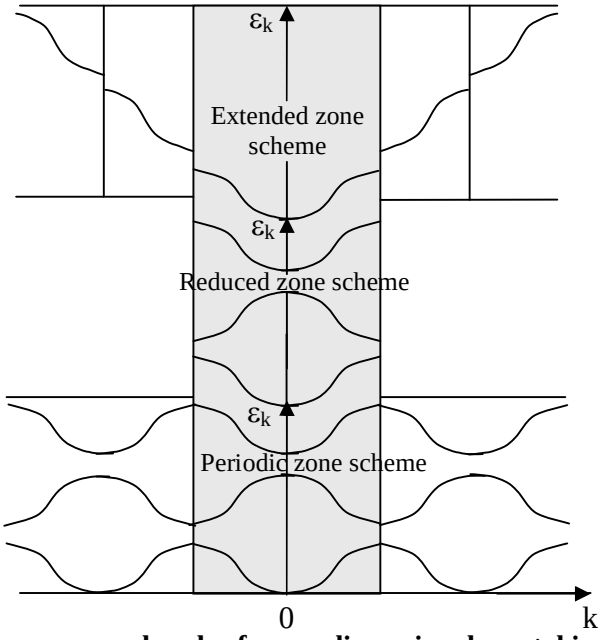


Fig 13.5 First three energy bands of a one-dimensional crystal in the extended, reduced and periodic zone schemes.

13.3 Energy Bands in a general Periodic Potential:

The wave equation

$$H\Psi_{(\mathbf{r})} = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r}) = \varepsilon\Psi_{(\mathbf{r})} \right] \quad \text{---------- (13.13)}$$

in principle can be replaced by

$$\left(\frac{\hbar^2 k^2}{2m} - \varepsilon \right) C_{\mathbf{k}} + \sum_{\mathbf{g}} V_{\mathbf{g}} C_{\mathbf{k}-g} = 0 \quad \text{---------- (13.14)}$$

with out any loss of generality. Hence, the process of treating a general periodic potential in continuum with (13.14) can be written as

$$(\lambda_{\mathbf{k}} - \varepsilon) C_{\mathbf{k}} + \sum_{\mathbf{g}} V_{\mathbf{g}} C_{\mathbf{k}-g} = 0 \quad \text{---------- (13.15)}$$

with

$$\lambda_{\mathbf{k}} = \frac{\hbar^2 k^2}{2m} \quad \text{---------- (13.16)}$$

Equation (13.15) is called the *central equation*. It represents a set of simultaneous linear equations that couple the coefficients $C_{\mathbf{k}-\mathbf{g}}$ for all reciprocal lattice vectors. The number of these equations equals the number of reciprocal lattice vectors. The equations are consistent if the determinant of the coefficients of *Cs* vanishes. Considering only one Fourier component $V_{\mathbf{g}}$, we write three consecutive equations which yield the following determinant

$$\begin{vmatrix} (\lambda_{\mathbf{k}-\mathbf{g}} - \varepsilon) & V & 0 \\ V & (\lambda_{\mathbf{k}} - \varepsilon) & V \\ 0 & V & (\lambda_{\mathbf{k}-\mathbf{g}} - \varepsilon) \end{vmatrix} \quad \text{---------- (13.16a)}$$

with

$$|V_{\mathrm{g}}| = |V_{-\mathrm{g}}| = V.$$

It is a small portion of a huge determinant that is evolved from the set (13.15). By equating (13.16a) to zero and solving for ε we get three roots that fall in three different energy bands at a certain value of *k*. The size of the determinant is chosen according to the extent of the energy spectrum that is required. But the size factor is no more a consideration now, with the availability of fast computers.

On choosing *k* that differs from a value in the first Brillouin zone by a reciprocal lattice vector, there occurs no change in the energy spectrum as the same set of equations in a different order appear giving the same roots of energy. Therefore, very often the *k*-values within the first zone alone are considered.

Roots that refer to a wave vector k are designed as ε_{nk} with n = 1, 2, 3… for the first, second, third…. Bands, respectively, forming the energy spectrum whose different levels belong to different energy bands. The roots ε_{nk} are evaluated by varying k over the allowed set of values. These when arranged on an energy scale produce the fall band structure. The pictures in the three dimensional space are fairly complex. On the account of *k* being direction dependent, the band structure in general appear differently in different directions.

13.3.1 Solution near the Zone Boundary:

It is important to study the solution of the central equation near the zone boundary because of the large deviations in the behavior of free electrons in this region. According to (13.15) if C_{k-g} is an important coefficient then C_k is equally important.

In the two-component approximation, the electron wave functions is taken as

$$\Psi = C_k \exp(i\mathbf{k}\cdot\mathbf{x}) + C_{k-g}\exp[(i(\mathbf{k}\,.\,\mathbf{g}))\cdot\mathbf{x}] \quad \text{----------} \quad (13.17)$$

The two equations coupling C_k and C_{k-g} are obtained from the central equation (13.15), we have a solution if

$$\begin{vmatrix} (\lambda_k - \varepsilon) & V \\ V & (\lambda_{k-g} - \varepsilon) \end{vmatrix} = 0$$

with

$$\lambda_{k-g} = \frac{\hbar^2}{2m} | k - g |^2$$

Solving it for ε, we obtain

$$\varepsilon_k(\pm) = \frac{1}{2}(\lambda_{\mathbf{k}-\mathbf{g}} + \lambda_{\mathbf{k}}) \pm \left[\frac{1}{4}(\lambda_{\mathbf{k}-\mathbf{g}} + \lambda_{\mathbf{k}})^2 + V^2\right]^{1/2} \quad \text{----------} \quad (13.18)$$

The roots ε_k (±) when plotted as a function of wavevector give dispersion curves for the first two energy bands as shown in fig 13.6.

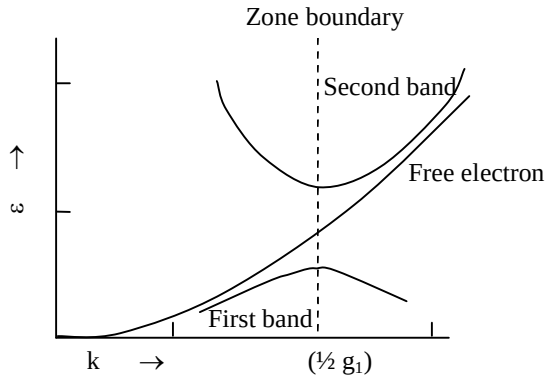


Fig 13.6 Qualitative form of solutions of (13.18) in the periodic zone scheme near the boundary of the first Brillouin zone. The free electron curve is drawn for comparison.

The corresponding roots at the zone boundary [k= $\frac{1}{2}$ g] may be written as

$$\varepsilon(\pm) = \lambda \pm V \qquad \text{---------- (13.19)}$$

where

$$\lambda = \frac{\hbar^2}{2m}\left(\frac{1}{2}g\right)^2 \qquad \text{---------- (13.20)}$$

The λ is the free electron energy at the zone boundary. We find that the band gap obtained from (13.19) is identical with the one given by (13.12). It is instructive however to express the roots near the zone boundary in terms of the wave vector $\tilde{k}$ as measured from the zone boundary. The *k* and $\tilde{k}$ one related by

$$\tilde{k} = k - \frac{1}{2}g \qquad \text{---------- (13.21)}$$

Using (13.21), (13.18) can be expressed as

$$\varepsilon_{\tilde{k}}(\pm) = \varepsilon(\pm) + \frac{\hbar^2 k^2}{2m}\left(1 \pm \frac{2\lambda}{V}\right) \qquad \text{---------- (13.22)}$$

Thus in this definition of the wave vector, the dependence of energy on wave vector is similar to that for the free electrons. When *V* is negative, $\varepsilon_{\tilde{k}}(-)$ corresponds to the upper band.

13.4 Insulators Semiconductors and Metals:

The Electronic transport in solids is found to be closely controlled by their band structures. With the knowledge of the band structure it is possible to predict whether a

solid is good conductor of electricity. By basing on qualitative difference in the band structure we distinguish between Insulator, semiconductor and metals.

We know that the number of energy states in a band is equal to the number of primitive cells (say N) in the crystal. Since each state can accommodate two electrons of opposite spins a number of ($2N$) electrons would be required to fill the band completely. Therefore if a primitive cell contributed an even number of electrons to a band, the band would be completely filled. For example, in a monoatomic crystal in which each atom has one valance electron, the band would be fully occupied if there are two atoms in the primitive cell. With two valence electrons per atom, only one atom in the cell will fill the band completely. Solids in which one or more bands are completely filled and all others are empty, behave as insulators at 0 **K**. The lowest allowed empty band is preceded by a region of forbidden energy gap, across which electrons need be excited for electric conductions. Since the usually applied electric field are not large enough to provide to this excitation, the material behaves as an insulator. When the energy gap is in the range of an eV, the material shows a conductivity intermediate to values of insulator and metals and the material is classified as a semiconductor. The difference between insulator and semiconductors is one of degree and not of type.

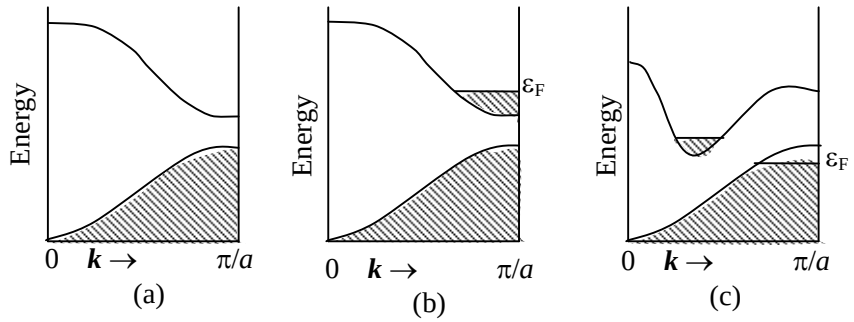


Fig 13.7 Qualitative band schemes for insulators and metals: (a) insulators – the lower band and all below it are completely filled; all higher bands are empty, (b) metals – a partially occupied band, and (c) metals of relatively low conductivity – overlap between, a filled band and an empty band.

On the other hand if the number of valence electrons per primitive cell in an odd number, the top most occupied band is only partially filled. Such a material shows the flow of current on the application of an electric field. These solids are thus good

conductors of electricity and are called metals. Alkali metals and noble metals are the best example of this class where one valence electron is contributed by each primitive cell and the highest occupied band is exactly half filled. Now take up few specific cases to see how far the actual band structures conform to the above principle.

***Sodium*: I**t belongs to the group of alkali metals all having the bcc structure with a rhombohedral primitive cell that contains effectively one atom. The electronic configuration of a sodium atom is $[1s^2\ 2s^2\ 2p^6]\ 3s^1$ since there is a correspondence between the discrete states of an isolated atom and the Brillouin zones in a solid, we expect that the 10 electrons in the closed inner shells form narrow bands in the sodium metal occupying the first five zones (1 for each s–shell and 3 for the p–shell) in the extended zone scheme .the single outermost ($3s^1$) electron produces the half filled band in the next zone. Accordingly the sodium is metallic and so are other alkali metals.

Magnesium: It is a member of alkaline earth metals all of which have two valence electron per primitive cell irrespective of the symmetry. The magnesium atom has electronic configuration $[1s^1\ 2s^2\ 2p^6]\ 3s^2$which would apparently give the insulating behaviour to the solid magnesium, contrary to the observed metallic character. The metallic character arises because of the overlapping of the empty 3p band with the filled 3s band. Thus the 3s electrons can be almost continuously excited to states in the 3p band. The overlap only saves the alkaline earth metals from being branded as insulator. They are not the same good conductors of electricity as the alkali metals. They are rather classified as semi-metals. A quantitative demonstration of the band overlap is shown in figure 13.7 (c). The Fermi energy ε_F in magnesium occurs at an energy which fills the 3s band about 90% with just a small % occupancy for the overlapping sp band.

Diamond:- The isolated carbon atoms have the electronic configuration $-1s^2\ 2s^2\ 2p^2$.

The mixing of 2s and 2p wave functions in the tetrahedrally bonded diamond crystal is well known to result in the sp^3 hybridization. The sp^3 hybrid band further splits into two because of the modification of the s- and p-levels in the crystal. Each of the two hybrid sub bands can accommodate 4 electrons. The 4 electrons belonging to the 2s- and 2p-states fill the lower sub-band, leaving the upper one empty. There exists a forbidden gap

E_g of 5eV at 0 K between the two sub-bands. These features that account for the insulating property of diamond are shown in fig13.8. Band structures of the semi conductors Ge and Si are characterized by similar feature.

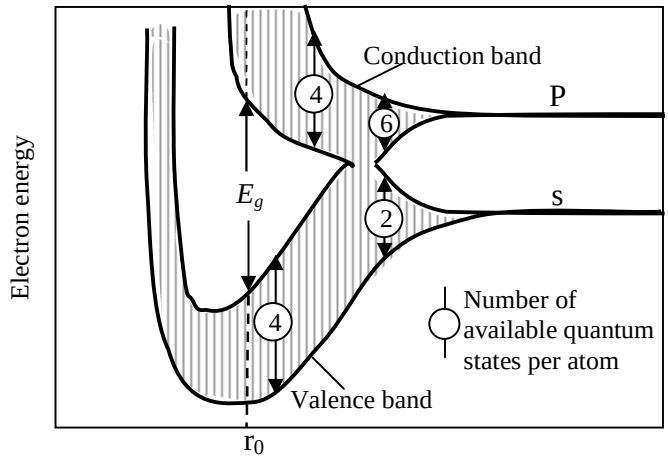


Fig 13.8 Behaviour of energy bands as a function of interatomic separation for diamond (C), Si and Ge. The band gap E_g is defined at the equilibrium separation r_0. The figure shows that the band gap is not tied to the periodicity of the lattice. Amorphous solids can also show a band gap.

13.5 The Tight-Binding Approximation:-

In the NFE model plane wave part exp (*i***k.r**), of the Bloch function [u_k(**r**) exp (*i***k.r**)], is emphasized and the atomic part is overlooked. But electrons in the low-lying inner core levels of a free atom are strongly localized in space. The property is largely retained by these electrons when atoms are assembled to form the solid. It points out to the inadequacy in describing every band character in terms of quasi-free electrons. In order to deal with localized electron, an alternative approach, is followed in which the atomic part of Bloch function is stressed. This approach is known as *tight binding approximation*. The single electron wave function in the crystal is expressed as a linear combination of the atomic orbitals (LCAO) that the electron occupies in a free atom. The forms of Bloch function in a linear crystal for $k = 0$ and $k \neq 0$ are drawn in figure 13.9.

The tight-binding approximation is ideally suited to deal with the inner core electrons. It has been successfully been applied to the d-electrons in transition metals and to the

valence electrons in the diamond live and inert gas crystals .We give below the simplest case of s-state electrons.

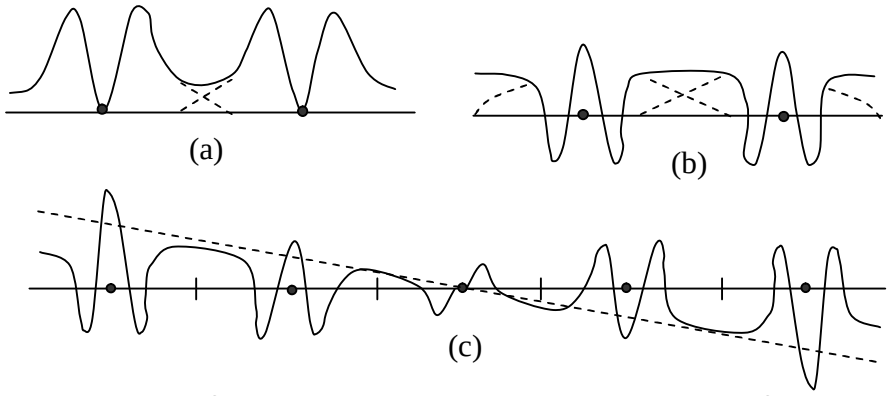


Fig 13.9 (a) Bloch functions corresponding to k = 0 state, built from atomic orbitals (dashed) with small overlap. (b) Bloch functions for a large overlap of the atomic orbitals (dashed). (c) Schematic representation of wave functions corresponding to k ≠ 0. Its form is approximately the product of (b) and cos k (shown as dashed). The real part of the wave function is shown.

Consider an s-electron in an isolated atoms positioned at $\mathbf{r}_n$ with the ground state wave functions $\phi(\mathbf{r}-\mathbf{r}_n)$ where $\mathbf{r}$ determines the electrons position in space. Then the one electron schrodinger wave equation for free atom is

$$H_0\phi(\mathbf{r} - \mathbf{r}_n) = \varepsilon_0\phi(\mathbf{r} - \mathbf{r}_n) \quad \text{---------- (13.23)}$$

Where H_0 and ε_0 are Hamiltonian and the ground state energy of the electron in a free atom. The Hamiltonian for an electron in the crystal is expressed as

$$H = H_0 + v(\mathbf{r}-\mathbf{r}_n) = \left[-\frac{\hbar^2}{2m}\nabla^2 + v_0(\mathbf{r}-\mathbf{r}_n) + v(\mathbf{r}-\mathbf{r}_n) \right] \quad \text{---------- (13.24)}$$

With

$$v(\mathbf{r}-\mathbf{r}_n) = \sum_{m\neq n} v_0(\mathbf{r}-\mathbf{r}_m) \quad \text{---------- (13.25)}$$

In the above equations, $v_0(\mathbf{r}-\mathbf{r}_n)$ denotes the potential energy of an electron when localized at the isolated atom positioned at $\mathbf{r}_n$. The influence of atoms in the vicinity of $\mathbf{r}_n$, where the electron in question is strongly localized on relative terms, is treated as a perturbation on H_0 and represented by $v(\mathbf{r}-\mathbf{r}_n)$. Now, our aim is to look for solutions of the following Schrödinger equation:

$$H\Psi_{\mathbf{k}}(\mathbf{r}) = \varepsilon_k \Psi_{\mathbf{k}}(\mathbf{r}) \quad \text{---------- (13.26)}$$

Where ε_k is the electron energy in the crystal and $\mathbf{\Psi_k(r)}$ the Bloch wavefunction.

In trying the solution

$$\mathbf{\Psi_k(r)} = \sum_n \exp(i\mathbf{k}.\mathbf{r}_n)\phi(\mathbf{r}-\mathbf{r}_n) \text{ ---------- (13.27)}$$

The above function satisfies the properties of a Bloch function. For example it is quite simple to check that the function is periodic in the *k*-space, useing

$$\exp(i\mathbf{g}\cdot\mathbf{r}_n) = 1$$

According to the perturbation theory, the first order energy is given by

$$\varepsilon_k = \frac{\int \Psi_k^*(r) H \Psi_k(r) dV}{\int \Psi_k^*(r)\Psi_k(r)dV} \text{ ---------- (13.28)}$$

Using (13.27), we have

$$\int \Psi_k^*(r)\Psi_k(r)dV = \sum_{n,m} \exp[ik\cdot(r_n - r_m)]\int \phi^*(r-r_m)\phi(r-r_n)dV \text{ ---------- (13.29)}$$

For a strongly localized electron, $\phi(r\text{-}r_m)$ is significant only in the proximity of r_n. Therefore, we evaluate (13.29) by putting m = n in the first approximation. If there are *N* atoms in the crystal, we have

$$\int \Psi_k^*(r)\Psi_k(r)dV = \sum_n 1 = N \text{ ---------- (13.30)}$$

Making use of (13.23),(13.24) and(13.30) electron energy ε_k is written as

$$\varepsilon_k \simeq \frac{1}{N}\sum_{n,m} \exp[ik\cdot(r_n - r_m)]\int \phi^*(r-r_m)[\varepsilon_0 + v(r-r_n)]\phi(r-r_n)dV \text{ ---------- (13.31)}$$

For the term containing ε_0, we again neglect the overlap between the nearest neighbors, putting m = n. Therefore

$$\frac{\varepsilon_0}{N}\sum_{n,m} \exp[ik\cdot(r_n - r_m)]\int \phi^*(r-r_m)\phi(r-r_n)dV = \varepsilon_0 \text{ ---------- (13.32)}$$

Including the overlap upto the nearest neighbors for the perturbation term, we write

$$\int \phi^*(r-r_n)v(r-r_n)\phi(r-r_n)dV = -\alpha \text{ ---------- (13.33)}$$

(on the same atom)

$$\int \phi^*(r-r_m)v(r-r_n)\phi(r-r_n)dV = -\gamma \text{ ---------- (13.34)}$$

(between the nearest neighbours)

All terms in the summation over n , each of which is evaluated over all m (either the same atom or the nearest neighbor) are equal in magnitude on the demand of periodicity.

Since the summation over *n* runs over all atoms in the crystal, the sum is simply *N* times the value of single term. This factor of *N* cancels with the factor *N* in the denominator. In view of this and relations (13.32) (13.33) and (13.34), the electron energy ε_k assumes the form

$$\varepsilon_k \sim \varepsilon_0 - \alpha - \gamma \sum_m e^{i\mathbf{k}\cdot(r_n - r_m)} \quad \text{----------} \quad (13.35)$$

The sum in (12.35) is carried only over the nearest neighbors. In a simple cubic crystal with lattice constant a the nearest neighbor atoms are at

$$r_n - r_m = (\pm a, 0, 0);\ (0, \pm a, 0)\ (0, 0, \pm a)$$

This gives s-state energy in the crystal as

$$\varepsilon_k \sim \varepsilon_0 - \alpha - 2\gamma(\cos k_x a + \cos k_y a + \cos k_z a) \quad \text{----------} \quad (13.36)$$

When the atoms are brought together to form a crystal, the single atomic energy level ε_0 broadens to constitute an energy band whose component levels are defined by (13.36). We can determine the bandwidth as follows.

Energy of the band at its bottom:

The bottom lies at $k = 0$, giving

$$\varepsilon_{bott} = \varepsilon_0 - \alpha - 6\gamma \quad \text{----------} \quad (13.37)$$

Energy of the band at its top:

The top occurs at $k = \pm \pi/a$, giving

$$\varepsilon_{top} = \varepsilon_0 - \alpha + 6\gamma \quad \text{----------} \quad (13.38)$$

Therefore, from (13.37) and (13.38),

$$\text{the bandwidth} = 12\gamma \quad \text{----------} \quad (13.39)$$

Thus the bandwidth is proportional to γ which represents the overlap of atomic orbitals (13.34). A qualitative illustration of the results of tight-binding calculation for a simple cubic crystal is made in fig 13.10. Where as γ determines the bandwidth α is interpreted as the lowering of the center of gravity of the free atomic level on forming the solid. As one proceeds from inner to outer shells in an atom, the width of the respective energy bands goes on increasing because of more and more overlap. This is consistent with the result of the Kronig-Penney model and also confirmed by experiments. It is amply clear from even Fig (13.10) where the second band appears as much wider.

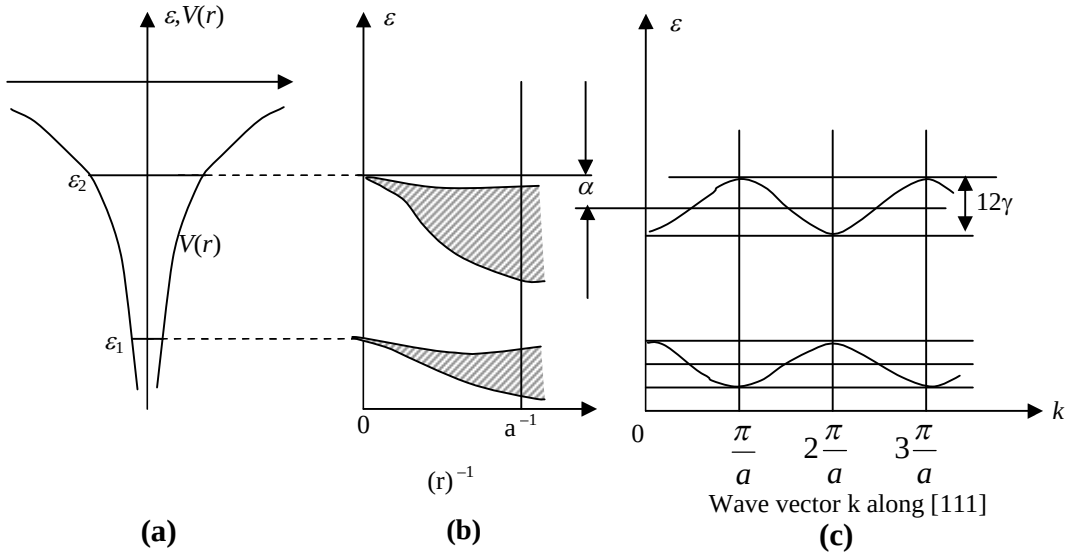


Fig 13.10 Illustration of the result of a tight-binding approximation for a primitive cubic lattice of lattice constant '*a*'. (a) Position of energy levels ε_1 and ε_2 in the potential $V(r)$ of the free atom. (b) Variation in broadening of the levels ε_1 and ε_2 as a function of the reciprocal atomic separation r^{-1}. (c) Variation in one electron energy ε as a function of the wavevector k (1, 1, 1) in the direction of the body diagonal [111].

Furthermore, the effect of tight binding on the energy surfaces of a simple cubic crystal can be examined with the help of (13.36). It is instructive to do this for the limiting *k*-values in the reduced zone scheme, i.e. near the centre and the boundary of the first zone. Near the centre $ka \leq 1$ and we expand the cosine function to obtain

$$\varepsilon_k \approx \varepsilon_0 - \alpha - 6\gamma + \gamma k^2 a^2 \quad \text{---------- (13.40)}$$

These values refer to the bottom region of the band and conform to the constant spherical energy surfaces. But as the wavevector increases, the shape gets distorted and deviates sufficiently from the spherical nature at large value of *k* [fig 13.10(a)].

In order to investigate the region near the zone boundary, we express *k* in terms of its value k' as measured from the zone boundary:

$$k = \frac{\pi}{a} - k' \quad \text{---------- (13.41)}$$

Substituting (13.41) in (13.40) and appreciating that $k'a \leq 1$ near the zone boundary, we obtain

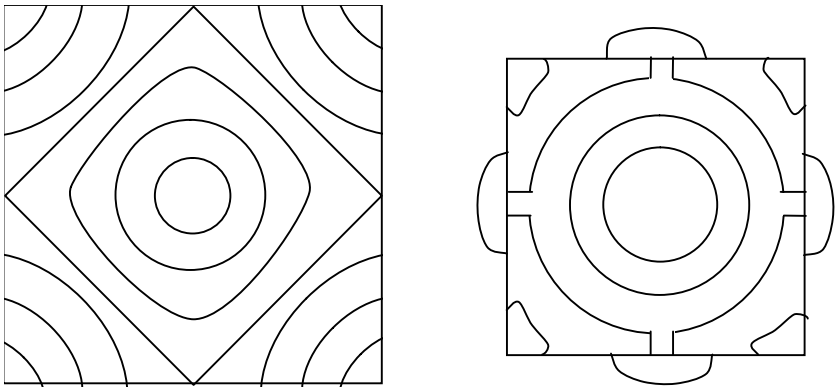
$$\varepsilon_k \approx \varepsilon_0 - \alpha + 6\gamma - \gamma k'^2 a^2 \quad \text{---------- (13.42)}$$

The above result is similar to that of (13.40) with the difference that the spherical surfaces are centred at the corners of the zone (fig 13.11a). We may compare the energy surfaces in the TBA with those in the NFE model, shown in (fig 13.11b). The main difference lies in the fact that the spherical shape is maintained to much larger values of *k* in the NFE model. This only shows that the results in the two models are almost similar for small wavevectors. In view of this fact, the *k*-dependent part of the energy dispersion (13.40) near the zone's centre is comparable with $\hbar^2k^2/2m$ (for the free electrons), and we have

$$\gamma k^2 a^2 = \frac{\hbar^2 k^2}{2m^*} \quad \text{----------} \quad (13.43)$$

giving

$$m^* = \frac{\hbar^2}{2\gamma a^2} \quad \text{----------} \quad (13.44)$$



(a) **(b)**

Fig 13.11 (a) Constant energy curves for the tight-binding approximation for a simple cube crystal in the plane $k_z = 0$. Energy surfaces are spherical around the zone centre for only small k-values. The surfaces are again spherical with centers at the corners of the zone for small k-values, measured from the corners. (b) Constant energy curves in the NFE model. Energy surfaces are spherical up to fairly large values of k. Notice the changes caused by the zone boundary.

where m* is identified as the effective mass of the electron, implying that electron mass should be treated as variable. Equation (13.44) predicts m* is larger for electrons in the inner shells whose overlap is far less. We consider this as an invaluable result of the theory of tight bounding approximation, especially because the concept of effective mass

has remarkably improved the understanding of several physical properties ranging from electronic conduction to the complex optical phenomena.

13.6 The Winger–Seitz Cellular Method:

The significance of the Wigner-Seitz model is substantiated by the impressive success it achieved in accounting for the band structure and the cohesive energy of alkali metals. The first systematic calculation of energy bands appeared in the form of this model. The technique of calculation, referred to as a *cellular method*, is based mainly on the symmetry properties of a certain primitive cell designed by Wigner and Seitz themselves. The method of construction of this cell has been described earlier. The alkali metals have the bcc Structure for which the Wigner-Seitz is a polyhedron and is shown in fig 13.12.

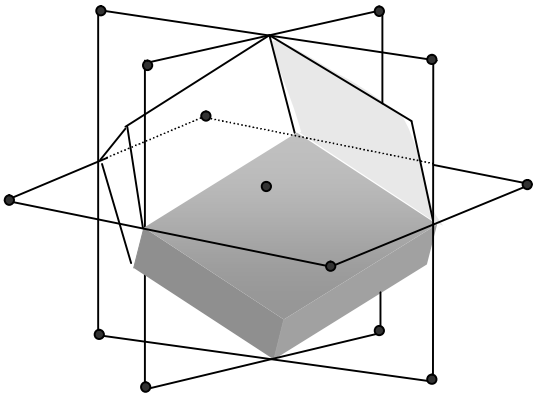


Fig 13.12 The first Brillouin zone of a BCC crystal. It is rhombododecahedral in shape.

The entire crystal volume is imagined to have been filled up with identical polyhedra, assuming that there is only one electron in a given polyhedron at a time, together with the positively charged metal ion at the centre. The polyhedra being neutral, the interactions among themselves are initially neglected and only the interaction within each polyhedron is considered. The electron is supposed to move in the spherically symmetric potential field of the ion. The field is assumed not to extend past the boundaries of the polyhedron.

We consider the extreme case of the k = 0 state. Then the Bloch wavefunction $\Psi_k(r)$ has the form

$$\Psi(r) = u_0(r) \quad \text{----------} \quad (13.45)$$

This wave function itself is periodic in the crystal i.e. the wave function remains unchanged when translated from one face to the opposite face of the cell. It requires that $\frac{\partial \Psi}{\partial \mathbf{n}} = 0$ at the boundary of the polyhedron, with **n** as a direction normal to any of the faces of the polyhedron. The above condition in the crystal replaces the free atom boundary condition, $\Psi(r) \to 0$ as $r \to \infty$, by

$$\left(\frac{\partial \Psi}{\partial r}\right)_{r=r_s} = 0 \quad \text{----------} \quad (13.46)$$

where r_s is the radius of a sphere to which the polyhedron may be approximated. Accordingly, the volume of the polyhedron may be given by $4\pi r_s^3/3$.

The solution of the one-electron Schrodinger wave equation is much easier for $k = 0$ than for any general k-value because $u_0(r)$ is non-degenerate and observes the full symmetry of the crystal. Wigner and Seitz gave an accurate estimate of $u_0(r)$. In view of the boundary condition (13.46), the exercise simply reduces to solving the radial Schrodinger wave equation:

$$\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{2m}{\hbar^2}(\varepsilon_0 - V_0(r))\right]\Psi(r) = 0 \quad \text{----------} \quad (13.47)$$

where $V_0(r)$ is the potential energy of an s-electron and ε_0 is the energy eigenvalue in the field of a crystalline ion within one of the polyhedra.

However, for a general wavevector we have to solve,

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V_0(r)\right]u_k(r)\exp(ik\cdot r) = \varepsilon_k u_k(r)\exp(ik\cdot r) \quad \text{----------} \quad (13.48)$$

But,

$$\nabla^2[u_k(r)\exp(ik\cdot r)] = \nabla[iku_k(r)\exp(ik\cdot r) + \exp(ik\cdot r)\nabla u_k(r)]$$
$$= \exp(ik\cdot r)[\nabla^2 u_k(r) - k^2 u_k(r) + 2ik\cdot\nabla u_k(r)] \text{----------} \quad (13.49)$$

Putting (13.49) in (13.48), we get

$$\left[-\frac{\hbar^2}{2m}(\nabla^2 + 2ik\cdot\nabla) + V_0(r)\right]u_k(r) = \left(\varepsilon_k - \frac{\hbar^2 k^2}{2m}\right)u_k(r) \quad \text{----------} \quad (13.50)$$

It must be observed that $u_0(r)$ is not an exact solution to (13.50) which is in fact satisfied by $u_k(r)$, the periodic part of the general Bloch function.

We treat the k·∇ terms as a perturbation and insist that $u_k(r)$ obey the boundary condition (13.46). This gives electron energies ε_k in the form.

$$\varepsilon_k = \varepsilon_0 + \frac{\hbar^2 k^2}{2m} \quad \text{---------- (13.51)}$$

which gives energies in the shape of a band as measured from the level of ε_0.

We must appreciate that these calculations depend only on the atomic volume and are independent of the crystal structure. Therefore, for a solid metal and a liquid of equal density we expect the same results in this model.

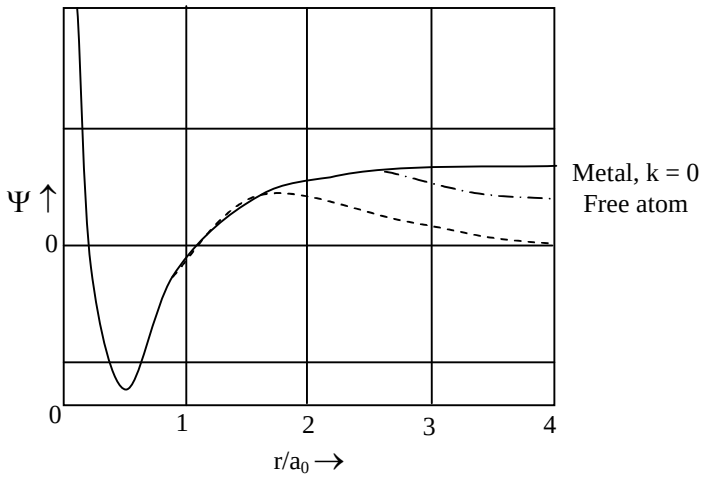


Fig 13.13 Variation of the radial wavefunction for the 3s electron in sodium. The continuous curve describes the Wigner-Seitz wavefunction (cellular wavefunction) at the Brillouin zone center (k = 0). The lower of the dashed curve represents the cellular wavefunction at the zone boundary. The dashed curve in between the cellular plots denotes the wavefunction in the free atom. The distance r of the electron from the centre of the atomic polyhedron is measured in units of the Bohr radius a_0

The 3s radial wave function in sodium metal as estimated in the Wigner-Seitz model at k = 0 (the Brillouin zone centre) and at the Brillouin zone boundary is plotted in Fig.13.13. The 3s atomic wave function is also drawn for comparison when the wave function is subject to the boundary conditions (13.46). Midway between the neighbouring atoms (r = r_s), the radial Schrodinger equation (13.47) for k = 0 yields the eigenvalue (ε_0) as –8.2 eV. This value is considerably lower than the ground state energy of the free atom (–5.15 eV), obtained by applying the boundary condition $d\Psi/dr \to 0$ as $r \to \infty$. The calculated

energy of 3s orbitals at the zone boundary is found to be +2.7 eV. It should be realized that these orbitals are empty because the 3s energy band in sodium is only half-filled and the corresponding states are located near the top of the band. The negligible amplitude of the Wigner-Seitz wavefunction (often referred to as the cellular wavefunction) at the zone boundary is consistent with this fact since it gives little probability for the states in this region to be occupied. It is significant to notice that all the three wavefunctions are identical in the core of the metal ion.

The shape of the cellular wavefunction at the zone centre (k = 0) carries the most vital information about the behaviour of the 3s electron in sodium metal. The plot in fig 13.13 shows the variation of the wavefunction as a function of the distance of the electron from the centre of the atomic polyhedron. The wavefunction is flat over about 90 percent of the atomic volume. The total charge distribution in the flat region corresponds to the charge on an electron. This takes us to the conclusion that $u_k(r)$ remains constant (u_0) over most of the atomic volume and the plane wave part of the wavefunction alone determines the electron motion in this region. Thus the valence electrons of sodium behave mostly as free electrons. This is found to be true for other alkali metals too. But the results for the noble metals, also monovalent, are on the other extreme. The ratio of the ionic to the atomic radius is close to unity making them act like hard spheres. Because of this reason, the noble metals in contrast to the alkali metals cannot be treated in the frame work of the free electron approximation.

13.6.1 Estimation of Cohesive Energy:

Concepts of cohesive energy were discussed earlier. We know that binding of atoms results in the lowering of their ground orbital energy. Thus the ground state energy of an electron in the crystal is lower than that in the free atom. This lowering in energy, taken as a measure of the cohesive energy, is a consequence replacing the Schrodinger boundary condition in the free atom by the periodic boundary condition (13.46). On the demand of energy conservation, an increase in the binding energy is offset by the Fermi energy contribution to the kinetic energy of valence electrons. The spherical

approximation of the Wigner-Seitz theory works satisfactorily in the bcc and fcc crystals. We take sodium metal, a bcc structure for our discussion.

The electron energy in the crystal is given by (13.47). The value of ε_0 is estimated from the Wigner-Seitz model. The second term in (13.47) denoting the average kinetic energy per electron, is obtained from the free electron theory. The ground state energy of the valance electrons in the crystal is then,

$$\varepsilon_k = \varepsilon_0 + \frac{\hbar^2 k^2}{2m} = \varepsilon_0 + \frac{3}{5}\varepsilon_F \quad \text{---------- (13.52)}$$

For sodium, $\varepsilon_0 = -8.2$ eV, $\varepsilon_F = 3.1$ eV (from tables). These values give, $\varepsilon_k = -6.34$ eV.

The above value of ε_k when subtracted from the corresponding value in the free atom (-5.15 eV) gives the cohesive energy as equal to 1.19 eV. The close approximation value (1.13 eV) demonstrates the success of the Wigner-Seitz approximation in alkali metals.

The model provides additional information from the ε_k versus r_s plot. The minimum value of ε_k (13.52) defines the theoretical lattice parameter. For this value of r_s, the cohesive energy and compressibility of alkali metals have been calculated and found in good agreement with the experiments.

13.7 Methods of band structure calculation in use: a qualitative view

We discussed two extreme cases of band structure calculation in the form of tight binding and Wigner-Seitz approximation. While one overstresses the atomic aspect, the other over stresses the plane wave aspect of the Bloch function. The tight binding approximation is useful for interpolation and the Wigner-Seitz approximation gives a good account of several properties of alkali metals. But the methods that actually work for a variety of solids are some what much different. Though these methods are mathematically tedious, the problem has eased considerably with access to modern computers. Of these the orthogonalized plane wave (OPW) and the augmented plane wave (APW) methods are most prominent and advanced. The pseudo potential method is also often used on account of its ability to predict the energy-wavevector relationship with acceptable accuracy.

The knowledge of atomic orbitals at each site readily enables us to distinguish between a plane wave and the complete wavefunction. Based on this idea Herring formulated the OPW method. The orthogonalized plane waves are in fact linear combination of plane waves and mixtures of atomic wave functions of the occupied states of the cores. This takes care of the electron behaviour both within and outside the core regions. The method has been applied to several metals and non-metals with reasonable success in getting the band shapes.

With a large value of the parameter *P*, the Kronig-Penney model changes over to a one-dimensional form of the APW method. It is another way to improve upon the NFE model by approximating the periodic potential suitably in the regions within and outside the cores. The potential within a sphere around each ion core is taken as the usual atomic potential and assumed constant outside the core regions. The Schrodinger wave equation, when solved in the two regions, yields two separate solutions that are matched on the spherical boundaries between the region. The wavefunction within the core is expanded in spherical harmonics. But outside the core region it is represented by a combination of plane waves, known as an augmented plane wave. The matching of the plane waves on to the atomic functions is the most difficult aspect of the exercise.

The Pseudopotential approach has its roots essentially in the effectiveness of the NFE model in many solids. The admixture of core states shows, surprisingly little effect on the energy of higher states in many metallic materials. In this approach the periodic potential energy function is replaced by a modified potential energy function with a few Fourier coefficients V_g that refer to only short reciprocal lattice vectors. The modified potential is called *Pseudopotential*. In the well-known Empirical Pseudopotential Method (EPM), the Fourier coefficients are deduced from the theoretical fits to the optical reflectance and absorption data of the crystal of interest. The potential is fairly smooth and free from the deep wells of the free potential. Success in the interpretation band structure not withstanding difficulties in the description of some other properties of electron come to surface. It is likely to happen since the method may yield an in correct wave function.

Such a problem is resolved by correcting the wave function for the due representation of the atomic component.

The band structure as derived from calculations are too complicated to be discussed.

13.8 Summary:

1. The statement of Bloch theorem is

$$\Psi_{\mathbf{k}}(\mathbf{r}) = u_{\mathbf{k}}(\mathbf{r}) \exp(\mathbf{ik}\cdot\mathbf{r})$$

where $\Psi_{\mathbf{k}}(\mathbf{r})$ denotes an electron wave describing an electron with wave vector $\mathbf{k}$ at $\mathbf{r}$ in the crystal and $u_{\mathbf{k}}(\mathbf{r})$ is the periodic in the crystal lattice. That is $u_{\mathbf{k}}(\mathbf{r} + \mathbf{t}_n) = u_{\mathbf{k}}(\mathbf{r})$ where t_n is an arbitrary translation vector. The function $\psi_{\mathbf{k}}(\mathbf{r})$ is called the Bloch function.

2. Bloch function $\Psi_{\mathbf{k}}(\mathbf{r})$ and their eigenvalues $\varepsilon_{\mathbf{k}}$ are periodic in the reciprocal lattice. That is,

$$\Psi_{\mathbf{k}+\mathbf{g}}(\mathbf{r}) = \Psi_{\mathbf{k}}(\mathbf{r})$$

$$\varepsilon_{\mathbf{k}+\mathbf{g}} = \varepsilon_{\mathbf{k}}$$

3. In a crystal of N primitive cells, there are $2N$ independent orbitals in an energy band.

4. Energy bands are separated by regions in which no solution to electron wave functions exist. These regions are called *band gaps*.

5. The bandwidth is proportional to the overlap of atomic orbitals.

6. The electron effective mass (m^*) is inversely proportional to overlap i.e. m^* is large for electrons in inner shells.

7. The cohesive energy of simple metals is estimated by calculating the lowering of the k=0 orbital in the conduction band. In the calculations, the boundary conditions on the wave function is changed from Schrodinger ($\Psi \rightarrow 0$) as ($r \rightarrow \infty$) to Wigner-Seitz $[(d\Psi/dr)_{r=rs} = 0]$ condition.

13.9 Keywords:

Fourier coefficients – Brillouin zone –NFE model – Central equation – Insulator – Metal – Semiconductor – Tight binding approximation – Effective mass of electron – The Wigner-Seitz cellular method – Cohesive energy – OPW method – APW method – EPM method.

13.10 Review Questions:

1. Describe the nearly free electron model and show that the band gap energy is equal to the twice the magnitude of Fourier coefficient of the crystal potential.

2. Define different zone schemes for energy bands and describe the energy bands in a general periodic potential and obtain the solution near the zone boundary.

3. Distinguish in between insulators, metals and semiconductors based on band theory of solids.

4. Explain the tight binding approximations and obtain the expression for the effective mass of the electrons.

5. Explain the Wigner-Seitz cellular method and obtain the energy $\varepsilon_{\mathbf{k}}$. Illustrate the plot of variation wave function versus r/a_0.

6. Explain how to estimate the cohesive energy. Give a brief account of the various methods for band structure calculation.

13.11 Text and Reference Books:

1. Elements of Solid State Physics by J.P. Srivatsava (PHI)
2. Solid State Physics by M.A. Wahab (Narosa)
3. Elements of Solid State Physics by Omar (Pearson education)
4. Solid State Physics by S.O.Pillai (New Age)
5. Solid State Physics by C.Kittel (Asia Publishing House)
6. Solid State Physics by S.L. Kakani and C. Hemrajani (S.Chand)
7. Solid State Physics by Saxena Gupta Saxena (Pragati Prakashan).
8. Solid State Physics by C.J. Dekker (Macmillan)

UNIT – IV **LESSON: 14**

INTRINSIC SEMICONDUCTORS

Aim: To learn about intrinsic semiconductors

Objectives:

- Classical semiconductors through their band structures.
- Deriving expressions for carrier densities.
- Studying the behaviour of Fermi level.

Structure of the Lesson:

14.1 Introduction:

The values of electrical conductivity of solids are spread over almost the widest range for any common physical property. Hence these form a basis for classifying solids. Solids characterized by extremely high and extremely low values of electrical conductivity are identified as metals and insulators, respectively. A pure metal at 1K may have a conductivity of the order of 10^{8} ohm^{-1} m^{-1} against a low of 10^{-20} ohm^{-1}m^{-1} for an extreme insulator. Materials with conductivity values intermediate to these extreme orders of magnitude are called *semiconductors*. Typical conductivity values for semiconductors lie in the range from 10^{-7} to 1 ohm^{-1} m^{-1}. The most useful feature of semiconductors is that their electrical conductivity generally decreases with increasing purification in contrast to metals where the conductivity always increases with increasing purification.

Notwithstanding the enormous technological importance of semiconductors, their study is even more crucial to the understanding of electronic properties of solids. We will see that it is possible to apply the Maxwell-Boltzmann statistics to deal with charge carriers in semiconductors. This gives exact analytical solutions of many problems that can be solved only by approximate or numerical methods in metals where the Fermi-Dirac statistics has to be used. Furthermore, with remarkable progress having been made in the technology of growing semiconductor crystals, the degree of purity and perfection achieved in growing single crystals of semiconductors is much higher than that in metals and insulators. It is a matter of absolute importance to the study of electronic properties, some aspects of which are obscured by effects arising because of the presence of impurities and crystalline imperfections. It simply amounts to say that the quantitative studies of some phenomena that are generally difficult or cannot be made accurately in metals, are easily carried out in semiconductors with the required precision. Thus the study of semiconductors is helpful in interpreting the electronic properties of solids in general. In this lesson we concentrate on shaping the basic theoretical ideas in the framework of the band theory of solids.

14.2 Semiconductors – classification:

The classification of solids made on the basis of their band structures gives a lead to exploiting the band theory for the interpretation of many solid state properties. We learnt in lesson 13 that completely filled and completely empty bands do not contribute to the

flow of current. The highest filled band (the valence band) in metals is only partially filled and the flow of current occurs on account of almost continuous excitation of valence band electrons to empty states of the band under the influence of an electric field. Therefore, a material that has only completely filled and completely empty bands behaves as a perfect insulator at absolute zero when there can hardly be found any electrons in the lowest empty band (the conduction band) as a result of the thermal excitation of electrons in the valence band. But if the difference between the upper edge of the valence band and the lower edge of the conduction band is not large and less than 2 eV, there is a finite probability for a small fraction of electrons occupying the uppermost states of the valence band to be thermally excited to the conduction band at moderate and high temperatures. At these temperatures the width of the energy range over which the Fermi distribution function rapidly changes is relatively substantial. This enables the consequences of the change in the distribution function easily observable. A representative band scheme for metals, semiconductors and insulators is drawn in fig 14.1. The figure portrays a qualitative difference in the band structures of these solids.

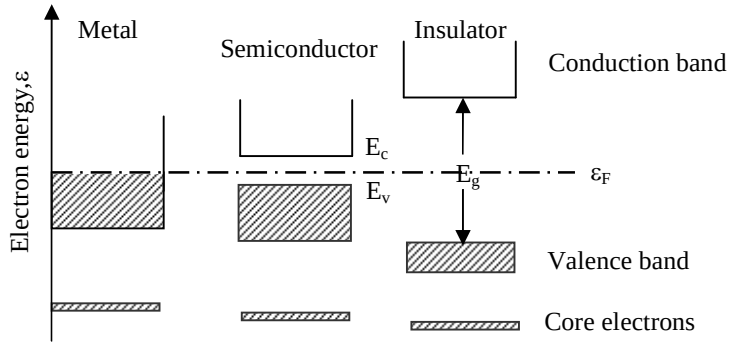


Fig 14.1 Comparative energy band schemes for metals, semiconductors and insulators at T = 0 K. The valence band is completely filled in semiconductors and insulators but only partially filled in metals. A relatively much smaller band gap E_g in semiconductors distinguishes them from insulators.

There is a long list of semiconductor materials. Only a few of them are elements. Si, Ge, grey Sn and grey Se are some examples. The first three are in Group IV and Se is in Group VI of the periodic table. Rest of them are mostly binary compounds, mainly of two types. In one type (e.g. GaAs, InSb, GaP) one element from Group III (e.g. B, Al, Ga, In) is combined with an element from Group V (e.g. N, P, As, Sb). The other type of binary compounds are formed with one element from Group II (e.g. Zn, Cd, Pb) and the

other from Group VI (e.g. S, Se, Te). Some important examples of this type are ZnS, CdSe, PbTe. SiC is the lone example of binary semiconductors whose both components are from Group IV. Some oxides also show semiconducting behaviour – TiO_2, Cu_2O and ZnO are prominent examples of this class. A list of technologically important semiconductor materials is given in Table 14.1 which also contains the relevant data on these materials.

Table 14.1 Data on the band gap of some important semiconductors (D = direct gap, I = indirect gap)

Crystal	Type of gap	E_g (eV)	
		0 **K**	300 **K**
Si	I	1.17	1.11
Ge	I	0.75	0.67
Grey Sn	D	0.00	0.00
GaAs	D	1.52	1.43
GaSb	D	0.81	0.68
GaP	I	2.32	2.25
InSb	D	0.23	0.17
InAs	D	0.43	0.36
InP	D	1.42	1.27
ZnS	—	3.91	3.6
SiC		—	2.86

Semiconductors are primarily of two types – *intrinsic* and *extrinsic*. The intrinsic semiconductors are usually pure monatomic or diatomic solids. An intrinsic material is converted into the extrinsic type (or the impurity type) by adding traces (~1 part in one million) of a suitable impurity with the aim to enhance the level of the charge carrier density. Values of electrical conductivity of intrinsic semiconductors lie far below the range that is useful for the purpose of applications. But semiconductors are gifted with the unique quality that their electrical conductivity can be increased by several orders of magnitude by mixing with them suitable impurities in small concentration. This has enormously increased their technological importance. Thus it is mostly the extrinsic semiconductors that form the basis for semiconductor devices.

14.3 Examples of band structure:

We know from Lesson 13 that the band structure of a solid is closely related with its crystal structure. Not many crystalline structures are favourable to the semiconducting behaviour. Most of the thoroughly investigated semiconductors have a diamond type lattice. We take examples of Si, Ge and GaAs which are distinguished for their large number of applications. The Si and Ge crystals have diamond structure. The crystal lattice of GaAs is zinc blende type which is only a modified diamond lattice. We describe below the band structures of these crystals.

14.3.1 Silicon and Germanium:

The outer electron configurations of Si and Ge are $3s^2 3p^2$ and $4s^2 4p^2$, respectively. The origin of band structure of these materials has been discussed in Lesson 13. The tetrahedral bonding orbitals (sp^3) are formed because of the mixing of s- and p- wave functions. Near the bonding distance, at equilibrium, these orbitals were shown to split into bonding and antibonding orbitals which constitute the valence band and the conduction band, respectively. All the four s- and p- electrons occupy states in the valence band, filling it completely. The completely empty conduction band should combine with this picture to produce an insulating behaviour. This is really the case with the diamond crystal (carbon) which has a similar band scheme. But on account of small band gap (E_g), Si and Ge show semiconducting properties.

Figure 13.8 in lesson 13 reveals an important feature of the band gap regarding its dependence on temperature. The observation that the size of the energy gap (or the splitting) between the valence and conduction bands decreases with increase in the interatomic separation, indicates that the gap is smaller at higher temperatures where the interatomic separation becomes larger because of thermal expansion. This fact is confirmed by the measured values of the band gap E_g at different temperatures (see table 14.1).

Band structure are calculated by fitting the measured physical quantities, such as the band gap, the positions of points of high symmetry in the Brillouin zone (critical points), and the curvature of energy surfaces (the effective mass). The calculated band structures of Si and Ge are shown in fig 14.2. The features of the two band schemes appear quite different in contrast to the qualitative similarity as expected on the basis of fig 13.8. This

is obviously the effect of the difference in electron wavefuctions associated with $3s^2 3p^2$ and $4s^2 4p^2$, configurations.

The symbols, Γ, X and L stand for positions of certain points of high symmetry in the Brillouin zone. They refer to the points at the zone centre (000), $\frac{2\pi}{a}$ (100) and $\frac{2\pi}{a}\left(\frac{1}{2}\frac{1}{2}\frac{1}{2}\right)$, respectively, where '*a*' is the lattice constant. The valence band maximum occurs at the zone centre, i.e. k = 0 for both Si and Ge. But the conduction band minimum occurs for k along the [100] direction in Si and for k along the [111] direction in Ge. This amounts to saying that electrons of the lowest energy in the conduction band have their wavevectors oriented along the [100] direction in Si and along the [111] direction in Ge. In both materials the valence band maximum and the conduction band minimum thus occur at different values of *k*. Semiconductors having this type of band structure are minimum thus occur at different values of *k*. Semiconductors having this type of band structure are called the *indirect gap* semiconductors and those for which the maximum and minimum in question fall at the same value of *k* are referred to as the *direct gap* semiconductors.

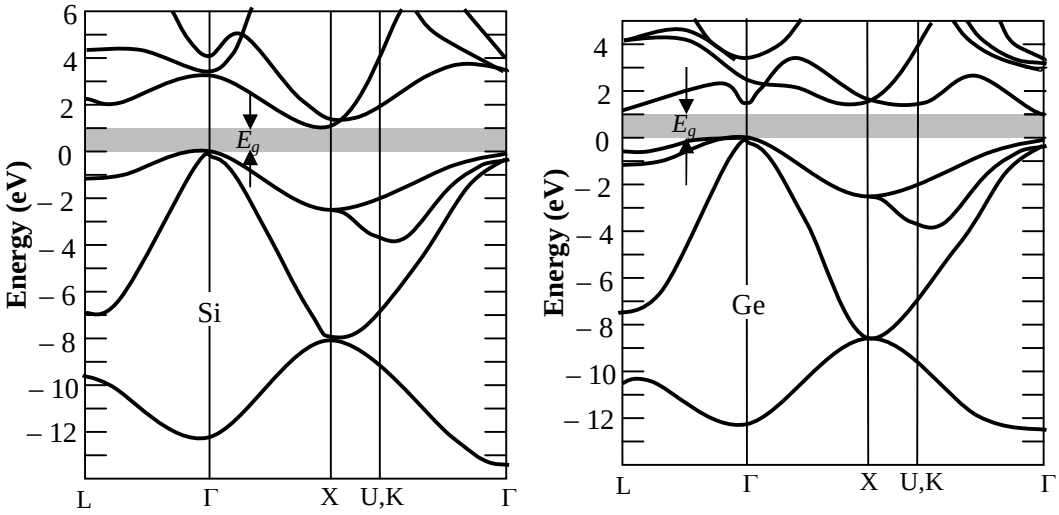


Fig 14.2 Calculated band structures of Si and Ge later J.R. Chelikowsky, M.L. Cohen, phys. Rev. B14, 556 (1976). The structures confirm the indirect gap nature for both Si and Ge. For Ge, the spin-orbit splitting is also considered.

The first Brillouin zone of Si and Ge crystals is a truncated octahedron appropriate to the fcc symmetry of their unit cell. In parabolic approximation (i.e. retaining terms up to

the order k^2 in the expression for ε_k), the surfaces of constant energy are ellipsoids as confirmed by cyclotron resonance studies

In Si there are six symmetry – related minima of the conduction band at points in the [100] directions. Each of the six ellipsoids is an ellipsoid of revolution about a cubic axis by symmetry. They appear as cigars elongated along the cube axes as shown in Fig 14.3(a). The electron has two effective masses – the longitudinal m_l^* (along the axis) and the transverse m_t^* (perpendicular to the axis). Their values in terms of the free electron mass m are given as - $m_l^* = 0.98m$ and $m_t^* = 0.19m$. The valance band shows two degenerate maxima both located at k = 0 with spherical symmetry within validity of the ellipsoidal expansion (fig 14.2). The two effective masses are 0.49m and 0.16m.

The conduction band minima in Ge occur at the zone boundaries in the [111] directions. The minima on parallel hexagonal faces of the zone correspond to the same energy levels giving four symmetry-related conduction band minima [fig 14.2b]. The constant energy surfaces are ellipsoids of revolution with elongation along the [111] directions and effective masses, $m_l^* = 1.57m$ and $m_t^* = 0.082m$. The two degenerate valence band maxima in this case give the characteristic effective masses of 0.28m and 0.44m.

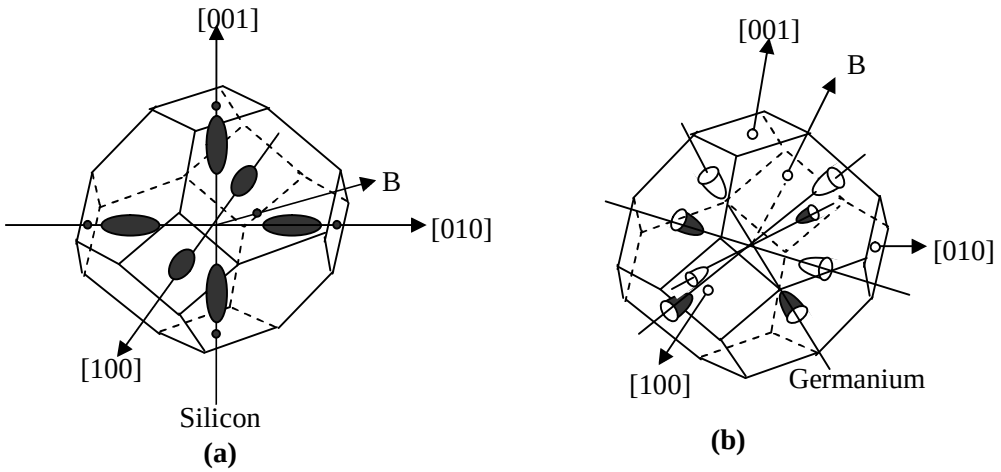


Fig 14.3 Constant energy surfaces for Si and Ge

Properties of holes are equally important as those of electrons for the study of semiconductors. Details of the cyclotron resonance studies show that the structure of the valence band maximum near Γ(k = 0) in Si and Ge is more complicated than what appears in Fig.14.2. In addition to the two degenerate bands at Γ, there is a third one that

is split off slightly by energy band Δ towards a lower energy band (fig 14.4). The splitting is caused by the spin-orbit interaction. On the basis of the effective mass values, the two degenerate bands are attributed to the light and heavy holes. The holes associated with the identity of the third band are named as 'split-off holes'. The split-off energy Δ has been estimated at 0.044 eV in Si and 0.29 eV in Ge.

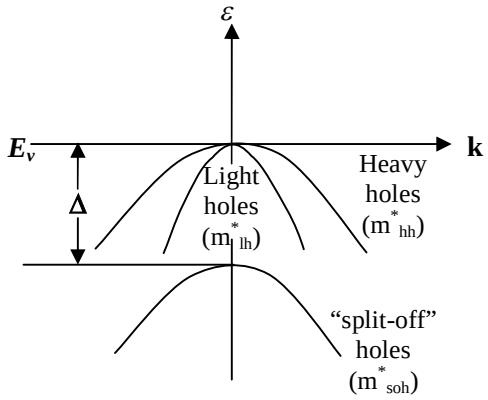


Fig 14.4 Structure of the valence band (qualitative) in Si or Ge near the top.

14.3.2 Gallium Arsenide:

GaAs crystal enjoys a special status, firstly on account of its direct-gap and secondly because of its band-gap energy being closely below the energy range of visible radiation. These properties render it most suitable for the fabrication of efficient optical devices.

The crystal has zinc blende (ZnS) structure. It is an example of the mixed ionic and covalent bonding. The chemical bonding is interpreted as the superposition of these two extreme cases of bonding. In ionic bonding, an electron is transferred from Ga to As to give the ionic structure Ga^+ As^-. On the other hand in the second extreme case, with the displacement of an electron from As to Ga, the number of electrons in the outer shell of both Ga and As atoms becomes four which results in the sp^3 hybridization as in the case of Si and Ge. The observed tetrahedrally coordinated ZnS structure of GaAs serves as a certain proof to the effect that the effects of covalent bonding dominate.

The band structure of GaAs is shown in Fig 14.5. All the valence band maxima and conduction band minima occur at Γ (k = 0) showing its direct gap nature with a gap of 1.43 eV at 300 K. The constant energy surfaces are accordingly spherical. The conduction band effective mass m_e^* is $0.07m$. There are three distinct valence bands

similar in form to those for Si and Ge at Γ. The three respective effective masses are given as $m^*_{lh} = 0.12m$, $m^*_{hh} = 0.68m$ and $m^*_{soh} = 0.2m$ with $\Delta = 0.34$ eV.

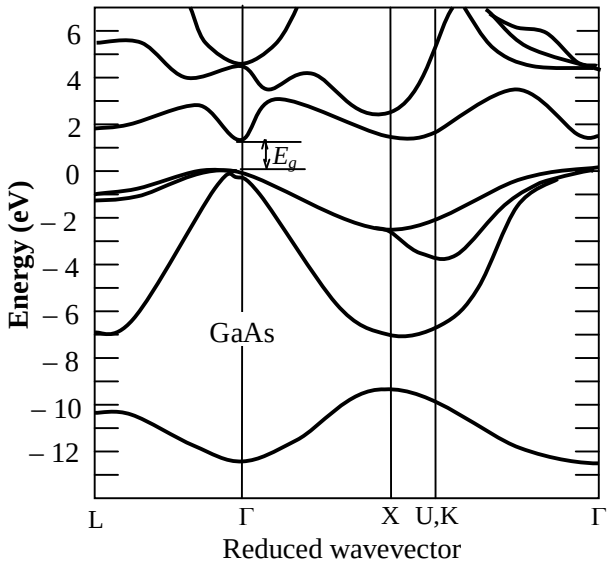


Fig 14.5 Calculated band structures of GaAs, a representative of III-V semiconductors.

14.3.3 Determination of Band Gap:

A number of methods are used to determine the band gap of semiconductors. The technique of continuous optical absorption is more often used on account of its accuracy and the important information gained about the band structure. An abrupt increase in the absorption of optical radiation takes place as soon as the energy of the incident radiation $\hbar\omega$ becomes large enough to exceed the energy gap.

In direct-gap semiconductors (e.g. GaAs, InSb), where the conduction band minimum and the valence band maximum occur at the same k-value in the k-space, the optical threshold at ω (= $E_g/\hbar$) directly gives the band gap [fig 14.6(a)]. But in indirect-gap materials (e.g. Si, Ge, GaP), the direct photon absorption accompanied by the transfer of an electron from the top of the valence band to the bottom of the conduction band would not conserve the crystal momentum because the initial and final points of the transition in the k-space do not have the same k-value. Hence such a direct transition is not allowed. The transition process will have to be indirect in which the absorption of an optical photon must be accompanied by some other process with whose involvement the

condition of momentum conservation may be satisfied. The intensity of a continuous absorption spectrum gets sufficient contribution from phonons. With the involvement of a phonon in the indirect transition under discussion, the sum of wavevectors before the transition becomes equal to their sum after the transition, showing the momentum conservation. There can be two possibilities–one in which a phonon is emitted (created) after the transition and the other in which a phonon is absorbed (destroyed) along with the optical photon to materialize the transition.

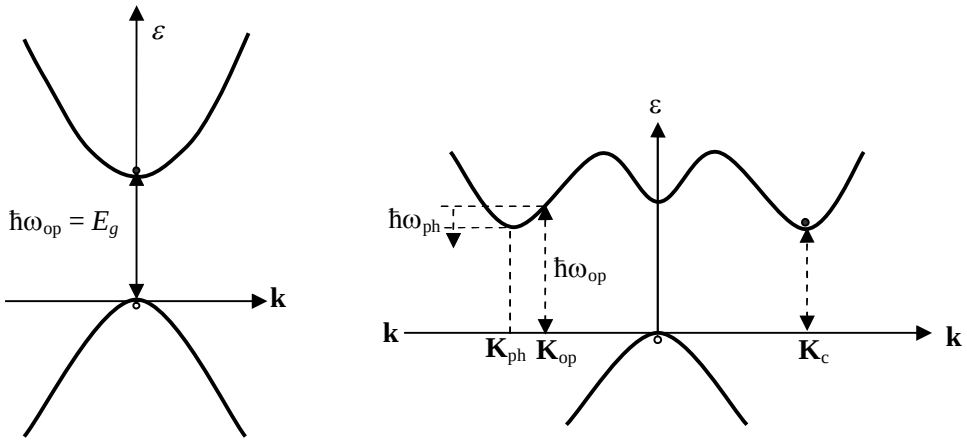


Fig 14.6 (a) Absorption of a photon in a direct-gap semiconductor.
(b) The photon absorption in an indirect-gap material.

Let us denote the wavevector of the optical photon by $\mathbf{K}_{op}$ and that of the phonon by $\mathbf{K}_{ph}$. If the wavevector at the conduction band minimum as measured from the valence band maximum be represented by $\mathbf{K}_{c,}$ then in the case of phonon emission [Fig 14.6(b)], we have

$$\underset{\text{(before the transition)}}{\mathbf{K}_{op}} = \underset{\text{(after the transition)}}{\mathbf{K}_c + \mathbf{K}_{ph}} \quad \text{----------} \quad (14.1)$$

But the photon wavevectros $\mathbf{K}_{op}$ are negligibly small in the range of energies concerned. Therefore, from (14.1),

$$\mathbf{K}_c \cong -\mathbf{K}_{ph} \quad \text{----------} \quad (14.2)$$

Relation (14.2) states that the increase of crystal momentum by $\hbar\mathbf{K}_c$ during the transition is offset by an equal amount owing to the emission of a phonon with wavevector ($-\mathbf{K}_c$). Since the phonon takes away a part of the energy of the incident photon (though very small), the optical threshold energy is greater than E_g [see fig 14.6(b)]. In this case,

$$\hbar\omega_{op} = E_g + \hbar\omega_{ph} \quad \text{----------} \quad (14.3)$$

At high temperatures a good number of phonons are present in the crystal. With the absorption of phonon during the transition, the momentum conservation is given by

$$\mathbf{K}_{op} + \mathbf{K}_{ph} = \mathbf{K}_{c} \text{ ---------- (14.4)}$$

(before the transition) (after the transition)

This gives the optical threshold as

$$\hbar\omega_{op} = E_g - \hbar\omega_{ph} \text{ ---------- (14.5)}$$

So, the optical threshold is lower than E_g. The change in optical threshold in these proceses is generally of little consequence because the phonon energies are characteristically small (~a few hundredths of an eV) as compared to the size of the band gap. This change, however, is of considerable value in semiconductors having a small band gap.

Since the indirect transitions generate heat when phonons are created, the direct-gap materials are preferred in order to have efficient devices. A mixture of the two types of materials is a desired manipulation in some devices such as injection lasers where it is practiced on account of the demand for higher power.

Another method commonly used for determining the band gap is based on the temperature dependent study of the electrical conductivity σ. The electrical conductivity of semiconductors depends on temperature according to the following proportionality:

$$\sigma \propto \exp(-E_g/2k_BT) \text{ ---------- (14.6)}$$

An approximate value of E_g is given by the $\ln(\sigma)$ versus $1/T$ graph. The gap is also determined from the intrinsic carrier concentrations derived from the experimental values of Hall coefficient. The value of band gap and nature of gap for a number of important semiconductors are given in Table 14.1.

14.4 Intrinsic semiconductor:

The difference between the electrical conductivity of semiconductors and metals in the form of former's strong dependence on temperature indicates that thermal excitations control the conductivity of semiconductors in a big way. The thermal excitation of an electron from the valence band across the forbidden energy gap E_g to the conduction band creates a hole in the valence band. The number of these carriers that contribute to the flow of electric current increases more and more with a continued thermal excitation. In

this section we calculate the density of these carriers in the state of thermal equilibrium in an intrinsic semiconductor considered as highly pure such that the contribution to the carrier concentrations from impurities may be neglected. Calculations based on appropriates statistics show that the electron and hole concentrations equal in the present case, are strongly temperature dependent which thus accounts for the conductivity behaviour. The conductivity of a semiconductor may be expressed as

$$\sigma = |e|(n\mu_n + p\mu_p) \quad \text{----------} \quad (14.7)$$

where n and p are electron and hole densities (per unit volume), respectively, and μ_n and μ_p represent the corresponding mobilities. The contribution from electrons and holes in (14.7) simply add up because of the opposite sign of their charge and opposite directions of their drift velocities.

It is easy to appreciate that in excitations of our interest the carriers near the band edges or the parabolic part of the valence and conduction bands are involved. In this region the effective mass can be treated as constant in the first approximation. This takes care of the neglect of energy dependence of mobility in (14.7). For calculations in semiconductors the chemical potential μ appearing in the Fermi distribution function is replaced by the Fermi energy ε_F. But the level of chemical potential always lies in the region of the forbidden energy gap where no single-electron energy levels exist. The very definition of the Fermi level which is considered as the highest occupied level at absolute zero, thus becomes redundant since no single-electron energy level is available to coincide with the Fermi level. Therefore, in the fitness of things the Fermi level in semiconductors should be interpreted as a synonym to the chemical potential.

14.5 Electron and hole densities in intrinsic semiconductors:

The occupancy of energy levels in semiconductors must be described the by Fermi-Dirac distribution function $f(\varepsilon, T)$ as in the other solids:

$$f(\varepsilon, T) = \frac{1}{1 + \exp[(\varepsilon - \varepsilon_F)/k_B T]} \quad \text{----------} \quad (14.8)$$

If $D_c(\varepsilon)$ and $D_v(\varepsilon)$ denote the density of states in the conduction and valence band, respectively, the charge carrier densities are usually written as

$$n = \int_{E_c}^{\infty} D_c(\varepsilon) f_e(\varepsilon, T) d\varepsilon \quad \text{---------- (14.9)}$$

with f_c as the electron occupancy and

$$p = \int_{-\infty}^{E_v} D_v(\varepsilon) f_h(\varepsilon, T) d\varepsilon = \int_{-\infty}^{E_v} D_v(\varepsilon)[1 - f_e(\varepsilon, T)] d\varepsilon \quad \text{---------- (14.10)}$$

with f_h as the hole occupancy.

Here E_c and E_v refer to the energy values at the edges of the conduction and valence band, respectively (fig 14.7). The electron energy as measured from the bottom (or edge) of the conduction band is $(\varepsilon - E_c)$ where ε is the absolute value of electron energy. Similarly, the hole energy when measured from the top (or edge) of the valence band is equal to $(E_v - \varepsilon)$. Applying the parabolic approximation, i.e. assuming the effective mass to remain constant, the two density of states are written as

$$D_c(\varepsilon) = \frac{1}{2\pi^2}\left(\frac{2m_e^*}{\hbar^2}\right)^{3/2} (\varepsilon - E_c)^{1/2} \quad \text{---------- (14.11)}$$

and

$$D_v(\varepsilon) = \frac{1}{2\pi^2}\left(\frac{2m_h^*}{\hbar^2}\right)^{3/2} (E_v - \varepsilon)^{1/2} \quad \text{---------- (14.12)}$$

When electron and hole energies are such that $|\varepsilon - \varepsilon_F| >> k_B T$, the respective distribution

$$f_e = \exp\left[\frac{-(\varepsilon - \varepsilon_F)}{k_B T}\right] \quad \text{---------- (14.13)}$$

(in the conduction band)

and

$$f_h = 1 - f_e \exp\left[\frac{-(\varepsilon_F - \varepsilon)}{k_B T}\right] \quad \text{---------- (14.14)}$$

(in the valence band)

In this condition the carrier density is not large and the semiconductor is called a non-degenerate semiconductor.

Substituting (14.13) in (14.9), the electron density in the state of thermal equilibrium at temperature T has the form,

$$n = \frac{1}{2\pi^2}\left(\frac{2m_e^*}{\hbar^2}\right)^{3/2} \exp\left(\frac{\varepsilon_F}{k_B T}\right) \int (\varepsilon - E_c)^{1/2} \exp\left(-\frac{\varepsilon}{k_B T}\right) d\varepsilon \quad \text{---------- (14.15)}$$

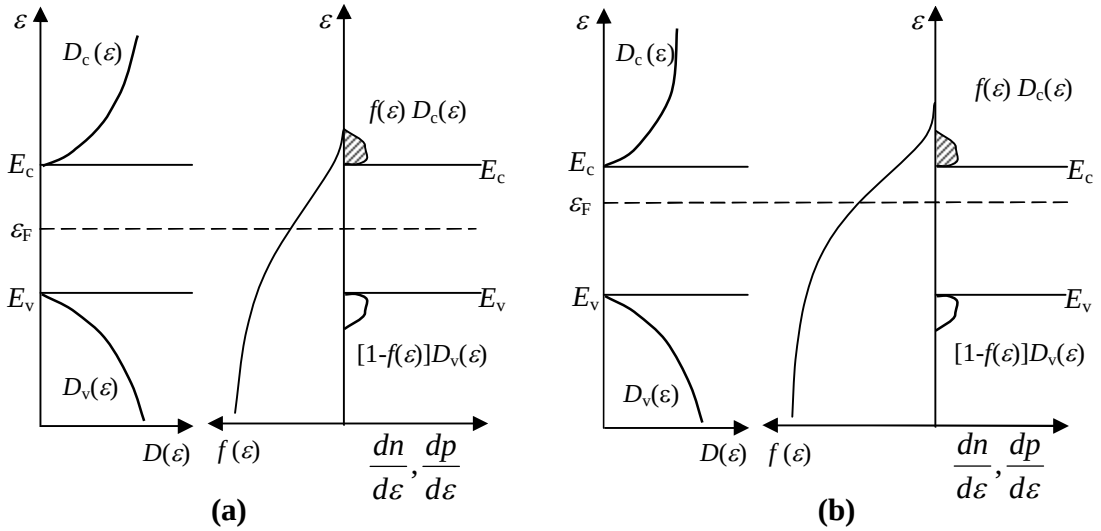


Fig 14.7 A schematic behaviour of the Fermi function *f*(ε), the density of states *D*(ε), and electron (*n*) and hole (*p*) densities in the conduction and valence bands when: (a)the densities of states in the conduction and valence bands are equal and (b) the densities of states in the conduction and valence bands are not equal.

Using the substitution $\left(\frac{\varepsilon-\varepsilon_c}{k_B T}\right) = x^2$, we get

$$n = \frac{1}{\pi^2}\left(\frac{2m_e^*}{\hbar^2}\right)^{3/2} (k_B T)^{3/2} \exp\left(\frac{-(E_c-\varepsilon_F)}{k_B T}\right)\int_0^\infty x^2 \exp(-x^2) d\varepsilon \quad \text{---------- (14.16)}$$

But,

$$\int_0^\infty x^2 \exp(-x^2) d\varepsilon = \frac{\sqrt{\pi}}{4} \quad \text{---------- (14.17)}$$

(from standard tables)

Using (14.17) in (14.16), we have

$$n = 2\left(\frac{m_e^* k_B T}{2\pi\hbar^2}\right)^{3/2} \exp\left(\frac{-(E_c-\varepsilon_F)}{k_B T}\right) \quad \text{---------- (14.18)}$$

Similarly for holes, using (14.14) and (14.10), we get

$$p = 2\left(\frac{m_h^* k_B T}{2\pi\hbar^2}\right)^{3/2} \exp\left(\frac{-(\varepsilon_F-E_v)}{k_B T}\right) \quad \text{---------- (14.19)}$$

By multiplying (14.18) by (14.19), we get the following generally valid relationship:

$$np = 4\left(\frac{k_B T}{2\pi\hbar^2}\right)^3 (m_e^* m_h^*)^{3/2} \exp\left(\frac{E_g}{k_B T}\right) \quad \text{---------- (14.20)}$$

where we have used $E_c - E_v = E_g$.

Relation (14.20) represents the law of mass action. It states that the product of electron and hole concentrations in a certain semiconductor that is completely characterized by its band gap E_g and effective masses m_e^* and m_h^* in the conduction and valence bands behaves as a function of temperature and as independent of the position of the Fermi level.

As already mentioned, the number of electrons is equal to the number of holes in an intrinsic semiconductor. So, in this case if n_i and p_i denote the respective concentrations, we get from (14.20)

$$n_i = p_i = 2\left(\frac{k_B T}{2\pi\hbar^2}\right)^{3/2} (m_e^* m_h^*)^{3/4} \exp\left(-\frac{E_g}{2k_B T}\right) \quad \text{---------- (14.21)}$$

The numbers as determined at 300 K in Si, Ge and GaAs are 2.4×10^{23}, 1.5×10^{10} and 5×10^{7} cm^{-3}, respectively.

14.6 Fermi level in Intrinsic Semiconductors:

From 14.20 it is evident that np = $np = n_i^2 = p_i^2 = \text{const.}\, T^3 e^{-E_g / k_B T}$.

Where n_i is a constant depending upon the temperature and width of the forbidden gap. It does not depend on the impurities introduced as long as the impurities do not change the width of the forbidden energy gap.

In intrinsic materials the Fermi level at a certain temperature adjusts its position that is required to maintain the charge neutrality condition consistent with (14.18) and (14.19). This demands that

$$n = p$$

giving

$$\exp\left(\frac{2\varepsilon_F}{k_B T}\right) = \left(\frac{m_h^*}{m_e^*}\right)^{3/2} \exp\left(\frac{E_v + E_c}{k_B T}\right) \quad \text{---------- (14.22)}$$

or

$$\varepsilon_F = E_v + \frac{1}{2}E_g + \frac{3}{4}k_B T \ln\left(\frac{m_h^*}{m_e^*}\right) \text{---------- (14.22)}$$

Thus the Fermi level shows a weak temperature dependence. But if the electron and hole effective masses are equal, $\varepsilon_F = E_v + \frac{1}{2}E_g$ and the Fermi level lies exactly in the middle of the gap. This is true at all temperatures.

But in general when the density of states of conduction and valence bands and the two effective masses are unequal, the Fermi level is asymmetrically placed with respect to the positions of E_c and E_v. It is demonstrated in fig (14.7(b)) where the two density of states are unequal.

Intrinsic conductivity:

From equations (14.7) and (14.21) the electrical conductivity in the intrinsic region will be

$$\sigma_i = 2|e|\left(\frac{k_B T}{2\pi\hbar^2}\right)^{3/2}(m_e^* m_h^*)^{3/4}\exp\left(\frac{E_g}{2k_B T}\right)(\mu_e + \mu_h)$$

As the mobilities are likely to depend on temperature only as a simple power law over an appropriate region, the temperature dependence of the conductivity will be dominated by the exponential dependence of the carrier concentration.

Mobility in the intrinsic region:

The mobility is defined as the drift velocity per unit electric field. In an ideal intrinsic semiconductor the mobility is determined by lattice scattering; that is, by collisions between lattice waves and electron waves. In actual intrinsic specimens there are always some impurity atoms which may dominate the scattering of electrons at low temperatures when the lattice waves are quiescent, but at higher temperatures the lattice scattering is dominant.

The mobility associated with lattice scattering in a non-polar (covalent) crystal such as diamond, silicon, or germanium has been calculated by Seitz and others. Seitz' finds

$$\mu = \frac{2^{1/2}6^{1/3}}{4\pi^{5/6}}\cdot\frac{N^{1/3}e\hbar^2k^2\Theta^2M}{m^{*5/8}C^2(kT)^{3/2}} \qquad \text{----------} \quad (14.23)$$

where Θ is the Debye temperature; k is the Boltzmann constant; N is the density of unit cells; m* is the effective mass; M is the atomic mass; and C is defined using the Bloch function $u(r)e^{i\mathbf{k}.\mathbf{r}}$ by

$$C = \frac{\hbar^2}{2m}\int |\text{grad } u|^2 \, dr \qquad \text{----------} \quad (14.24)$$

and is treated as an unknown parameter of value of the order of 1 to 10 eV.

Table 14.2 Typical order of carrier mobilities at room temperature in some important semiconductors.

Crystal	*Mobility* (cm^2/V. s)		*Crystal*	*Mobility* (cm^2/V. s)	
	Electrons	*Holes*		*Electrons*	*Holes*
Si	1350	480	InSb	800	450
Ge	3600	1800	InAs	30,000	450
GaAs	8000	300	InP	4500	100
GaSb	5000	1000	PbSe	1020	930
AlAs	280	—	PbTe	2500	1000
AlSb	900	400	SiC	100	10 – 20

Experimental values of the mobility at room temperature are given in table 14.2. In most substances the quoted values are probably representative of lattice scattering. We notice a tendency for crystal with small energy gaps to have high values of the electron mobility. This is because small gaps imply small effective masses and (14.23) shows that small masses favor high mobilities. The low mobilities generally characteristic of holes in diatomic crystals are believed to be connected with the complex degenerate forms of the energy surfaces in such crystals at the top of the valence band. By comparison the mobility of metallic copper is 35 cm^2/v-sec at room temperature.

14.7 Summary

Semiconductors are primarily of two types – *intrinsic* and *extrinsic*. The intrinsic semiconductors are usually pure monatomic or diatomic solids.

Values of electrical conductivity of intrinsic semiconductors lie far below the range that is useful for the purpose of applications

Not many crystalline structures are favourable to the semiconducting behaviour. Most of the thoroughly investigated semiconductors have a diamond type lattice.

The outer electron configurations of Si and Ge are $3s^2 3p^2$ and $4s^2 4p^2$, respectively.

All the four s- and p- electrons occupy states in the valence band, filling it completely. The completely empty conduction band should combine with this picture to produce an insulating behaviour. This is really the case with the diamond crystal (carbon) which has a similar band scheme. But on account of small band gap (E_g), Si and Ge show semiconducting properties.

The valence band maximum occurs at the zone centre, i.e. k = 0 for both Si and Ge. But the conduction band minimum occurs for k along the [100] direction in Si and for k along the [111] direction in Ge. This amounts to saying that electrons of the lowest energy in the conduction band have their wavevectors oriented along the [100] direction in Si and along the [111] direction in Ge. In both materials the valence band maximum and the conduction band minimum thus occur at different values of *k*. Semiconductors having this type of band structure are minimum thus occur at different values of *k*. Semiconductors having this type of band structure are called the *indirect gap* semiconductors and those for which the maximum and minimum in question fall at the same value of *k* are referred to as the *direct gap* semiconductors.

In Si there are six symmetry – related minima of the conduction band at points in the [100] directions.

The conduction band minima in Ge occur at the zone boundaries in the [111] directions.

Properties of holes are equally important as those of electrons for the study of semiconductors.

GaAs crystal enjoys a special status, firstly on account of its direct-gap and secondly because of its band-gap energy being closely below the energy range of visible radiation. These properties render it most suitable for the fabrication of efficient optical devices.

GaAs crystal has Zincblende (ZnS) structure. It is an example of the mixed ionic and covalent bonding. All the valence band maxima and conduction band minima occur at Γ (k = 0) showing its direct gap nature with a gap of 1.43 eV at 300 K. The constant energy surfaces are accordingly spherical.

A number of methods are used to determine the band gap of semiconductors. The technique of continuous optical absorption is more often used on account of its accuracy and the important information gained about the band structure. An abrupt increase in the

absorption of optical radiation takes place as soon as the energy of the incident radiation $\hbar\omega$ becomes large enough to exceed the energy gap.

Since the indirect transitions generate heat when phonons are created, the direct-gap materials are preferred in order to have efficient devices. A mixture of the two types of materials is a desired manipulation in some devices such as injection lasers where it is practiced on account of the demand for higher power.

The law of mass action states that the product of electron and hole concentrations in a certain semiconductor that is completely characterized by its band gap E_g and effective masses m_e^* and m_h^* in the conduction and valence bands behaves as a function of temperature and as independent of the position of the Fermi level.

The Fermi level shows a weak temperature dependence. But if the electron and hole effective masses are equal, $\varepsilon_F = E_v + \frac{1}{2}E_g$ and the Fermi level lies exactly in the middle of the gap. This is true at all temperatures.

But in general when the density of states of conduction and valence bands and the two effective masses are unequal, the Fermi level is asymmetrically placed with respect to the positions of E_c and E_v.

In a semiconductor at a given temperature, the product np of the electron and hole concentrations is constant and independent of purity of crystal.

Effective density of states in the conduction band

$$N_c = 2\left(\frac{2\pi m_e^* k_B T}{\hbar^2}\right)^{3/2}$$

and the effective density of states in the valence band

$$N_v = 2\left(\frac{2\pi m_h^* k_B T}{\hbar^2}\right)^{3/2}$$

Electrical conductivity of an intrinsic semiconductor.

$\sigma = |e| (n\mu_c + p\mu_h)$

14.8 Key words

Valence band – Conduction band – Indirect gap semiconductors – Direct gap semiconductors – Band gap – Mobilities of charge carriers – Fermi level – Optical absorptions.

14.9 Review questions

1. What are intrinsic semiconductors? Obtain expression for the intrinsic carrier concentration in an intrinsic semiconductor. Under what condition will Fermi level be in the middle of the forbidden gap.
2. Derive the law of mass action in semiconductor.
3. Describe the band structure of Germanium and Silicon. Compare the band structure of these with that of Gallium Arsenide.
4. Describe how you determine the band gap in indirect gap semiconductor.
5. Derive an expression for electrical conductivity of intrinsic semiconductors.

14.10 Text and Reference Books:

1. Elements of Solid State Physics by J.P. Srivatsava (PHI)
2. Solid State Physics by M.A. Wahab (Narosa)
3. Elements of Solid State Physics by A. Omar (Pearson education)
4. Solid State Physics by S.O. Pillai (New Age)
5. Solid State Physics by C. Kittel (Asia Publishing House)
6. Solid State Physics by S.L. Kakani and C. Hemrajani (S.Chand)
7. Solid State Physics by Saxena Gupta Saxena (Pragati Prakashan).
8. Solid State Physics by C.J.Dekker (Macmillan)

UNIT – IV **LESSON: 15**

EXTRINSIC SEMICONDUCTORS

Aim: To estimate the population of donor and acceptor levels and to study the Extrinsic carrier densities and temperature dependence of electrical conductivity.

Objectives:

- To learn about the properties of extrinsic semiconductors.
- To obtain an expression to the population of donor and acceptor levels.
- To obtain an expression for the electrical conductivity of extrinsic semiconductor.
- To obtain an expression for the Fermi level in extrinsic semiconductors.

Structure of the Lesson:

15.1 Extrinsic Semiconductors:

As mentioned in lesson 14, the extrinsic semiconductors were developed because the electrical conductivity of intrinsic semiconductor is, generally, not large enough to meet the requirement of devices. Besides, it is hard to imagine of an absolute intrinsic material. Even the purest single crystals have impurity contents to a certain degree. But these impurity contents normally do not increase the carrier concentration to a useful level. GaAs is, however, an exception. The purest available single crystals of GaAs show a carrier density of about 10^{16} cm^{-3} which is enormously high compared to the intrinsic value (5×10^{7} cm^{-3}). In general, the standard method of increasing the conductivity of an intrinsic material is to add to it a suitable impurity or electrically active element in small concentration. The method is known as *doping*. Impurities that enhance the carrier density by contributing additional electrons to the conduction band are called *donors* and those which create additional holes in the valence band are known as *acceptors*.

For example, let us consider the electrically active elements suitable for doping in Si and Ge crystals. When pure crystals of Si and Ge are doped with any element from Group III (e.g. B, Al, In), *p*-type semiconductors are formed.

15.2 The *n*-type Semiconductors:

Suppose a pure Ge crystal is doped with As, an immediate neighbour to the right of Ge in the periodic table. The arsenic atom may enter the germanium crystal lattice either by replacing a germanium atom (i.e.substitutionally) or by occupying a position where no germanium atom is supposed to be located in a pure and perfect crystal (i.e. interstitially). Data on lattice constant measurements show that the arsenic atom enters the crystal substitutionally. The electronic measurements show that the arsenic atom enters the crystal substitutionally. The electronic configurations in the outermost shells of germanium and arsenic atoms are $3s^2$ $3p^2$ and $4s^2$ $4p^2$, respectively. The germanium crystal has a diamond structure in which each atom forms tetrahedral bonds with its four neighbours. When an arsenic atom that substitutes a germanium atom finds itself surrounded by four germanium atoms, its four electrons get engaged in tetrahedral bonds with all four neighbouring atoms as depicted in fig 15.1. The arsenic atom is left with an extra loosely bound valence electron which may be easily freed and made available in the

conduction band for the purpose of conduction. On account of having these additional electrons for conduction, the doped crystal is called an *n*-type semiconductor where *n* stands for electrons.

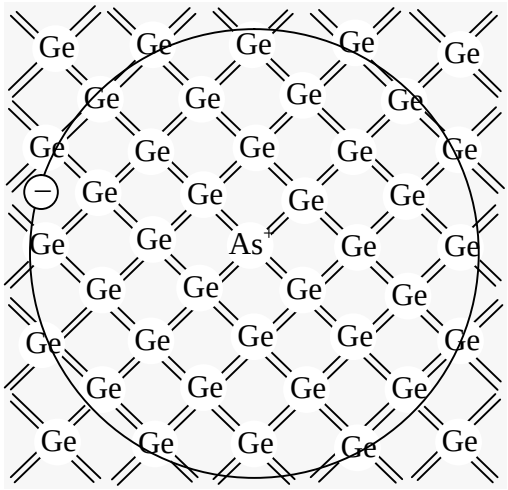


Fig 15.1 A donor impurity atom As is shown to substitute a Ge atom in the germanium crystal.

According to a model proposed to interpret the replacement of a germanium atom by an arsenic atom the arsenic atom is pictured as an occasional germanium atom with an additional positive charge of *e* fixed at its core to which an additional electron is bound. Thus the donor impurity can be described as a hydrogen-like centre in which the Coulomb attraction between the core and the valence electron is screened by the neighbouring germanium electrons. The centre being a bound system is characterized by a set of quantized energy levels whose scheme is qualitatively similar to that of the hydrogen atom. The ionization energies of the hydrogen and the donor atoms are expressed as

$$E_{\mathrm{H}} = \frac{me^4}{2(4\pi\varepsilon_0\hbar)^2} \quad \text{---------- (15.1)}$$

for the hydrogen atom

and

$$E_{\mathrm{d}} = \frac{m_{\mathrm{e}}^* e^4}{2(4\pi\varepsilon_{\mathrm{m}}\hbar)^2} = \frac{m_{\mathrm{e}}^* E_{\mathrm{H}}}{m\varepsilon_{\mathrm{s}}^2} \quad \text{---------- (15.2)}$$

for the donor atom

Here ε_0 is the permittivity of vaccum space and ε_s is the relative permittivity or the static dielectric constant of the medium of the germanium crystal. The constant ε_m (= ε_0 ε_s)

represents the permittivity of the medium of the crystal. The values of the static dielectric constant of some important semiconductors are given in Table15.1.

For calculating E_d we take the effective mass of the conduction electron in germanium as $m_e^* = 0.12m$ where m is the free electron mass. Using this substitution in (15.2), we get

$$E_d = \frac{(0.12)E_H}{\varepsilon_s^2} \quad \text{---------- (15.3)}$$

With E_H = 13.6 eV and ε_0 = 15.8 for Ge, E_d is found out to be about 6.5 meV. For Si, $m_e^* = 0.3m$ and ε_s = 11.7 and, therefore, E_d comes to about 30 meV. The position of the donor's ground energy level E_D with respect to the conduction and valence band edges is shown in Fig.15.2(a). Its energy as measured from the conduction band edge is E_d. The energy continuum of the energy level scheme begins at the conduction band edge.

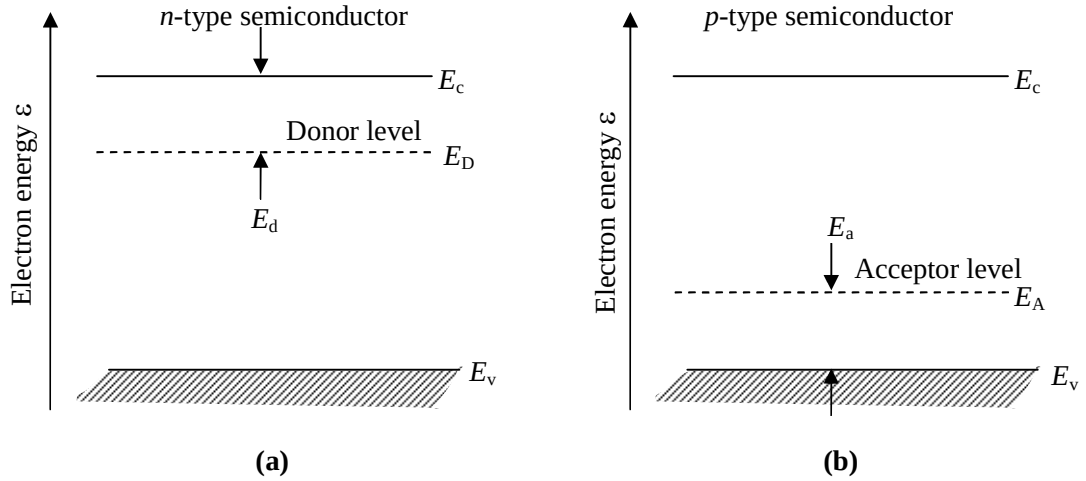


Fig 15.2 (a) A qualitative picture showing the position of the ground donor level E_D relative to the conduction and valence band edges E_c and E_v. The E_d denotes the ionization energy of the donor atom. (b) A qualitative picture showing the position of the ground acceptor level E_A relative to the conduction and valence band edges E_c and E_v. The E_a denotes the ionization energy of the acceptor atom

Therefore, the electron from a donor atom can be transferred to the conduction band by simply ionizing the donor atom at the expense of only a few meV (~ 6.5 meV in Ge and ~ 30 meV in Si) and not of a few eV of energy as is the case with the intrinsic material. This makes abundantly clear why the conductivity of an *n*-type semiconductor is several orders bigger than that of the corresponding intrinsic material.

Table 15.1 Values of static dielectric constant ε_s for some semiconductors

Crystal	ε_s
Si	11.7
Ge	15.8
AlAs	10.1
AlSb	10.3
GaAs	13.13
GaSb	15.69
InSb	17.88
InAs	14.55
InP	12.37
SiC	10.2

An odd feature of the model described above is that it gives the same value of E_d for all donor impurities such as P, As and Sb in a semiconductor host. The experimental values, however, show a slight variation (see Table 15.2). The description of screening in terms of the dielectric constant is rather crude and it restricts the operational domain of the model. The limitations are reflected in the model's failure to account for some of the subtle consequences of the atomic effects. The effect of screening on the Bohr radius is further revealing. The Bohr radius of the donor impurity atom is written as

$$r_d = \frac{4\pi\varepsilon_0\varepsilon_s\hbar^2}{m_e^* e^2} = \left(\frac{m\varepsilon_s}{m_e^*}\right) r_0$$

where r_0 is the hydrogen Bohr radius (0.53 Å).

Table 15.2 Ionization energies for a few donors and acceptors in silicon and germanium (E_d: ionization energy of donors and E_a: ionization energy of acceptors)

Impurity	E_d (meV)		*Impurity*	E_a (meV)	
	Si	Ge		Si	Ge
P	45	12	B	45	10.4
As	49	12.7	Al	57	10.2
Sb	39	9.6	Ga	65	10.8
			In	16	11.2

Using the respective values of ε and m_e^*, we get

$r_d \simeq 70$ Å in Ge and $\simeq 20$ Å in Si

So we see that screening increases the Bohr radius enormously. Increasing the Bohr radius is the same thing as loosening the binding which is reflected in drastically reduced values of the ionization energy E_d in proportion to that of hydrogen (13.6 eV). The calculated large values of the Bohr radius r_d are exhibition of the fact that the bound valence electron of the donor impurity is smeared over thousands of lattice sites. This is shown in fig 15.1, though not to scale. Thus the first Bohr orbits of impurity levels is formed. It is referred to as the impurity band. Conduction occurs even in the *impurity band* as soon as ionized donors are available to initiate the hopping of electrons from donor to donor. It is a well-established mechanism of conduction in extrinsic semiconductors and is known as *hopping conduction*.

15.3 The *p*-type Semiconductors:

The formation of p-type materials may be conveniently discussed by taking examples of valence four elemental semiconductors, i.e. Si and Ge. The elements used for doping to convert these materials into p-type are from Group III (e.g. B, Al, Ga, In) with sp^3 electronic configuration of the outermost shell. The substitutional impurity in this case completes only three of the four characteristic tetrahedral bonds of the host crystal. The doped impurity atom lacks one electron to complete the bonding with its all the neighbours. A valence electron of the host material may meet this requirement by ionizing the impurity atom negatively and creating a hole in the vicinity. Since the impurity atom is willing to accept an electron, the impurity atom is called an acceptor. The acceptor model is similar to the donor model. For example, a boron atom substituted in the germanium crystal is pictured as a germanium atom with a charge of (-e) fixed at its core and a hole of charge (e) bound to it. The impurity acts as the centre of a bound system whose energy level scheme is similar to that of the hydrogen atom. The ground acceptor level E_A lies close to the valence band edge as shown in fig 15.2(b). The ionization energy E_a as measured from the valence band edge is again very small as compared to the size of E_g. The values of E_a for some acceptors in Si and Ge are given in Table 15.2. Arguments concerning large values of conductivity and hopping conduction follow the same lines as those for n-type materials. With the latest methods of doping, the lowest impurity concentrations that can be obtained in semiconductors are of the order of 10^{12} cm^{-3}. Hence Si that has an intrinsic concentration of 1.5×10^{10} cm^{-3} at

300**K** does not show intrinsic conductivity though Ge does, on account of its higher intrinsic concentration (2.4×10^{13} at 300**K**).

15.4 Population of Donor and acceptor levels in the state of thermal equilibrium

Let us first define the symbols that will be used in our future descriptions:

N_D/N_A : density of all available donors/acceptors.

N_D^0 / N_A^0 : density of the neutral donors/acceptors.

N_D^+ / N_A^- : density of ionized donors/acceptors.

n_D/p_A : density of electrons bound to donors/density of holes bound to acceptors.

n : density of electrons in the conduction band.

p : density of holes in the valence band.

For small impurity concentrations (i.e. for non-degenerate semiconductors) the occupancy of the conduction and valence bands is as usual described by the Boltzmann distribution function. Therefore, at these concentrations the law of mass action (14.20) in lesson 14 which was derived on the basis of the above conditions must apply even to extrinsic semiconductors. But in the extrinsic case the value of the Fermi energy, not figuring in the law of mass action, depends on a more complicated charge neutrality condition. For homogeneous doping, the neutrality condition can be expressed as

$$n + N_A^- = p + N_D^+ \quad \text{----------} \quad (15.5)$$

with

$$N_D = N_D^0 + N_D^+ \quad \text{----------} \quad (15.6a)$$

$$N_A = N_A^0 + N_A^- \quad \text{----------} \quad (15.6b)$$

The above terminology may be appreciated better with the help of fig 15.3.

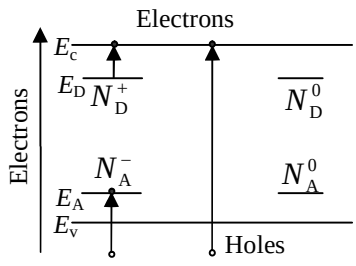


Fig 15.3 Electrons in the conduction band and holes in the valence band made available either by inter-band excitation or by impurity ionization.

Let us first consider the calculation of population of the donor levels. At low impurity concentrations (up to 10^{17} cm^{-3}), the interaction between electrons bound to separate donors may be neglected. We further assume for simplicity of calculations that the impurity introduces only a single one-electron level. Under these approximations the level could either be empty, or contain one electron of either spin, or contain two electrons with opposite spins. The double occupation is not favoured on account of its high energy arising from the Coulomb repulsion between two localized electrons. In the state of thermal equilibrium, the mean number of electrons in a system is expressed as

$$\langle n \rangle = \frac{\sum_j N_j \exp[-(\varepsilon_j - \varepsilon_F N_j)/k_B T]}{\sum_j \exp[-(\varepsilon_j - \varepsilon_F N_j)/k_B T]} \quad \text{---------- (15.7)}$$

where the sum is over all states of the system; the ε_j and N_j denote the energy and the number of electrons in state j. With a single impurity we have just three states: one with no electrons, thereby making no contribution to the energy, and two distinguishable spin states with a single electron of energy E_D. Hence from (15.7), we have

$$\langle n \rangle = \frac{2\exp[-(E_D - \varepsilon_F)/k_B T]}{1 + 2\exp[-(E_D - \varepsilon_F)/k_B T]} = \frac{1}{1 + \frac{1}{2}\exp[(E_D - \varepsilon_F)/k_B T]} \quad \text{---------- (15.8)}$$

The factor of ½ serves to modify the Fermi-Dirac distribution in this case. However, this factor is generally ignored which amounts to treating two single electron spin states of equal energy as one state. The population density of the donor level in the state of thermal equilibrium is given by

$$n_D = N_D \langle n \rangle$$

or

$$n_D = \frac{N_D}{1 + \exp[(E_D - \varepsilon_F)/k_B T]} \qquad \text{(ignoring the factor of ½)} \quad \text{---------- (15.9)}$$

Also note that,

$$n_D = N_D^0 \quad \text{---------- (15.10)}$$

Similarly, it can be shown that the population density of the acceptor level in thermal equilibrium is given by

$$p_D = \frac{N_A}{1 + \exp[(\varepsilon_F - E_A)/k_B T]} \quad \text{---------- (15.11)}$$

and also,

$$p_A = N_A^0 \quad \text{---------- (15.12)}$$

15.5 Extrinsic carrier densities:

The case of an extrinsic semiconductor in which both donors and acceptors are present is difficult to handle. Because of this reason we consider a pure n-type semiconductor and calculate its carrier concentration that is contributed by donors alone. Generally, every n-type material has a few acceptors and every p-type material has a few donors because of practical limitations on growing 100 per cent pure crystals. But the concentration of naturally present impurities in a pure crystal is expected to be negligibly small compared to the concentration of the impurity doped. Therefore, the relation (15.5) in the present case is rewritten as

$$n = N_D^+ + p_i \quad \text{---------- (15.13)}$$

The net electron density n has contributions of electrons from donors and from the valence band (see fig 15.3). The density of the latter type (n_i, being intrinsic) is equal to the density of holes (p_i). Normally $N_D^+ >> n_i$ (or p_i). For example, in Si this condition is satisfied even at low levels of the donor concentration since the intrinsic concentration is mere 1.5×10^{10} cm^{-3} around room temperature. Therefore for our n-type material, p_i in (15.13) may be dropped giving,

$$n \simeq N_D^+ = N_D - N_D^0 \quad \text{[using (15.5)]} \quad \text{---------- (15.14)}$$

or $$n = N_D - n_D \quad [\text{since } n_D = N_D^0]$$

$$= N_D\left(1 - \frac{1}{1+\exp[(E_D - \varepsilon_F)/k_B T]}\right) \quad \text{[using 15.9]}$$

$$= \frac{N_D}{1+\exp[-(E_D - \varepsilon_F)/k_B T]} \quad \text{---------- (15.15)}$$

As already mentioned in the beginning of Section 15.4, the relations for n, p and np (the law of the mass action) as derived in lesson 14 are applicable to extrinsic materials with low doping levels. We rewrite the relation (14.18) as

$$n = N(c)\exp\left[\frac{-(E_c - \varepsilon_F)}{k_B T}\right] \text{---------- (15.16a)}$$

with

$$N(\mathrm{c}) = 2\left(\frac{m_e^* k_B T}{2\pi\hbar^2}\right)^{3/2} \quad \text{---------- (15.16b)}$$

and (14.19)as

$$p = N(\mathrm{v})\exp\left[\frac{-(\varepsilon_F - E_v)}{k_B T}\right] \quad \text{---------- (15.17a)}$$

with

$$N(\mathrm{v}) = 2\left(\frac{m_h^* k_B T}{2\pi\hbar^2}\right)^{3/2} \quad \text{---------- (15.17b)}$$

From (15.16a), we have

$$\exp\left(\frac{\varepsilon_F}{k_B T}\right) = \frac{n\exp(E_c / k_B T)}{N(\mathrm{c})} \quad \text{---------- (15.18)}$$

Eliminating ε_F in (15.15) with the help of (15.18), we get

$$n = \frac{N_D}{1 + \dfrac{n\exp(E_d / k_B T)}{N(\mathrm{c})}} \qquad (\text{with } E_c - E_D = E_d)$$

or

$$n^2\left(\frac{\exp(E_d / k_B T)}{N(\mathrm{c})}\right) + n - N_D = 0 \quad \text{---------- (15.19)}$$

Using the substitution,

$$\frac{\exp(E_d / k_B T)}{N(\mathrm{c})} = X \quad \text{---------- (15.20)}$$

we get

$$Xn^2 + n - N_D = 0 \quad \text{---------- (15.21)}$$

The physically meaningful solution to the above equation is

$$n = \frac{-1 + \sqrt{1 + 4N_D X}}{2X} \quad \text{---------- (15.22)}$$

Rationalizing (15.22), we get

$$n = 2N_D\left[1 + \sqrt{1 + 4N_D X}\right]^{-1}$$

Using (15.20), we hve

$$n = 2N_D\left[1 + \sqrt{1 + 4N_D \frac{\exp(E_d / k_B T)}{N(\mathrm{c})}}\right]^{-1} \quad \text{---------- (15.23)}$$

There are three limiting cases of the above expression as discussed below:

Case I. At low temperatures,

$$4\left(\frac{N_D}{N(c)}\right)\exp\left(\frac{E_d}{k_B T}\right) >> 1 \quad \text{---------- (15.24)}$$

This transforms (15.23) to the following form,

$$n = \sqrt{N_D N(c)}\exp\left(\frac{-E_d}{2k_B T}\right) \quad \text{---------- (15.25)}$$

At these low temperatures a large number of donors stay in the unionized state. The range of temperatures over which this condition exists is known as the freeze-out range. We can notice the similarities between (15.25) and (14.21) that expresses the intrinsic carrier density. Both depend exponentially on temperature since the exponential dependence dominates the other dependence entering through $N(c)$. In the donors' case a much smaller quantity E_d appears in the exponential as against E_g in the intrinsic case. This accounts for a larger carrier concentration in the n-type material.

Case II. For temperatures at which

$$4\left(\frac{N_D}{N(c)}\right)\exp\left(\frac{E_d}{k_B T}\right) << 1 \quad \text{---------- (15.26)}$$

the relation (15.22) reduces to

$$n \simeq N_D \quad \text{---------- (15.27)}$$

i.e. a constant.

This means that at these temperatures all donors are ionized and the electron density reaches its maximum with the excitation of electrons from the valence band considered negligible in the first approximation. This is referred to as the *saturation range* or *exhaustion range* of carriers.

Case III. When temperatures is still higher, the concentration of conduction electrons contributed by the valence band becomes appreciable. Since the concentration of donor electron in the conduction band no more increases on account of saturation, the intrinsic electron density overtakes the density of donor electrons at some stage. In this condition an n-type semiconductor is characterized by intrinsic behaviour and we speak of the intrinsic range of carriers. The variations of electron density and energy as functions of temperature are sketched in fig 15.4, identifying the three ranges.

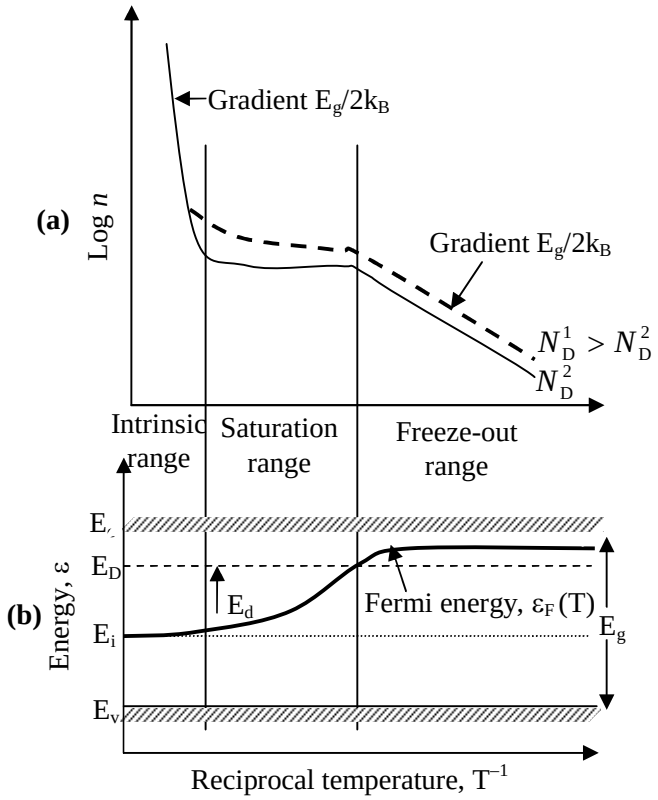


Fig 15.4 (a) A qualitative illustration of the variation of electron concentration (n) in the conduction band with change in temperature in two samples of different donor concentrations N_D^1 and N_D^2. (b) A qualitative illustration of the temperature dependence of the Fermi energy ε_F in the same semiconductor. The E_i represents the Fermi energy in the instrinsic material. All other symbols have their usual meaning.

The calculation of electron density in the conduction band as carried out above is based on the assumption that there are no acceptor impurities in the n-type semiconductor under consideration. Experiments, however, confirm that traces of acceptors are always present. Therefore, quantitative estimates made on the basis of the above theory may differ slightly from the experimental values.

There are several applications in which there is a need to monitor the electron and hole concentrations. Since at low doping levels the law of mass action (14.20) is valid, it is exploited to manipulate the carrier concentration in non-degenerate semiconductors. When the electron concentration *n* is increased by adding the trace of a suitable donor impurity, the hole concentration *p* goes down so that the product (*np*) remains constant at a temperature in accordance with the law of mass action. This way it is possible to reduce the sum (*n* + *p*) considerably. This method of reduction is in wide practice and

known as compensation. The sign of majority carriers is quickly determined by detecting the sign of Hall voltage.

15.6 Fermi level in extrinsic semi conductor

We have seen that Fermi level is situated at the middle of the band gap in an intrinsic semiconductor as the electron and hole densities are equal. When the intrinsic semiconductor is doped the carrier densities change, consequently the position of the Fermi level also changes. The shift in the position of the Fermi level can easily be related to the majority carrier density in an extrinsic semiconductor if it is assumed that the addition of the impurities do not affect the densities of energy states in the conduction and valence bands.

Let N_e and N_v denotes the density of states in the conduction band and density of states in the valence band, respectively. We have for an intrinsic semiconductor.

$$\left.\begin{aligned} n &= N_c \exp\left(\frac{E_{fi} - E_e}{k_B T}\right) \\ \text{and} \qquad p &= N_v \exp\left(\frac{E_v - E_{fi}}{k_B T}\right) \end{aligned}\right\}$$

Here E_{fi} is the energy associated with the Fermi level in an intrinsic semiconductor. For an intrinsic semiconductor, we have n = p and therefore from the above expressions,

$$\frac{N_c}{N_v} = \exp\left(\frac{E_c + E_v - 2E_{fi}}{k_B T}\right)$$

Let E_{fn} be the energy associated with the Fermi level in an n-type semiconductor having an electron density n, we have

$$n = N_c \exp\left(\frac{E_{fn} - E_c}{k_B T}\right)$$

and

$$p = N_v \exp\left(\frac{E_v - E_{fn}}{k_B T}\right)$$

Therefore,

$$\frac{n}{p} = \frac{n^2}{n_i^2} = \frac{N_c}{N_v} = \exp\left[\frac{2E_{fn} - E_c - E_v}{k_B T}\right] = \exp\left[\frac{2(E_{fn} - E_{fi})}{k_B T}\right]$$

or

$$n = n_i \exp\left(\frac{E_{fn} - E_{fi}}{k_B T}\right)$$

Similarly for a p-type semiconductor

$$p = n_i \exp\left(\frac{E_{fi} - E_{fp}}{k_B T}\right)$$

Thus the shift in the Fermi level in the n and p types of semiconductor can be expressed as

$$\left.\begin{aligned} E_{fn} - E_{fi} &= k_B T \ln \frac{n}{n_j} \\ \text{and} \qquad E_{fi} - E_{fp} &= k_B T \ln \frac{p}{n_i} \end{aligned}\right\}$$

Fig 15.4.1 represents the shift in the Fermi level in the n and p type semiconductors.

Fig 15.4.1 Shift in the Fermi level in the n- and p-types of semiconductors.

15.7 Temperature dependence of electrical conductivity:

We refer back to relation (15.5) according to which the conductivity depends on the carrier density and the mobility as

$$\sigma = |e| (n\mu_n + p\mu_p) \quad \text{----------} \quad (15.28)$$

The aim of this section can be fulfilled by establishing how the concentration and the mobility of electrons and holes vary with temperature. The temperature dependence of carrier concentration was treated in Section (15.5). Here, we set out to determine the temperature dependence of mobility. Then, the temperature dependence of conductivity will be interpreted by comparing the two contributions. The conduction in semiconductors differs from the metallic conduction in the sense that electrons near the Fermi level contribute to the metallic conductivity whereas in semiconductors electrons near the bottom of the conduction band and holes near the top of the valence band take part in the process. There are no charge carriers near the Fermi level in semiconductors. In the approximation where the Fermi statistics can be replaced by Boltzmann statistics for non-degenerate semiconductors, the Boltzmann equation is solved to get to the expression for mobility of charge carriers. On demand of simplicity we present here only a qualitative description of the scattering processes in which electrons and holes by and

large behave similarly. Several simplifications of the exact expression for mobility lead to the result.

$$\mu \propto \tau \qquad \text{---------- (15.29)}$$

where τ is the relaxation time defined by the relation $\tau = \frac{\Lambda}{v_{th}}$ where $\Lambda$is the electron mean free path and v_{th} is thermal velocity of electron.

Since τ is proportional to the average time between successive collisions,

$$\frac{1}{\tau} \propto \langle v \rangle \Sigma \quad \text{---------- (15.30)}$$

where Σ denotes the scattering cross-section of electrons and holes at a scattering centre. Relation (15.30) gives a measure of the scattering probability with $\langle v \rangle$ treated as thermal average over all electron or hole velocities in the lower conduction band or upper valence band. The use of Boltzmann statistics in semiconductors gives

$$\langle v \rangle \propto \sqrt{T} \quad \text{---------- (15.31)}$$

Phonons happen to be the prominent source of carrier scattering in crystals. It is simpler to calculate the phonon scattering cross-section Σ_L for acoustic phonons, ($T >> \theta_D$), where θ_D is the Debye temperature. This gives the dependence,

$\Sigma_L \propto T$ (L stand for lattice or phonons) ---------- (15.32)

Making use of (15.30), (15.31) and (15.32), we have from (15.29)

$$\mu_L \propto T^{-3/2} \quad \text{---------- (15.33)}$$

In semiconductors, centers of ionized donors and acceptors serve as another important source of scattering. As an electron or a hole approaches such a centre, it experiences a Coulomb force and suffers scattering similar to Rutherford scattering. A rigorous treatment shows that

$\Sigma_1 \propto \langle v \rangle^{-4}$ ---------- (15.34)

where I stands for ionized impurity. Therefore, with the use of (15.30), (15.31) and (15.34) in (15.29), we get

$\mu_I \propto T^{-3/2}$ ---------- (15.35)

For one type of carriers, say electrons, we can write contributions of phonons and ionized impurities to resistivity as

$$\rho_L = \frac{1}{\sigma_L} = \frac{1}{ne\mu_L} \text{ and } \rho_I = \frac{1}{\sigma_I} = \frac{1}{ne\mu_I} \quad \text{---------- (15.36)}$$

Applying the Matthiessen's rule, the total resistivity is written as

$$\rho = \frac{1}{ne\mu} = \frac{1}{ne}\left[\frac{1}{\mu_L} + \frac{1}{\mu_I}\right] \quad \text{---------- (15.37)}$$

or

$$\frac{1}{\mu} = \frac{1}{\mu_L} + \frac{1}{\mu_I} \quad \text{---------- (15.38)}$$

A qualitative display of the temperature dependence of mobility in an extrinsic semiconductor is made in fig 15.5.

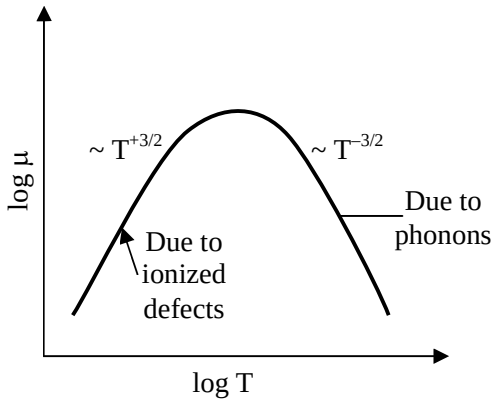


Fig 15.5 Qualitative temperature dependence of the mobility contributions from charged impurities and phonons.

The variation of conductivity σ as a function of temperature is plotted for six samples of *n*-type germanium in fig 15.6. The conductivity shows a maximum in the exhaustion range where the electron concentration n becomes almost constant. The behaviour of σ in this range is governed by the characteristic temperature dependence of mobility as shown in fig 15.5. In the intrinsic range (at high temperatures) and in the freeze-out (at low temperatures), the variation in σ is effectively controlled by the exponential dependence of concentration n on temperature. The temperature dependence of mobility has been derived from the curves of fig 15.6 and the Hall measurements. The donor concentration N_D in six samples ranges from 10^{13} to 10^{18} cm^{-3}. The mobility in the purest crystal with $N_D \simeq 10^{13}$ cm^{-3} shows ideal $T^{-3/2}$ dependence whereas in a sample with increasing values of N_D the mobility approaches the expected $T^{3/2}$ behaviour. Typical orders of electron and hole mobilities in some important semiconductors are given in Table 15.3.

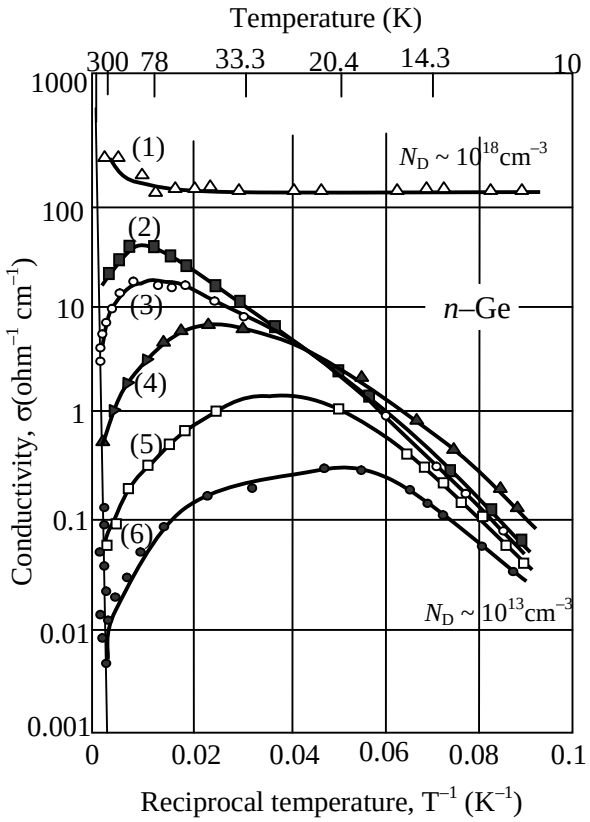


Fig 15.6 Observed conductivity behaviour as a function of temperature for six samples of n-type germanium.

Table 15.3 Typical order of carrier mobilities at room temperature in some important semiconductors.

Crystal	*Mobility* (cm^2/V. s)		*Crystal*	*Mobility* (cm^2/V. s)	
	Electrons	*Holes*		*Electrons*	*Holes*
Si	1350	480	InSb	800	450
Ge	3600	1800	InAs	30,000	450
GaAs	8000	300	InP	4500	100
GaSb	5000	1000	PbSe	1020	930
AlAs	280	—	PbTe	2500	1000
AlSb	900	400	SiC	100	10 – 20

15.8 Hall Effect

Suppose a semiconductor crystal carries a current of density $\mathbf{j}_x$ along the x-direction under the action of a steady electric field $\mathbf{E}_x$ directed along the x-direction. When a

steady magnetic field $\mathbf{B}_z$ is applied along the z-direction, a Hall voltage develops between the crystal's faces along the y-direction due to the deflection of charge carriers by the Lorentz force. The geometry of the experimental set-up is drawn in fig 15.7. If the corresponding Hall electric field be denoted by $\mathbf{E}_y$, the Hall coefficient is expressed as

$$R_H = \frac{E_y}{j_x B_z} \quad \text{---------- (15.39)}$$

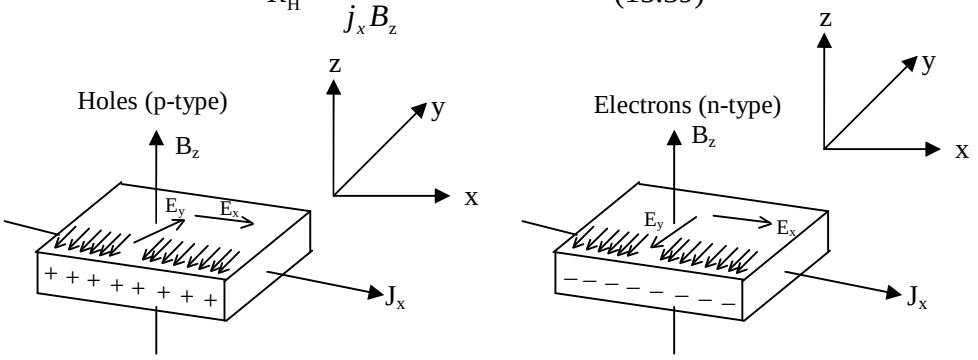


Fig 15.7 Directions of the applied electric and magnetic fields (E and B respectively) relative to the direction of current flow in two samples of a semiconductor. The current density j is along the x-direction. Note that the polarity of the Hall voltage is opposite in p- and n-type samples.

For metals, where the current is generally carried by electrons, R_H in small magnetic fields is related to the election density n by

$$R_H = \frac{r}{ne} \quad \text{---------- (15.40)}$$

where r is called the Hall factor. Its value is usually close to unity. It is exactly equal to 1 if all electrons move at the same speed as is true in respect of the Drude model, or in respect of the degenerate electron gas in which all electrons move with Fermi velocity. The value of r was found as $3\pi/8$ (= 1.18) in the Lorentz model.

In semiconductors, things are somewhat different as every crystal in principle can have both types of charge carriers. Nevertheless, the Hall factor remains close to unit for both electrons and holes. The n-type and p-type crystals are quickly differentiated since the sign of the Hall voltage is opposite in the two cases as may easily be made out from fig 15.7. When both electrons and holes are present simultaneously, a partial recombination of electrons and holes occurs on the crystal face onto which both are reflected. As a result of this, the Hall voltage is reduced and its sign refers to the sign of majority carriers. From the measured value of R_H on the basis of (15.39), we can

estimate the concentration of majority carriers using (15.40). However, the most general expression for R_H in semiconductors is found as

$$R_H = \frac{p\mu_p^2 r_p - n\mu_n^2 r_n}{e(p\mu_p + n\mu_n)} \quad \text{----------} \quad (15.41)$$

where all symbols have their usual meaning.

Table 15.4 Hall coefficient and mobilities for some metals at 300 K

Metal	$R_n(Vm^3 A^{-1}wb^{-1})$ in10^{-10}	μ ($m^2 V^{-1}s^{-1}$)
Silver	–0.84	0.0056
Copper	–0.55	0.0032
Gold	–0.71	0.0030
Sodium	–2.50	0.0052
Aluminium	–0.31	0.0012
Lithium	–1.70	0.0018
Zinc	+0.30	0.0060
Cadmium	+0.60	0.0080

15.9 Summary

The standard method of increasing the conductivity of an intrinsic material is to add to it a suitable impurity or electrically active element in small concentration. The method is known as *doping*. Impurities that enhance the carrier density by contributing additional electrons to the conduction band are called *donors* and those which create additional holes in the valence band are known as *acceptors*.

When pure crystals of Si and Ge are doped with any element from Group III (e.g. B, Al, In), *p*-type semiconductors are formed.

The energy of the impurity atom $E_n = -\frac{m^* e^4}{2k_B n^2 \hbar^2}$

The conduction electron concentration (n) in extrinsic semiconductor at high temperature is equal to the donor concentration (N_d).

$n = n_0 \exp[(E_f - E_g)/k_B T] \cong N_d$

and at low temperatures

$$n = (N_d)^{1/2}(n_0)^{1/2} \exp\left[\frac{E_2 - E_g}{2k_B T}\right].$$

Conduction occurs even in the *impurity band* as soon as ionized donors are available to initiate the hopping of electrons from donor to donor. It is a well-established mechanism of conduction in extrinsic semiconductors and is known as *hopping conduction*.

The elements used for doping to convert these materials into p-type are from Group III (e.g. B, Al, Ga, In) with sp^3 electronic configuration of the outermost shell.

The case of an extrinsic semiconductor in which both donors and acceptors are present is difficult to handle. Because of this reason we consider a pure n-type semiconductor and calculate its carrier concentration that is contributed by donors alone. Generally, every n-type material has a few acceptors and every p-type material has a few donors because of practical limitations on growing 100 per cent pure crystals. But the concentration of naturally present impurities in a pure crystal is expected to be negligibly small compared to the concentration of the impurity doped.

the conductivity depends on the carrier density and the mobility as

$$\sigma = |e| (n\mu_n + p\mu_p)$$

Donor impurities lead to an excess of electrons over holes and acceptor impurities lead to an excess of holes over electrons. In an instrinsic semiconductor the number of electrons is equal to the number of holes.

15.10 Key words

Doping – Donors – Acceptors – Pentavalent – Impurity – n-type semiconductor – p-type semiconductor – Hopping conduction – Inter band excitation – Localized electrons – Extrinsic carrier densities – Electron and hole mobilities.

15.11 Review questions

1. Indicate with the help of diagram how Fermi level changes as a function of temperature in an n-type semiconductor.
2. Derive an expression relating to shift in the Fermi level and carrier density in an extrinsic semiconductor.
3. In an n-type semiconductor the Fermi level lies 0.3 eV, below the conduction band at 300 K. With the temperature is increased to 330 K, find the new position of Fermi level.
4. Assuming a valence band above which there are n_a acceptor levels per unit volume, derive an expression for the Fermi level and for the density of free holes in the valence band as function of *T*.
5. If the Fermi level in a semiconductor lies more than a few kT below the bottom of the conduction band and more than a few kT above the top of the valence band,

show that the product of the number of free electrons and the number of free holes per cm^3 is given by $n_e n_{\hbar} = 2.33\times10^{31} T^3 e^{-E_g/kT}$ where the E_g is the gap width. Note that this holds irrespective of the presence of donors or acceptors in the gap, as long as the condition imposed on the Fermi level is satisfied.

6. Prove that the minimum conductivity of an extrinsic semiconductor is given by

$$\sigma = 2n_i(\mu_n\mu_p)^{1/2}$$

7. Show that the conductivity minimum occurs when

$$N_A - N_D = n_i\left[\left(\frac{\mu_n}{\mu_p}\right)^{1/2} - \left(\frac{\mu_p}{\mu_n}\right)^{1/2}\right]$$

8. (a) For an n-type semiconductor, show that

$$E_F = E_c - k_B T \ln\left(\frac{n_0}{N_d^+}\right)$$

(b) For a p-type semiconductor, show that

$$E_F = E_v + k_B T \ln\left(\frac{p_0}{N_a^-}\right)$$

9. (a) Describe any one experimental technique for determining the band gap of a p-type semiconductor.

(b) State the conditions under which the Fermi level of an extrinsic semiconductor can move inside either valence or conduction band.

(c) What are degenerate semiconductors? How does their conductivity vary with temperature?

10. Discuss the Hall effect. Explain how the measurement of Hall coefficient helps one to determine the mobility and sign of charge carriers.

15.12 Text and Reference Books:

1. Elements of Solid State Physics by J.P. Srivatsava (PHI)
2. Solid State Physics by M.A. Wahab (Narosa)
3. Elements of Solid State Physics by A. Omar (Pearson education)
4. Solid State Physics by S.O. Pillai (New Age)
5. Solid State Physics by C. Kittel (Asia Publishing House)
6. Solid State Physics by S.L. Kakani and C.Hemrajani (S.Chand)
7. Solid State Physics by Saxena Gupta Saxena (Pragati Prakashan).
8. Solid State Physics by C.J. Dekker (Macmillan)

UNIT – IV **LESSON: 16**

JUNCTION DEVICES

Aim: To study the properties of p-n junction and understand the working of other junction devices

Objectives:

- To understand the working of p-n junction and junction based semiconductor devices.
- To understand the principle of quantum hall effect

Structure:

Introduction:

The comparative ease with which the electrical characteristic of semiconductors can be monitored by controlling the traces of impurities being doped is simply striking. Because of this property, semiconductors have emerged as the indispensable material base for the development of Solid State Electronics. The dramatic and extensive technological consequences of this property have given a tremendous boost to the commercial interest in semiconductors.

16.1 The *p-n* Junction:

It is of significance to note that semiconductor devices generally exploit the characteristics of inhomogeneous semiconductors in which the donor and acceptor concentrations are not uniform. A common form in which they are used contains two separate *n* and *p* type regions with an abrupt partition boundary (a junction) within a single crystalline sample. This junction is known as the *p-n* junction. There can be more than one junction in some devices. Characteristics of the *p-n* junction are crucial to the fabrication and working of most of the semiconductor devices. It is an essential component of the present electronics circuitry ranging from a simple rectifier circuit to transistor circuits and integrated circuit used in sophisticated appliances like modern computer. On the other hand, optical applications such as photo cell, LEDs and injection lasers (laser diodes) are equally important today in view of the rapidly upcoming field of opto-electronics. In this lesson we learn about the salient features of the working of a selected few devices. Consider a *p-n* junction, for example of a Si crystal whose left half is *p*-type and the right side is *n*-type. A *p-n* junction is never made by simply joining the two types of a material. In actual practice a p-n junction is formed during the crystal-growing process, diffused or alloyed. Other methods of actual junction formation are also used.

Nonetheless, the supposition that a piece of *n*-type and a piece of *p*-type material are brought into intimate contact helps in determining the electrostatic conditions present at a junction. When the two regions are treated as isolated, the levels of Fermi energy in the two regions are at different positions on the common energy scale [fig 16.1(a)]. Since the two regions are parts of the same crystal, the Fermi energy (representing the electrochemical potential in semiconductors) must have the same value in both halves in

the state of thermal equilibrium. Near $T \simeq 0$ K the position of the Fermi level is near the acceptor level in the *p*-region and near the donor level in the *n*-region.

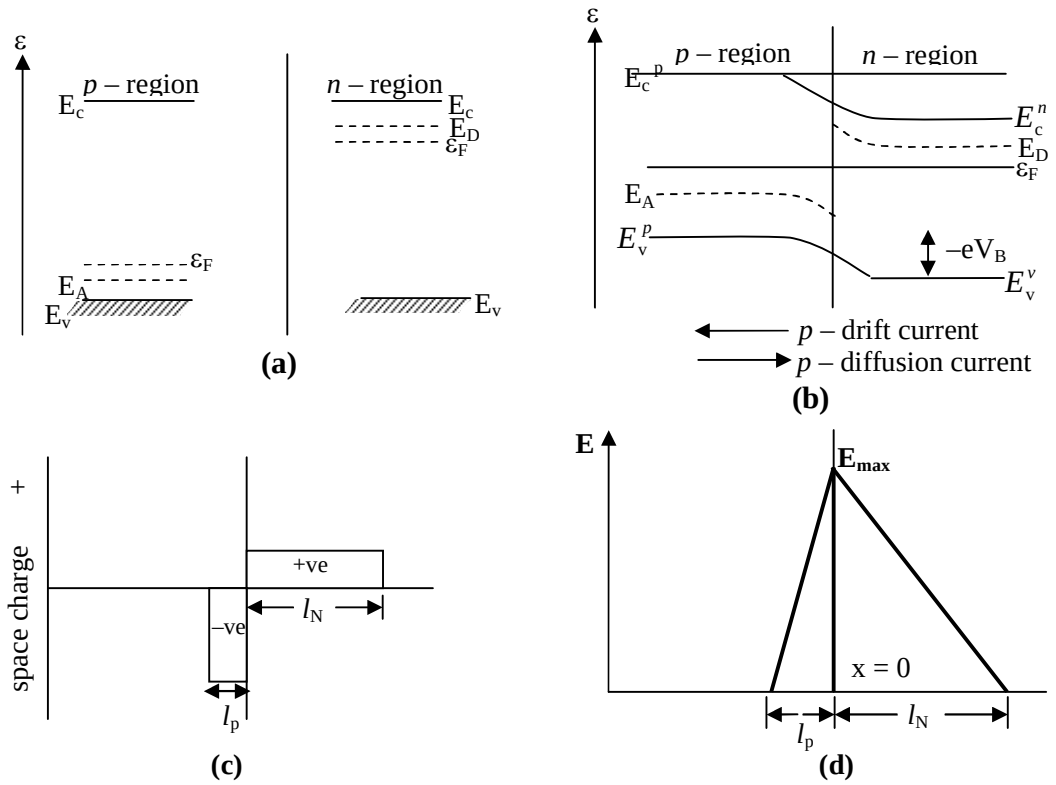


Fig 16.1 (a) Relative positions of the conduction and valence band edges, the donor and acceptor levels; and the Fermi level in a p-n junction in the state of thermal equilibrium when p-and n-type regions are treated as isolated. (b) All the above positions when the p- and n-type regions are brought in contact. (c) The space charge distribution at p-n junction. When the effect of the mobile carriers is neglected, i.e. $l_N N_D = l_p N_A$. (d) Variation of the space charge electric field E across the junction plane (x = 0).

Because the impurity levels are at the extremes of the gap, the Fermi level would not have maintained a constant level unless the bending of conduction and valence band edges occurred in the transition zone. [fig 16.1(b)]. Immediately after the junction comes into existence, electrons from the n-region begin to diffuse into the *p*-region and combine with the holes present in the vicinity of acceptors. This leaves behind the positively charged donors in the *n*-region and produces negatively charged acceptors in the *p*-region. The immobile ionized impurities are considered to form a charged double layer of space charges across the junction in the doping profile [fig 16.1(c)]. The charged donors and acceptors create an electric field across the transition region whose direction is such as to oppose the diffusion of free charge carriers. Therefore, in the absence of any

external field the diffusion current stops as soon as the field of ionized impurities becomes big enough just to keep the diffusion in complete check. The electric field in question is basically interpreted as the consequence of a macroscopic potential $V(x)$ varying over the transition region. In a one-dimensional model with an abrupt change from *p*-type to *n*-type at $x = 0$, the potential $V(x)$ is related to the space charge by the Poisson equation.

$$\left.\begin{aligned} \frac{d^2V}{dx^2} &= \frac{-e}{\varepsilon_0\varepsilon_s}(N_D^+ + p_n - n_n) \quad \text{for } 0 < x < l_N \,(n-\text{type region}) \\ \frac{d^2V}{dx^2} &= \frac{e}{\varepsilon_0\varepsilon_s}(N_A^- + n_p - p_p) \quad \text{for } -l_N < x < 0 \,(p-\text{type region}) \end{aligned}\right\} \text{---------- (16.1)}$$

where the subscripts n and p refer to concentrations in the n- and p-type regions, respectively, and $-l_p < x < l_N$ defines the space charge region.

The exact solution of (16.1) is almost impossible, since the carrier concentrations depend on position. To make the solution feasible, we simplify (16.1) by assuming that the electric field is strong enough to keep all mobile free carriers away from the space charge region. The variation of electric field **E** across the junction is shown in Fig 16.1(d). It is maximum at x = 0. Thus the density of space charge is given by simply N_D^+ in the n-region and by N_A^- in the p-region. This approximation is good except at the boundary of the space charge region for a small current across the junction. In this assumption the space charge region is depleted of free carriers and, therefore, is also identified as the *depletion layer*. The relation (16.1) is solved under the boundary conditions:

$$\frac{dV}{dx} = 0 \text{ at } x = l_N \text{ and } x = -l_p$$

with $\frac{dV}{dx}$ continuous at x = 0 and $V(l_N) - V(-l_p) = V_B$, where V_B is the height of the potential step, known as the barrier height. It is also called the *diffusion voltage*.

Well outside the space charge zone, N_D^+ and N_A^- are compensated by equally large free carrier concentrations n_n and p_p, respectively. In accordance with the type of doping, electrons serve as majority carriers in the n-type region and holes as majority carriers in the p-type region. There may always be present small concentration of holes in the n-

region (p_n) and of electrons in p-region (n_p), contributed by the naturally occurring impurities in the host crystal.

An estimate of V_B is made in terms of carrier concentrations in the state of thermal equilibrium. These concentrations are given by (15.16a) and (15.17a) which we rewrite below:

$$n_n = N(\mathrm{c})\exp\left[\frac{-(E_\mathrm{c}^n - \varepsilon_\mathrm{F})}{k_\mathrm{B}T}\right] \quad \text{---------- (16.2a)}$$

$$p_p = N(\mathrm{v})\exp\left[\frac{-(\varepsilon_\mathrm{F} - E_\mathrm{v}^p)}{k_\mathrm{B}T}\right] \quad \text{---------- (16.2a)}$$

Their relationship with the intrinsic concentration n_i is then given by

$$n_i^2 = n_n p_n N(\mathrm{v})N(\mathrm{c})\exp\left[\frac{-(E_\mathrm{c}^n - E_\mathrm{v}^n)}{k_\mathrm{B}T}\right] \quad \text{---------- (16.3)}$$

Here $E_\mathrm{c}^n, E_\mathrm{v}^n, E_\mathrm{v}^p$ are conduction band edges (in the n-region) and valence band edges (in the n- and p-regions), respectively [see fig 16.2(b)].

From (16.2b), we have

$$E_\mathrm{v}^p - \varepsilon_\mathrm{F} = k_\mathrm{B}T\ln\left(\frac{p_p}{N(\mathrm{v})}\right) \quad \text{---------- (16.4)}$$

likewise,

$$E_\mathrm{v}^n - \varepsilon_\mathrm{F} = k_\mathrm{B}T\ln\left(\frac{p_n}{N(\mathrm{v})}\right) \quad \text{---------- (16.5)}$$

Subtracting (16.5) from (16.4) and using (16.3), we get

$$E_\mathrm{v}^p - E_\mathrm{v}^n = k_\mathrm{B}T\ln\left(\frac{p_p n_n}{n_i^2}\right) \quad \text{---------- (16.6)}$$

But V_B is equal to the difference between the maximum and minimum of the macroscopic potential V(x). Hence,

$$eV_\mathrm{B} = -(E_\mathrm{v}^n - E_\mathrm{v}^p) = k_\mathrm{B}T\ln\left(\frac{p_p n_n}{n_i^2}\right)$$

or

$$V_\mathrm{B} = \frac{k_\mathrm{B}T}{e}\ln\left(\frac{p_p n_n}{n_i^2}\right) \quad \text{---------- (16.7)}$$

Other important properties of a p-n junction are the maximum value of the electric field $\mathbf{E}_x$ and the space charge width $l(=l_N + l_p)$. The expressions for them can be obtained from (16.1).

When an external steady voltage is applied across a p-n junction, the barrier step height changes. In a forward-biased configuration it decreases to value,

$$V_{BF} = V_B - V_{ext} \quad \text{----------} \quad (16.8)$$

Where V_{ext} denotes the applied voltage.

The 'flow of current changes' in the band edges and the Fermi level are depicted in fig 16.2(a).

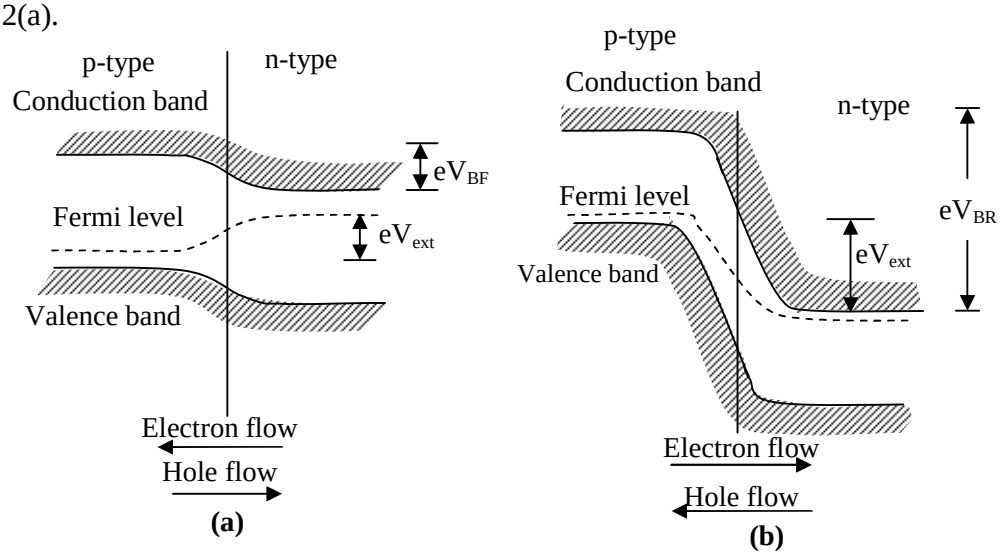


Fig 16.2 Qualitative behaviour of the Fermi level and the electron potential energy at points near the p-n junction when (a) a forward bias of V_{ext} volt is applied, (b) a reverse bias of V_{ext} volt is applied.

On the other hand, for a reverse-biased connection in contrast to the forward-biased case, the direction of the applied field does not favour the flow of majority carriers across the junction, thus increasing the barrier height to

$$V_{BR} = V_B + V_{ext} \quad \text{----------} \quad (16.9)$$

the electronic conduction and changes in the band scheme structure are illustrated in fig 16.2(b).

In the forward-biased configuration the current is carried by majority carriers and is of the order of a few mA which is much higher than that in the reverse-biased case (~μA). The current in a reverse-biased junction is low as it is contributed by the flow of minority carriers that are relatively very small in number and are required to cross a higher potential barrier. The property that a p-n junction favours the flow of current in one direction across the junction forms the basis for using it as a rectifier. The knowledge of voltage-current characteristics of a p-n junction is crucial in order to understand the behaviour of the p-n junction for the desired application.

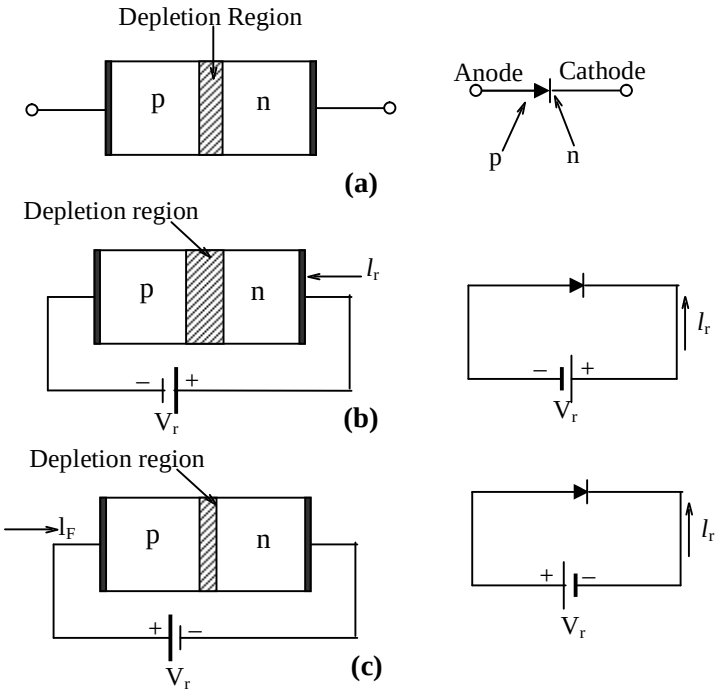
It may again be reminded that there are numerous devices that make use of homogeneous semiconductors. Perhaps, the simplest of them is a thermistor which is used as a standard temperature measuring device.

16.2 Some examples of p-n junction:

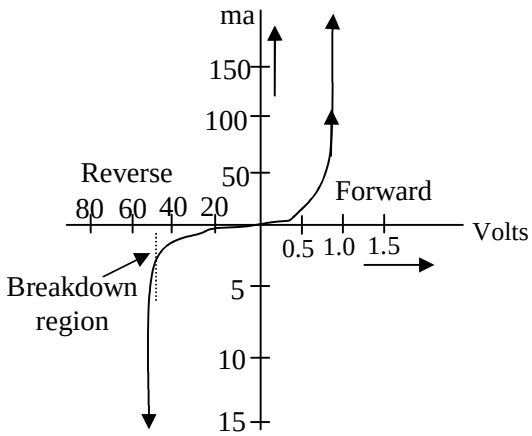
The p-n junction based devices have acquired such a tremendous technological potential that even their listing is quite a formidable task. We now study the salient features of some important devices..

16.2.1 Junction rectifier

A p-n junction acts as a rectifier. A large current will flow if we apply a voltage across the junction in one direction, but if the voltage is in the opposite direction only a small current will flow. If an alternating voltage is applied across the junction, the current will flow in one direction, i.e the junction has rectified the current. The biasing arrangements and the variation of depletion region with biasing voltage is shown in fig 16.3. The current – voltage (I-V) characteristic is shown in fig 16.4. We may see from the curve that it depicts a nearly saturation current in the reverse bias condition and a current which rises with the rise of potential in the forward bias condition.



16.3 The p-n junction (a) No external bias (b) Reverse bias (c) Forward bias, Schematic representation, circuit diagram and potential diagram.



16.4 Current voltage characteristics of a p-n junction rectifier.

Let us first consider that no external potential is applied across the junction and it is in thermal equilibrium. In the thermal equilibrium there will be a small flow of electrons J_{nr} from the n region into the p region where they get anhilated by combining with holes. This recombination current is balanced by a current J_{ng} of electrons generated thermally in the p region and which diffuse to the n- region

$$J_{np} = -J_{ng}$$

Otherwise electrons would pile up on one side of the barrier. A similar argument holds for hole fluxes J_{pr} and J_{pg}.

Let us now calculate the current across the p-n junction in the reverse and forward bias cases. Assuming the Maxwell-Boltzmann statistics are valid, one can easily write the expression for electrons having sufficient energy to surmount the junction potential barrier as proportional to

$$\mathbf{e}^{-e(\phi_0 - V_0)/k_B T} = (\text{const}) \exp\left[\frac{eV_0}{k_B T}\right]$$

Accordingly, one expects J_{nr} to be proportional to the Boltzmann factor exp $[eV_0/k_BT]$. The recombination current is reduced by the Boltzmann factor, i.e.

$$J_{nr} = J_{ng} \exp [eV_0/k_BT]$$

Similarly, for holes

$$J_{pr} = J_{pg} \exp [eV_0/k_BT]$$

The Boltzmann factor controls the number of electrons with enough energy to get over the barrier. We must remember that the holes prefer to be at the top of the potential barrier whereas the electrons prefer to be at the bottom of the barrier. Now

$$J_n = J_{ng} - J_{nr} = -J_{ng}\left[\exp\left(\frac{eV_0}{k_BT}\right)-1\right]$$

$$J_p = J_{pr} - J_{pg} = J_{pg}\left[\exp\left(\frac{eV_0}{k_BT}\right)-1\right]$$

The total current density due to both the electrons and the holes, across the p-n junction is thus given by

$$I = e(J_p - J_n) = (J_{pg} + J_{ng})\left[\exp\left(\frac{eV_0}{k_BT}\right)-1\right] = I_0\left[\exp\left(\frac{eV_0}{k_BT}\right)-1\right] \quad \text{---------- (16.10)}$$

Here $I_0 = e(J_{pg} + J_{ng})$ is called the maximum or the saturation current. Eqn. (16.10) is well satisfied for p-n junctions in germanium and is called the rectifier equation. From (16.10) one can easily write the current relations for forward and reverse bias cases, i.e.

$$I = I_0\left[\exp\left(\frac{eV_0}{k_BT}\right)-1\right] \qquad \text{(Forward bias)}$$

$$I = I_0\left[\exp\left(-\frac{eV_0}{k_BT}\right)-1\right] \qquad \text{(Reserve bias)} \quad \text{---------- (16.11)}$$

In the reverse bias case, however, $\exp\left(-\frac{eV_0}{k_BT}\right) \ll 1$ for $eV_0 \geq 4k_BT$, and clearly $I \rightarrow I_0$, the saturation value. It is of interest to note that the saturation current is of the order of a few milli-amperes per sq. cm of the junction area and remains almost steady at about this value until the voltage reaches a value sufficient to cause the breakdown of the function as shown in fig. 16.4. The breakdown may not occur until the reverse bias voltage reaches some hundreds of volts. In certain functions breakdown may occur even earlier. The maximum permissible forward current in a p-n junction is many amperes per square cm and occurs for a voltage of only a few volts.

16.2.2 Junction transistor:

A junction transistor is a three-terminal device, analogous to a vacuum triode. Like a vacuum triode, the transistor utilizes one of the elements to control or modulate the flow

of charges through the whole device. The three elements in the transistor are *emitter*, base and *collector*. The names are suggestive of the functions each performs. The emitter injects charge carriers into the base, which in turn controls the number of these carriers that are eventually gathered up by the *collector*.

This action is based on the fact that the transistor is actually a sandwich of two p-n junctions: the emitter base junction and the collector base junction. The emitter and collector are always doped by the same impurity. They may both be p-type or both be n-type, while the base always has opposite doping as shown in fig. 16.5. In npn transistors, the emitter and collector are n-type and the base is p-type; while in pnp transistors the emitter and collector are p-type and the base is n-type. This is a direct consequence of the fact that a semi-conductor material is capable of sustaining a current due to the motion of both positive as well as negative charge carriers. It should be noted that in schematic symbols for the npn and pnp transistors the only difference is the direction of the arrow on the emitter terminal. This arrow identifies the actual direction of the emitter current, and is always from p-to n-material. Therefore, it is a means by which one may identify two types of transistors.

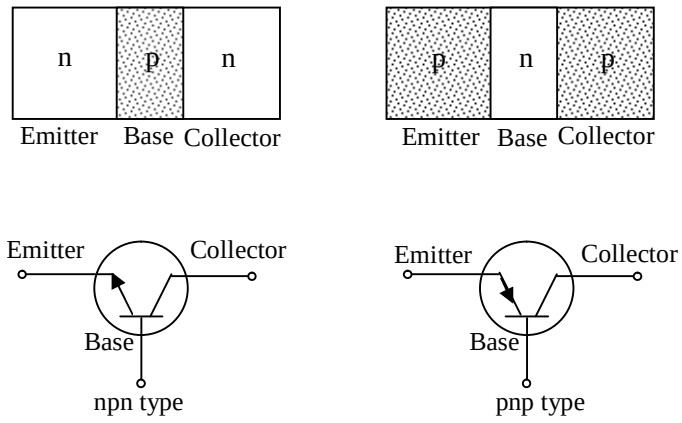


Fig 16.5 An npn and pnp transistor with their circuit symbols.

Fig 16.5(a) and 16.5(b) shows an n-p-n transistor and its energy band model when unbiased respectively. Fig 16.6(a) and (b) shows an n-p-n transistor biased as an amplifier and its energy band model respectively.

Transistor amplification is possible because a small signal applied to the base of the transistor can control a large current flow between the emitter and the collector. The

simple common base configuration as shown in fig 16.6 is the best for analyzing transistor amplification effect. A positive voltage called forward bias is applied to the left hand n-p junction by a small battery of about 1.5V through a resistor R_i of perhaps 50Ω across which a signal can be developed. A negative voltage called reverse bias is applied to the right hand junction by a battery, perhaps 9V, through a large resistor R_c of some 10,000 Ω, across which amplified signals can be developed. From our knowledge of the junction, it follows that electrons will flow from the emitter into the base material; but electrons cannot flow from the base terminal to the collector. Now, if the base is thick, 0.25 cm or more, these electrons would flow about the emitter circuit and will not be sensed by the collector circuit. But if the base is thin enough, then what follows is the principle of transistor amplification.

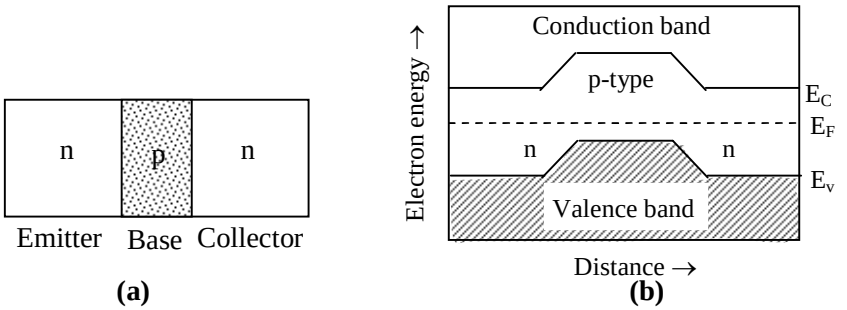


Fig 16.5(a) An n-p-n transistor, (b) Energy band model of an n-p-n transistor when unbiased .

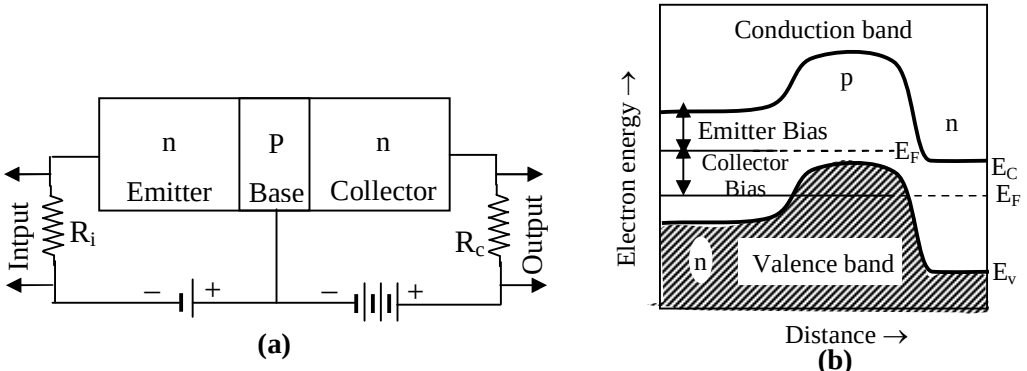


Fig 16.6 (a) An n-p-n transistor biased as an amplifier,
(b) Energy band model of n-p-n transistor biased as an amplifier.

Electrons emitted into the base find themselves in a region with small electric field; they diffuse rather randomly and may ultimately find themselves in the strong field of the reverse-biased junction. At this point they are swept through a collector junction and drawn through the collector circuit, producing a voltage across the output resistor R_c,

which is an amplified replica of the signal applied to the base. For a sufficiently thin base, a large fraction of the electrons (~99% or more) emitted by the forward-bias junction pass through the collector, and the efficiency of the device is high.

The gain of the transistor can be obtained as follows. Since the forward resistance of the emitter is low compared with R_i, any input signal V_i developed across R_i produces a forward current $i_i = V_i/R_i$. The collector current i_c, passing through the resistor R_c, produces an output voltage $V_c = i_c R_c$. If $i_c = i_i$ *i.e.*, if the current efficiency is 100%, then the voltage gain V_c/V_i is given by R_c/R_i. We must note that the high resistance of the reverse biased collector does not enter into the calculation; and the power generated in the output resistor is derived from the collector-biasing battery.

A p-n-p transistor operates in the same way as the n-p-n transistor we have just described, but the bias voltage is reversed. Since the transistor requires biasing, it does consume some power, but the consumption is much less than that of a vacuum tube.

Heavily doped p-n junctions:

16.2.3 The Tunnel Diode:

Tunneling is a well-founded concept in quantum mechanics. The electron tunneling at a p-n junction implies that electrons reach the other side of the junction by penetrating through the potential wall at the junction. It should not be confused with the act of crossing or jumping over the potential hill. It is purely a quantum statistical phenomenon based on the concept that the electron has a finite probability of being found on either side of the junction.

For tunneling to occur, the width of the space charge layer (or the depletion layer) should be of the order of, or less than, the electron mean free path. The width of the space charge can be reduced to this value by doping both regions of the p-n junction heavily. A p-n junction that is doped heavily enough to make the tunnel current greater than the usual diffusion current under certain conditions is called a tunnel diode. The width of the space charge layer in these diodes is usually less than 100 Å. The impurity concentrations are in the range 10^{19}–10^{20}cm^{-3}, whereas in ordinary p-n junctions they vary between 10^{14} and 10^{18}cm^{-3}.

The material of a tunnel diode behaves as a degenerate semiconductor. The Fermi level ε_F no more lies in the gap. Instead, it lies within the valence band of the p-region

and the conduction band of the n-region. We say that the Fermi level lies in a hybrid impurity-intrinsic band. Though a small width for the space charge layer is essential, it is not a sufficient condition for tunneling. There must be an unoccupied state on the other side of the junction into which an electron could tunnel at constant energy.

Figure 16.7(a) shows the state of an unbiased junction where the Fermi level has the same energy in both the regions of the junction. Let us first examine tunneling in a reverse-biased junction as shown in fig 16.7(b). When a small reverse bias voltage V_{ext} is applied, the height of the barrier step becomes much more than V_B. On account of this, the Fermi level has different values in the regions and at large values of V_{ext} a large number of occupied states of the valence band on the p-side of the depletion region lie opposite even to a larger number of empty states of the conduction band on n-side. The band picture is similar to that in a metal. When l_T is of the order of the electron wavelength or less, the junction has ideal conditions for electron tunneling. At higher values of V_{ext} the tunnel current can be quite large because of the enhanced level of overlap between the occupied and empty states as referred above. The probability of tunneling is given as

$$P_t \sim \exp(-kl_T) \quad \text{----------} \quad (16.12)$$

where k is the electron wavevector.

For a simple band structure of geometry as shown in fig (16.7(b)),

$$l_T = \frac{E_g}{eE} \quad \text{----------} \quad (16.13)$$

Substituting this value of l_T in (16.12), we get

$$P_t = \exp\left[\frac{-E_g(2m^*\varepsilon)^{1/2}}{eE\hbar}\right] \quad \text{----------} \quad (16.14)$$

Here E is an average value of the electric field at the junction which generally does not deviate much from E_{max}. Owing to the tunnel current, there is a quick onset of a critical reverse voltage. This property enables a reverse-biased tunnel diode to be used as a voltage regulator. These tunnel diodes are known as Zener diodes. The relation (16.14) provides a good description of tunneling in a Zener diode.

Now, we proceed to describe the other important tunnel diode that is used in the forward-biased configuration [Fig 16.7(c)]. This is again a heavily doped p-n junction,

known as the Esaki diode. If one approaches the forward-biased state from the reverse-biased condition, by decreasing V_{ext}, the tunnel current is on decrease and becomes zero at $V_{ext} = 0$.

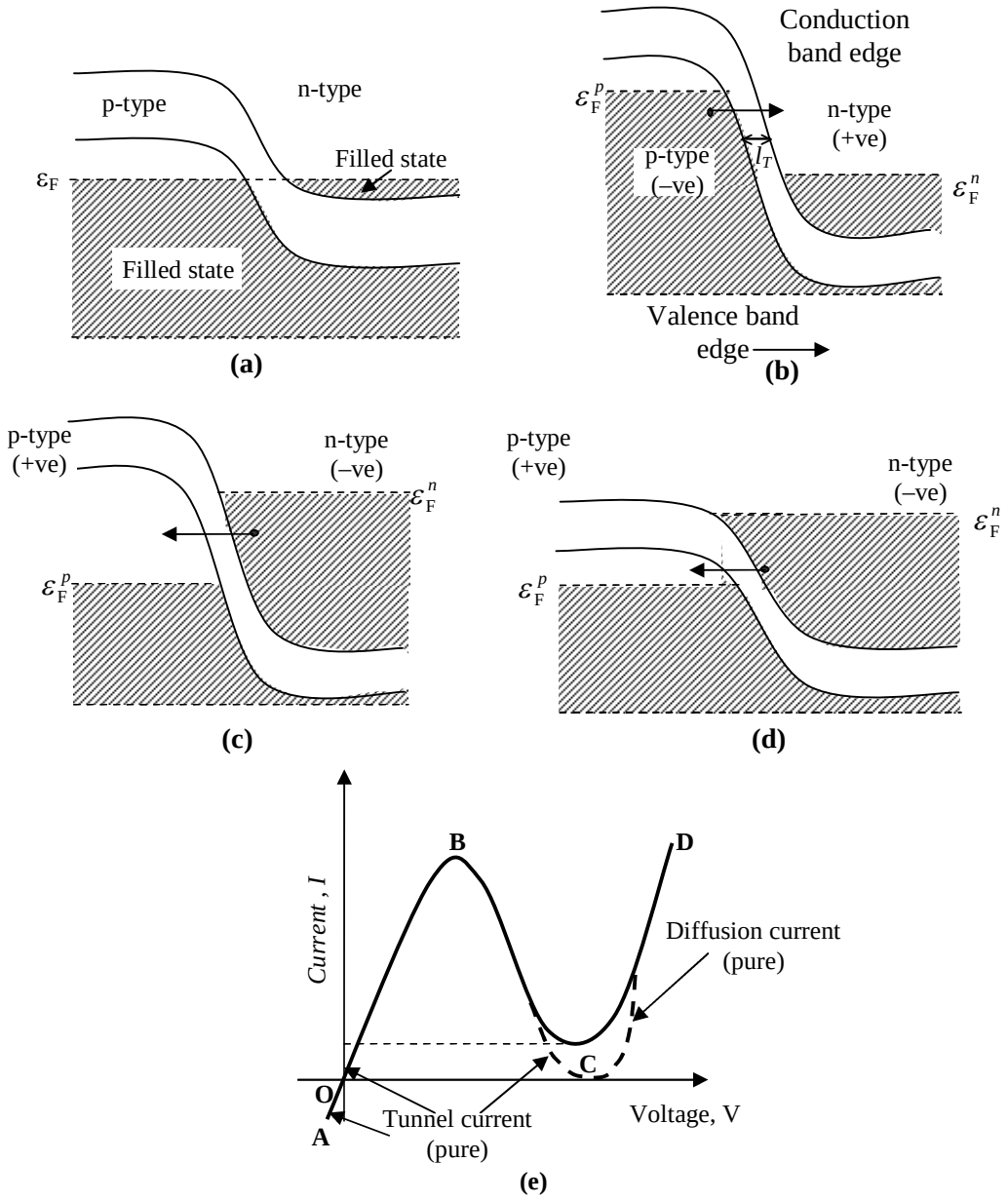


Fig 16. 7 (a) An unbiased tunnel diode in the state of thermal equilibrium. (b) A reverse-biased tunnel diode with different values of Fermi energy on the two sides of the junction.(c) A forward – biased tunnel diode.(d) The decrease in tunnel current on increasing the bias voltage further refers to the region of the negative resistance of the Esaki diode. (e) Voltage current characteristics of a tunnel diode. The portion AO represents the reversed-biased state; OB – current increases when forward biased; BC – negative resistance region when forward biased; CD – purely diffusion current.

The decrease in current occurs because of the reducing level of overlap at lower values of V_{ext} between the regions of occupied states in the valence band of the p-side and the unoccupied states in the conduction band of the n-side [Fig 16.7(b)]. When unbiased ($V_{ext} = 0$) there is no current at T = 0. At non-zero temperatures in this condition almost an equal small number of electrons tunnel from both sides of the junction keeping the current at zero value. The V-I characteristics of a tunnel diode are drawn in Fig 16.7(e). The portion AO refers to a reverse biased junction.

The characteristic curve beyond the point O describes the V-I relationship in an Esaki diode. As the forward voltage grows from zero to a certain value, the current keeps on increasing and approaches a maximum value. During this rise in the applied voltage, the Fermi level ε_F^P dips towards the conduction band edge of the n-side (E_c^n) and ε_F^n rises towards the valence band edge on p-side (E_v^p) as shown in fig 16.7(c). In the state of the maximum overlap of interest, the maximum tunnel current results. With a further increase in the applied voltage, the Fermi levels maintain the trend of dipping and rising in p- and n-regions, respectively. For the present band structure, the above change in Fermi levels is accompanied by reduction in the degree of overlap under question [fig 16.7 (d)] and therefore the reduction in tunnel current.

The main features of the forward-biased tunnel diode (the Esaki diode) are that the tunnel current passed through a sharp maximum and then drops to zero [see fig 16.7(e)]. The current is maximum when the forward voltage is about $(\varepsilon_F^P + \varepsilon_F^n)/2e$ and vanishes at about twice this value. In the V-I characteristic curve [Fig 16.7(e)], the portion OB represents the region of increasing current and the portion BC shows that the current decreases even when the voltage is being further increased. This is an indication of the negative resistance. We say that the tunnel diode has a negative resistance in this region. Beyond the point C there is no tunnel current. The observed rising current is totally contributed by the normal diffusion current. The current in practice does not drop to zero but a valley is observed. The most likely source of this excess current lies in the losses incurred during collisions of electrons with the lattice.

The negative resistance part of the V-I characteristics has been exploited to use an Esaki diode as a power amplifier or source. In conjunction with a capacitor and an

inductance it can be made to work as an oscillator and a switch. Switching times are limited by high capacitance C_j of the thin junction. The times are usually of the order of nanosecond (10^{-9} s). The switching time RC_j can be brought down to the picosecond range (10^{-12} s) by heavy doping that lowers the resistance *R* considerably.

16.2.4 Photodiodes:

Photodiodes convert light energy incident on them to electrical energy and the effect is termed as *Photovoltaic effect*. Photodiodes are used as sources of power (solar cell) and also as photo detectors.

The principle underlying the photovoltaic effect is straight-forward. Absorption of a photon in the region of the p-n junction leads to creation of an electron and a hole. Since the field within the junction is from n-side to p-side, the excess minority carries thus generated diffuse to the junction where they are carried across and become majority carriers – the electrons generated on p-side move towards n-side and holes generated on the n-side move towards p-side. Consequently, for the electrons and holes to recombine the electrons must go through the external circuit. If there is one external current path, majority carrier excess charge will be built up on both sides of the junction, with the result that the step in the built up potential will be reduced as shown in fig 16.8. The open circuit voltage can never be greater than the energy gap.

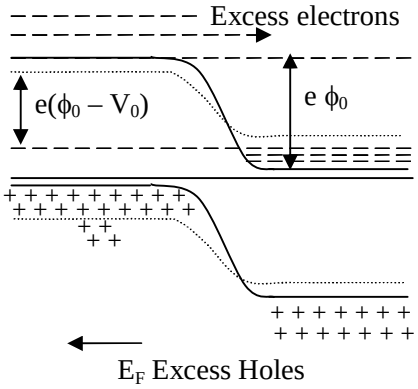


Fig 16.8 Band model picture of a photodiode (……) lines shows the band edge when it is illuminated and (——) show the band edges when it is in darkness.

If the external circuit is closed, the current will therefore flow in the circuit. The current in the circuit will flow so long as the semiconductor regions are illuminated i.e. due to absorption of photons diffusion of excess electrons from n-side and excess holes

from p-side will continue. This explains clearly how the incident light sets up the current flow in the external circuit.

The ability of an illuminated junction i.e., photodiodes to deliver power is made use of in a solar cell, used to convert solar energy into electrical energy. Such a solar cell has indeed been built and operated. The considerations involved in the operation of the solar cell are (i) power delivered (ii) efficiency of conversion. Presently large area p-n junctions of silicon are used to convert solar photons to electrical energy. The efficiency of these solar cells is not as much as one would wish. The conversion efficiency of solar cells is about 15%. The main problem is of fabricating high quality junctions of large areas necessary to intercept the maximum amount of incident light. Extensive research is being made for the search of proper materials.

Photodiodes are also used as radiation detectors. Diodes operated as Detectors of radiation are all reverse biased. When radiation is incident on the diode, the photo-generated carriers in the depletion region and in the bulk semiconductors on the either side of the depletion region move under the influence of the electric field present in the depletion layer. The increase in reverse current is proportional to the intensity of the incident radiation.

16.2.5 The Injection Laser:

The laser in which the lasing action is achieved by passing a current through a forward-biased p-n junction is called an *injection laser*. It is also referred to as a laser diode. The injection lasers have the distinction of meeting the requirements of fibre optic communication systems in the most convenient and effective manner. This has created tremendous interest in their study and development.

The principle behind the emission of light from a semiconductor is that of the recombination of electrons and holes at a p-n junction when a current is injected through the diode. As the current is passed through a forward-biased junction, the injected electrons and holes increase the density of electrons in the bottom of the conduction band and holes at the top of the valence band, simultaneously in the same region of space. A spontaneously emitted photon resulting from electron-hole recombination is fed back into the system by cleaving or polishing the ends of the junction diode. Under the state of population inversion the photon interacts with electrons in the conduction band to

produce simulated emission. The amplification is accomplished by the multiple reflections at the diode ends. At a certain value of the current, the electron and hole densities are so large that the rate of stimulated emission becomes greater than the rate of absorption and thus the process of amplification starts. When the current is increased further, at a certain threshold value the amplification exceeds the cavity losses and a coherent radiation is obtained in the output.

The threshold current density in the p-n junction based lasers is inconveniently high (~ 50,000 **A** cm^{-2}). This is drastically reduced in practical lasers that employ a double heterojunction. A heterojunction has two different semiconductor materials on its two sides. In a double heterojunction a lasing semiconductor is sandwiched between two wider-gap semiconductors of opposite doping. In a heterojunction laser developed by Kresel and Butler, the active layer is an undoped $Ga_{1-x}\,Al_x\,As$ (x = 0.05 – 0.1) which is embedded between n- and p-type layers of $Ga_{1-y}\,Al_y\,As$ (y = 0.3 – 0.4). The n- and p-type layers have the same band gap and the same refractive index. The band gap of the active layer is lower and characterized by a higher value of the refractive index. This ensures that the radiation emitted in the active layer is confined within this layer because of the total internal reflection occurring at its interfaces with the n- and p-type materials. Furthermore, the electrons from the n-side and the holes from the p-side that enter the active layer are not allowed by the potential barriers at the junctions to escape into the p- and n-sides (regions), respectively. This confinement of the light and the carriers results in amplification and wave guidance. For the high efficiency, materials only with the direct band gap are selected for fabrication. One of the biggest advantages with this heterostructure is that the frequency of the laser output can be finely monitored by varying the composition of the compound. For the present heterostructure there are two quasi-Fermi levels ε_F^c and ε_F^v referred to the conduction and valence band edges. The ε_F^c in the optically active layer is on level with the ε_F^n in the n-region and ε_F^v with ε_F^P in the p-region. The population inversion is achieved by applying a forward voltage that exceeds the voltage equivalent of the band gap of the active layer. Once the electrons and holes enter the active layer from n- and p-type regions respectively, they remain confined in the active layer as explained in the preceding paragraph. The emission of radiation occurs in the plane of the active layer when an electron from the conduction

band combines with a hole in the valence band as shown in fig 16.9. The process of stimulated emission and amplification proceeds the same way as in a p-n junction based device. The frequency of the emitted radiation is temperature dependent. A typical GaAs laser emits in the near infrared around 8380 Å.

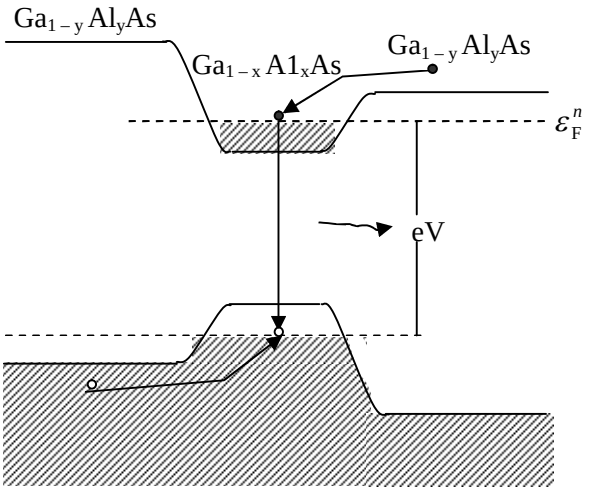


Fig 16.9 Double heterojunction GaAs-Al laser.

In the above device structure the carriers and the light are confined only along one direction. The confinement can also be provided in the lateral direction by surrounding the optically active layer with higher band gap materials on all sides, leaving a window for the output

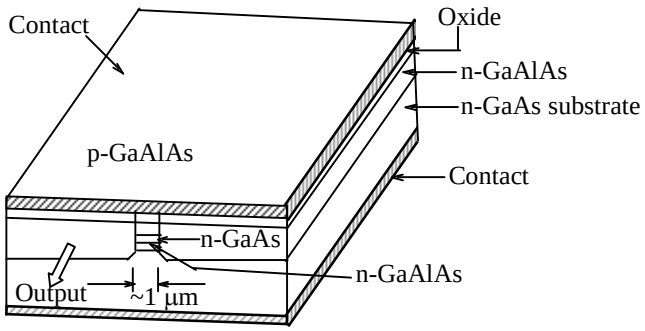


Fig 16. 10 Design structure of a buried Heterojunction laser.

. Such a heterostructure laser is known as the *buried heterostructure la*ser and shown in fig 16.10. The required threshold current densities (2000 – 4000 A cm^{-2}) are achieved in this structure for a significantly reduced value of current (< 50 mA), whereas for a stripe-

geometry laser with a standard size cavity the threshold current shoots a little over one ampere.

The heterostructures described above are fabricated by the successive deposition of multiepitaxial layers on an appropriate crystalline base (called the substrate). For GaAs laser, Kressel and Butler used a Si-doped n-type GaAs crystal as the substrate. The molecular beam epitaxy (MBE) and the metal organic chemical vapour deposition (MOCVD) techniques are used for the epitaxial growth because the level of precision obtained with them is rated as the highest toady. The injection lasers as described in this section are by no means the examples of most recent and versatile lasers, some of which have very complicated structures such as multiple quantum well structures. Our object was limited to show that the semiconductor devices have given birth to an altogether new field (called Photonics) whose promise to effectively replaced Electronics in near future to a large extent is exemplified by the great efficiency achieved in the field of fibre optic communication. Though still in its infancy, the subject has a great deal of potential for the future.

16.3 Thermoelectric effects:

Semiconductors have thermoelectric response many times stronger than metals. The three main thermoelectric parameters, namely the thermoelectric power S, the Thomson coefficient μ_T and the Peltier coefficient Π in semiconductors may be positive or negative, depending on whether the semiconductor is of p-type or n-type. We may recall that these parameters can have only negative sign in ideal conductors. Consider a semiconductor rod whose one end is hot and the other cold. As far as the algebraic sign is concerned, the case of an n-type semiconductor is similar to that of ideal conductors. But electrons at the hot end in a p-type semiconductor occupy acceptor levels on being excited from the valence band. This enables electrons at the cold end near the top of the valence band to lower their energy by moving into holes created in the valence band at the hot end. Thus the hot end becomes negatively charged and the cold end positively charged. Hence in a p-type semiconductor the parameters S, μ_T and Π are all positive. These arguments lead to a quick method of determining whether a specimen is of n-type or p-type. All that is needed is to measure the sign of the voltage across the given specimen whose one end is hot.

Now, we first consider the above mentioned semiconductor rod in an open circuit, ensuring that a constant temperature difference is maintained between its ends. In this condition a constant difference of potential is observed between its end. This potential difference can be interpreted to be associated with an electric field,

$$E = \frac{dV}{dx} = \frac{dV}{dT}\frac{dT}{dx} = S\frac{dT}{dx} \quad \text{----------} \quad (16.15)$$

Next, we consider a semiconductor rod kept at a constant temperature in a closed circuit. If an electric current is forced into it at one end from a certain external source, the current density in an n-type semiconductor may be written as

$$j_e = n(-e)(-\mu_e)E = ne\mu_e E \quad \text{----------} \quad (16.16)$$

where μ_e is the electron mobility in the direction opposite to that of the electric field E from the external source.

It is appropriate to measure the electron energies relative to the Fermi energy ε_F because the two materials in contact have a common Fermi level. Therefore, the average energy transported by electrons is equal to

$$(E_c - \varepsilon_F) + \frac{3}{2}k_B T \quad \text{----------} \quad (16.17)$$

where E_c denotes the energy at the conduction band edge.

The heat current density j_Q associated with the electric current density j_e is given by

$$j_Q = -n\left(E_c - \varepsilon_F + \frac{3}{2}k_B T\right)\mu_e E \quad \text{----------} \quad (16.18)$$

an electric current density j_x through a conductor along its length at constant temperature always has a heat current density j_Q associated with it such that

$$j_Q \propto j_x \text{ or } j_Q = \pi j_x$$

where π is called Peltier coefficient. This relation can be adopted to the present case as

$$j_Q = \Pi_e j_e$$

Therefore, the Peltier coefficient of an n-type semiconductor is

$$\Pi_e = -\frac{E_c - \varepsilon_F + \frac{3}{2}k_B T}{e} \quad \text{----------} \quad (16.19)$$

Similarly, for a p-type semiconductor, we have

$$j_h = pe\mu_h E \quad \text{----------} \quad (16.20)$$

$$j_Q = p\left(\varepsilon_F - E_v + \frac{3}{2}k_B T\right)\mu_h E \quad \text{----------} \quad (16.21)$$

and

$$\Pi_h = \frac{\varepsilon_F - E_v + \frac{3}{2}k_B T}{e} \quad \text{---------- (16.22)}$$

The Peltier coefficient can be easily determined by measuring the absolute thermoelectric power to which it is linked by the Kelvin relation Π = ST. The Π versus T plots for n-type and p-type specimens of silicon are displayed in fig 16.11. The significant observation is that the specimens behave as intrinsic semiconductors above 600 K.

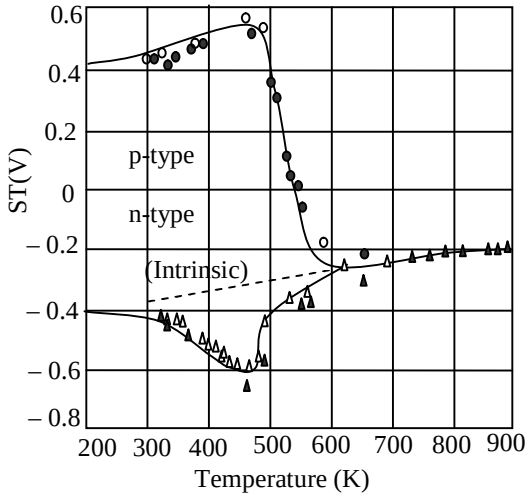


Fig 16.11 Variation of the Peltier coefficient (= ST, with S as the absolute thermoelectric power) of n-type and p-type specimens of silicon as a function of temperature.

Being many times intense than in metals, the thermoelectric effects in semiconductors are of added interest. For example, the thermoelectric powers of semiconductors are two orders of magnitude larger than those for non-ferromagnetic metals. The reasons ascribed to this huge quantitative difference in response are mainly twofold. Firstly, the carrier density of semiconductors is sensitive to temperature owing to which the hot end has more conduction electrons or holes per unit volume, depending on whether the material is of p-type or n-type. Secondly, and more importantly the presence of a forbidden energy gap in semiconductors proves to be a major cause of the observed behaviour.

To elaborate on the second reason we examine the quantity of Peltier heat evolved or absorbed at a metal semiconductor junction. Let a current be flown from an n-type semiconductor to a metal in contact forming a junction as shown in fig 16.12(a). In an n-

type semiconductor the majority carriers (electrons) are in the conduction band placed well above the Fermi level ε_F. On the other hand, most of the electrons in a metal are near the Fermi level. Because of being in contact, the Fermi energy would be at the same level in the metal and the semiconductor. Figure 16.12(a) clearly indicates that the average energy of conduction electrons in the electric current decreases as the same electric current enters the metal. The difference of electron energies in the two regions is lost in the form of heat at the junction. Similarly, when a p-type material is in contact with a metal, heat would be absorbed for the shown direction of electric current because the majority carriers (holes) in the semiconductor would occupy energy states in the valence band that are much below the Fermi level (the approximate level of electric current in metal).

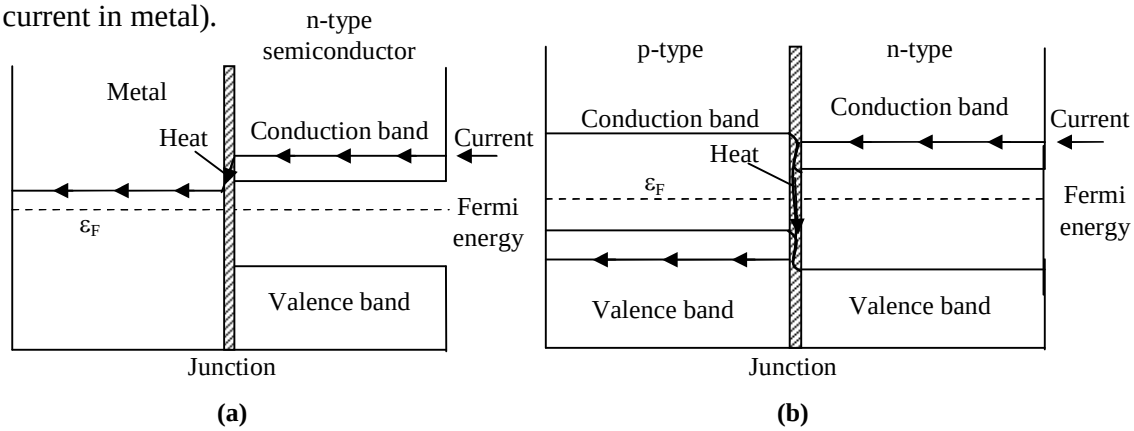


Fig 16.12 (a) Peltier effect at a metal–semiconductor junction. Heat is evolved at the junction when an electric current flows from an n-type semiconductor to a metal. (b) Peltier effect at a p-n junction. Relatively large amount of heat is evolved at the junction in the case of an electric current flowing from the n-region to the p-region.

Application:

It is obvious from the above discussion that a thermocouple junction can function as a heat pump or a refrigerator when an electric current is forced through the thermocouple. Under the flow of an electric current the thermocouple pumps heat from one junction to the other. We show below that a p-n junction can serve as a better heat pump or refrigerator than any common thermocouple junction consisting of two semimetals or even a junction such as that shown in Fig 16.12(a).

Consider a p-n junction through which an electric current flows as demonstrated in fig 16.12(b). On the n-side the majority carriers are at much higher energy levels than the Fermi level whereas on the p-side the majority carriers occupy energy states that are

much below the Fermi level. Hence, with the flow of electric current in this condition from the n-region to the p-region a relatively far more energy would appear in the form of heat at the junction, as compared to that in a junction shown in fig 16.12(a).

In principle, a Peltier heater can be more efficient than an electrical resistance heater. For the requirement of each **kW** of heat, an electrical resistance heater must consume one **kW** of electrical power that is dissipated. On the other hand, an ideal Peltier heater needs electrical power only to pump the heat energy from one junction (cold) to the other (hot) like a refrigerator or heat engine and thus operates at a relatively much smaller power. But in practice the Peltier pumps are not as efficient as the conventional mechanical heat pumps. The inefficiency is attributed partly to the heat loss through the thermocouples from the hot to the cold side and partly to the Joule heating of the thermocouples.

16.4 Quantum Hall effect:

Although the treatment of Hall effect given in Lesson 15 is based on purely classical considerations, it gives a good account of the electrical transport in metals and semiconductors. But the classical magnetoconducting scenario undergoes a spectacular transformation under quantum conditions of temperature and magnetic field in a two-dimensional conductivity channel. K. von Klitzing, Dorda and Pepper observed that such a channel is formed at the oxide interface in a metal-oxide-semiconductor (MOS) transistor when a gate voltage is applied between the metal and the semiconductor, as shown in fig 16.13. The fascinating aspect of their observation is that the Hall resistance ρ_{H} varies with the magnetic field according to the following rule.

$$\rho_{\mathrm{H}} = \frac{h}{ie^2} \quad \text{----------} \quad (16.23)$$

where i is an integer (= 1, 2, 3, ……).

The phenomenon expressed by this rule, where the Hall conductance is quantized in units of e^2/h, is called the *Integral Quantum Hall Effect* (IQHE).

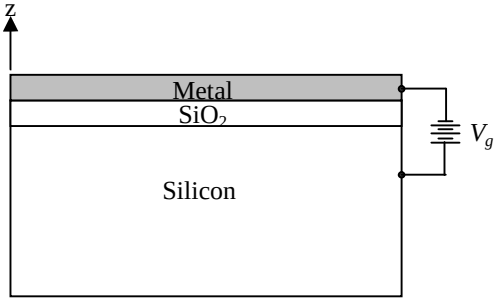


Fig 16.13 A metal-oxide-semiconductor (MOS) transistor.

In the original experiment a constant current of 1 μA was forced to flow between the source and the drain in the presence of a magnetic field of 18 tesla at 1.5 K. The results of this experiment are graphically displayed in fig 16.14 and the experimental geometry is illustrated in fig 16.15(a).

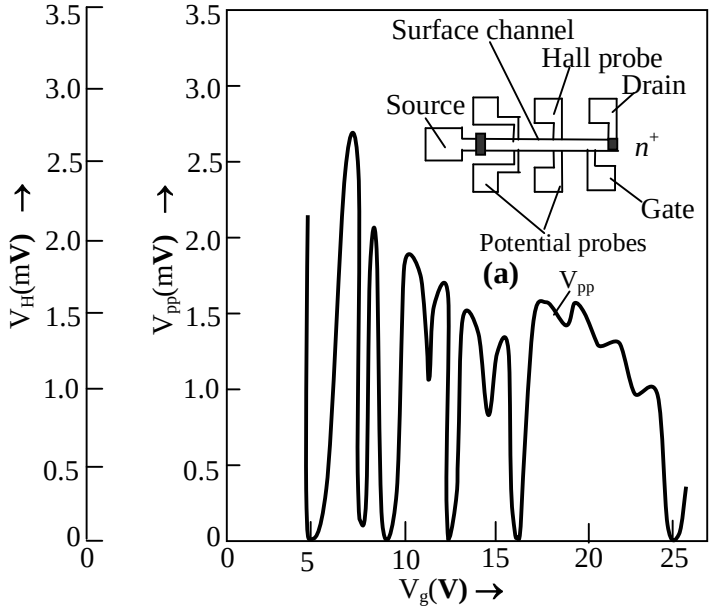


Fig 16.14 The IQHE voltage V_{pp} and the Hall voltage V_H as functions of the gate voltage V_g. The MSO transistor set-up for the IQHE measurements is shown above the plots.

The Hall field versus magnetic field plot in this case is not a straight line (as in the classical Hall effect). Instead, it shows plateaus with steps in between at certain values of the magnetic field [see fig 16.15b].

First we show how the observed minima in the longitudinal voltage V_{pp} can be explained and relation (16.23) be derived, using only a crude model. Then, we give a qualitative analysis of real two-dimensional systems in the framework of a general theory.

Consider a surface current density $\mathbf{J}_x$ in the x-direction defined as the current crossing a line of unit length in the y-direction on the oxide interface (the xy-plane):

$$I_x = j_x L_y \quad \text{----------} \quad (16.24)$$

If n is the total electron density per unit area of the xy-plane.

$$j_x = nev_d \quad \text{----------} \quad (16.25)$$

with v_d as the drift velocity of electrons in the x-direction.

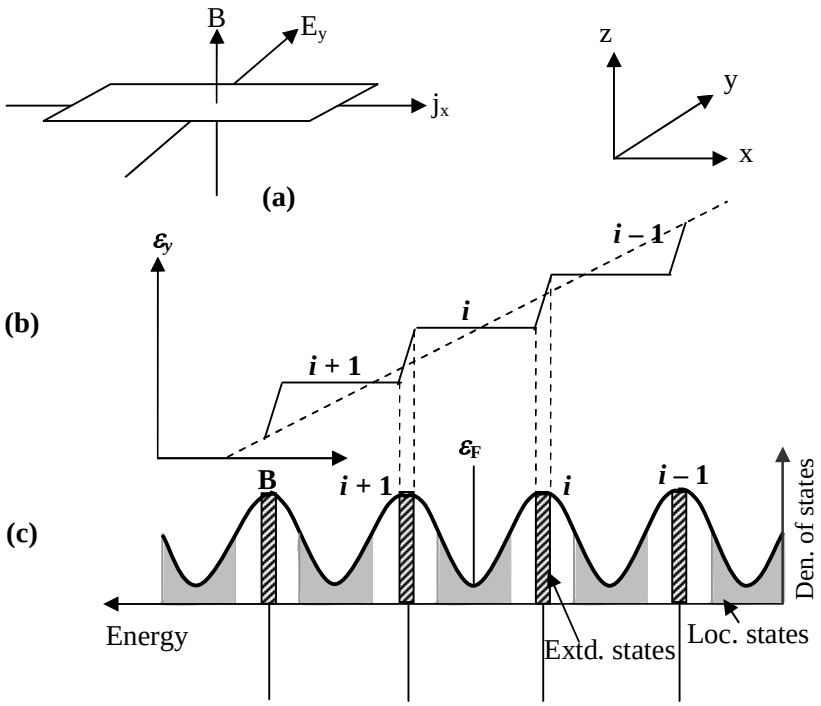


Fig 16.15 (a) The IQHE geometry. (b) The IQHE plateaus in the Hall field for a fixed system current j_x. (c) Density of states in a two-dimensional real system in a strong magnetic field.

In an alternative description of the Hall effect, the current in a specimen with mobile charge carriers is produced when the specimen is placed in a region of mutually crossed (perpendicular) electric and magnetic fields. The flow of current in a direction orthogonal to both the fields is detected on closing the circuit. If the electric field ($\mathbf{E}_y$) and the magnetic field **B** act along y- and z-directions, respectively

$$\mathbf{j}_x = \sigma_{xx}\mathbf{E}_x + \sigma_{xy}\mathbf{E}_y = \sigma_{xy}\mathbf{E}_y \quad \text{----------} \quad (16.26)$$

(since $\mathbf{E}_x = 0$)

and

$$v_d = \frac{E_y}{B} \quad \text{----------} \quad (16.27)$$

Here σ_{xy} denotes the conductivity tensor in the plane of the two-dimensional channel. Interpreting the resistivity ρ_{xy} as the Hall resistance ρ_H.

$$\rho_H = \frac{V_H}{I_x} = \frac{E_y L_y}{j_x L_y} = \frac{B}{ne} \quad \text{----------} \quad (16.28)$$

As follows from relation (16.24), ρ_H represents the resistance of a channel of unit thickness. In the commonly used geometry for the Hall effect, a current is flown along

the x-direction in the presence of a magnetic field along the z-direction and the Hall voltage V_H measured along the y-direction in the specimen. The description contained in relation (16.28) is consistent with this experimental geometry.

We now know that the areas of the successive electron orbits in the k-space in the presence of a magnetic field B differ by an amount $2\pi eB/\hbar$. Considering a square of side L on the xy-plane, the number of states in this area are estimated as

$$\left(\frac{2\pi eB}{\hbar}\right)\left(\frac{L}{2\pi}\right)^2 = \frac{eL^2B}{h} \quad \text{---------- (16.29)}$$

It gives the number of electron levels that coalesce into a single magnetic level as soon as even a small magnetic field is switched on. This magnetic level defines a Landau level whose energies in the xy-plane are quantized as $(i + ½)\ \hbar\omega_c$ with $\omega_c = eB/m^*$ (the cyclotron frequency). Relation (16.29) gives essentially the measure of degeneracy of a Landau level. In the present context we define the degeneracy per unit area of the xy-plane as

$$D(B) = \frac{eB}{h} \quad \text{---------- (16.30)}$$

Let the applied magnetic field be so strong that $\hbar\omega_c >> k_BT$. It is then reasonable to talk in terms of completely filled or completely empty Landau levels. Suppose B_i is the critical field at which no level is partly filled and I is the magnetic quantum number of the highest occupied level. When the electron density on the oxide interface is adjusted by varying the gate voltage so that the Fermi level coincides with level *i*,

$$n = i \cdot D(B_i) \text{ ---------- (16.31)}$$

Under these conditions electrons can undergo neither elastic nor inelastic collisions. The elastic collisions would involve scattering of electrons from one state to the other in the same Landau level. But this is not permitted by the Pauli principle, since all possible final states of equal energy are occupied. The inelastic collisions can be possible with the scattering of electrons to a vacant Landau level by absorbing the required energy from some source, most likely phonons. But in the experimental conditions of law temperature and $\hbar\omega_c >> k_BT$ as established here, there are hardly any phonons whose energy could compare with the large energy interval $\hbar\omega_c$. Therefore, the inelastic collisions are too

ruled out and the electron mean free paths are greatly enhanced. This results in the occurrence of the voltage minima in V_{pp} (or the longitudinal resistance minima).

Placing the value of D(B) from (16.30) in (16.31), we obtain (16.9):

$$\rho_H = \frac{h}{ie^2} \quad \text{(in ohms)}$$

Analysis in real systems

The experimental evidence for the Hall resistance being accurately quantized at h/ie^2 ohms might apparently suggest that IQHE is independent of purity and crystallinity, simply because the theory predicting this quantization does not take these aspects into consideration. But the presence of impurities or microcrystallinity produces disorder, rendering the crystal potential irregular as a result of which the sharp levels in ideal systems are broadened into bands in real systems [fig 16.15(c)]. This affects the Hall resistance to such an extent that its linear variation changes to develop plateaus. In a two-dimensional system that concerns the present discussion all the electron states are predicted as localized at any disorder. To the credit of this prediction the IQHE actually approaches this limit as the magnetic field goes to zero. Therefore, it is logical to believe that there can exist both the extended and localized types of carrier states in a band. As per the latest concept of localization, the extended and localized states cannot coexist at the same energy. The localized states occupy the region of the lowest density of states forming the mobility gap and do not contribute to the flow of electronic current. The extended states, on the other hand, appear around the peaks of the density of states [see fig 16.15(c)].

In light of his thought experiment on a two-dimensional system, Laughlin has interpreted the IQHE in real systems as a consequence of the principle of gauge invariance. By analogy with the flux quantization in a superconductor (where the unit of charge is 2e), the flux quantization in the IQHE (with e as the unit of charge) is discussed. For a certain increase δB in the magnetic field, there is an addition of one flux quantum that enhances the degeneracy of each Lamdau level by one. Suppose i denotes the magnetic quantum number of the highest completely filled Landau level. If all electrons cannot be accommodated up to this level at absolute zero, the Fermi level ε_F will coincide with the level $(i + 1)$which is only partially filled.

With an increase of one flux quantum, each of the Landau levels will have one additional level in its subband. Consequently, the level (i + 1) will be vacated, since the electrons in this level can now be accommodated in the newly created lower energy levels (i in total). Over a fixed range δB (equivalent to one flux quantum), during which the Fermi level remains in a level, a plateau in the Hall resistance is observed. The increase δB equal a magnitude for which at the stage under consideration all the states in the level (i + 1) are vacated and the Fermi level drops to coincide with the level i. Thus for each increase of δB, a jump to the next plateau takes place. The next plateau refers to a level with the next lower magnetic quantum number [see fig 16.15(c)].

It is important to assert the role of disorder and localization in the IQHE for real systems. Given next is a brief discussion conducted in this approach.

The Hall electric field in an ideal system is

$$E_y(B) = j_x B_z R_{\mathrm{H}} = v_d B_z = \frac{D(B) v_d h}{e} \quad \text{---------- (16.32)}$$

where R_H denotes the Hall coefficient.

In a disordered system the degeneracy of states and the electron densities may be split as

$$D(B) = D^{E}(B) + D^{L}(B) \quad \text{---------- (16.33)}$$

$$n = n^{E}(B) + n^{L}(B) \quad \text{---------- (16.34)}$$

where E and L refer to the extended and localized states, respectively.

Then, the Hall field in a disordered system may be expressed as

$$\varepsilon_y(B) = \frac{D^{\mathrm{E}}(B) v_d(B) h}{e} \quad \text{---------- (16.35)}$$

Where $v_d(B)$ stands for the drift velocity in the disordered system.

As observed in a typical IQHE experiment, the current density j_x carried by electrons in the extended states must remain unchanged at the value nev_d. Hence,

$$v_{\mathrm{d}}(B) = \left(\frac{n}{n^{E}(B)} \right) v_{\mathrm{d}} \quad \text{---------- (16.36)}$$

It implies that n^E electrons per unit area carry the current (of density nev_d) with a higher drift velocity to compensate for the loss of current because of the localization of n^L electrons per unit area.

The $D^E(B)$ in a particular band always increases with B, though non-monotonically. It increases by one, only when a δB–increase of B creates an extended state in this band. We can check that this occurs with probability 1/(v + 1), where

$$v = \frac{D^L(B)}{D^E(B)}$$

From (16.21) we see that $\varepsilon_y(B)$ can remain unchanged (forming a plateau) with B whenever $v_d(B)$ decreases, provided

$D^E(B)\, v_d(B)$ = a constant

But it should be observed that $v_d(B)$ increases as well as decreases with B depending on where the Fermi level ε_F lies in the mobility gap, we will have either (i) $v_d(B + \delta B) < v_d(B)$, when an extended state is produced and D^E increases in the subbands below ε_F or (ii) $v_d(B + \delta B) = v_d(B)$, in between the above events where the Hall field $\varepsilon_y(B)$ remains unchanged and the plateau occurs.

On the other hand, when ε_F falls in the band of extended state we always have $v_d(B + \delta B) < v_d(B)$ because D^E decreases on account of the downward movement of ε_F. For a complete treatment, the reader is referred to calculations made elsewhere. The calculations show that a plateau in the Hall field $\varepsilon_y(B)$ is formed within an accuracy of a few parts in $10^6 D(B)/(v + 1)$.

In the extreme quantum conditions, i.e. at extremely low temperatures and extremely high magnetic fields a QHE has been observed for fractional values of *i* in relation (16.23). In this limit the lowest Landau level is only partially occupied and the IQHE is not expected. Some of the fractional values of *i* for which the Hall resistance has been observed to be quantized are $\frac{1}{3}, \frac{2}{3}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}$ and $\frac{2}{7}$. At these occupancies the longitudinal Hall resistance ρ_{xx} is found to vanish. This phenomenon is called the *Fractional Quantum Hall Effect* (FQHE).

16.5 Summary

Semiconductors have emerged as the indispensable material base for the development of Solid State Electronics.

Semiconductor devices generally exploit the characteristics of inhomogeneous semiconductors in which the donor and acceptor concentrations are not uniform. A

common form in which they are used contains two separate *n* and *p* type regions with an abrupt partition boundary (a junction) within a single crystalline sample. This junction is known as the *p-n* junction

Characteristics of the *p-n* junction are crucial to the fabrication and working of most of the semiconductor devices. It is an essential component of the present electronics circuitry ranging from a simple rectifier circuit to transistor circuits and integrated circuit used in sophisticated appliances like modern computer.

Unbiased p-n junction Barrier step height $V_B = \frac{k_B T}{e} \ln\left(\frac{p_p n_n}{n_i^2}\right)$

When an external steady voltage is applied across a p-n junction, the barrier step height changes. In a forward-biased configuration it decreases to value,

$$V_{BF} = V_B - V_{ext}$$

Where V_{ext} denotes the applied voltage.

For a reverse-biased connection in contrast to the forward-biased case, the direction of the applied field does not favour the flow of majority carriers across the junction, thus increasing the barrier height to

$$V_{BR} = V_B + V_{ext}$$

The property that a p-n junction favours the flow of current in one direction across the junction forms the basis for using it as a rectifier.

In a junction rectifier the current voltage relationship has the form

$$I = I_0\left(e^{eV_0/k_BT} - 1\right)$$

when V_0 – is the bias voltage. When this voltage is large in forward direction $V_0 > 0$ the current is large and increases rapidly with the voltage. But for reverse bias $V_0 < 0$ and $e^{eV_0/k_BT} \ll 1$ and $I = -I_0$. The current is now small, and indent of voltage.

A junction transistor comprises of two junctions connected back-to-back. One, called the emitter, is forward biased and the other, called the collector, is reverse biased.

When an electric signal is applied at the emitter, a corresponding carrier pulse passes through the base and the collector, and the amplified signal is picked up at a load resistor inserted into collector circuit.

Photo diodes convert light energy into electrical energy and are used as sources of power (solar cell) and also as photo detectors.

A p-n junction that is doped heavily enough to make the tunnel current greater than the usual diffusion current under certain conditions is called a **tunnel diode**.

The material of a tunnel diode behaves as a degenerate semiconductor. The Fermi level ε_F no more lies in the gap. Instead, it lies within the valence band of the p-region and the conduction band of the n-region.

This property enables a reverse-biased tunnel diode to be used as a voltage regulator. These tunnel diodes are known as Zener diodes.

The main features of the forward-biased tunnel diode (the Esaki diode) are that the tunnel current passed through a sharp maximum and then drops to zero.

The laser in which the lasing action is achieved by passing a current through a forward-biased p-n junction is called an *injection laser*. It is also referred to as a laser diode. The injection lasers have the distinction of meeting the requirements of fibre optic communication systems in the most convenient and effective manner. This has created tremendous interest in their study and development.

Semiconductors have thermoelectric response many times stronger than metals.

In a p-type semiconductor the parameters S, μ_T and Π are all positive.

A p-n junction can serve as a better heat pump or refrigerator than any common thermocouple junction consisting of two semimetals.

Magnetoconducting scenario undergoes a spectacular transformation under quantum conditions of temperature and magnetic field in a two-dimensional conductivity channel.

Under quntum conditions the Hall resistance ρ_H varies with the magnetic field according to the following rule.

$$\rho_H = \frac{h}{ie^2}$$

where *i* is an integer (= 1, 2, 3, ……).

The phenomenon expressed by this rule, where the Hall conductance is quantized in units of e^2/h, is called the *Integral Quantum Hall Effect* (IQHE).

16.6 Key words

Depletion layer – Barrier height – Forward bias – Reverse bias – Tunnel diode – Esaki diode – Zener diode – Injection laser – Stimulated emission – Electron recombination – Hetero structures – Buried hetero structure laser – Photonics – Thermo electric effects –

Thermo electric power – Thomson coefficient – Peltier coefficient – Integral Quantum Hall effect – Flux quantization – Fractional Quantum Hall effect

16.7 Review questions

1. Describe the effect on space charge width at a p-n junction when the junction is a) forward biased b) when reverse biased.
2. Describe transistor action with the help of band diagram.
3. Describe the volt-ampere characteristic of Tunnel diode.
4. Explain the mechanism whereby a p-n junction when illuminated delivers electrical energy to a circuit.
5. Explain why Isaki diodes do not show high electrical conductivity in spite of having very large carrier concentrations.
6. Describe briefly Hall effect, Integral Quantum Hall effect and fractional Hall effect.

16.8 Text and reference books

1. Elements of Solid State Physics by J.P. Srivatsava (PHI)
2. Solid State Physics by M.A. Wahab (Narosa)
3. Elements of Solid State Physics by A. Omar (Pearson education)
4. Solid State Physics by S.O. Pillai (New Age)
5. Solid State Physics by C. Kittel (Asia Publishing House)
6. Solid State Physics by S.L. Kakani and C. Hemrajani (S.Chand)
7. Solid State Physics by Saxena Gupta Saxena (Pragati Prakashan).
8. Solid State Physics by C.J. Dekker (Macmillan)