

PHS 319

(SOLID STATE PHYSICS I)

BY

OLOPÄDE M.A.

TABLE OF CONTENT

	PAGE
Unit 1 Crystal Structure of solids	3
Unit 2 Crystal Binding	20
Unit 3 X-ray diffraction in crystals and Applications	36
Unit 4 Thermal properties of the crystal lattice	53
Unit 5 Elastic properties	69
Unit 6 Lattice Vibrations and Phonon	82
Unit 7 Free electron theory of Metals	98
Unit 8 Motion of electrons in periodic fields	110
Unit 9 Energy bands	125
Unit 10 Semiconductors I.....	144
Unit 11 Semiconductors II	159
Unit 12 Hall Effect	173
Unit 13 Superconductivity	179
Solutions to Self Assessment questions (SAQ)	185

Study Session 1 Crystal Structure of Solids

Introduction

A crystal is a solid composed of atoms or other microscopic particles arranged in an orderly repetitive array. That is about the shortest definition of a crystal. It may not be a complete definition, yet it is a true description.

There are three general types of solids – amorphous, polycrystalline and single crystal which are distinguished by the size of ordered regions within the materials. The arrangement of atoms in amorphous solids is limited to a few molecular distances. In polycrystalline materials, the solid is made up of grains which are highly ordered crystalline regions of irregular size and orientation. Single crystals have long-range order. Many important properties of materials are found to depend on the structure of crystals and on the electron states within the crystals. At the beginning of the study of crystals it was their external form which was related to the physical properties. In this way, only a limited success was achieved. In the middle of the last century, a deeper understanding was developed regarding the correlation of the structure of crystals and mechanical, thermal, electrical, and magnetic properties of solids. This is primarily due to the advances in the band theory of electron states and in the theory of bonding in solids. This knowledge has led to the development of newer and better materials for electrical, electronic and structural engineering. The study of crystal physics aims to interpret the macroscopic properties in terms of the properties of the microscopic particles of which the solid is composed. The study of the geometric form and other physical properties of crystalline solids by using x-rays, electron beams and neutron beams constitute the science of crystallography or crystal physics.

Learning Outcomes for Study Session 1

When you have studied this session, you should be able to:

- 1.1 Define crystal and understand the three types of solids.
- 1.2 Explain what is meant by a unit cell, Bravais lattice and non-bravais lattice.
- 1.3 Describe the different kinds of crystal symmetries.

1.1 Main Content

A solid is said to be a crystal if its atoms are arranged in such a way that their positions are exactly periodic. This concept is illustrated in Figure 1.1 using a two-dimensional (2D) structure.

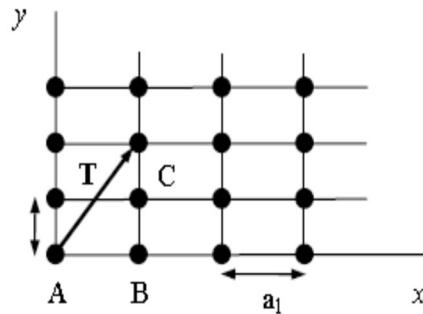


Figure 1.1 Crystal Structure showing the positions of the atoms

A perfect crystal maintains this periodicity in both the x and y directions from $-\infty$ to $+\infty$. Following from this periodicity, the atoms A, B, C, etc. are *equivalent*. In other words, for an observer located at any of these atomic sites, the crystal appears exactly the same. The same idea can be expressed by saying that a crystal possesses a *translational symmetry*. The translational symmetry means that if the crystal is translated by any vector joining two atoms, say \mathbf{T} in Figure 1.1, the crystal appears exactly the same as it did before the translation. In other words the crystal remains *invariant* under any such translation. The structure of all crystals can be described in terms of a *lattice*, with a group of atoms attached to every lattice point. For example, in the case of the structure shown in Fig.1.1, if we replace each atom by a geometrical point located at the equilibrium position of that atom, we obtain a crystal lattice. The crystal lattice has the same geometrical properties as the crystal, but it is devoid of any physical contents. There are two classes of lattices: the *Bravais* and the *non-Bravais*. In a Bravais lattice all lattice points are equivalent and hence by necessity all atoms in the crystal are of the same kind. On the other hand, in a non-Bravais lattice, some of the lattice points are non-equivalent.

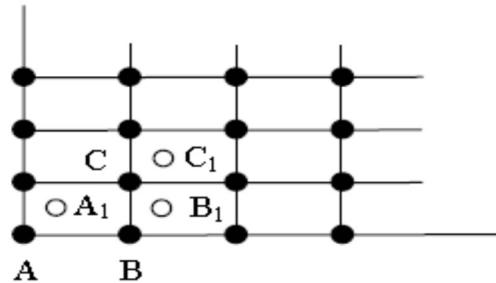


Figure 1.2 Representation of atomic lattice sites

In Figure 1.2 the lattice sites A, B, C are equivalent to each other. Also the sites A₁, B₁, C₁, are equivalent among themselves. However, sites A and A₁ are not equivalent: the lattice is not invariant under translation AA₁.

Non-Bravais lattices are often referred to as a ***lattice with a basis***. The basis is a set of atoms which is located near each site of a Bravais lattice. Thus, in Figure 1.2 the basis is represented by the two atoms A and A₁. In a general case, crystal structure can be considered as

$$\text{Crystal structure} = \text{lattice} + \text{basis}.$$

The lattice is defined by fundamental translation vectors. For example, the position vector of any lattice site of the two dimensional lattice in Figure 1.3 can be written as

$$\mathbf{T} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 \quad (1.1)$$

where \mathbf{a}_1 and \mathbf{a}_2 are the two vectors shown in Fig.3, and n_1, n_2 is a pair of integers whose values depend on the lattice site.

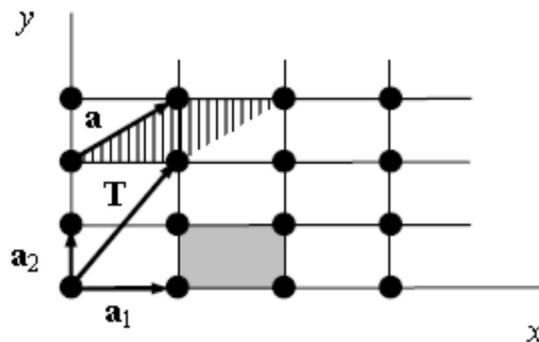


Figure 1.3 Position vector of lattice sites in 2-D

So, the two non-collinear vectors \mathbf{a}_1 and \mathbf{a}_2 can be used to obtain the positions of all lattice points

which are expressed by Eq.(1.1). The set of all vectors \mathbf{T} expressed by this equation is called the ***lattice vectors***. Therefore, the lattice has a translational symmetry under displacements specified by the lattice vectors \mathbf{T} . In this sense the vectors \mathbf{a}_1 and \mathbf{a}_2 can be called ***the primitive translation vectors***.

The choice of the primitive translations vectors is not unique. One could equally well take the vectors \mathbf{a}_1 and $\mathbf{a} = \mathbf{a}_1 + \mathbf{a}_2$ as primitive translation vectors (see Figure 1.3). This choice is usually dictated by convenience.

Unit cell: In the case of a rectangular two dimensional lattice the unit cell is the rectangle, whose sides are the vectors \mathbf{a}_1 and \mathbf{a}_2 . If the unit cell is translated by all the lattice vectors expressed by Eq. (1.1), the area of the whole lattice is covered once and only once. A ***primitive unit cell*** is the unit cell with the smallest area which produces this coverage. In the two dimensional case the area of the unit cell is given by $S=|\mathbf{a}_1 \times \mathbf{a}_2|$.

The choice of the unit cell is not unique. For example, the parallelogram formed by the vectors \mathbf{a}_1 and \mathbf{a} in Figure 1.3 is also an acceptable unit cell. The choice is again dictated by convenience. The area of the unit cell based on vectors \mathbf{a}_1 and \mathbf{a}_2 is the same as that based on vectors \mathbf{a}_1 and \mathbf{a} .

In-text question

1. Q: Define the following terms in crystal structure study: (i) Lattice (ii) Primitive cell
(iii) Unit cell.

A: see pages 4-6.

Some important crystal structure terms are defined below:

Coordination number (N): The coordination number is defined as the number of equidistant nearest neighbours that an atom has in the given structure. The greater the Coordination number, the more closely packed up will be the structure.

Atomic packing factor: This the fraction of the space occupied by atoms in a unit cell.

Atomic radius (r): Atomic radius is defined as half the distance between nearest neighbours in a crystalline solid without impurity.

1.1.1. Wigner-Seitz unit cell:

The primitive cell may be chosen as shown in Figure 1.4.

- (i) Draw lines to connect a given lattice point to all nearby lattice points.
- (ii) At the midpoint and normal to these lines, draw new lines (planes in 3D). The smallest volume enclosed is the Wigner-Seitz primitive cell. All the space of the crystal may be filled by these primitive cells, by translating the unit cell by the lattice vectors.

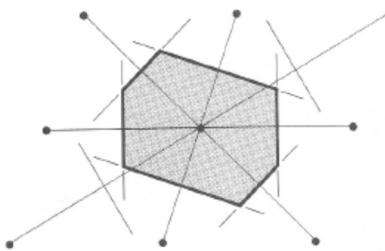


Figure 1.4 Wigner-Seitz primitive cell

The unit cell can be **primitive** and **non-primitive (or conventional)**. The unit cell discussed above is primitive. However, in some cases it is more convenient to deal with a unit cell which is larger; however, as it exhibits the symmetry of the lattice more clearly.

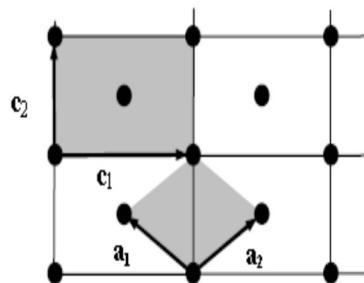


Figure 1.5 Showing primitive translation vectors for a lattice site

Vectors \mathbf{a}_1 and \mathbf{a}_2 can be chosen as primitive translation vectors for the lattice shown in Fig.1.5. In this case, the unit cell is a parallelogram. However, the lattice can also be regarded as adjacent

rectangles, where the vectors \mathbf{c}_1 and \mathbf{c}_2 can be considered as primitive translation vectors. The unit cell in this case is larger; however it exhibits the rectangular symmetry more clearly. In the first case, we have just one atom in a unit cell, whereas in the second case, we have a lattice with a basis. The basis consists of the two atoms: one atom is located in the corner of the unit cell and another atom in the center of the unit cell. The area of the conventional unit cell is larger by a factor of two than the area of the primitive unit cell.

1.2 Crystal symmetry

Crystal lattices are classified according to their symmetry properties, such as inversion, reflection and rotation.

Inversion center: A cell has an inversion center if there is a point at which the cell remains invariant under transformation $\mathbf{r} \rightarrow -\mathbf{r}$. All the Bravais lattices are inversion symmetric. Non-Bravais lattices may or may not have an inversion center depending on the symmetry of the basis.

Reflection plane: A cell has a reflection plane if it remains invariant when a mirror reflection in this plane is performed.

Rotation axis: This is an axis such that, if the cell rotated around the axis through some angle, the cell remains invariant. The axis is called n -fold if the angle of rotation is $2\pi/n$. Only 2-, 3-, 4-, and 6-fold axes are possible.

There are five Bravais lattice types in two dimensions shown in Figure 1.6. For each of them, the rotation axes and/or mirror planes occur at the lattice points. However, there are other locations in the unit cell with comparable or lower degrees of symmetry with respect to rotation and reflection.

For non-Bravais lattice, we have to take into account the symmetry of the basis which is referred to as **point-group symmetry**. The point group symmetry includes all possible rotations, reflections and inversion, which leave the basis invariant. Point groups are denoted by a numerical and “ m ”. The numerical indicates how many positions within the basis are equivalent by rotation symmetry. A single m shows that the basis has mirror plane symmetry. (In two dimensions – it is mirror axis). E.g., $3m$ means that there are 3 equivalent sites within the unit cell and there is one mirror plane. In two dimensions there are 10 point groups. When we combine the rotation symmetry of the point group with the translational symmetries, we obtain a **space-group symmetry**.

In-text question

1. Q: Explain the following terms in crystallography:

- (i) Space-group symmetry (ii) Wigner-Seitz primitive cell (iii) Packing factor.

A: See pages 6-8 .

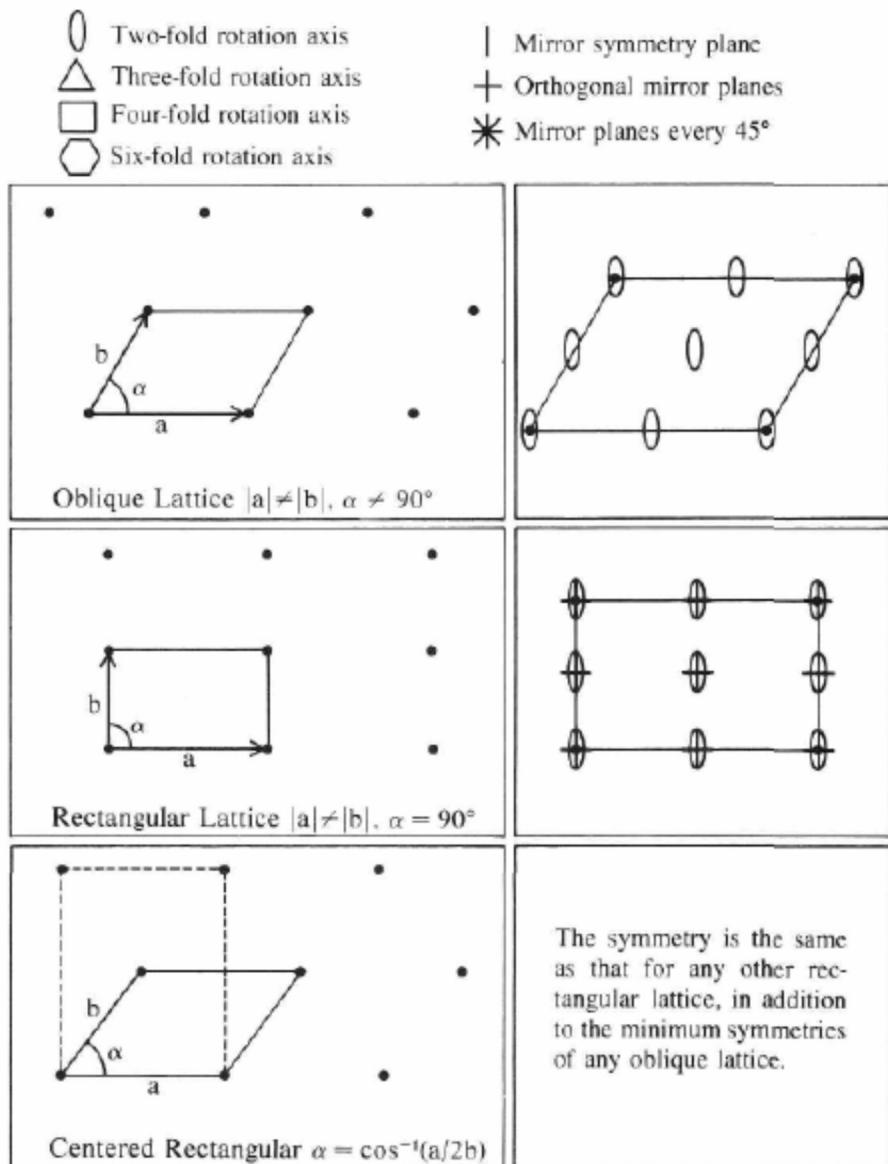


Figure 1.6a Bravais lattice types in 2-D

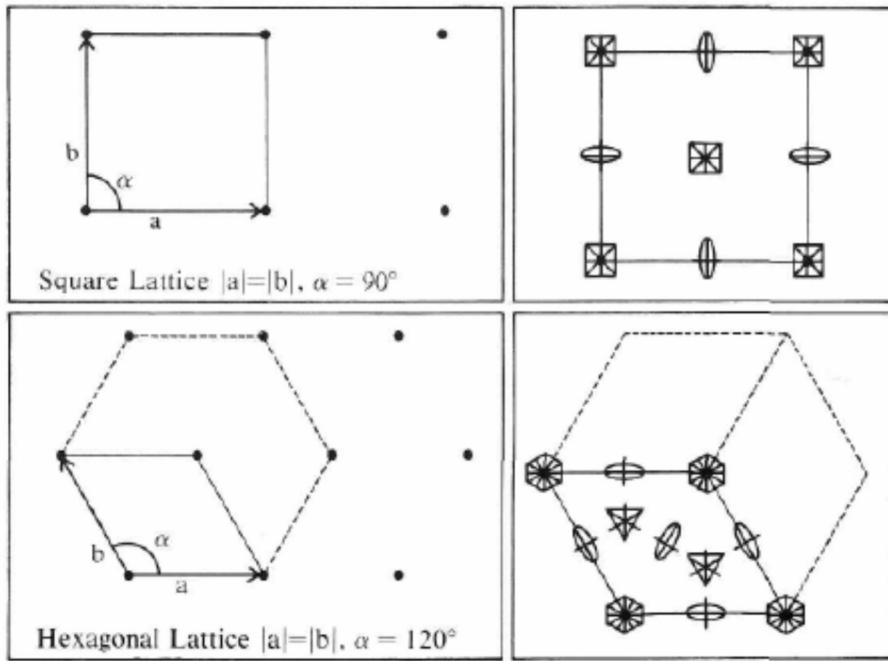


Figure 1.6b Bravais lattice types in 2-D

All the lattice properties we discussed for two dimensions can be extended to three dimensions. The lattice vectors are in this case,

$$\mathbf{T} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3 \quad (1.2)$$

where \mathbf{a}_1 , \mathbf{a}_2 and \mathbf{a}_3 are the primitive translation vectors, and (n_1, n_2, n_3) are a triplet of integers whose values depend on a particular lattice site. The unit cell in three dimensions is a parallelepiped; whose sides are the primitive translation vectors (see Figure 1.7). Here again the choice of the unit cell is not unique, although all primitive unit cells have equal volumes. The unit cell fills all space by the repetition of crystal translation operations. The volume of the unit cell represented by a parallelepiped with sides \mathbf{a}_1 , \mathbf{a}_2 and \mathbf{a}_3 is given by

$$V = |\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3| \quad (1.3)$$

Also, it is sometimes more convenient to deal with non-primitive or conventional cells, which have additional lattice sites either inside the cell or on its surface.

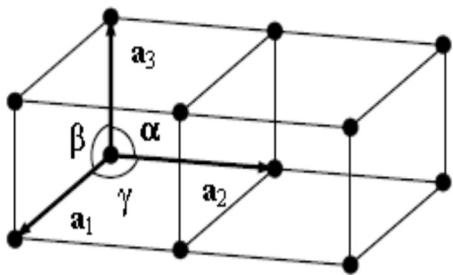


Figure 1.7 The unit cell in 3-D

In three dimensions there are 14 different Bravais crystal lattices which belong to 7 crystal systems. These systems are triclinic, monoclinic, orthorhombic, tetragonal, cubic, hexagonal and trigonal. The crystal lattices are shown in Fig.1.8. In all the cases, the unit cell represents a parallelepiped whose sides are \mathbf{a}_1 , \mathbf{a}_2 and \mathbf{a}_3 . The opposite angles are called α , β and γ . The relationship between the sides and the angles determines the crystal system. A simple lattice has sites only at the corners, a body-centered lattice has one additional point at the center of the cell, and a face-centered lattice has six additional points, one on each side. Note that in all the non-simple lattices the unit cells are non-primitive. The volume of the primitive unit cell is equal to the volume of the conventional unit cell divided by the number of sites.

Each of the 14 lattices has one or more types of symmetry properties with respect to reflection and rotation.

Reflection: The triclinic structure has no reflection plane; the monoclinic has one plane midway between and parallel to the basis plane, and so forth. The cubic cell has nine reflection planes: three parallel to the faces, and six others, each of which passes through two opposite edges.

Rotation: The triclinic structure has no axis of rotation (do not take into account 1-fold axis), the monoclinic has a 2-fold axis normal to the base. The cubic cell has three 4-fold axis normal to the faces and four 3-fold axis, each passing through two opposite corners.

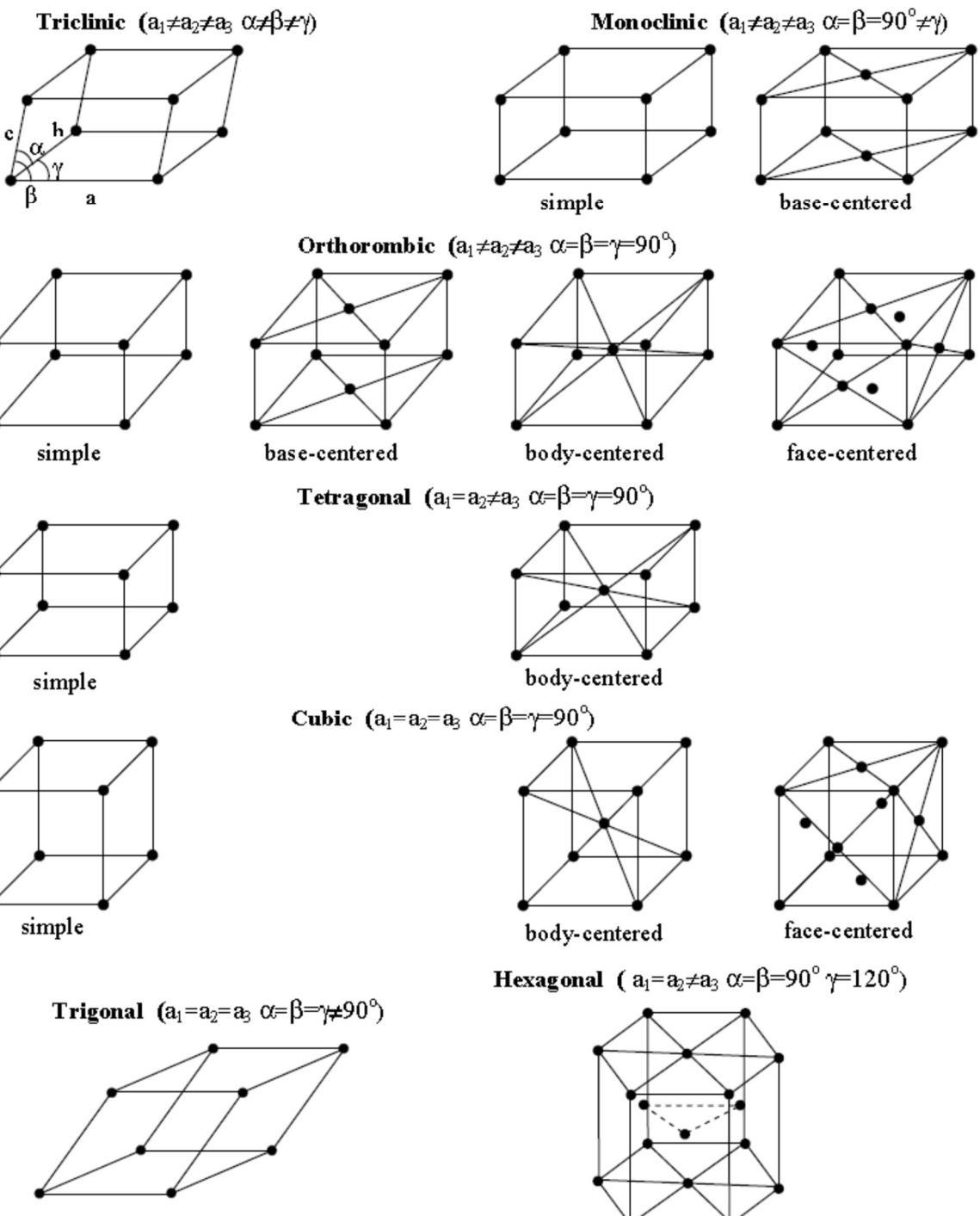
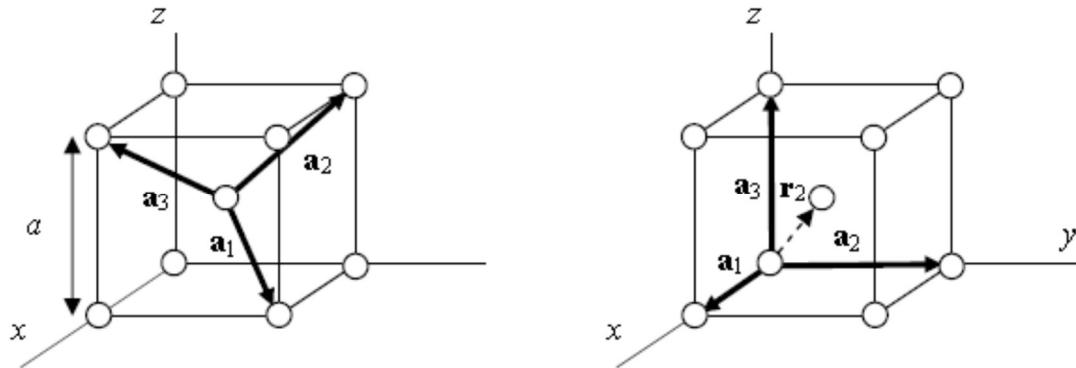


Figure 1.8

1.3 Most common crystal structures

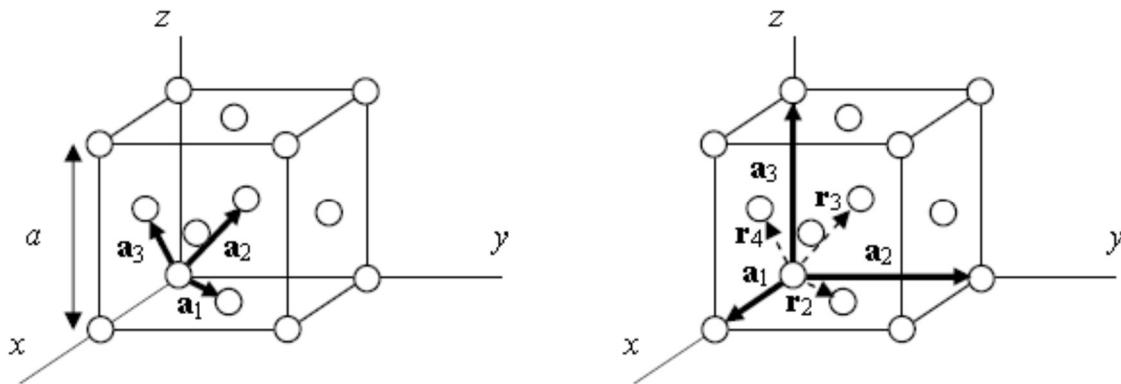
Body-centered cubic (bcc) lattice:



Figures showing the bcc lattice crystal structure

Primitive translation vectors of the bcc lattice (in units of lattice parameter a) are $\mathbf{a}_1 = \frac{1}{2}\frac{1}{2}-\frac{1}{2}$; $\mathbf{a}_2 = -\frac{1}{2}\frac{1}{2}\frac{1}{2}$; $\mathbf{a}_3 = \frac{1}{2}-\frac{1}{2}\frac{1}{2}$. The primitive cell is the rhombohedron. The packing ratio is 0.68, defined as the maximum volume which can be filled by touching hard spheres in atomic positions. Each atom has 8 nearest neighbors. The conventional unit cell is a cube based on vectors $\mathbf{a}_1 = 001$; $\mathbf{a}_2 = 010$; $\mathbf{a}_3 = 001$. It is twice as big compared to the primitive unit cell and has two atoms in it with coordinates $\mathbf{r}_1 = 000$ and $\mathbf{r}_2 = \frac{1}{2}\frac{1}{2}\frac{1}{2}$. The alkali metals (Na, Li, K, Rb, Cs), magnetic metals (Cr and Fe) and refractory metals (Nb, W, Mo, Ta) exhibit the bcc lattice structure.

Face-centered cubic (fcc) lattice:



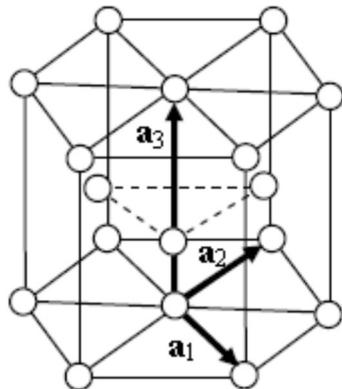
Figures showing the fcc lattice crystal structure

Primitive translation vectors of the bcc lattice (in units of lattice parameter a) are $\mathbf{a}_1 = \frac{1}{2}\frac{1}{2}0$; $\mathbf{a}_2 = 0\frac{1}{2}\frac{1}{2}$; $\mathbf{a}_3 = \frac{1}{2}0\frac{1}{2}$. The primitive cell is the rhombohedron. The packing ratio is 0.74. Each atom has 12 nearest neighbors.

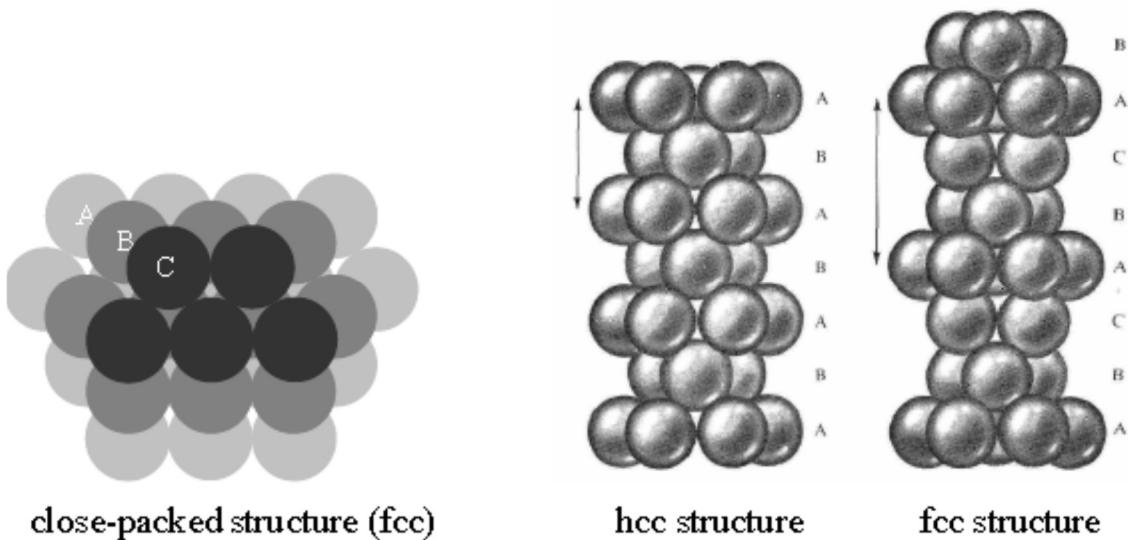
The conventional unit cell is a cube based on vectors $\mathbf{a}_1 = 001$; $\mathbf{a}_2 = 010$; $\mathbf{a}_3 = 001$. It is 4 times bigger than the primitive unit cell and has 4 atoms in it with coordinates $\mathbf{r}_1 = 000$; $\mathbf{r}_2 = \frac{1}{2}\frac{1}{2}0$; $\mathbf{r}_3 = 0\frac{1}{2}\frac{1}{2}$; $\mathbf{r}_4 = \frac{1}{2}0\frac{1}{2}$.

The fcc lattice has noble metals such as Cu, Ag, Au, common metals such as Al, Pb, Ni and inert gas solids such as Ne, Ar, Kr, Xe.

Hexagonal closed-packed (hcp) lattice:



The hcp structure has $a_1=a_2\neq a_3$, $\alpha=\beta=90^\circ$ and $\gamma=120^\circ$ with a basis of two atoms, one at 000 and the other at $\frac{2}{3}\frac{1}{3}\frac{1}{2}$. Along with the fcc structure, the hcp structure maximizes the packing ratio, making it 0.74.

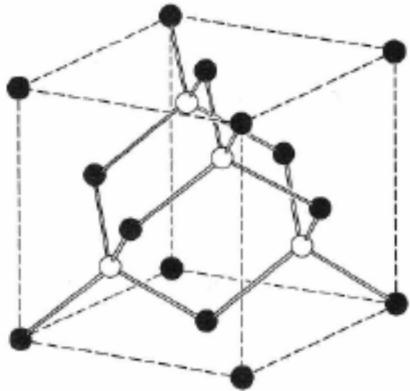


A closed-packed structure is created by placing a layer of spheres B on top of identical close-packed layer of spheres A. There are two choices for a third layer. It can go in over A or C. If it goes in over A the sequence is ABABAB... and the structure is hcp. If the third layer goes in over C the sequence is ABCABCABC... and the structure is fcc.

In perfect hcp structure the ratio of the height of the cell to the nearest neighbor spacing is $(8/3)^{1/2}$. In practice the (a_3/a_1) ratio is larger than 1.633 for most hexagonal crystals. Examples of nominally hcp crystals include the elements from Column II of the Periodic Table: Be, Mg, Zn, and Cd. Hcp is also the stable structure for several transition elements, such as Ti and Co.

Diamond structure is adopted by solids with four symmetrically placed covalent bonds. This is the situation in silicon, germanium, and grey tin, as well as in diamond. Diamond has the translational symmetry of fcc lattice with a basis of two atoms, one at 000 and the other at $\frac{1}{4}\frac{1}{4}\frac{1}{4}$.

Diamond structure represents two inter-penetrating fcc sublattices displaced from each other by one quarter of the cube diagonal distance.



In-text question

1. Q: Define the term Coordination number.

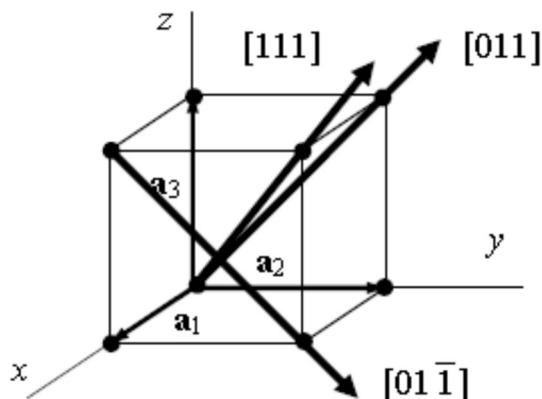
A: This is the number of nearest neighbours to a given atom in a crystal

2. Q: Describe briefly the seven systems of crystals.

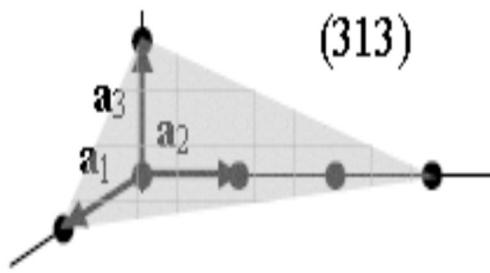
A: See pages 13-15.

1.3.1 Index system for crystal directions and planes

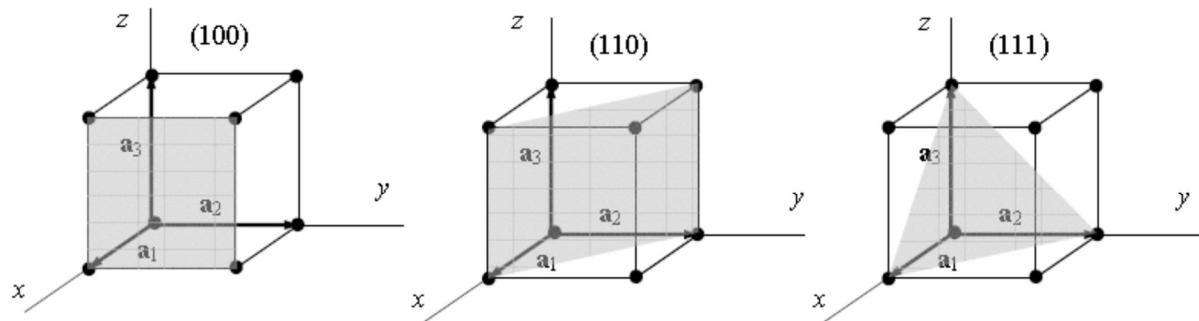
Crystal direction: Any lattice vector can be written as that given by Eq. (1.2). The direction is then specified by the three integers $[n_1 n_2 n_3]$. If the numbers $n_1 n_2 n_3$ have a common factor, this factor is removed. For example, [111] is used rather than [222], or [100], rather than [400]. When we speak about directions, we mean a whole set of parallel lines, which are equivalent due to transnational symmetry. Opposite orientation is denoted by the negative sign over a number. For example:



Crystal planes: The orientation of a plane in a lattice is specified by **Miller indices**. They are defined as follows. We find intercept of the plane with the axes along the primitive translation vectors \mathbf{a}_1 , \mathbf{a}_2 and \mathbf{a}_3 . Let these intercepts be x , y , and z , so that x is a fractional multiple of a_1 , y a fractional multiple of a_2 and z a fractional multiple of a_3 . Therefore we can measure x , y , and z in units a_1 , a_2 and a_3 respectively. We then have a triplet of integers $(x\ y\ z)$. Then we invert it $(1/x\ 1/y\ 1/z)$ and reduce this set to a similar one having the smallest integers by multiplying by a common factor. This set is called Miller indices of the plane (hkl) . For example, if the plane intercepts x , y , and z in points 3, 1, and 3, the index of this plane will be (313) .



The Miller indices specify not just one plane but an infinite set of equivalent planes. Note that for cubic crystals the direction $[hkl]$ is perpendicular to a plane (hkl) having the same indices, but this is not generally true for other crystal systems. Examples of the planes in a cubic system:



In-text question

1. Q: What are Miller indices? Draw neat diagrams to indicate Miller indices of the important plane systems in a simple cubic crystal.

A: Miller indices are used to specify the orientation of a plane in a lattice. See page 17.

Summary of Study Session 1

In this study session 1, you have learnt:

1. What is meant by a crystal structure, the definition of terms such as lattice points, space lattice, basis, unit cell, lattice parameters, and primitive cell.
2. That there are 32 classes of crystal systems based on the geometrical considerations, that it is common practice to divide all the crystal systems into 7 groups or basic systems (cubic, tetragonal, orthorhombic, monoclinic, triclinic, trigonal and hexagonal).
3. About crystal symmetries and the different types of symmetry operations that can be performed on crystals and the Wigner Seitz cell, and miller indices.

Self-Assessment Questions for Study Session 1

Now that you have completed this study session, you can assess how well you have achieved its learning outcomes by answering the following questions. Write your answers in your Study Diary and discuss them with your Tutor at the next Study Support Meeting.

SAQ 1.1

The packing ratio is defined as the fraction of the total volume of the cell that is filled by atoms. Determine the maximum values of this ratio for equal spheres located at the points of simple-cubic, body-centered cubic (bcc), face-centered cubic (fcc) and diamond-structure crystals.

SAQ 1.2

With reference to crystals, state what you understand by the terms:

- (i) Unit cell (ii) Single crystal (iii) coordination number (iv) packing factor

SAQ 1.3

What do you understand by Miller indices of crystal plane? Show that in a cubic crystal the spacing between consecutive parallel planes of Miller indices $(h k l)$ is given by:

$$d_{hkl} = \frac{a}{\sqrt{h^2+k^2+l^2}}$$

SAQ 1.4

Calculate the number of atoms per unit cell of a metal having a lattice parameter 0.29nm and density of 7870 kgm^{-3} . Take the atomic weight of the metal is 55.85.

References/Further Readings

Richard Turton (2000). The Physics of Solids: Oxford University Press.

Study Session 2 Crystal Binding

Introduction

Many solids are aggregates of atoms. The arrangement of atoms in any solid materials is determined by the character, strength and directionality of the binding forces, cohesive forces or chemical bonds. The bonds are made of attractive and repulsive forces that tend to hold the adjacent atoms or atomic units at a particular spacing such that the opposite forces just balance; and the process of holding them together is known as **bonding**. Since the particular type of bonding within a material plays a major role in determining the physical, chemical and electrical properties of materials, engineers must possess a sound working knowledge of the type of bonding that exist in materials.

Learning Outcomes for Study Session 2

When you have studied this session, you should be able to:

- 2.1 Explain interatomic forces.
- 2.2 Distinguish the different kinds of bonding between atoms of solids such as ionic, metallic, covalent and van-der waals forces.

2.1 Main Content

2.1.1 Interatomic forces

Solids are stable structures, and therefore there exist interactions holding atoms in a crystal together. For example a crystal of sodium chloride is more stable than a collection of free Na and Cl atoms. This implies that the Na and Cl atoms attract each other, i.e. there exist an attractive Interatomic force, which holds the atoms together. This also implies that the energy of the crystal is lower than the energy of the free atoms. The amount of energy which is required to pull the crystal apart into a set of free atoms is called the **cohesive energy** of the crystal.

$$\text{Cohesive energy} = \text{energy of free atoms} - \text{crystal energy}$$

The magnitude of the cohesive energy varies for different solids from 1 to 10 eV/atom, except inert gases in which the cohesive energy is of the order of 0.1eV/atom (see table 1 below). The cohesive energy controls the melting temperature (compare table 1 and table 2).

Table 1 Cohesive energies

	Energy required to form separated neutral atoms in their ground electronic state from the solid at 0 K at 1 atm. The data were supplied by Prof. Leo Brewer in units kcal per mole, revised to May 4, 1977, after LBL Report 3720 Rev.												
	B	C	N	O	F								Ne
Li	Be	320.											1.92
158.	3.32												0.020
1.63													
37.7	76.5												0.46
Na	Mg	145.											
107.	1.51												
1.113													
25.67	34.7												
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge
90.1	178.	376	468.	512.	395.	282.	413.	424.	428.	336.	130	271.	285.3
0.934	1.84	3.90	4.85	5.31	4.10	2.92	4.28	4.39	4.44	3.49	1.35	2.81	3.85
21.54	42.5	89.9	111.8	122.4	94.5	67.4	98.7	101.3	102.4	80.4	31.04	64.8	88.8
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn
82.2	166	422	603.	730.	658.	661.	650.	554.	376.	284.	112	243.	303.
0.852	1.72	4.37	6.25	7.57	6.82	6.85	6.74	5.75	3.89	2.95	1.16	2.52	3.14
19.64	39.7	100.8	144.2	174.5	157.2	158.	155.4	132.5	89.8	68.0	26.73	58.1	72.4
Cs	Ba	La	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb
77.6	183.	431.	621.	782.	859.	775.	788.	670.	564.	368.	65.	182.	196.
0.804	1.90	4.47	6.44	8.10	8.90	8.03	8.17	6.94	5.84	3.81	0.67	1.88	2.03
18.54	43.7	103.1	148.4	186.9	205.2	185.2	188.4	160.1	134.7	87.96	15.5	43.4	46.78
Fr	Ra	Ac	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er
	160.	410	4.25	417.	357.	328.	206	179.	400	391.	294.	302.	317.
	1.66		98.	4.32	3.70	3.40	2.14	1.86	414.	4.05	3.04	3.14	3.29
	38.2			99.7	85.3	78.5	49.3	42.8	95.5	93.4	70.2	72.3	75.8
	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No
	598.		536.	456	347.	264.	385.						
	6.20		5.55	4.73	3.60	2.73	3.99						
	142.9		128.	109.	83.0	63	92.1						
	Lr												

Table 2 Melting points, in K.
(After R. H. Lamoreaux, LBL Report 4995)

Li 453.7	Be 1562	B 2365	C C	N 63.15	O 54.36	F 53.48	Ne 24.56										
Na 371.0	Mg 922	Al 933.5	Si 1687	P w 317 r 863	S 388.4	Cl 172.2	Ar 83.81										
K 336.3	Ca 1113	Sc 1814	Ti 1946	V 2202	Cr 2133	Mn 1520	Fe 1811	Co 1770	Ni 1728	Cu 1358	Zn 692.7	Ga 302.9	Ge 1211	As 1089	Se 494	Br 265.9	Kr 115.8
Rb 312.6	Sr 1042	Y 1801	Zr 2128	Nb 2750	Mo 2895	Tc 2477	Ru 2527	Rh 2236	Pd 1827	Ag 1235	Cd 594.3	In 429.8	Sn 505.1	Sb 903.9	Te 722.7	I 386.7	Xe 161.4
Cs 301.6	Ba 1002	La 1194	Hf 2504	Ta 3293	W 3695	Re 3459	Os 3306	Ir 2720	Pt 2045	Au 1338	Hg 234.3	Tl 577	Pb 600.7	Bi 544.6	Po 527	At 1	Rn
Fr	Ra 973	Ac 1324	Ce 1072	Pr 1205	Nd 1290	Pm 1346	Eu 1091	Gd 1587	Tb 1632	Dy 1684	Ho 1745	Er 1797	Tm 1820	Yb 1098	Lu 1938		
	Th 2031		Pa 1848	U 1406	Np 910	Pu 913	Am 1449	Bk 1613	Cf 1562								Lw

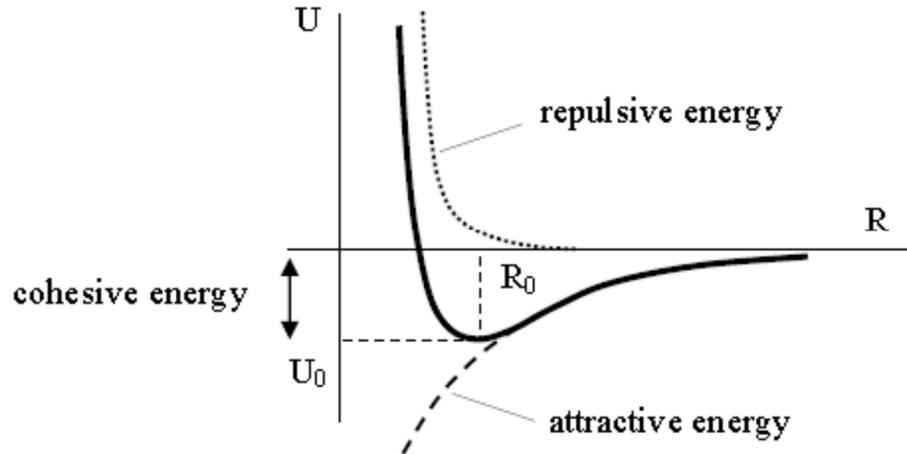


Figure 2.1 A typical binding energy curve

A typical curve for the potential energy (binding energy) representing the interaction between two atoms is shown in Figure 2.1. It has a minimum at some distance $R=R_0$. For $R>R_0$ the potential increases gradually, approaching 0 as $R\rightarrow\infty$, while for $R<R_0$ the potential increases very rapidly, tending to infinity at $R=0$. Since the system tends to have the lowest possible energy, it is most stable at $R=R_0$, which is *the equilibrium interatomic distance*. The corresponding energy U_0 is the cohesive energy. A typical value of the equilibrium distance is of the order of a few angstroms (e.g. 2-3Å), so that the forces under consideration are short range.

The interatomic force is determined by the gradient of the potential energy, so that

$$F(R) = -\frac{\partial U}{\partial R} \quad (2.1)$$

If we apply this to the curve in Fig.2.1, we see that $F(R)<0$ for $R>R_0$. This means that for large separations the force is *attractive*, tending to pull the atoms together. On the other hand $F(R)>0$ for $R<R_0$, i.e. the force becomes *repulsive* at small separations of the atoms, and tends to push the atoms apart. The repulsive and attractive forces cancel each other exactly at the point R_0 , which is the point of equilibrium.

The attractive interatomic forces reflect the presence of *bonds* between atoms in solids, which are responsible for the stability of the crystal. There are several types of *bonding*, depending on the physical origin and nature of the bonding force involved. The four main types are: *Van der Waals* (or *molecular*) bonding, *ionic* bonding, *covalent* bonding and *metallic* bonding.

Although the nature of the *attractive energy* is different in different solids, the origin of the *repulsive energy* is similar in all solids. The origin of the repulsive force is mainly due to the *Pauli exclusion principle*. The elementary statement of this principle is that two electrons cannot occupy the same orbital. As ions approach each other close enough, the orbits of the electrons begin to overlap, i.e. some electrons attempt to occupy orbits already occupied by others. This is, however, forbidden by the Pauli exclusion principle. As a result, electrons are excited to unoccupied higher energy states of the atoms. Thus, the electron overlap increases the total energy of the system and gives repulsive contribution to the interaction. The repulsive interaction is not easy to treat analytically from first principles. In order to make some quantitative estimates it is often assumed that this interaction can be described by a central field repulsive potential of the form $\lambda \exp(-r/\rho)$, where λ and ρ are some constants or of the form A/R^n , where n is sufficiently large and B is some constant.

In-text question

1. Q: Discuss the term “cohesive energy”.

A: The amount of energy which is required to pull the crystal apart into a set of free atoms is called the *cohesive energy* of the crystal.

2.1.2 Van der Waals (molecular) bonding:

First, we consider crystals of inert gases, which are characterized by *van der Waals (or molecular)* bonding. The electron distribution in such crystals is very close to that in free atoms. The noble gases such as neon (Ne), argon (Ar), krypton (Kr) and xenon (Xe) are characterized by filled electron shells and a spherical distribution of electronic clouds in the free atoms. In the crystal, the inert gas atoms are packed together within the cubic fcc structure.

What holds atoms in an inert gas crystal together? Consider two inert gas atoms (1 and 2) separated by distance R . The average charge distribution in a single atom is spherically symmetric, which implies that the average dipole moment of atom 1 is zero: $\langle \mathbf{d}_1 \rangle = 0$. Here, the brackets denote the time average of the dipole moment. However, at any moment of time, there may be a non-zero dipole moment caused by fluctuations of the electronic charge distribution. We denote this dipole moment by \mathbf{d}_1 . According to electrostatics, this dipole moment produces an

electric field, which induces a dipole moment on atom 2. This dipole moment is proportional to the electric field, which is in turn proportional to the d_1/R^3 so that

$$d_2 \sim E \sim \frac{d_1}{R^3} \quad (2.2)$$

The dipole moments of the two atoms interact with each other. The energy is therefore reduced due to this interaction. The energy of the interaction is proportional to the product of the dipole moments and inversely proportional to the cube of the distance between the atoms, so that

$$-\frac{d_1 d_2}{R^3} \sim -\frac{d_1^2}{R^6} \quad (2.3)$$

So, we see that the coupling between the two dipoles, one caused by a fluctuation, and the other induced by the electric field produced by the first one, results in the attractive force, which is called the *van der Waals force*. The time averaged potential is determined by the average value of $\langle d_1^2 \rangle$ which does not vanish, even though $\langle \mathbf{d}_1 \rangle$ is zero.

$$U \sim \frac{\langle d_1^2 \rangle}{R^6} = -\frac{A}{R^6} \quad (2.4)$$

The respective potential decreases as R^6 with the separation between the atoms.

Van der Waals bonding is relatively weak; the respective cohesive energy is of the order of 0.1 eV/atom.

This attractive interaction, described by Eq.(2.4), holds only for a relatively large separation between atoms. At small separations a very strong repulsive force caused by the overlap of the inner electronic shells starts to dominate. It appears that for inert gases this repulsive interaction can be fitted quite well by the potential of the form B/R^{12} , where B is a positive constant. Combining this with Eq. (2.4) we obtain the total potential energy of two atoms at separation R which can be represented as

$$U = 4\epsilon \left[\left(\frac{\sigma}{R}\right)^{12} - \left(\frac{\sigma}{R}\right)^6 \right], \quad (2.5)$$

where $4\epsilon\sigma^6 \equiv A$ and $4\epsilon\sigma^{12} \equiv B$. This potential is known as the Lennard-Jones potential.

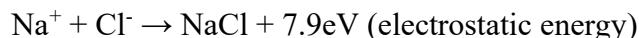
In-text question

1. Q: Briefly describe the Van der Waals forces of interaction in atoms.

A: See pages 22-24.

2.1.3 Ionic bonding

The ionic bond results from the electrostatic interaction of oppositely charged ions. Let us take sodium chloride as an example. In the crystalline state, each Na atom loses its single valence electron to a neighboring Cl atom, producing Na^+ and Cl^- ions which have filled electronic shells. As a result an ionic crystal is formed containing positive and negative ions coupled by a strong electrostatic interaction.



The cohesive energy with respect to neutral atoms can be calculated as $7.9\text{eV} - 5.1\text{eV} + 3.6\text{eV}$, i.e.



The structure of NaCl is two interpenetrating fcc lattices of Na^+ and Cl^- ions as shown in Figure 2.2

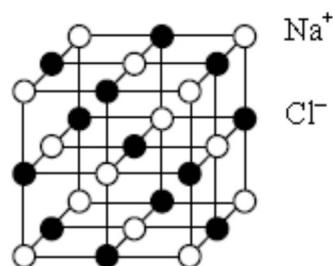


Figure 2.2 Structure of the NaCl crystal

Thus, each Na^+ ion is surrounded by 6 Cl^- ions and vice versa. This structure suggests that there is a strong attractive Coulomb interaction between nearest-neighbors ions, which is responsible for the ionic bonding.

To calculate binding energy we need to include Coulomb interactions with all atoms in the solid. Also, we need to take into account the repulsive energy, which we assume to be exponential. Thus, the interaction between two atoms i and j in a lattice is given by

$$U_{ij} = \lambda e^{-r_{ij}/\rho} \pm q^2/r_{ij} \quad (2.6)$$

Here r_{ij} is the distance between the two atoms, q is the electric charge on the atom, the (+) sign is taken for like charges and the (−) sign for unlike charges.

The total energy of the crystal is the sum over i and j so that

$$U = \frac{1}{2} \sum_{i,j} U_{ij} = N \sum_j (\lambda e^{-r_{ij}/\rho} \pm q^2/r_{ij}) \quad (2.7)$$

In this formula, $\frac{1}{2}$ is due to the fact that each pair of interactions should be counted only once. The second equality results from the fact in the NaCl structure the sum over j does not depend on whether the reference ion i is positive or negative, which gives the total number of atoms. The latter divided by two gives the number of molecules N , composed of a positive and a negative ion. We assume, for simplicity, that the repulsive interaction is non-zero only for the nearest neighbors (because it drops down very quickly with the distance between atoms). In this case we obtain

$$U = N(z\lambda e^{-R/\rho} - \alpha q^2/R) \quad (2.8)$$

Here, R is the distance between the nearest neighbours; z is the number of the nearest neighbors, and α is the *Madelung constant*:

$$\alpha = \sum_{j \neq 1} \frac{(\pm 1)}{P_{ij}}, \quad (2.9)$$

where P_{ij} is defined by $r_{ij} \equiv P_{ij}R$. The value of the Madelung constant plays an important role in the theory of ionic crystals. In general it is not possible to compute the Madelung constant analytically. A powerful method for calculation of lattice sums was developed by Ewald, which is called *Ewald summation*.

Example:

A one-dimensional lattice of ions of alternating sign as shown in Figure 2.3.

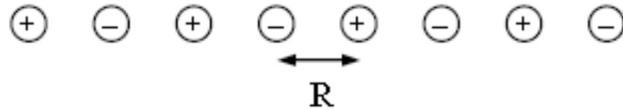


Figure 2.3 One-dimensional lattice of ions

In this case

$$\alpha = 2 \left[1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} + \dots \right] = 2 \ln 2 , \quad (2.10)$$

where we took into account the logarithm expansion into series

$$\ln(1+x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n} .$$

In three dimensions, calculation of the series is much more difficult and cannot be performed so easily. The values of the Madelung constants for various solids are calculated, tabulated and can be found in various solid-state books.

Now we calculate the equilibrium distance between the nearest neighbors for the NaCl type lattice using Eq.(3.8). At the equilibrium, the derivative $dU/dR=0$, so that

$$-\frac{z\lambda}{\rho} e^{-R_0/\rho} + \frac{\alpha q^2}{R_0^2} = 0 \quad (2.11)$$

and therefore

$$R_0^2 e^{-R_0/\rho} = \frac{\alpha \rho q^2}{z\lambda} \quad (2.12)$$

This relationship determines the equilibrium separation R_0 in terms of the parameters ρ and λ of the repulsive potential. Using Eq. (2.8) and Eq. (2.12) the cohesive energy per atom of the ionic solid can be written as:

$$U_0 = \frac{\alpha N \rho q^2}{R_0^2} - \frac{\alpha N q^2}{R_0} = -\frac{\alpha N}{R_0}^2 \left(1 - \frac{\rho}{R_0}\right) \quad (2.13)$$

Let us estimate the magnitude of the cohesive energy in NaCl. The Madelung constant is $\alpha=1.75$. The interatomic distance is $R_0=a/2\approx2.8\text{\AA}$. The charge $q=e$. The repulsive interaction of atoms has a very short range of the order of $\rho=0.1R_0$. As follows from Eq.(2.13)

$$\frac{U_0}{N} \approx -\frac{\alpha}{(R_0/a_0)} \frac{e^2}{a_0} \left(1 - \frac{0.1R_0}{R_0}\right) \approx -\frac{1.8}{6} 27 \cdot 0.9eV \approx -8eV \quad (2.14)$$

We see that the typical value of the binding energy per pair of atoms is about 8eV. This implies that ionic bond is very strong. Experimentally, this strength is characterized by the relatively high melting temperatures. For example, the melting temperature of NaCl is about 1100K, while the melting temperature for the Na metal is about 400K.

2.1.4 Covalent bonding

The covalent bond is another important type of bond which exists in many solids. The covalent bond between two atoms is usually formed by two electrons, one from each atom participating in the bond. The electrons forming the bond tend to be partly localized in the region between the two atoms joined by the bond. Normally the covalent bond is strong: for example, it is the bond which couples carbon atoms in diamond. The covalent bond is also responsible for the binding of silicon and germanium crystals.

We illustrate the appearance of the covalent bond by considering two atoms (e.g., hydrogen atoms), which are described by orbitals ψ_1 and ψ_2 . The molecular orbital of the two atoms is a linear combination of the two orbitals. There are only two possibilities,

$$\Psi_b = \Psi_1 + \Psi_2 \quad (2.15)$$

Or

$$\Psi_a = \Psi_1 - \Psi_2, \quad (2.16)$$

because symmetry considerations preclude any other linear combinations, since the distribution of electron charge must be symmetric with respect to the two atoms.

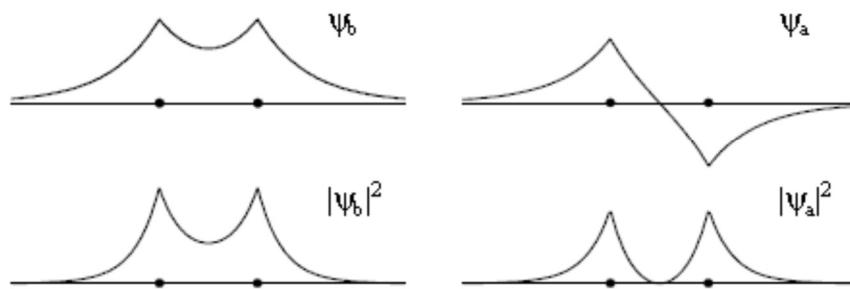


Figure 2.4 Molecular orbital and charge distribution in covalent bonds

The molecular orbitals are sketched in Figure 2.4. This figure also shows the charge distribution given by $|\psi_b|^2$ and $|\psi_a|^2$. It can be seen that there is a sizable contribution to the charge density in the region between the nuclei for the symmetric orbital, while there is a zero density between the nuclei for the antisymmetric orbital. The two orbitals have different energies as is illustrated in Figure 2.5, which shows the energy as a function of the interatomic distance.

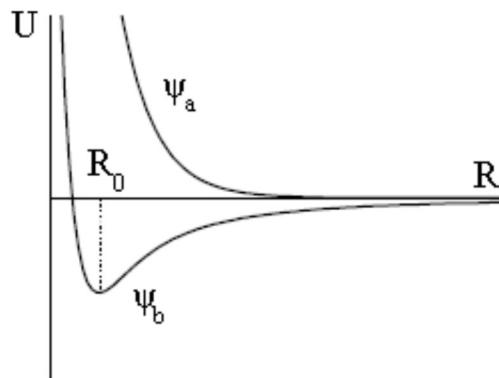


Figure 2.5 Energy as a function of Interatomic distance in the two orbitals

We see that the symmetric orbital has a minimum of energy at certain distance and has a lower energy than antisymmetric orbital. Thus, this is a *bonding* orbital which leads to a stable state of the molecule. The other orbital is called *anti-bonding* orbital, which has a minimum of energy at infinite separation of the atoms. This is a simple example of the covalent bonding between two atoms.

Note that spins of the two electrons which participate in bonding are anti-parallel. This is a consequence of the Pauli exclusion principle which requires the total wave function of the system of electrons to be antisymmetric with respect to any interchange of the coordinates of two electrons. In the case of the bonding state the orbital wave function is symmetric and therefore the spin contribution has to be asymmetric which means that the spins are anti-parallel. On the other hand, the spins are parallel for the anti-bonding orbital. We see that The Pauli principle modifies the distribution of charge depending on the spin orientation of electrons. This spin-dependent contribution to the Coulomb energy is called *exchange interaction*. The simplest example of the covalent bond is a hydrogen molecule.

The covalent bond in solids has strong directional properties. For example, carbon has four valence electrons $1s^22s^22p^2$ and forms tetrahedral bonds with nearest neighbors, resulting in the diamond type structure. The carbon atom is positioned in the center of the tetrahedron, the neighboring carbon atoms being at the vertices of the tetrahedron (Figure 2.6). Since there are four bonds joining the central atom to its neighbors, each C atom surrounds itself with eight valence electrons, which is a stable structure because the second shell is now completely full. Such tetrahedral coordination also occurs for the Si and Ge – those elements which can be found in the fourth column of the periodic table.

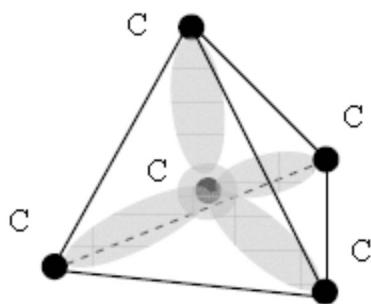


Figure 2.6 Diamond Structure

To explain the tetrahedron arrangement in diamond, we note that each C atom has four electrons in the second shell: two 2s electrons and two 2p electrons ($2s^22p^2$). The s states are spherically symmetric, whereas the p states represent charge distributions lying along x, y, and z coordinates. The energy difference between these states is not very big. It appears that it is energetically

favorable to excite one of the s electrons to p states so that the electronic configuration becomes $2s2p^3$. We can now construct the linear combinations of atomic orbitals:

$$\begin{aligned}\Psi_1 &= 1/2(s + p_x + p_y + p_z) \\ \Psi_2 &= 1/2(s + p_x + p_y - p_z) \\ \Psi_3 &= 1/2(s + p_x - p_y - p_z) \\ \Psi_4 &= 1/2(s - p_x - p_y - p_z)\end{aligned}\tag{2.17}$$

The densities corresponding to these orbitals are oriented along the tetrahedral directions.

These orbitals are therefore a better representation of the electron states i.e s, p_x, p_y, p_z orbitals. The mixing of the s and p states in Eq. (2.17) is referred to as the sp-hybridization. The particular type of hybridization in diamond is known as sp^3 hybridization. The sp^3 hybridization occurs also in Si and Ge. In Si one 3s and three 3p states hybridize to form tetrahedral bonds. In Ge the sp^3 hybridization involves one 4s and three 4p electrons.

Concluding the discussion about ionic and covalent bonds, we note that there is a continuous range of crystals between the ionic and covalent limits. In many cases, it is important to estimate the extent a given bond is ionic or covalent. There are modern theoretical approaches that allow us to quantify the degree of ionicity and covalence in many solids.

In-text question

1. Q: Describe ionic, covalent and Van der Waals bondings with examples.
A: See pages 24-27
2. Q: Calculate the binding (lattice energy) of the NaI for which the nearest neighbor distance is 0.324 nm. Express the energy in eV/molecule and also in KJ per kmol. Madelung constant for NaI= 1.748 and n=9.5.

Solution

From the equation $U_0 = -\left[\frac{Ae^2N_A}{4\pi\varepsilon_0 r_0}\right] \left[\frac{n-1}{n}\right]$ and substituting the values given in the question above (i.e. A=1.748, n=9.5, $r_0=0.324\text{nm}$ and $N_A=\text{Avogadro's constant}$) gives the answer below:

A: 6.9eV or $667 \times 10^3 \text{ kJ/kmol}$.

2.1.5 Metallic bonding

Metals are characterized by a high electrical conductivity, which implies that a large number of electrons in a metal are free to move. The electrons capable to move throughout the crystal are called the *conduction electrons*. Normally the valence electrons in atoms become the conduction electrons in solids. The main feature of the metallic bond is the lowering of the energy of the valence electrons in metal as compared to the free atoms. Below, some qualitative arguments are given to explain this fact.

According to the Heisenberg uncertainty principle, the indefiniteness in coordinate and in the momentum are related to each other such that $\Delta x \Delta p \sim \hbar$. In a free atom the valence electrons are restricted by a relatively small volume. Therefore, Δp is relatively large, which makes the kinetic energy of the valence electrons in a free atom large. On the other hand in the crystalline state the electrons are free to move throughout the whole crystal, the volume of which is large. Therefore, the kinetic energy of the electrons is greatly reduced, which leads to diminishing the total energy of the system in the solid. This mechanism is the source of the metallic bonding. Figuratively, the negatively charged free electrons in a metal serve as glue that holds positively charged ions together.

The metallic bond is somewhat weaker than the ionic and covalent bond. For instance the melting temperature of metallic sodium is about 400K which is smaller than 1100K in NaCl and about 4000K in diamond. Nevertheless, this type of bond should be regarded as strong.

In transition metals like Fe, Ni, Ti, and Co, the mechanism of metallic bonding is more complex. This is due to the fact that in addition to s electrons which behave like free electrons, there are 3d electrons which are more localized. Hence, the d electrons tend to create covalent bonds with nearest neighbors. The d electrons are normally strongly hybridized with s electrons, making the picture of bonding much more complicated.

In-text question

1. Q: What are the distinguishing characteristics of metallic bonding?

A: See pages 31-32.

Summary of Study Session 2

In this study session, you have learnt that:

1. Bonding between atoms is as a result of interatomic forces or bonds.
2. There are three strong principal types of primary bonds: Ionic, covalent and metallic. These bonds are distinguished on the basis of the positions assumed by the bond electrons during the formation of the bond. Van der Waals and hydrogen bonds are typical examples of secondary bonds and they result from intermolecular attraction.
3. Generally, the stronger the bond, the higher the melting and boiling points.

Self-Assessment Questions for Study Session 2

Now that you have completed this study session, you can assess how well you have achieved its learning outcomes by answering the following questions. Write your answers in your Study Diary and discuss them with your Tutor at the next Study Support Meeting.

SAQ 2.1

How is a covalent bond formed in a hydrogen molecule?

SAQ 2.2

Give two examples for: (i) ionic solid (ii) covalent solid.

SAQ 2.3

Describe the salient features of covalent and metallic bonded crystals

SAQ 2.4

If the ionic radius of Na decreases by 0.88 and that of Cl increases by 0.89, calculate the binding energy of NaCl. Madelung constant for NaCl is 1.75 and n=9 for ionic crystals. Express your result in kJ/kmol [$r_{\text{Cl}}=0.0905 \text{ nm}$ and $r_{\text{Na}}=0.186 \text{ nm}$].

SAQ 2.5

The ionic radii of Cs and Cl are 0.165 nm and 0.181 nm respectively, and their atomic weights respectively 133 and 35.5. Calculate the density of CsCl.

References/Further Readings

Richard Turton, (2000). The Physics of Solids: Oxford University Press.

Study Session 3 X-ray Diffraction in Crystals, Applications.

Introduction

X-rays are electromagnetic waves like ordinary light; therefore, they should exhibit interference and diffraction. The wavelength of X-rays is of the order of 0.1 nm, so that ordinary devices such as ruled diffraction gratings do not produce observable effects with X-rays. In 1912, German physicist Laue suggested that a crystal which consisted of a three-dimensional array of regularly spaced atoms could serve the purpose of grating. The crystal differs from the ordinary grating in the sense that the diffracting centres in the crystal are not in one plane. Hence the crystal acts as a space grating rather than a plane grating.

On the suggestions of Laue, his associates, Friedrich and Knipping succeeded in diffracting X-rays by passing them through a thin crystal of zinc blende. The diffraction pattern obtained consists of a central spot and a series of spots arranged in a definite pattern around the central spot. This symmetrical pattern of spots is known as Laue pattern, and it proves that X-rays are electromagnetic radiation. A simple interpretation of the diffraction pattern was given by W.L. Bragg (1912). According to him, the spots are produced due to the reflection of some of the incident X-rays from the various sets of parallel crystal planes (called Bragg's planes) which contain a large number of atoms.

Learning Outcomes for Study Session 2

When you have studied this session, you should be able to:

- 3.1 Explain the genesis of X-ray diffraction
- 3.2 Describe Brillouin zones and reciprocal lattices
- 3.3 Describe X-ray diffraction patterns and uses of x-ray diffraction.

3.1. Main Content

3.1.1 Bragg Law

Most methods for determining the atomic structure of crystals are based on the idea of scattering of radiation. X-rays is one of the types of radiation which can be used for such a study. The

wavelength of the radiation should have a wavelength comparable to a typical interatomic distance which is in solids of a few angstroms (10^{-8} cm). The x-ray wavelength λ can be estimated as follows

$$E = h\nu = \frac{hc}{\lambda} \implies \lambda(\text{\AA}) = \frac{12.4}{E(\text{KeV})} \quad (3.1)$$

Therefore, x-rays of energy 2-10 keV are suitable for studying the crystal structure.

X-rays interact with electronic shells of atoms in a solid. Electrons absorb and re-radiate x-rays which can then be detected. Nuclei are too heavy to respond. The reflectivity of x-rays is of the order of 10^{-3} - 10^{-5} , so that the penetration in the solid is deep. Therefore, x-rays serve as a *bulk* probe.

In 1913 Bragg found that crystalline solids have remarkably characteristic patterns of reflected x-ray radiation. In crystalline materials, for certain wavelengths and incident directions, intense peaks of scattered radiation were observed. Bragg accounted for this by regarding a crystal as made out of parallel planes of atoms, spaced by distance d apart. The conditions for a sharp peak in the intensity of the scattered radiation were that:

- (1) the x-rays should be specularly reflected by the atoms in one plane;
- (2) the reflected rays from the successive planes interfere constructively.

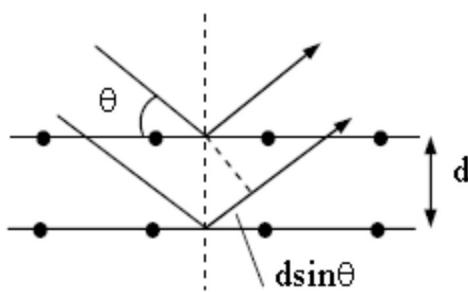


Figure 3.1 X-ray reflected from planes

Figure 3.1 shows x-rays which are specularly reflected from adjacent planes. The path difference between the two x-rays is equal to $2ds\sin\theta$. For the x-rays to interfere constructively this difference must be an integral number of wavelengths. This leads to the Bragg condition:

$$2ds\sin\theta = m\lambda \quad (3.2)$$

The integer m is known as the *order* of the corresponding reflection (or order of interference).

There are a number of various setups for studying crystal structure using x-ray diffraction. In most cases, the wavelength of radiation is fixed, and the angle is varied to observe diffraction peaks corresponding to reflections from different crystallographic planes. Using the Bragg law one can then determine the distance between the planes.

The Bragg law is greatly oversimplified (but it works). It

- (i) says nothing about intensity and width of x-ray diffraction peaks;
- (ii) neglects differences in scattering from different atoms;
- (iii) neglects the distribution of charge around atoms.

In-text question

1. Q: Explain Bragg's law for x-ray diffraction in crystals.

A: See pages 35-36

3.1.2. Diffraction condition and reciprocal lattice

Later, Von Laue (1914) introduced a different approach for x-ray diffraction. He regarded a crystal as composed of identical atoms placed at the lattice sites \mathbf{T} and assumed that each atom could reradiate the incident radiation in all directions. Sharp peaks were observed only in the directions and at wavelengths for which the x-rays scattered from all lattice points interfered constructively.

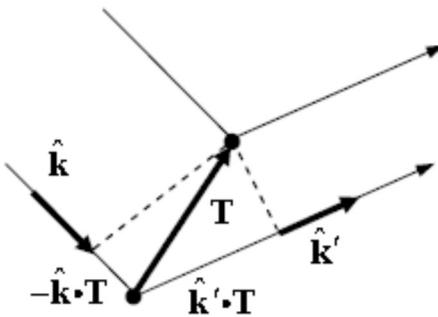


Figure 3.2 Constructive interference from reflected x-rays

To find the condition of constructive interference we consider two scatterers (Figure 3.2) separated by a lattice vector \mathbf{T} . Let x-rays be incident from infinity, along direction \mathbf{k} with wavelength λ and wave vector $\mathbf{k}=2\pi\mathbf{k}/\lambda$. We assume that the scattering is *elastic*, i.e. the x-rays are scattered in \mathbf{k}' direction with *same* wavelength λ , so that the wave vector ($k'=2\pi k'/\lambda$). The path difference between the x-ray scattered from the two atoms should be an integral number of wavelengths. Therefore, as is seen from Fig.3.2, the condition for constructive interference is

$$(k' - k) \cdot T = m\lambda, \quad (3.3)$$

where m is an integer. Multiplying both sides of Eq.(3.3) by $2\pi/\lambda$ leads to a condition on the incident and scattered wave vectors:

$$(k' - k) \cdot T = 2\pi m. \quad (3.4)$$

Defining the *scattering wave vector* $\Delta\mathbf{k} = \mathbf{k}' - \mathbf{k}$, the diffraction condition can be written as

$$\Delta\mathbf{k} = \mathbf{G}, \quad (3.5)$$

where \mathbf{G} is, by definition, such a vector for which

$$\mathbf{G} \cdot \mathbf{T} = 2\pi m. \quad (3.6)$$

A set of vectors \mathbf{G} which satisfies this condition form a *reciprocal lattice*. Vectors \mathbf{G} are called *reciprocal lattice vectors*.

A reciprocal lattice is defined with reference to a particular Bravais lattice which is determined by a set of lattice vectors \mathbf{T} . The Bravais lattice that determines a particular reciprocal lattice is referred as the *direct lattice*, when viewed in relation to its reciprocal.

There is an algorithm for constricting the reciprocal lattice from the direct lattice. Let \mathbf{a}_1 , \mathbf{a}_2 , and \mathbf{a}_3 be a set of primitive vectors of the direct lattice. Then the reciprocal lattice can be generated using primitive vectors

$$\mathbf{b}_1 = \frac{2\pi}{V} \mathbf{a}_2 \times \mathbf{a}_3, \quad \mathbf{b}_2 = \frac{2\pi}{V} \mathbf{a}_3 \times \mathbf{a}_1, \quad \mathbf{b}_3 = \frac{2\pi}{V} \mathbf{a}_1 \times \mathbf{a}_2, \quad (3.7)$$

where $V = \mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)$ is the volume of the unit cell, so that

$$\mathbf{G} = m_1 \mathbf{b}_1 + m_2 \mathbf{b}_2 + m_3 \mathbf{b}_3 \quad (3.8)$$

In order to prove that the vectors built in this way satisfy condition (3.6), we first note that the \mathbf{b}_i satisfy the condition

$$\mathbf{b}_i \cdot \mathbf{a}_j = 2\pi \delta_{ij} \quad (i,j=1,2,3) \quad (3.9)$$

where δ_{ij} is the Kronecker symbol defined by

$$\begin{aligned} \delta_{ij} &= 0, \text{ if } i \neq j \\ \delta_{ij} &= 1, \text{ if } i = j. \end{aligned} \quad (3.10)$$

Indeed, since $\mathbf{T} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3$ we obtain

$$\mathbf{G} \cdot \mathbf{T} = 2\pi(n_1 m_1 + n_2 m_2 + n_3 m_3), \quad (3.11)$$

which proves Eq. (3.6).

Examples: reciprocal lattices for 1D and 2D-rectangular structures. Note: Eqs.(3.9) rather than Eqs.(3.7) should be used in 1D and 2D cases.

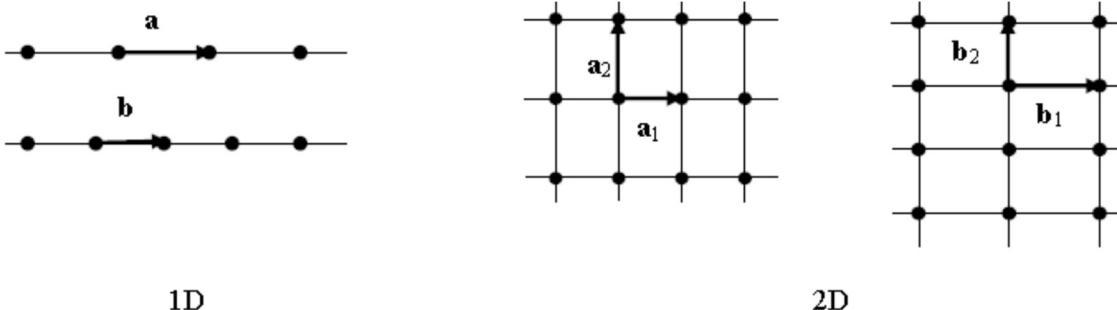


Figure 3.3 Reciprocal lattices for 1D and 2D

Reciprocal lattice types for some 3D lattices:

Direct lattice	Reciprocal lattice
sc	sc
bcc	fcc
fcc	bcc
hcp	hcp

Coming back to the diffraction condition (3.5), we can say that constructive interference occurs provided that the scattering wave vector is a vector of the reciprocal lattice.

It is sometimes more convenient to give a different formulation of the diffraction condition. In elastic scattering the photon energy is conserved, so that the magnitudes of k and k' are equal, and therefore $k^2 = k'^2$. Therefore, it follows from Eq.(3.5) that

$$k'^2 = (\mathbf{G} + \mathbf{k})^2 \Rightarrow 0 = \mathbf{G}^2 + 2\mathbf{k} \cdot \mathbf{G}. \quad (3.12)$$

By replacing \mathbf{G} with $-\mathbf{G}$, which is also a reciprocal lattice vector, we arrive at

$$2\mathbf{k} \cdot \mathbf{G} = G^2, \quad (3.13)$$

that can be considered as the diffraction condition.

Equation (3.13) is another statement of the Bragg law (3.1). We prove this in three steps.

(1) We show that the reciprocal lattice vector $\mathbf{G} = h\mathbf{b}_1 + k\mathbf{b}_2 + l\mathbf{b}_3$ is orthogonal to the plane represented by Miller indices (hkl) .

Consider the plane (hkl) which intercepts axes at points x, y , and z given in units a_1, a_2 and a_3 :

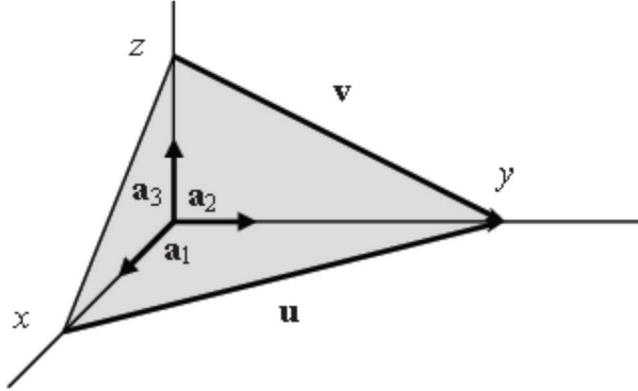


Figure 3.4

By the definition of the Miller indices we can always find such interceptions that

$$(h, k, l) = \left(\frac{1}{x}, \frac{1}{y}, \frac{1}{z} \right). \quad (3.14)$$

As we know, any plane can be defined by two non-collinear vectors lying within this plane. We can choose vectors \mathbf{u} and \mathbf{v} shown in Fig.3.4. They are given by $\mathbf{u} = ya_2 - xa_1$ and $\mathbf{v} = ya_2 - za_3$. To prove that the reciprocal vector \mathbf{G} is normal to the plane (hkl) , it is sufficient to prove that this vector is orthogonal to \mathbf{u} and \mathbf{v} , i.e. $\mathbf{u} \cdot \mathbf{G} = 0$ and $\mathbf{v} \cdot \mathbf{G} = 0$. We have

$$\mathbf{u} \cdot \mathbf{G} = (ya_2 - xa_1) \cdot (h\mathbf{b}_1 + k\mathbf{b}_2 + l\mathbf{b}_3) = 2\pi(yk - xh) = 0, \quad (3.15)$$

where the second equation follows from the orthogonality condition of the vectors of the direct and reciprocal lattices (3.9) and the last equation follows from Eq.(3.14). In the same manner we can show that \mathbf{G} is orthogonal to \mathbf{v} . We have proved, therefore, that vector \mathbf{G} is orthogonal to the plane (hkl) .

(2) Now we prove that the distance between two adjacent parallel planes of the direct lattice is $d=2\pi/G$.

First, we note that the nearest plane which is parallel to the plane (hkl) goes through the origin of the Cartesian coordinates in Fig.3.4. Therefore, the interplanar distance is given by the projection of one of the vectors $x\mathbf{a}_1$, $y\mathbf{a}_2$, $z\mathbf{a}_3$, to the direction normal to the (hkl) plane. This direction is given by the unit vector \mathbf{G}/G , since we have already established that \mathbf{G} is normal to the plane. Therefore

$$d = x\mathbf{a}_1 \cdot \mathbf{G}/G = 2\pi x h/G = 2\pi/G. \quad (3.16)$$

The connection between reciprocal vectors and crystal planes is now clear. The reciprocal vector $\mathbf{G}(hkl)$ is associated with the crystal planes (hkl) and is normal to these planes. The separation between these planes is 2π times the inverse of G .

(3) Now we are ready to show that the diffraction condition (3.13) is equivalent to the Bragg law (3.2). It follows from Eqs. (3.13) and (3.16) that

$$2(2\pi/\lambda) \sin \theta = 2\pi/d, \quad (3.17)$$

or $2d \sin \theta = m\lambda$, where θ is the angle between the incident beam and the crystal plane. The integers hkl that define \mathbf{G} are not necessarily identical with the indices of the actual plane, because hkl may contain a common factor m , whereas in the definition of the Miller indices the common factor has been eliminated. Therefore, we can substitute $m\mathbf{G}$ for \mathbf{G} and obtain the Bragg result

$$2d \sin \theta = m\lambda. \quad (3.18)$$

Brillouin zones. Brillouin gave another statement of the diffraction condition.

Consider a two-dimensional lattice in the reciprocal space (Fig.3.5a). Let O be the origin of this lattice. Consider a reciprocal lattice vector, which connects points O and another reciprocal lattice site. Now draw the line (in three dimensions it would be a plane), which is orthogonal to this vector and intercepts it in the midpoint. The x-ray will be diffracted if its wavevector \mathbf{k} has the

magnitude and direction that is required by the condition (3.13), which can be rewritten in the following way

$$\mathbf{k} \cdot \left(\frac{1}{2}\mathbf{G}\right) = \left(\frac{1}{2}\mathbf{G}\right)^2 , \quad (3.19)$$

It is easy to see that any \mathbf{k} vector connecting the origin and the plane will satisfy the diffraction condition.

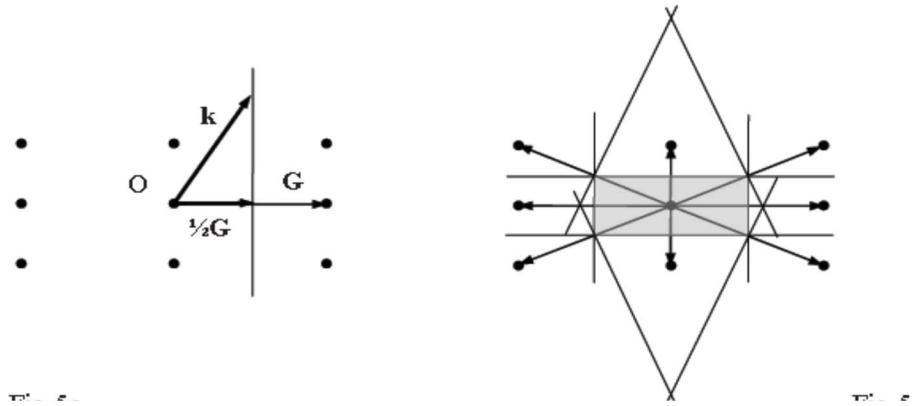


Figure 3.5a&b Two dimensional lattices in reciprocal lattice

In a similar way we can draw other lines (planes), which satisfy the diffraction condition. This is shown in Fig.3.5b. So the Brillouin construction exhibits all the wave vectors \mathbf{k} which can be Bragg-reflected by the crystal.

The first Brillouin zone is the smallest volume entirely enclosed by the planes that are perpendicular bisectors of the reciprocal lattice vectors. In this construction it is the rectangle about the origin. The first Brillouin zone is the Wigner-Seitz primitive cell in the reciprocal lattice.

Two examples of the first Brillouin zone :

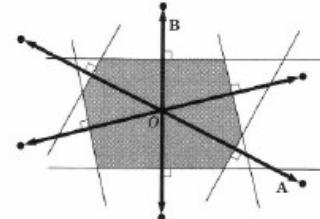
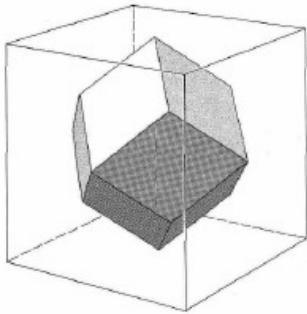


Figure 3.6

First Brillouin zone of the bcc lattice (rhombic-Dodecahedron). **First Brillouin zone of an oblique lattice in two dimensions.**

Another alternative expression for the diffraction conduction can be given in terms of *Laue equations*. Taking a dot product of Eq. (3.5) with $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ we obtain:

$$\begin{aligned} \mathbf{a}_1 \cdot \Delta\mathbf{k} &= 2\pi m_1, \\ \mathbf{a}_2 \cdot \Delta\mathbf{k} &= 2\pi m_2, \\ \mathbf{a}_3 \cdot \Delta\mathbf{k} &= 2\pi m_3, \end{aligned} \tag{3.20}$$

These conditions state that the allowed scattering vectors, $\Delta\mathbf{k}$, should lie at the intersections of cones around each lattice vector.

In-text question

1. Q: Briefly explain the conditions necessary for diffraction to take place.

A: See pages 35-36

2. Q: Explain the following terms: (i) Brillouin zone (ii) reciprocal lattice.

A: (i) See page 37 (ii) See page 42

3.1.3. Diffraction amplitude:

So far we have not discussed the amplitude and the width of diffraction peaks, which play an important role in the interpretation of x-ray diffraction data. This requires a more sophisticated analysis which we outline now.

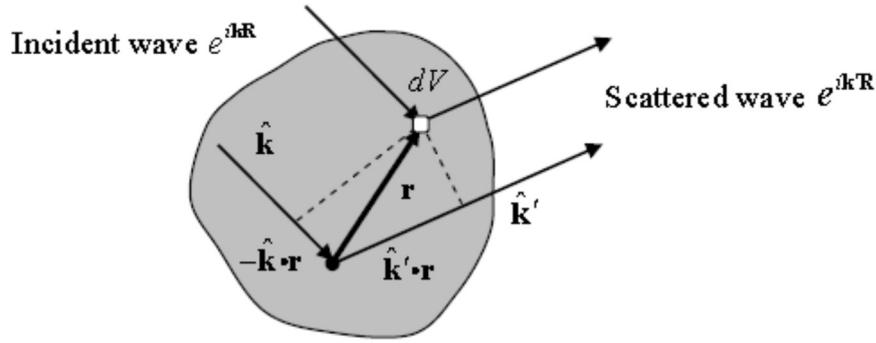


Figure 3.7 Scattering of x-ray in a solid

We consider scattering of x-rays by a solid as is shown in Fig.3.7. An incident plane wave e^{ikR} (wave vector \mathbf{k}) is scattered, and a scattered wave $e^{ik'R}$ (wave vector ' \mathbf{k}') is detected. Scattering occurs due to the interaction of the incident x-rays with the electron charge distributed in a solid with charge density $n(\mathbf{r})$. The amplitude of scattering by an infinitesimal volume dV is proportional to the charge at this point, i.e. $n(\mathbf{r})dV$, and a phase factor $e^{i\Delta\varphi}$ acquired by the scattered wave. The phase shift is equal to

$\Delta\varphi = \mathbf{K} \cdot \mathbf{r} - \mathbf{K}' \cdot \mathbf{r} = -(\mathbf{K}' - \mathbf{K}) \cdot \mathbf{r} = -\Delta\mathbf{k} \cdot \mathbf{r}$. The field at the detector is the superposition of the waves scattered from all points of the sample, and, therefore, the total amplitude of scattering is given by

$$F = \int_{solid} dV n(r) e^{-i\Delta k \cdot r} \quad (3.21)$$

The intensity of detected radiation is proportional to $|F^2|$, and therefore F determines the intensity and the width of diffraction peaks.

3.1.4. Scattering from a lattice with basis.

If the crystal structure represents a lattice with a basis, then we should take into account scattering by the atoms which have non-equivalent positions in a unit cell. The intensity of radiation scattered in a given Bragg peak will depend on the extent to which the rays scattered from these basis sites interfere with one another. To take into account basis atoms, first, let us rewrite Eq. (3.21) for the scattering amplitude at the diffraction condition, in terms of the integral over a unit cell:

$$F = \sum_T \int_{cell} dV n(\mathbf{r} + \mathbf{T}) e^{-i\mathbf{G} \cdot (\mathbf{r} + \mathbf{T})}, \quad (3.22)$$

where the sum is taken over all the lattice vectors \mathbf{T} . Taking into account that $n(\mathbf{r} + \mathbf{T}) = n(\mathbf{r})$, and that $e^{i\mathbf{G} \cdot \mathbf{T}} = 1$, we obtain

$$F = \sum_{\mathbf{T}} \int_{cell} dV n(\mathbf{r}) e^{-i\mathbf{G} \cdot \mathbf{r}} = N \int_{cell} dV n(\mathbf{r}) e^{-i\mathbf{G} \cdot \mathbf{r}} = NS_G. \quad (3.23)$$

where N is the number of cells in the solid, and we define the *structure factor*

$$S_G = \int_{cell} dV n(\mathbf{r}) e^{-i\mathbf{G} \cdot \mathbf{r}}. \quad (3.24)$$

Assuming that we have s atoms in a unit cell located at $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots$ it is convenient to write charge density as the superposition of charge densities n_j associated with each atom j of the basis:

$$n(\mathbf{r}) = \sum_{j=1}^s n_j(\mathbf{r} - \mathbf{r}_j). \quad (3.25)$$

The structure factor may now be written as

$$S_G = \sum_{j=1}^s \int dV n_j(\mathbf{r} - \mathbf{r}_j) e^{-i\mathbf{G} \cdot \mathbf{r}} = \sum_{j=1}^s e^{-i\mathbf{G} \cdot \mathbf{r}_j} \int dV n_j(\mathbf{p}) e^{-i\mathbf{G} \cdot \mathbf{p}}. \quad (3.26)$$

where $\mathbf{p} \equiv \mathbf{r} - \mathbf{r}_j$. The integral in Eq.(3.26) is known as the *atomic form factor*:

$$f_j(\mathbf{G}) = \int dV n_j(\mathbf{p}) e^{-i\mathbf{G}\cdot\mathbf{p}}, \quad (3.27)$$

which is determined by the charge density of atom j in the basis. The structure factor is then

$$S_{\mathbf{G}} = \sum_{j=1}^s f_j(\mathbf{G}) e^{-i\mathbf{G}\cdot\mathbf{r}_j}. \quad (3.28)$$

Example: structure factor of bcc lattice. A conventional cell of the bcc lattice contains two identical atoms with coordinates: $\mathbf{r}_1 = 0$ and $\mathbf{r}_2 = (a/2)(x + y + z)$. Since the atoms are identical, the atomic form factors are same, i.e. $f_1 = f_2 = f$. The reciprocal unit cell is cubic with a cell side of $2\pi/a$, and the reciprocal vector is given by

$$\mathbf{G} = \frac{2\pi}{a} (m_1 \hat{\mathbf{x}} + m_2 \hat{\mathbf{y}} + m_3 \hat{\mathbf{z}}). \quad (3.29)$$

The structure factor is, then,

$$\begin{aligned} S_{\mathbf{G}} / f &= \exp(-i\mathbf{G}\cdot\mathbf{r}_1) + \exp(-i\mathbf{G}\cdot\mathbf{r}_2) = 1 + \exp\left[-i\frac{2\pi}{a}(m_1 \hat{\mathbf{x}} + m_2 \hat{\mathbf{y}} + m_3 \hat{\mathbf{z}}) \cdot \frac{a}{2}(\hat{\mathbf{x}} + \hat{\mathbf{y}} + \hat{\mathbf{z}})\right] = \\ &= 1 + \exp[-i\pi(m_1 + m_2 + m_3)] = 1 + (-1)^{m_1+m_2+m_3} \end{aligned} \quad (3.30)$$

Therefore $S_G = 2f$, if $m_1 + m_2 + m_3$ is even, and $S_G = 0$, if $m_1 + m_2 + m_3$ is odd.

Thus, diffraction peaks will be observed, e.g., from the (110), (200), (211) planes, but not from the (100), (111), (210) planes. The later fact is due to the destructive interference from the basis atoms which cancel some peaks. For example, as is seen from Fig.3.8 for the (100) plane, the phase difference between successive planes is π , so that the reflected amplitude from two adjacent planes is $1 + e^{-i\pi} = 0$.

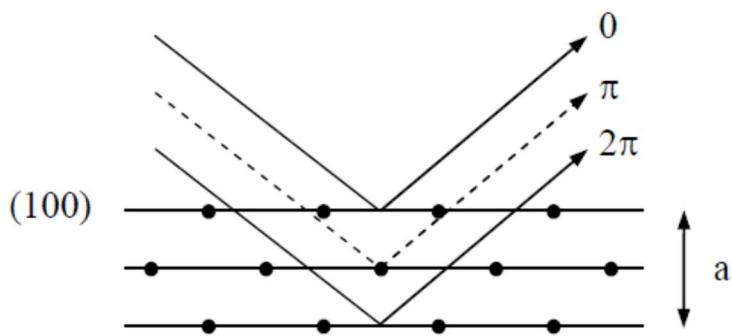


Figure 3.8 Reflection of x-ray from two adjacent planes

Structure factor of fcc lattice. The basis of the fcc structure referred to the cubic cell has identical atoms at $\mathbf{r}_1 = 000$; $\mathbf{r}_2 = \frac{1}{2}\frac{1}{2}0$; $\mathbf{r}_3 = 0\frac{1}{2}\frac{1}{2}$; $\mathbf{r}_4 = \frac{1}{2}0\frac{1}{2}$. Therefore,

$$S_G / f = 1 + e^{-i\pi(m_2+m_3)} + e^{-i\pi(m_1+m_3)} + e^{-i\pi(m_1+m_2)}, \quad (3.31)$$

which is non-zero only if all the indices are even or all the indices are odd. Allowed peaks are, e.g., (111), (200), (222), (220), (131).

Atomic form factor: The atomic form factor is defined by Eq.(3.27). It depends on the number and distribution of atomic electrons, and on the wavelength and angle of scattering of the radiation. It measures the scattering power of the j-th atom in the unit cell. An example of the importance of the atomic form factor is given below.

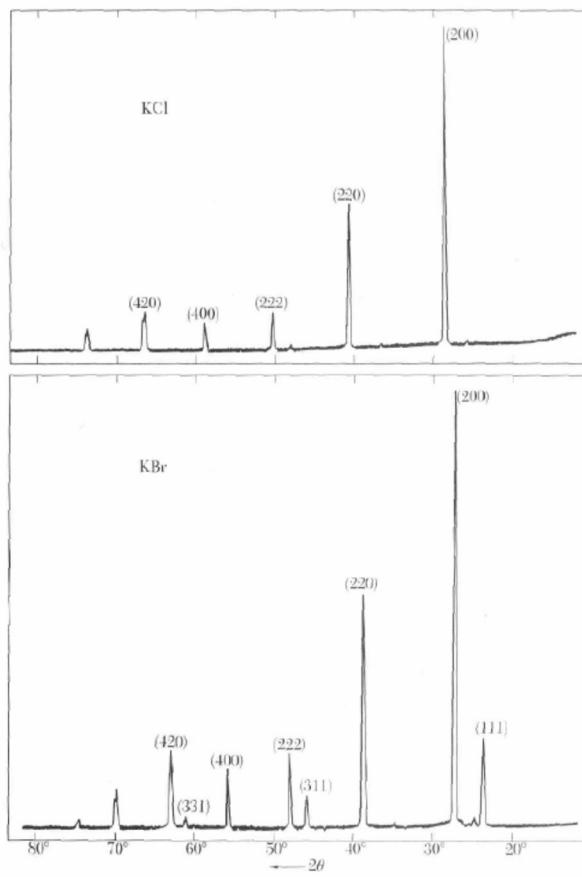


Fig. 3.9 Comparison of x-ray reflections from KCl and KBr powders

Both KCl and KBr have *sodium chloride structure*. In this structure the two types of atoms are arranged alternatively at the lattice sites of a simple cubic lattice. The space lattice is fcc with a basis of two non-equivalent atoms at 000 and $\frac{1}{2}\frac{1}{2}\frac{1}{2}$.

In KCl, the numbers of electrons of K^+ and Cl^- ions are equal and the charge distribution is similar. Therefore, the form factors for K^+ and Cl^- are almost exactly equal, so that the crystal looks to x-rays as if it were a monatomic simple cubic lattice of lattice constant $a/2$. Only even integers occur in the reflection indices when these are based on a cubic lattice of lattice constant a . In KBr the form factor of Br^- is quite different than that of K^+ , and all reflections of the fcc lattice are present.

3.1.5 Uses of X-ray Diffraction

- (i) Electron diffraction is particularly useful in exploring the structure of thin surface layers such as oxide layers on metal surfaces.
- (ii) Electron diffraction helps us to study orientation, lattice parameters and perfection of evaporated thin films.
- (iii) Electron diffraction helps us to see regularities of the atomic arrangement in a thin film which varies over regions of crystal imperfection.

In-text question

1. Q: Discuss briefly the structure factor and atomic force factor.
A: See pages 46-47
2. Q: Give three uses of X-ray diffraction.
A: See page 49

Summary of Study Session 3

In this study session, you have learnt:

1. The basis of the study of x-ray diffraction is the Bragg's reflection.
2. The brillouin zones have effect on the study of x-ray diffraction.
3. about the scattering of x-rays from a lattice with a basis.
4. about the uses of x-ray diffraction.

Self- Assessment Questions for Study Session 3

Now that you have completed this study session, you can assess how well you have achieved its learning outcomes by answering the following questions. Write your answers in your Study Diary and discuss them with your Tutor at the next Study Support Meeting.

SAQ 3.1

What is the de Broglie wavelength of an electron moving with velocity of $\frac{3c}{5}$?(take the value of the rest mass of the electron as 9.11×10^{-31}).

SAQ 3.2

A certain orthorhombic crystal has a ratio a: b: c = 0.429: 1: 0.377. Find the Miller indices of the faces whose intercepts are:

- 0.214:1:0.188
- 0.858:1:0.754
- 0.429: ∞ :0.126

SAQ 3.3

What is the de Broglie wavelength of neutrons at room temperature? Can they be used to study crystal structure?

SAQ 3.4

Deduce Bragg's law in X-ray diffraction. Describe Bragg's spectrometer and explain how it is used to determine the wavelength of x-rays.

SAQ 3.5

Derive Bragg's law of x-ray diffraction in crystals. Give an account of powder method of crystal structure analysis.

References/Further Readings

Ashcroft N. W. and Mermin N.D. (1976), Solid state Physics, Holt, Rinehart, Winston.

Study Session 4 Thermal Properties of Crystal Lattice

Introduction

Heat capacity per unit mass of a substance is known as specific heat. Specific heat is really a measure of the number of degrees of freedom of a system. Since degrees of freedom imply freedom to absorb potential or kinetic energy, the question to be answered is how many ways energy can be given to a system. The system we would concern with is the oscillating lattice, but its analysis is quite difficult, particularly in three dimensions. Instead, we would carry out the development of the theory in a series of warming-up exercises, each of which bears a little more to reality. Atoms vibrate about their mean equilibrium lattice sites in solids. These vibrations occur at any temperature, even near absolute zero. They are almost entirely responsible for the thermal properties- heat capacity, thermal conductivity, thermal expansion, etc., of insulators and contribute the greater of the heat capacity of metals. (The conduction electrons contribute only a small part of the heat capacity of metals but are almost entirely responsible for thermal conductivity).

Learning Outcomes for Study Session 4

When you have studied this session, you should be able to:

- 4.1 Describe the behavior of lattices of crystals to heat.
- 4.2 Explain the assumption made through classical and quantum theories.
- 4.3 Describe Einstein's and Debye's theories of the specific heat of crystal lattices
- 4.4 Discuss the successes and failures of these theories.
- 4.5 Explain thermal conductivity with respect to crystal lattices.

4.1 Main Content

4.1.1 Heat Capacity

There are two contributions to thermal properties of solids: one comes from *phonons* (or lattice vibrations) and another from *electrons*. This section is devoted to the thermal properties of solids due to lattice vibrations (the contribution from electrons in metals will be considered separately).

First, we consider the *heat capacity* of the *specific heat*. The heat capacity C is defined as the heat ΔQ which is required to raise the temperature by ΔT , i.e.

$$C = \frac{\Delta Q}{\Delta T} \quad (4.1)$$

If the process is carried out at constant volume V , then $\Delta Q = \Delta E$, where ΔE is the increase in internal energy of the system. The heat capacity at constant volume C_V is therefore given by

$$C_V = \left(\frac{\partial E}{\partial T} \right)_V. \quad (4.2)$$

The contribution of the phonons to the heat capacity of the crystal is called the *lattice heat capacity*. The total energy of the phonons at temperature T in a crystal can be written as the sum of the energies over all phonon modes, so that

$$E = \sum_{qp} \langle n_{qp} \rangle \hbar \omega_p(\mathbf{q}), \quad (4.3)$$

Where $\langle n_{qp} \rangle$ is the thermal equilibrium occupancy of phonons of wavevector \mathbf{q} and mode p ($p = 1\dots 3s$, where s is the number of atoms in a unit cell). The angular brackets denote the average in thermal equilibrium. Note that we assume here that the zero-point energy is chosen as the origin of the energy, so that the ground energy lies at zero. Now we calculate this average.

Consider a harmonic oscillator in a thermal bath. The probability of finding this oscillator in an excited state, which is characterized by a particular energy E_n is given by the Boltzmann distribution:

$$P_n = P_0 e^{-n\hbar\omega/k_B T}, \quad (4.4)$$

where the constant P_0 is determined from the normalization condition

$$\sum_{n=0}^{\infty} P_n = 1 \quad (4.5)$$

So that

$$P_0 = \left(\sum_{n=0}^{\infty} e^{-n\hbar\omega/k_B T} \right)^{-1}. \quad (4.6)$$

The average excitation number of the oscillator is given by

$$\langle n \rangle = \sum_{n=0}^{\infty} n P_n = \frac{\sum_{n=0}^{\infty} n e^{-n\hbar\omega/k_B T}}{\sum_{n=0}^{\infty} e^{-n\hbar\omega/k_B T}}. \quad (4.7)$$

The summation in the numerator can be performed using the known property of geometrical progression:

$$\sum_{n=0}^{\infty} x^n = 1 \quad (4.8)$$

Using this property we find:

$$\sum_{n=0}^{\infty} n x^n = x \frac{d}{dx} \sum_{n=0}^{\infty} x^n = x \frac{d}{dx} \frac{1}{1-x} = \frac{x}{(1-x)^2}, \quad (4.9)$$

Where $x = e^{-\hbar\omega/k_B T}$. Thus we obtain

$$\langle n \rangle = \frac{x}{(1-x)} = \frac{1}{x^{-1} - 1} = \frac{1}{e^{\hbar\omega/k_B T} - 1}. \quad (4.10)$$

The distribution given by Eq. (4.10) is known as the Planck distribution. Coming back to the expression for the total energy of the phonons, we find that

$$E = \sum_{\mathbf{q}, p} \frac{\hbar\omega_p(\mathbf{q})}{e^{\hbar\omega_p(\mathbf{q})/k_B T} - 1} \quad (4.11)$$

Usually it is convenient to replace the summation over \mathbf{q} by an integral over frequency. In order to do this we need to introduce the *density of modes* or the *density of states* $D_p(\omega)$. $D_p(\omega)d\omega$

represents the number of modes of a given number s in the frequency range $(\omega, \omega + d\omega)$. Then the energy is

$$E = \sum_p \int d\omega D_p(\omega) \frac{\hbar\omega}{e^{\hbar\omega/k_B T} - 1} \quad (4.12)$$

The lattice heat capacity can be found by differentiation of this equation with respect to temperature, so that

$$C_V = \frac{\partial E}{\partial T} = k_B \sum_p \int d\omega D_p(\omega) \frac{\left(\frac{\hbar\omega}{k_B T}\right)^2 e^{\hbar\omega/k_B T}}{(e^{\hbar\omega/k_B T} - 1)^2}. \quad (4.13)$$

We see that the central problem is to find the density of states $D_p(\omega)$, the number of modes per unit frequency range.

In-text question

1. Q: In Aluminium, $v_l = 6.32 \times 10^3 \text{ m/s}$ and $v_t = 3.10 \times 10^3 \text{ m/s}$. The density of aluminium is $2.7 \times 10^3 \text{ kg/m}^3$ and its atomic weight is 26.97. Calculate the Debye cut-off frequency for aluminium from these data.

solution

From $n = N/V = \rho/\mu = \text{density of material/atomic weight}$.

Using the formula, $\nu_D^3 = \left(\frac{9N}{4\pi V}\right) \left[\frac{2}{v_t^3} + \frac{1}{v_l^3}\right]^{-1}$ and substituting the parameters given in the question to have

A: debye frequency = $8.44 \times 10^{12} \text{ Hz}$.

2. Q: Give the assumptions of classical theory of lattice specific heat.

A: see pages 5.3-56

4.1.2 Density of States

Consider the longitudinal waves in a long bar. The solution for the displacement of atoms is given by

$$u(x) = A e^{iqx} \quad (4.14)$$

where we omitted a time-dependent factor as irrelevant for the present discussion. We shall now consider the effects of the boundary conditions on this solution. These boundary conditions are determined by the external constraints applied to the ends of the bar. The most convenient type of boundary condition is known as the *periodic boundary condition*. By this we mean that the right end of the bar is constrained in such a way that it is always in the same state of oscillation as the left end. It is as if the bar were deformed into a circular shape so that the right end joined the left. Given that the length of the bar is L , if we take the origin as being at the left end, the periodic condition means that

$$u(x = 0) = u(x = L), \quad (4.15)$$

where u is the solution given by Eq.(4.14). If we substitute (4.14) into (4.15), we find that

$$e^{iqL} = 1. \quad (4.16)$$

This equation imposes a condition on the admissible values of q i.e.

$$\cos qL + i \sin qL = 1$$

Equating the real and imaginary parts,

$$\cos qL = 1, \sin qL = 0,$$

This implies $qL = 2n\pi$,

$$\text{and } q = n \frac{2\pi}{L} \quad (4.17)$$

where $n = 0, +1, \pm 2$, etc. When these values are plotted along a q -axis, they form a one-dimensional mesh of regularly spaced points. The spacing between the points is $2\pi/L$. When the bar length is large, the spacing becomes small and the points form a quasi-continuous mesh.

Each q -value of (4.17) represents a *mode* of vibration. Let us choose an arbitrary interval dq in q -space, and look for the number of modes whose q 's lie in this interval. We assume here that L is large, so that the points are quasi-continuous, which is true for the macroscopic objects. Since the spacing between the points is $2\pi/L$, the number of modes is

$$= \frac{L}{2\pi} dq \quad (4.18)$$

We are interested in the number of modes in the frequency range $d\omega$ lying between $(\omega, \omega + d\omega)$. The *density of states* $D(\omega)$ is defined such that $D(\omega)d\omega$ gives this number. Comparing this definition with

(4.18), one may write $D(\omega)d\omega = (L/2\pi) dq$, or $D(\omega) = (L/2\pi)/(d\omega/dq)$. We note from Fig. 4.1, however, that in calculating $D(\omega)$ we must include the modes lying in the negative q -region as well as in the positive region. The effect is to multiply the above expression for $D(\omega)$ by a factor of two. That is,

$$D(\omega) = \frac{L}{\pi} \frac{1}{d\omega/dq} \quad (4.19)$$

We see that the density of states $D(\omega)$ is determined by the dispersion relation $\omega=\omega(q)$.

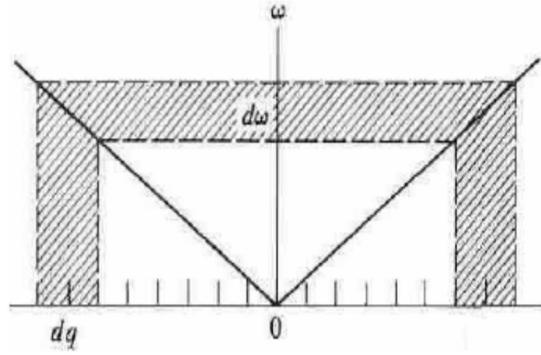


Figure 4.1 Density of states at different modes

Now we extend these results to the 3D case. The wave solution analogous to (4.14) is

$$\mathbf{u} = \mathbf{A} e^{i(q_x x + q_y y + q_z z)}, \quad (4.20)$$

where the propagation is described by the wave vector $\mathbf{q} = (q_x, q_y, q_z)$, whose direction specifies the direction of wave propagation. Here again we need to take into account the boundary conditions. For simplicity, we assume a cubic sample whose edge is L . By imposing the periodic boundary conditions, we find that the allowed values of \mathbf{q} must satisfy the condition

$$e^{iq_x L} = e^{iq_y L} = e^{iq_z L} = 1. \quad (4.21)$$

Therefore, the values are given by

$$(q_x, q_y, q_z) = \left(l \frac{2\pi}{L}, m \frac{2\pi}{L}, n \frac{2\pi}{L} \right), \quad (4.22)$$

where l, m, n are some integers.

If we plot these values in a q -space, as in Fig. 4.2, we obtain a three-dimensional cubic mesh. The volume assigned to each point in this q -space is $(2\pi/L)^3$.

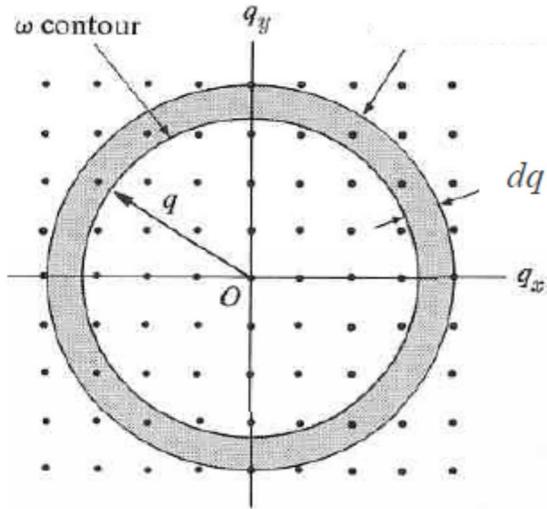


Figure 4.2 Allowed values of q for a wave traveling in 3 dimensions.

Only the cross section in the (q_x, q_y) plane is shown. The shaded circular shell is used for counting the modes.

Each point in Fig. 4.2 determines one mode. We now wish to find the number of modes lying in the spherical shell between the radii q and $q + dq$, as shown in Fig. 4.2. The volume of this shell is $4\pi q^2 dq$ and since the volume per point is $(2\pi/L)^3$, it follows that the number we seek is

$$\left(\frac{L}{2\pi}\right) 4\pi q^2 dq = \frac{V}{(2\pi)^3} 4\pi q^2 dq \quad (4.23)$$

where $V = L^3$ is the volume of the sample. By definition of the density of modes, this quantity is equal to $D(\omega)d\omega$. Thus, we arrive at

$$D(\omega) = \frac{V\omega^2}{2\pi^2 v^3}, \quad (4.24)$$

We note that Eq. (4.24) is valid only for an *isotropic solid*, in which the vibrational frequency, ω , does not depend on the direction of \mathbf{q} . Also we note that in the above discussion we have associated a single mode with each value of q . This is not quite true for the 3D case, because for each q there are actually three different modes, one longitudinal and two transverse, associated with the same value of q . In addition, in the case of non-Bravais lattice we have a few sites, so that the number of modes is $3s$, where s is the number of non-equivalent atoms. This should be taken into account by index $p = 1\dots 3s$ in the density of states as was done before, because the dispersion relations for the longitudinal and transverse waves are different, and acoustic and optical modes are different.

In-text question

1. Q: Derive the density of modes of atoms in a crystal when subjected to thermal effects.
A: See pages 54-58

4.1.3 Debye Model

The Debye model assumes that the acoustic modes give the dominant contribution to the heat capacity. Within the Debye approximation, the velocity of sound is taken as a constant independent of polarization as it would be in a classical elastic continuum. The dispersion relation is written as

$$\omega = vq \quad (4.25)$$

Where v is the velocity of sound.

In this approximation, the density of states is given by

$$D(\omega) = \frac{V\omega^2}{2\pi^2 v^3}, \quad (4.26)$$

i.e., the density of states increases quadratically with the frequency.

The normalization condition for the density of states determines the limits of integration over ω . The lower limit is obviously $\omega=0$. The upper limit can be found from the condition that the number of vibrational modes in a crystal is finite and is equal to the number of degrees of freedom of the lattice. Assuming that there are N unit cells in the crystal, and there is only one atom per cell (so that there are N atoms in the crystal), the total number of phonon modes is $3N$. Therefore, we can write

$$\sum_p \int_0^{\omega_D} D(\omega) d\omega = 3N \quad (4.27)$$

where the cutoff frequency ω_D is known as Debye frequency. Assuming that the velocity of the three acoustic modes is independent of polarization and substituting (4.26) in (4.27), we obtain:

$$\omega_D = \left(\frac{6\pi^2 v^3 N}{V} \right)^{1/3}. \quad (4.28)$$

The cutoff wavevector which corresponds to this frequency is given by

$$q_D = \frac{\omega_D}{v} = \left(\frac{6\pi^2 N}{V} \right)^{1/3}, \quad (4.29)$$

so that modes of wavevector larger than q_D are not allowed. This is due to the fact that the number of modes with $q \leq q_D$ exhausts the number of degrees of freedom of the lattice.

The thermal energy is given by Eq. (4.12), so that

$$E = 3 \int_0^{\omega_D} d\omega \frac{V\omega^2}{2\pi^2 v^3} \frac{\hbar\omega}{e^{\hbar\omega/k_B T} - 1}, \quad (4.30)$$

where a factor of 3 is due to the assumption that the phonon velocity is independent of polarization. This leads to

$$E = \frac{3V\hbar}{2\pi^2 V^3} \int_0^{\omega_D} d\omega \frac{\omega^3}{e^{\hbar\omega/k_B T} - 1} = \frac{3Vk_B^4 T^4}{2\pi^2 V^3 \hbar^3} \int_0^{x_D} dx \frac{x^3}{e^x - 1}, \quad (4.31)$$

Where $x \equiv \hbar\omega/k_B T$ and

$$x_D \equiv \hbar\omega_D/k_B T \equiv \theta_D/T \quad (4.32)$$

The latter expression defines the Debye temperature

$$\theta_D = \frac{\hbar V}{k_B} \left(\frac{6\pi^2 N}{V} \right)^{1/3}. \quad (4.33)$$

The total phonon energy is then

$$E = 9Nk_B T \left(\frac{T}{\theta_D} \right)^3 \int_0^{x_D} dx \frac{x^3}{e^x - 1} \quad (4.34)$$

where N is the number of atoms in the crystal and $x_D \equiv \theta_D/T$

The heat capacity is most easily found by differentiating the middle expression of (4.31) with respect to the temperature (in Eq.(4.34) we will have to differentiate the upper limit) so that

$$C_V = \frac{3V\hbar^2}{2\pi^2 V^3 k_B T^2} \int_0^{\omega_D} d\omega \frac{\omega^4 e^{\hbar\omega/k_B T}}{\left(e^{\hbar\omega/k_B T} - 1\right)^2} = 9Nk_B \left(\frac{T}{\theta_D} \right)^3 \int_0^{x_D} dx \frac{x^4 e^x}{\left(e^x - 1\right)^2}. \quad (4.35)$$

In the limit $T \gg \theta$, we can expand the expression under the integral and obtain: $C_V = 3Nk_B$. This is exactly the classical value for the heat capacity, which is known from the elementary physics. Recall that according to the elementary thermodynamics, the average thermal energy per degree of freedom is equal to $k_B T$. Therefore for a system of N atoms $E = 3Nk_B T$ which results in $C_V = 3Nk_B$. This is known as the *Dulong-Petit* law.

Now consider an opposite limit, i.e. $T \ll \theta_D$. At very low temperatures we can approximate (4.34) by letting the upper limit go to infinity. We obtain

$$E = 9Nk_B T \left(\frac{T}{\theta_D} \right)^3 \int_0^\infty dx \frac{x^3}{e^x - 1} = 9Nk_B T \left(\frac{T}{\theta_D} \right)^3 \frac{\pi^4}{15} = \frac{3\pi^4}{5} Nk_B T \left(\frac{T}{\theta_D} \right)^3, \quad (4.36)$$

and therefore

$$C_V = \frac{12\pi^4}{5} Nk_B \left(\frac{T}{\theta_D} \right)^3. \quad (4.37)$$

We see that within the Debye model at low temperatures, the heat capacity is proportional to T^3 .

The cubic dependence may be understood from the following qualitative argument. At low temperature, only a few modes are excited. These are the modes whose quantum energy $\hbar\omega$ is less than $k_B T$. The number of these modes may be estimated by drawing a sphere in the q -space whose frequency $\omega = k_B T / \hbar$, and counting the number of points inside, as shown in Fig. 4.3. This sphere may be called the *thermal sphere*, in analogy with the Debye sphere discussed above. The number of modes inside the thermal sphere is proportional to $q^3 \sim \omega^3 \sim T^3$. Each mode is fully excited and has an average energy equal to $k_B T$. Therefore the total energy of excitation is proportional to T^4 , which leads to a specific heat proportional to T^3 , in agreement with (4.37).

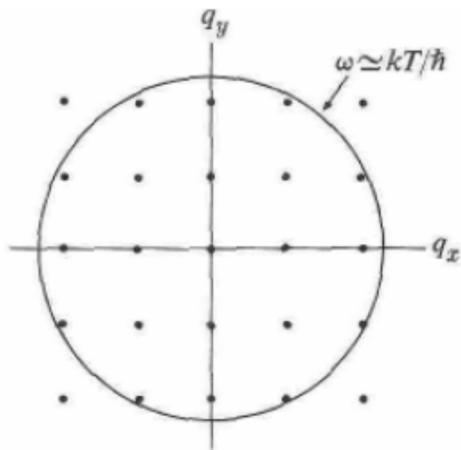


Figure 4.3 The thermal sphere which is the frequency contour $\omega = k_B T / \hbar$

To compare these predictions with experimental results one should know the Debye temperature. This temperature is normally determined by fitting experimental data. Fig.4.4 shows the fitted data versus the reduced temperature T/θ_D . You see that the curve is *universal*; it is the same for different substances. The agreement between the calculated and experimental data is remarkable.

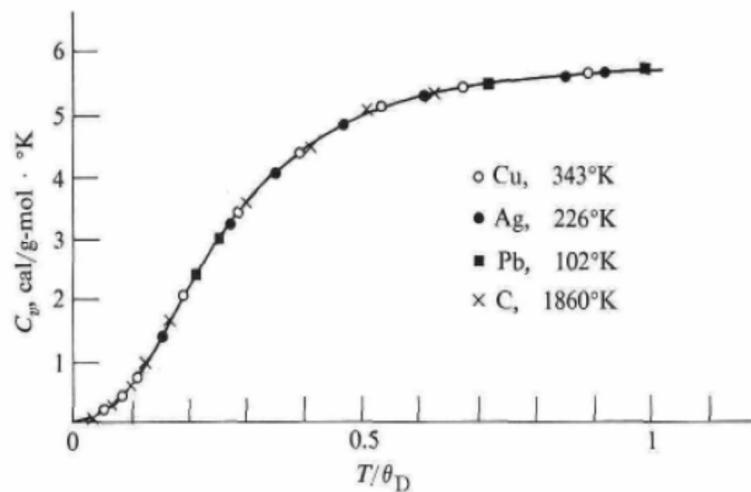


Figure 4.4. Specific heats versus reduced temperature for four substances.

In-text question

- Q: At what temperature does the electronic contribution to the specific heat of silver become identical with the Dulong-Petit value? $E_f = 5\text{eV}$.

Solution

From $E_f = 2k_B T$, where $k_B = 8.6173324(78) \times 10^{-5}\text{eVK}^{-1}$,

A: $3.125 \times 10^4 \text{ K}$

- Q: Write the expression for the classical value of molar lattice specific heat.

A: $3R_u$

4.1.4 Einstein Model

Within the Einstein model the density of states is approximated by a delta function at some frequency ω_E , i.e

$$D(\omega) = N\delta(\omega - \omega_B) \quad (4.38)$$

where N is the total number of atoms (oscillators). ω_E is known as the Einstein frequency. The thermal energy of the system (4.12) is then

$$E = \frac{3N\hbar\omega_E}{e^{\hbar\omega_E/k_B T} - 1} \quad (4.39)$$

where a factor of 3 reflects the fact that there are three degree of freedom for each oscillator. The heat capacity is then

$$C_V = \frac{\partial E}{\partial T} = 3Nk_B \left(\frac{\hbar\omega_E}{k_B T} \right)^2 \frac{e^{\hbar\omega_E/k_B T}}{(e^{\hbar\omega_E/k_B T} - 1)^2}. \quad (4.40)$$

The high temperature limit for the Einstein model is the same as that for the Debye model, i.e. $C_v = 3Nk_B$, which is the Dulong-Petit law. At low temperatures however (4.40) decreases as $C_v \sim e^{-\hbar\omega_E/k_B T}$, which is different from the Debye T^3 law. The reason for this disagreement is

that at low temperatures only acoustic phonons are populated and the Debye model is much better approximation than the Einstein model. The Einstein model is often used to approximate the optical phonon part of the phonon spectrum.

Concluding our discussion about the heat capacity of solids, we note that a real density of vibrational modes could be much more complicated than those described by the Debye and Einstein models. Fig.4.5 shows the density of states for Cu. The dash line is the Debye approximation, which has the same area (under the curve) as the solid curve. The Einstein approximation would have a delta peak at some frequency. At low frequencies the density of states varies quadratically with the frequency, which is due to acoustic modes and similar to that within the Debye approximation. At higher frequencies, there is a peak which is due to optical modes. This density of states has to be included in order to obtain a quantitative description of experimental data.

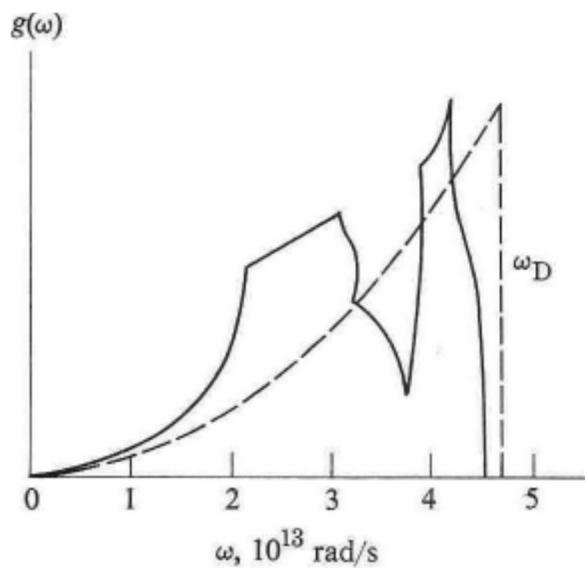


Figure 4.5. Total density of states for Cu, as deduced from data on neutron scattering.

In-text question

1. Q: State Dulong-Petit's law and show how the departure from this law at lower temperatures has been explained by Einstein's theory.

A: See page 62-65

2. Q: Discuss the variation of specific heat capacity of solids with temperature and use Einstein's theory to explain it.

A: See pages 63-64

Summary of Study Session 4

In this study session, you have learnt that:

1. At high temperatures, the experimental deduction on specific heat is in accordance with the classical theory on specific heat (Dulong-Petit theory) while at low temperatures, the classical theory fails.
2. Einstein's approximation shows that lattice specific heat varies with temperature.
3. At high temperatures, Einstein's result agrees with the classical and experimental results because there is increase in the lattice vibration and also due to high quantum number (correspondence principle).
4. Debye's theory agrees with the experimental deductions both at low and high temperatures.
5. Debye's theory gives a perfect insight into the lattice specific heat of solids.

Self-Assessment Questions for Study Session 4

Now that you have completed this study session, you can assess how well you have achieved its learning outcomes by answering the following questions. Write your answers in your Study Diary and discuss them with your Tutor at the next Study Support Meeting.

SAQ 4.1

Show that the kinetic energy of a three-dimensional electron gas of N electrons at zero temperature is $U=3/5NE_F$.

SAQ 4.2

Show that the density of states of a free-electron gas in two dimensions is independent of energy.

SAQ 4.3

Fcc Au (cubic lattice parameter $a=4.08\text{\AA}$) has electrical resistivity $\rho=2.2\mu\Omega\text{cm}$ at room temperature. Using a free-electron model and assuming one valence electron per atom calculate:

- a) the concentration of the conduction electrons n ,
- b) the Fermi energy,
- c) the relaxation time τ ,
- d) the Fermi velocity v_F ,
- e) the mean free path l ,
- f) the electronic heat capacity per atom at room temperature.

SAQ 4.4

The residual resistivity for 1 atomic percent of As impurities in Cu is $6.8\mu\Omega\text{cm}$. Calculate the cross section for the scattering of an electron by one As impurity in Cu. Use a free-electron model assuming that Cu has the fcc structure with the cubic lattice parameter $a=3.62\text{\AA}$ and one valence electron per atom.

References/Further Readings

Ashcroft N. W. and Mermin N.D., (1976), Solid state Physics, Holt, Rinehart, Winston.

Study Session 5 Elastic Properties

Introduction

All materials in nature are elastic. Thus, solids stretch in the direction in which they are pulled. Any force or load applied on the material will result in stress and strain in the material. Stress represents the intensity of the reaction force at any point in the body as imposed by service load, assembly conditions, fabrication and thermal changes.

Learning Outcomes for Study Session 5

When you have studied this session, you should be able to:

- 5.1 Explain the effect of external influences like pressure on the atoms of solid state materials.
- 5.2 Describe what stress and strain depict in the elastic properties of materials.
- 5.3 Calculate the elastic constants of materials

5.1 Main Content

5.1.1 Elastic Constants

Elastic properties of solids are determined by interatomic forces acting on atoms when they are displaced from the equilibrium positions. At small deformations these forces are proportional to the displacements of atoms. As an example, consider a 1D solid. A typical binding curve has a minimum at the equilibrium interatomic distance R_0 :

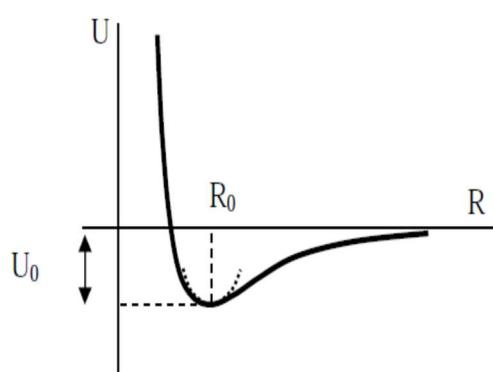


Figure 5.1 Graph of potential energy against interatomic distance

Expanding the energy at the minimum in the Taylor series we find:

$$U(R) = U_0 + \frac{\partial U}{\partial R} \Big|_{R_0} (R - R_0) + \frac{\partial^2 U}{\partial R^2} \Big|_{R_0} (R - R_0)^2 + \dots \quad (5.1)$$

At equilibrium $\frac{\partial U}{\partial R} \Big|_{R_0} = 0$, so that

$$U(R) = U_0 + \frac{1}{2}ku^2, \quad (5.2)$$

where we defined $\frac{1}{2}k = \frac{\partial^2 U}{\partial R^2} \Big|_{R_0}$ and $u = R - R_0$ is the displacement of an atom from equilibrium position R_0 . Differentiating Eq.(5.2), $F = -\frac{\partial U}{\partial R}$, we obtain the force F acting on an atom:

$$F = -ku \quad (5.3)$$

The constant k is an interatomic force constant. Eq.(5.3) represents the simplest expression for the *Hooke's law* showing that the force acting on an atom, F , is proportional to the displacement u . This law is valid only for small displacements and characterizes a *linear region* in which the restoring force is linear with respect to the displacement of atoms.

The elastic properties are described by considering a crystal as a homogeneous continuum medium rather than a periodic array of atoms. In a general case the problem is formulated as follows:

- (i) Apply forces, which are described in terms of *stress* σ , and determine displacements of atoms which are described in terms of *strain* ε .
- (ii) Define elastic constants C relating stress σ and strain ε , so that $\sigma = C\varepsilon$.

Example: In 1D case, $F = -ku$, where u is a change in the crystal length under applied force F . We can therefore write

$$\sigma = \frac{F}{A} = \left(\frac{-kL}{A}\right) \left(\frac{u}{L}\right) = C\varepsilon, \quad (5.4)$$

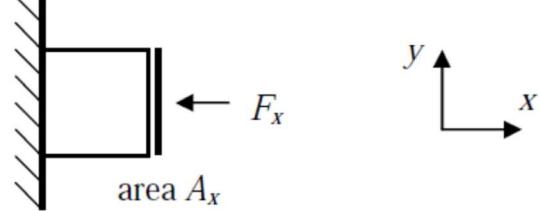
where A is the area of the cross section, and L is the equilibrium length of the 1D crystal. The stress σ is defined as the force per unit area and the strain ε is the dimensionless constant which describes the relative displacement (deformation).

In a general case of a 3D crystal the stress and the strain are tensors which are defined as follows.

Stress has the meaning of locally applied “pressure”. It has components σ_{ij} , showing that the force can be applied along 3 directions “ i ” and 3 faces “ j ”. The stress is defined locally, so that $\sigma_{ij} = \sigma_{ij}(\mathbf{r})$.

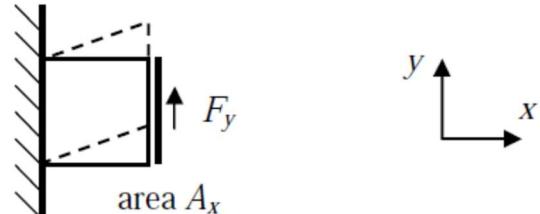
Compression stress (σ_{xx} σ_{yy} σ_{zz}):

$$\sigma_{xx} = \frac{F_x}{A_x},$$

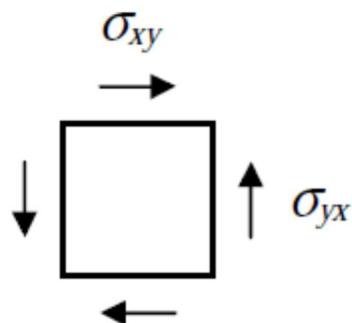


Shear stress (σ_{xy} σ_{yx} σ_{xz} σ_{zx} σ_{yz} σ_{zy}):

$$\sigma_{yx} = \frac{F_y}{A_x},$$



Shear forces must come in pairs to conserve angular acceleration inside the crystal:



That makes the stress tensor diagonal, i.e.

$$\sigma_{ij} = \sigma_{ji}. \quad (5.5)$$

Strain determines relative atomic displacement:

$$\varepsilon_{ij}(\mathbf{r}) = \frac{du_i}{dx_j}, \quad (5.6)$$

where u_i is displacement in “ i ” direction and x_j is the direction along which u_i may vary.

Compression strain (ε_{xx} ε_{yy} ε_{zz}):

$$\varepsilon_{xx} = \frac{du_x}{dx}, \quad (5.7)$$

In a homogeneous crystal ε_{xx} is a constant $\varepsilon_{xx} = \frac{u}{L}$, where u is the change in the crystal length L .

Shear strain (ε_{xy} ε_{yx} ε_{xz} ε_{zx} ε_{yz} ε_{zy}):

$$\varepsilon_{yx} = \frac{du_y}{dx}, \quad (5.8)$$

Since σ_{ij} and σ_{ji} must always be applied together, we can define shear strains symmetrically:

$$\varepsilon_{ij} = \varepsilon_{ji} = \frac{1}{2} \left(\frac{du_i}{dx_j} + \frac{du_j}{dx_i} \right). \quad (5.9)$$

Elastic constants C relate the strain and the stress in a linear fashion:

$$\sigma_{ij} = \sum_{kl} c_{ijkl} \varepsilon_{kl} \quad (5.10)$$

Eq.(5.10) is a general form of the Hooke's law. The matrix C in a most general form has $3 \times 3 \times 3 \times 3 = 81$ components. However, due to the symmetrical form of σ_{ij} and ε_{ij} - each of them

has 6 independent components, we need only 36 elastic constants. There is a convention to denote these constants by C_{mn} , where indices m and n are defined as $1 = xx$, $2 = yy$, $3 = zz$ for the compression components and as $4 = yz$, $5 = zx$, $6 = xy$ for the shear components. For example, $C_{11} = C_{xxxx}$, $C_{12} = C_{xxyy}$, $C_{44} = C_{yzyz}$.

Therefore, the general form of the Hooke's law is given by

$$\begin{array}{c}
 \text{compression} & \text{mixed} \\
 \left(\begin{array}{c} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \hline \sigma_{yz} \\ \sigma_{zx} \\ \sigma_{xy} \end{array} \right) = \left(\begin{array}{ccc|ccc} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{21} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{31} & C_{32} & C_{33} & C_{34} & C_{35} & C_{36} \\ \hline C_{41} & C_{42} & C_{43} & C_{44} & C_{45} & C_{46} \\ C_{51} & C_{52} & C_{53} & C_{54} & C_{55} & C_{56} \\ C_{61} & C_{62} & C_{63} & C_{64} & C_{65} & C_{66} \end{array} \right) \left(\begin{array}{c} \epsilon_{xx} \\ \epsilon_{yy} \\ \epsilon_{zz} \\ \hline \epsilon_{yz} \\ \epsilon_{zx} \\ \epsilon_{xy} \end{array} \right)
 \end{array} \quad \begin{array}{l} \text{stress} \quad \text{mixed} \quad \text{shear} \quad \text{strain} \\ \text{compression} \quad \text{shear} \end{array} \quad (5.11)$$

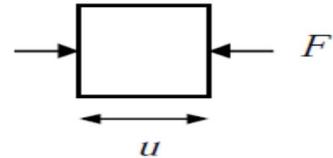
All 36 elastic constants are independent. However in crystals many of them are the same due to symmetry. In particular, in cubic crystals $C_{11} = C_{22} = C_{33}$, $C_{12} = C_{21} = C_{23} = C_{32} = C_{13} = C_{31}$, $C_{44} = C_{55} = C_{66}$ due to the fact that x , y , and z axes are identical by symmetry. Also, the off diagonal shear components are zero, i.e. $C_{45} = C_{54} = C_{56} = C_{65} = C_{46} = C_{64} = 0$, and mixed compression/shear coupling does not occur, i.e. $C_{14} = C_{41} = \dots = 0$. Therefore, the cubic elasticity matrix has the form

$$\begin{pmatrix} C_{11} & C_{12} & C_{12} & 0 \\ C_{12} & C_{11} & C_{12} & 0 \\ C_{12} & C_{12} & C_{11} & \\ & & C_{44} & 0 \\ 0 & & 0 & C_{44} \\ & & 0 & 0 \\ & & & C_{44} \end{pmatrix}. \quad (5.12)$$

We have only 3 independent constants.

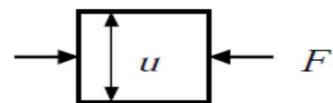
Longitudinal compression (Young's modulus):

$$C_{11} = \frac{\sigma_{xx}}{\varepsilon_{xx}} \\ = \frac{F/A}{u/L}$$



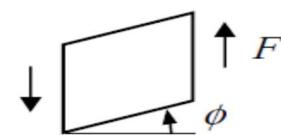
Transverse expansion:

$$C_{12} = \frac{\sigma_{xy}}{\varepsilon_{yy}}$$



Shear modulus:

$$C_{44} = \frac{\sigma_{xy}}{\varepsilon_{xy}} \\ = \frac{F/A}{\phi}$$



In-text question

1. Q: Consider a longitudinal wave $U_n = A\cos(qna - \omega t)$ in a monoatomic linear lattice of atoms of mass M , spacing a and nearest-neighbour interaction C .

(i) Show that the total energy of the wave is

$$E = \frac{1}{2} M \sum_n \left(\frac{du_n}{dt} \right)^2 + \frac{1}{2} C \sum_n (u_{n+1} - u_n)^2,$$

Where n runs over all atoms.

A: The total energy of the wave is the sum of the kinetic energy E_{kin} and the potential energy E_{pot} . The kinetic energy is the sum of the kinetic energies of all atoms, i.e.

$$E_{kin} = \frac{1}{2} M \sum_n \left(\frac{du_n}{dt} \right)^2$$

where M is the mass of atoms and $\frac{du_n}{dt}$ is the velocity of n -th atom. The potential energy is the potential energy of all the “springs” connecting atoms. For two atoms, this energy is the same as that for a harmonic oscillator, i.e. $\frac{1}{2}Cx^2$, where C is interatomic force constant and x is the change in distance between the atoms from the equilibrium distance. Therefore, for atoms n and $n+1$ having displacements u_n and u_{n+1} respectively this energy is $\frac{1}{2}C(u_{n+1} - u_n)^2$. The total potential energy is given by

$$E_{pot} = \frac{1}{2} C \sum_n (u_{n+1} - u_n)^2$$

The total energy of the wave is therefore

$$E = E_{kin} + E_{pot} = \frac{1}{2} M \sum_n \left(\frac{du_n}{dt} \right)^2 + \frac{1}{2} C \sum_n (u_{n+1} - u_n)^2.$$

5.1.2 Elastic waves

So far, we have assumed that atoms were at rest at their lattice sites. Atoms, however, are not quite stationary, but can oscillate around their equilibrium positions (e.g., as a result of thermal energy). This leads to lattice vibrations.

When considering lattice vibrations three major approximations are made:

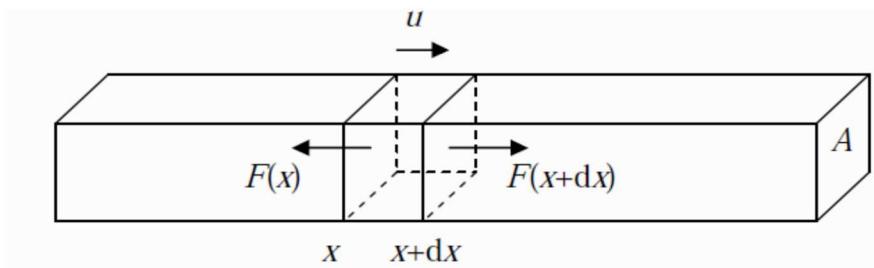
- (i) It is assumed that displacements of atoms are small, i.e. $u \ll a$, where a is a lattice parameter.
- (ii) Forces acting on atoms are assumed to be *harmonic*, i.e. proportional to the displacements: $F = -Cu$. This is the same approximation which is used to describe a harmonic oscillator.
- (iii) It is assumed that the adiabatic approximation is valid – electrons follow atoms, so that the nature of bond is not affected by vibrations.

The discreteness of the lattice must be taken into account in the discussion of lattice vibrations.

However, when the wavelength is very long, i.e. $\lambda \ll a$, one may disregard the atomic nature and treat the solid as a continuous medium. Such vibrations are referred to as *elastic waves*.

We consider an elastic wave in a long bar of cross-sectional area A and mass density $\rho = M/V$.

- (1) First, we consider a *longitudinal wave* of compression/expansion.



We look at a segment of width dx at the point x and denote the elastic displacement by u . According to the Newton's second law

$$m \frac{d^2u}{dt^2} = \sum F, \quad (5.13)$$

Which implies that

$$(\rho Adx) \frac{d^2u}{dt^2} = F(x + dx) - F(x), \quad (5.14)$$

$$\rho \frac{d^2u}{dt^2} = \frac{1}{A} \frac{dF}{dx} = \frac{d\sigma_{xx}}{dx}, \quad (5.15)$$

where we introduce the compression stress σ_{xx} . Assuming that the wave propagates along the [100] direction, we can write the Hooke's law in the form

$$\sigma_{xx} = c_{11} \varepsilon_{xx}, \quad (5.16)$$

where C_{11} is Young's modulus. Since $\varepsilon_{xx} = \frac{du}{dx}$, this leads to the *wave equation*

$$\frac{d^2u}{dt^2} = \left(\frac{c_{11}}{\rho}\right) \frac{d^2u}{dx^2}. \quad (5.17)$$

A solution of the wave equation has the form of a propagating *longitudinal plane wave*

$$\mathbf{u}(x, t) = A e^{i(qx - \omega t)} \mathbf{x}, \quad (5.18)$$

where q is the wave vector,

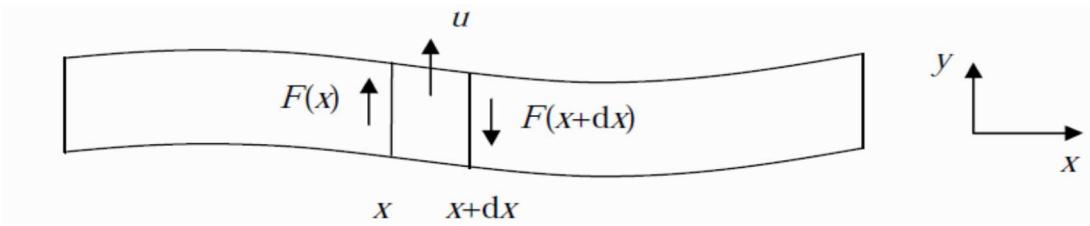
$$\omega = V_L q, \quad (5.19)$$

is the frequency, and

$$V_L = \sqrt{\frac{c_{11}}{\rho}} \quad (5.20)$$

is the *longitudinal velocity of sound*.

(2) Now we consider a *transverse wave* which is controlled by shear stress and strain.



In this case

$$\rho \frac{d^2u}{dt^2} = \frac{d\sigma_{xy}}{dx}, \quad (5.21)$$

where the shear stress σ_{xy} is determined by the shear modulus C_{44} and shear strain $\varepsilon_{xy} = \frac{du}{dx}$

$$\sigma_{xy} = C_{44} \varepsilon_{xy} \quad (5.22)$$

Therefore Eq.(5.21) takes the form

$$\rho \frac{d^2u}{dt^2} = C_{44} \frac{d\varepsilon_{xy}}{dx} = C_{44} \frac{d^2u}{dx^2}, \quad (5.23)$$

resulting in the wave equation

$$\frac{d^2u}{dt^2} = \left(\frac{C_{44}}{\rho}\right) \frac{d^2u}{dx^2} \quad (5.24)$$

This is the equation for the transverse plane wave, which has displacements in the y direction but propagates in the x direction:

$$\mathbf{u}(x, t) = A e^{i(qx - \omega t)} \mathbf{y}, \quad (5.25)$$

where q is the wave vector,

$$\omega = V_T q,$$

in the frequency, and

$$V_T = \sqrt{\frac{C_{44}}{\rho}} \quad (5.26)$$

is the *transverse velocity of sound*. Note that there are two linear independent transverse modes characterized by the displacements in y and in z directions. For the [100] direction, by symmetry the velocities of these modes are the same and given by Eq.(5.26).

Normally $C_{11} > C_{44}$ and therefore $V_L > V_T$.

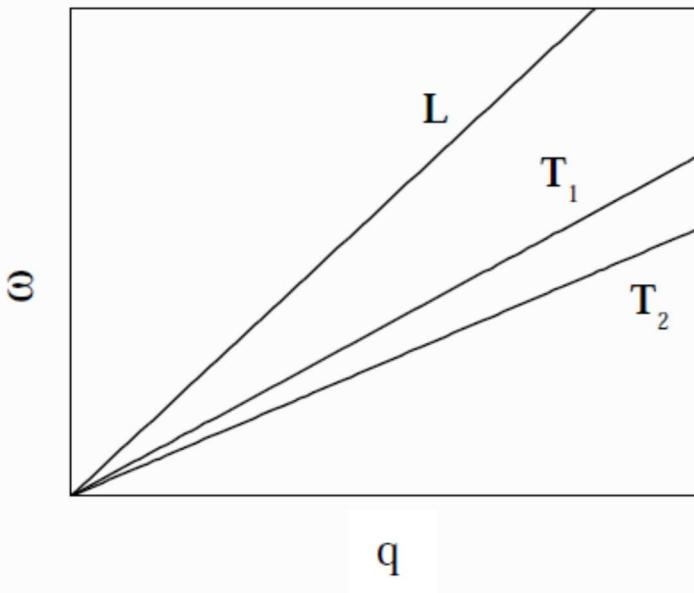
Waves we have considered are in [100] direction, i.e. $\mathbf{q} \parallel [100]$. In other directions, the sound velocity depends on combinations of elastic constants:

$$V = \sqrt{\frac{C_{eff}}{\rho}}, \quad (5.27)$$

where C_{eff} is an effective elastic constant which is given for cubic crystals in the table:

Mode	$\mathbf{q} \parallel [100]$	$\mathbf{q} \parallel [110]$	$\mathbf{q} \parallel [111]$
L	C_{11}	$\frac{1}{2}(C_{11}+C_{12}+2C_{44})$	$\frac{1}{3}(C_{11}+2C_{12}+4C_{44})$
T_1	C_{44}	C_{44}	$\frac{1}{3}(C_{11}-C_{12}+C_{44})$
T_2	C_{44}	$\frac{1}{2}(C_{11}-C_{12})$	$\frac{1}{3}(C_{11}-C_{12}+C_{44})$

The relation connecting the frequency ω and the wave vector q is known as the *dispersion relation*. For elastic waves, ω is proportional to q , and the ratio ω/q gives a constant velocity. The figure below shows the dispersion relation for elastic waves. There are three modes – one longitudinal and two transverse, which represent straight lines whose slopes are equal to the respective velocities of sound. For the [100] and [111] directions the two transverse modes are degenerate, i.e. have the same v_T .



In-text question

1. Q: List the major approximations considered when dealing with lattice vibrations.

A: See page 74

Summary of Study Session 5

In this study session, you have learnt that:

1. The general Hooke's law has been transformed into 36 elastic constants.
2. All these 36 elastic constants are independent. However in crystals many of them are the same due to symmetry. In particular, in cubic crystals $C_{11} = C_{22} = C_{33}$, $C_{12} = C_{21} = C_{23} = C_{32} = C_{13} = C_{31}$, $C_{44} = C_{55} = C_{66}$ due the fact that x , y , and z axes are identical by symmetry. Also the off diagonal shear components are zero, i.e. $C_{45} = C_{54} = C_{56} = C_{65} = C_{46} = C_{64} = 0$, and mixed compression/shear coupling does not occur, i.e., $C_{14} = C_{41} = \dots = 0$.
3. These 36 elastic constants have been reduced to 3 independent constants.
4. Approximation were made assuming the atoms were not at rest at their lattice sites and these led to two major vibrational modes of the atoms (optical and acoustical mode).

Self Assessment Questions for Study Session 5

Now that you have completed this study session, you can assess how well you have achieved its learning outcomes by answering the following questions. Write your answers in your Study Diary and discuss them with your Tutor at the next Study Support Meeting.

SAQ 5.1

Consider a longitudinal wave $U_n = A\cos(qna - \omega t)$ in a mono-atomic linear lattice of atoms of mass M , spacing a and nearest-neighbour interaction C .

- (a) Show that the time-averaged kinetic energy is equal to the time-averaged potential energy.
- (b) Show that the total time-averaged energy per atom is equal to $\frac{1}{2}MA^2\omega^2$.

SAQ 5.2

Consider a linear chain in which alternative ions have masses M_1 and M_2 and only nearest neighbours interact.

- (a) Discuss the form of the dispersion relation and the nature of the vibrational modes when $M_1 \gg M_2$.
- (b) Show that for $M_1=M_2$ the dispersion relation becomes identical to that for the monoatomic lattice.

SAQ 5.3

Consider the normal modes of a linear chain in which the force constants between nearest-neighbour atoms are alternatively C and $10C$. Assuming that the masses are equal and the nearest neighbour separation is $a/2$. Find $\omega(q)$ at $q = 0$ and $q = \pi/a$. Sketch the dispersion curve. This problem simulates a crystal of diatomic molecules such as H_2 .

References/Further Readings

Kittel C., (1963), Quantum theory of solids, Wiley.

Study Session 6 Lattice Vibration

Introduction

So far we have been discussing equilibrium properties of crystal lattices. When the lattice is at equilibrium each atom is positioned exactly at its lattice site. Now suppose that an atom is displaced from its equilibrium site by a small amount. Due to external force acting on this atom, it will tend to return to its equilibrium position. This results in lattice vibrations. Due to interactions between atoms, various atoms move simultaneously, so we have to consider the motion of the entire lattice.

Learning Outcomes for Study Session 6

When you have studied this session, you should be able to:

- 6.1 Explain the classical approach to the lattice vibrations in terms of 1-d, 2-d, 3-d.
- 6.2 Discuss the treatment of lattice vibrations as phonons using the quantum approach.

6.1 Main Content

6.1.1 One-dimensional lattice

For simplicity we consider, first, a one-dimensional crystal lattice and assume that the forces between the atoms in this lattice are proportional to relative displacements from the equilibrium positions.

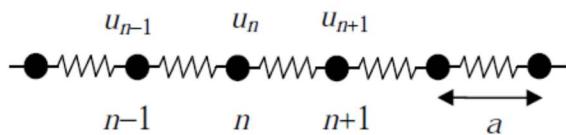


Figure 6.1 Monatomic Lattice

This is known as the *harmonic approximation*, which holds well provided that the displacements are small. One might think about the atoms in the lattice as interconnected by elastic springs. Therefore, the force exerted on n -th atom in the lattice is given as

$$F_n = C(U_{n+1} - U_n) + C(U_{n-1} - U_n), \quad (6.1)$$

where C is the interatomic force (elastic) constant. Applying Newton's second law to the motion of the n -th atom we obtain

$$M \frac{d^2 U_n}{dt^2} = F_n = C(U_{n+1} - U_n) + C(U_{n-1} - U_n) = -C(2U_n - U_{n+1} - U_{n-1}), \quad (6.2)$$

where M is the mass of the atom. Note that we neglected here by the interaction of the n -th atom with all but its nearest neighbours. A similar equation should be written for each atom in the lattice, resulting in N coupled differential equations, which should be solved simultaneously (N is the total number of atoms in the lattice). In addition, the boundary conditions applied to the end atom in the lattice should be taken into account.

Now let us attempt a solution of the form

$$U_n = A e^{i(qx_n - \omega t)} \quad (6.3)$$

where x_n is the equilibrium position of the n -th atom so that $x_n = na$. This equation represents a travelling wave, in which all the atoms oscillate with the same frequency ω and the same amplitude A and have wavevector q . Note that a solution of the form (6.3) is only possible because of the translational symmetry of the lattice.

Now substituting Eq.(6.3) into Eq.(6.2) and cancelling the common quantities (the amplitude and the time-dependent factor) we obtain

$$M(-\omega^2)e^{iqna} = -C[2e^{iqna} - e^{iq(n+1)a} - e^{iq(n-1)a}] \quad (6.4)$$

This equation can be further simplified by cancelling the common factor e^{iqna} , which leads to

$$M\omega^2 = C(2 - e^{iqa} - e^{-iqa}) = 2C(1 - \cos qa) = 4C \sin^2 \frac{qa}{2}. \quad (6.5)$$

We find therefore the dispersion relation for the frequency

$$\omega = \sqrt{\frac{4C}{M} \left| \sin \frac{qa}{2} \right|}, \quad (6.6)$$

which is the relationship between the frequency of vibrations and the wavevector q . This dispersion relation has a number of important properties.

(i) Reducing to the first Brillouin zone. The frequency (6.6) and the displacement of the atoms (6.3) do not change when we change q by $q+2\pi/a$. This means that these solutions are physically identical. This allows us to set the range of independent values of q within the first Brillouin zone, i.e.

$$-\frac{\pi}{a} \leq q \leq \frac{\pi}{a}. \quad (6.7)$$

Within this range of q the ω versus q is shown in Figure 6.2.

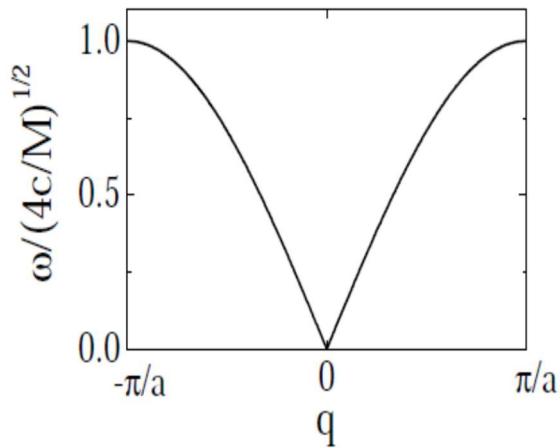


Figure 6.2 Dispersion curve in the first brillouin zone

The maximum frequency is $\sqrt{4C/M}$. The frequency is symmetric with respect to the sign change in q , i.e. $\omega(q) = \omega(-q)$. This is not surprising because a mode with positive q corresponds to the

wave travelling in the lattice from the left to the right and a mode with a negative q corresponds to the wave travelling from the right to the left. Since these two directions are equivalent in the lattice the frequency does not change with the sign change in q .

At the boundaries of the Brillouin zone $q = \pm\pi/a$ the solution represents a standing wave

$$U_n = A(-1)^n e^{-i\omega t}:$$

(ii) Phase and group velocity. The phase velocity is defined by

$$v_p = \frac{\omega}{q} \quad (6.8)$$

and the group velocity by

$$v_g = \frac{d\omega}{dq} \quad (6.9)$$

The physical distinction between the two velocities is that v_p is the velocity of the propagation of the plane wave, whereas the v_g is the velocity of the propagation of the wave packet. The latter is the velocity for the propagation of energy in the medium.

For the particular dispersion relation (6.6) the group velocity is given by

$$v_g = \sqrt{\frac{Ca^2}{M}} \cos \frac{qa}{2}. \quad (6.10)$$

As is seen from Eq. (6.10) the group velocity is zero at the edge of the zone where $q=\pm\pi/a$. Here the wave is a standing wave and therefore the transmission velocity for the energy is zero.

(iii) Long wavelength limit. The long wavelength limit implies that $\lambda \gg a$. In this limit $qa \ll 1$. Expanding the sine in Eq.(6.6) to obtain the positive frequencies:

$$\sin qa - \frac{(qa)^3}{3!} + \dots$$

for $qa \ll 1$, $\sin qa \approx qa$, then

$$\omega = \sqrt{\frac{c}{M}} qa . \quad (6.11)$$

We see that the frequency of vibration is proportional to the wave vector. This is equivalent to the statement that velocity is independent of frequency. In this case

$$v_p = \frac{\omega}{q} = \sqrt{\frac{c}{M}} a \quad (6.12)$$

This is the velocity of sound for the one dimensional lattice which is consistent with the expression we obtained earlier for elastic waves.

In-text question

1. Q: Briefly explain the harmonic approximation and use it to deduce the dispersion relation.

A: See page 80.

2. Q: Highlight the important properties of the dispersion relation.

A: See pages 82-84.

6.1.2 Diatomic 1D lattice

Now we consider a one-dimensional lattice with two non-equivalent atoms in a unit cell. It appears that the diatomic lattice exhibits important features different from the monoatomic case. Fig. 6.3 shows a diatomic lattice with the unit cell composed of two atoms of masses M_1 and M_2 with the distance between two neighbouring atoms a .

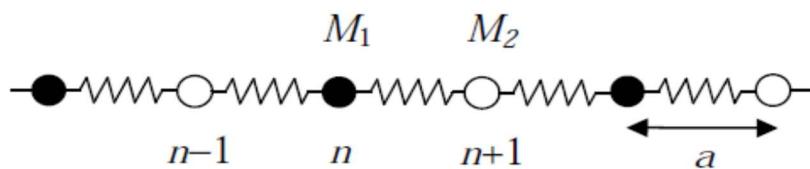


Figure 6.3 Diatomic lattice

We can treat the motion of this lattice in a similar fashion as for monoatomic lattice. However, in this case because there are two different kinds of atoms, we write two equations of motion:

$$M_1 \frac{d^2 u_n}{dt^2} = -C(2u_n - u_{n+1} - u_{n-1}) \quad (6.13)$$

$$M_2 \frac{d^2 u_{n+1}}{dt^2} = -C(2u_{n+1} - u_{n+2} - u_n)$$

In analogy with the monoatomic lattice we are looking for the solution in the form of travelling mode for the two atoms:

$$\begin{bmatrix} u_n \\ u_{n+1} \end{bmatrix} = \begin{bmatrix} A_1 e^{iqna} \\ A_2 e^{iq(n+1)a} \end{bmatrix} e^{-i\omega t}, \quad (6.14)$$

which is written in the obvious matrix form. Substituting this solution into Eq.(6.13) we obtain

$$\begin{bmatrix} 2C - M_1 \omega^2 & -2C \cos qa \\ -2C \cos qa & 2C - M_2 \omega^2 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = 0. \quad (6.15)$$

This is a system of linear homogeneous equations for the unknown quantities A_1 and A_2 . A nontrivial solution exists only if the determinant of the matrix is zero. This leads to the secular equation

$$(2C - M_1 \omega^2)(2C - M_2 \omega^2) - 4C^2 \cos^2 qa = 0. \quad (6.16)$$

This is a quadratic equation, which can be readily solved:

$$\omega^2 = C \left(\frac{1}{M_1} + \frac{1}{M_2} \right) \pm \sqrt{\left(\frac{1}{M_1} + \frac{1}{M_2} \right)^2 - \frac{4 \sin^2 qa}{M_1 M_2}}. \quad (6.17)$$

Depending on signs in this formula there are two different solutions corresponding to two different dispersion curves, as is shown in Figure 6.4:

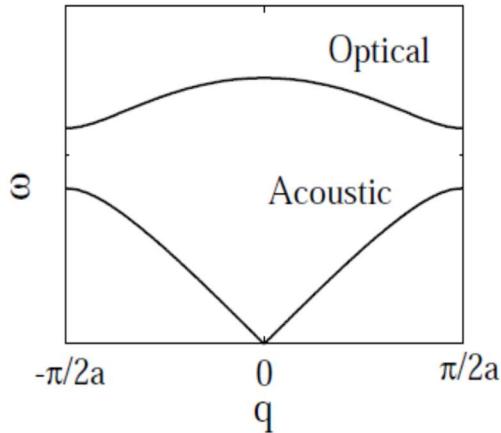


Figure 6.4 Dispersion curves

The lower curve is called the *acoustic branch*, while the upper curve is called the *optical branch*. The optical branch begins at $q = 0$ and $\omega = 0$. Then, with increasing q , the frequency increases in a linear fashion. This is why this branch is called acoustic: it corresponds to elastic waves or sound. Eventually this curve saturates at the edge of the Brillouin zone. On the other hand, the optical branch has a non-zero frequency at zero q .

$$\omega_0 = \sqrt{2C \left(\frac{1}{M_1} + \frac{1}{M_2} \right)}. \quad (6.18)$$

and does not change much with q .

The distinction between the acoustic and optical branches of lattice vibrations can be seen most clearly by comparing them at $q = 0$ (infinite wavelength). This follows from Eq. (6.15), for the acoustic branch $\omega = 0$ and $A_1 = A_2$. So in this limit the two atoms in the cell have the same amplitude and phase. Therefore, the molecule oscillates as a rigid body, as shown in Fig. 6.5 for the acoustic mode.

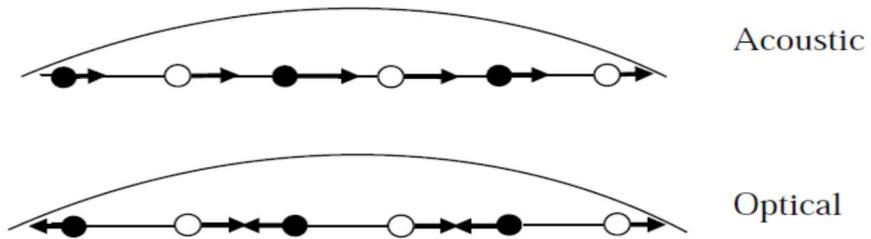


Figure 6.5

On the other hand, for the optical vibrations, substituting eq.(6.18) to eq.(6.15), we obtain for $q = 0$:

$$M_1 A_1 + M_2 A_2 = 0. \quad (6.19)$$

It implies that the optical oscillation takes place in such a way that the centre of mass of a molecule remains fixed. The two atoms move in out of phase as shown in Fig. 6.5. The frequency of these vibrations lies in the infrared region which is the reason for referring to this branch as optical.

In-text question

1. Q: Highlight the differences between the monoatomic and diatomic atoms as regards lattice vibration.
A: See pages 80-87.
2. Q: Explain briefly the dispersion curves observed in a diatomic lattice.
A: See pages 84-86.

6.1.3 Three Dimensional Lattice

These considerations for monoatomic and diatomic lattice can be extended to the three-dimensional lattice. To avoid mathematical details, we shall present only a qualitative discussion. Consider, first, the monatomic Bravais lattice, in which each unit cell has a single atom. The equation of motion of each atom can be written in a manner similar to that of Eq.(6.2). The

solution of this equation in three dimensions can be represented in terms of *normal modes*. Type equation here.

$$\mathbf{u} = \mathbf{A} e^{i(\mathbf{qr}-\omega t)} \quad (6.20)$$

where the wave vector \mathbf{q} specifies both the wavelength and direction of propagation. The vector \mathbf{A} determines the amplitude as well as the direction of vibration of the atoms. Thus this vector specifies the *polarization* of the wave, i.e., whether the wave is *longitudinal* (\mathbf{A} parallel to \mathbf{q}) or *transverse* (\mathbf{A} perpendicular to \mathbf{q}).

When we substitute Eq.(6.20) into the equation of motion, we obtain three simultaneous equations involving A_x , A_y , and A_z , the components of \mathbf{A} . These equations are coupled together and are equivalent to a 3×3 matrix equation. The roots of this equation lead to three different dispersion relations, or three dispersion curves, as shown in Fig. 6.6. All three branches pass through the origin, which means all the branches are acoustic. This is of course to be expected, since we are dealing with a *monatomic* Bravais lattice.

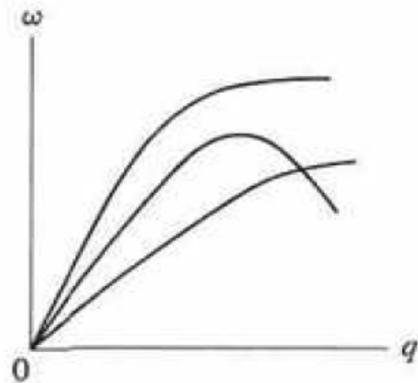


Figure 6.6 Dispersion curves

The three branches in Figure 6.6 differ in their polarization. When \mathbf{q} lies along a direction of high symmetry - for example, the [100] or [110] directions – these waves may be classified as either pure longitudinal or pure transverse waves. In that case, two of the branches are transverse and one is longitudinal. One usually refers to these as the **TA** - *transverse acoustic* and **LA** – *longitudinal acoustic* branches, respectively. However, along non-symmetry directions the waves may not be pure longitudinal or pure transverse, but have a mixed character.

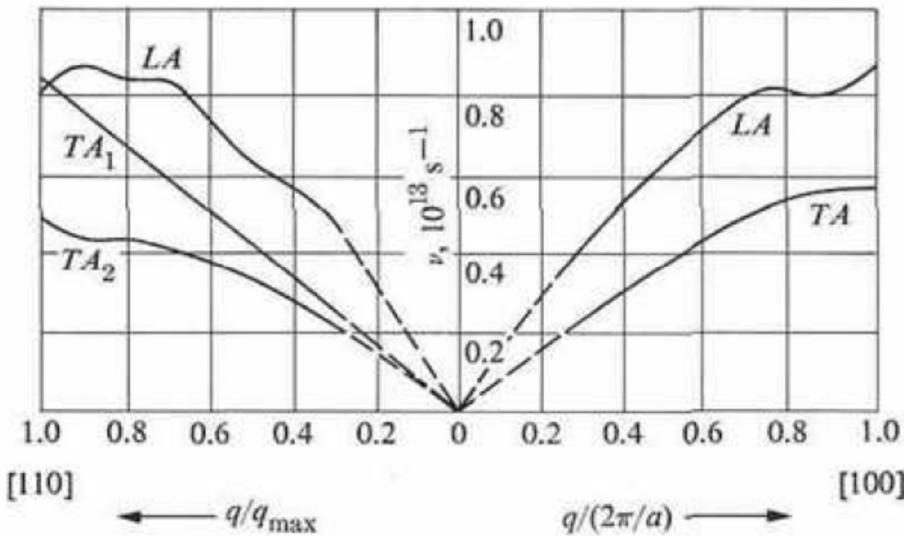


Figure 6.7

Figure 6.7 The dispersion curves for Al in the [100] and [110] directions.

In certain high-symmetry directions, such as the [100] in Al, the two transverse branches coincide. The branches are then said to be *degenerate*.

We turn our attention now to the non-Bravais three-dimensional lattice. Here the unit cell contains two or more atoms. If there are s atoms per cell, then on the basis of our previous experience we conclude that there are $3s$ dispersion curves. Of these, *three* branches are acoustic, and the remaining $(3s - 3)$ are optical. The mathematical justification for this assertion is as follows: We write the equation of motion for each atom in the cell, which results in s equations. Since these are vector equations, they are equivalent to $3s$ scalar equations, which have $3s$ roots. It can be shown that three of these roots always vanish at $q = 0$, which results in three acoustic branches. The remaining $(3s - 3)$ roots, therefore, belong to the optical branches, as stated above.

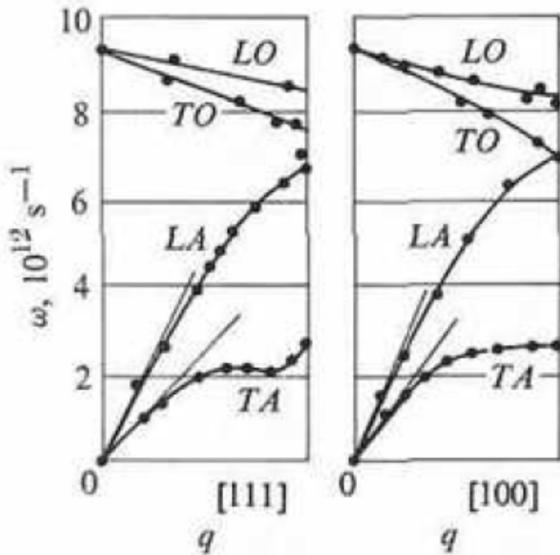


Figure 6.8 The dispersion curves for Ge

The acoustic branches may be classified, as before, by their polarizations as TA_1 , TA_2 , and LA . The *optical* branches can also be classified as longitudinal or transverse when q lies along a high symmetry direction, and one speaks of LO and TO branches. As in the one-dimensional case, one can also show that, for an optical branch, the atoms in the unit cell vibrate out of phase relative to each other. As an example of a non-Bravais lattice, the dispersion curves for Ge are shown in Fig. 6.8. Since there are two atoms per unit cell in germanium, there are six branches: three acoustic and three optical. Note that the two transverse branches are degenerate along the [100] direction, as indicated earlier.

6.1.4 Phonons

So far we discussed a classical approach to the lattice vibrations. As we know from quantum mechanics, the energy levels of the harmonic oscillator are quantized. Similarly the energy levels of lattice vibrations are quantized. The quantum of vibration is called a *phonon* in analogy with the photon, which is the quantum of the electromagnetic wave.

We know that the allowed energy levels of the harmonic oscillator are,

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega \quad (6.21)$$

where n is the quantum number. A normal vibration mode in a crystal of frequency ω is given by Eq.(6.20). If the energy of this mode is given by Eq.(6.21) we can say that this mode is occupied by n phonons of energy $\hbar\omega$. The term $\frac{1}{2}\hbar\omega$ is the zero point energy of the mode.

Let us now make a comparison between the classical and quantum solutions in the one-dimensional case. Consider a normal vibration

$\mathbf{q} \cdot \mathbf{r} = qx$ in this case

$$u = Ae^{i(qx - \omega t)} \quad (6.22)$$

where u is the displacement of an atom from its equilibrium position x and A is the amplitude. The energy of this vibrational mode averaged over time is

$$E = \frac{1}{2}M\omega^2A^2 = \left(n + \frac{1}{2}\right)\hbar\omega = E_n \quad (6.23)$$

We see that there is a relationship between the amplitude of vibration and the frequency and the phonon occupation of the mode. In classical mechanics any amplitude of vibration is possible, whereas in quantum mechanics only discrete values are allowed. This is illustrated in Figure 6.9

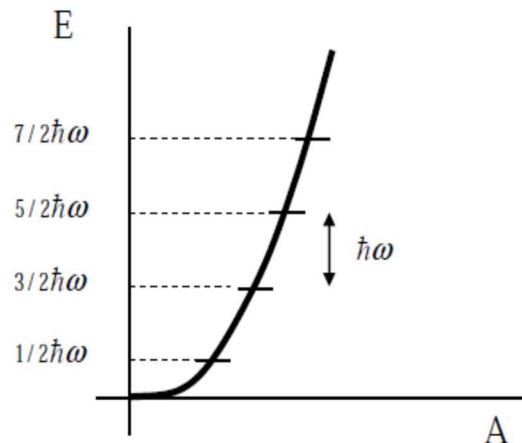


Figure 6.9 Amplitude of vibration in Quantum mechanics

The lattice with s atoms in a unit cell is described by $3s$ independent oscillators. The frequencies of normal modes of these oscillators will be given by the solution of $3s$ linear equations as we discussed before. They are $\omega_p(\mathbf{q})$, where p denotes a particular mode, i.e. $p = 1, \dots, 3s$. The energy of this mode is given by

$$E_{qp} = (n_{qp} + 1/2)\hbar \omega_p(\mathbf{q}). \quad (6.24)$$

where n_{qp} is the occupation number of the normal mode and is an integer. A vibrational state of the entire crystal is specified by giving the occupation numbers for each of the $3s$ modes. The total vibrational energy of the crystal is the sum of the energies of the individual modes, so that

$$E = \sum_{qp} E_{qp} = \sum_{qp} (n_{qp} + 1/2)\hbar \omega_p(\mathbf{q}) . \quad (6.25)$$

Phonons can interact with other particles such as photons, neutrons and electrons. This interaction occurs such as if photon had a momentum $\hbar\mathbf{q}$. However, a phonon does not carry real physical momentum. The reason is that the center of mass of the crystal does not change its position under vibrations (except at $q = 0$).

In crystals there exist selection rules for allowed transitions between quantum states. We saw that the elastic scattering of an x-ray photon by a crystal is governed by the wave vector selection rule $\mathbf{k}' = \mathbf{k} + \mathbf{G}$, where \mathbf{G} is a vector in the reciprocal lattice, \mathbf{k} is the wave vector of the incident photon and \mathbf{k}' is the wave vector of the scattered photon. This equation can be considered as the condition for the conservation of the momentum of the whole system, in which the lattice acquires a momentum $-\hbar\mathbf{G}$.

If the scattering of photon is inelastic and is accompanied by the excitation or absorption of a phonon the selection rule becomes

$$\mathbf{k}' = \mathbf{k} \pm \mathbf{q} + \mathbf{G} , \quad (6.26)$$

where sign (+) corresponds to creation of phonon and sign (–) corresponds to absorption of phonon.

Phonon dispersion relations $\omega_p(\mathbf{q})$ can be determined by the inelastic scattering of neutrons with emission or absorption of phonons. In this case in addition to the condition of momentum conservation, there is the requirement of conservation of energy. The latter condition can be written as

$$\frac{\hbar^2 k^2}{2M} = \frac{\hbar^2 k'^2}{2M} \pm \hbar\omega \quad (6.27)$$

where M is the mass of the neutron and $\hbar\omega$ and $\hbar k'$ are respectively the momenta of the incident and scattered

neutron. Once we know in experiment the kinetic energy of the incident and scattered neutrons from Eq.(6.27), we can determine the frequency of the emitted or absorbed phonon. Then experimentally, we need to determine those directions, which are characterized by the highest intensity of the scattered beam. For these directions the conditions (6.26) are satisfied and therefore from Eq.(6.26), we can find the wave vector of the phonon. Therefore, this is the way to obtain the dispersion conditions for the frequency of phonons, which we have discussed before.

In-text question

1. Q: Briefly explain what you understand by “degenerate branches” in three dimensional lattices.

A: See pages 88-89.

2. Q: Derive the phonon dispersion relations from the Neutron inelastic scattering method.

A: See pages 92-93.

Summary of Study Session 6

In Study Session 6, you have learnt that:

1. The quantum unit of a lattice vibration is a phonon. If the angular frequency is ω , the energy of the phonon is $\hbar\omega$.
2. When a phonon of wave vector \mathbf{K} is created by the inelastic scattering of a photon or neutron from wave vector \mathbf{k} to \mathbf{k}' , the wave vector selection rule that governs the process is $\mathbf{k} = \mathbf{k}' + \mathbf{K} + \mathbf{G}$, where \mathbf{G} is a reciprocal lattice vector.
3. All lattice waves can be described by wave vectors that lie within the first Brillouin zone in the reciprocal space.
4. If there are p atoms in the primitive, the phonon dispersion relation will have 3 acoustical phonon branches and $3p-3$ optical phonon branches.

Self Assessment Questions for Study Session 6

Now that you have completed this study session, you can assess how well you have achieved its learning outcomes by answering the following questions. Write your answers in your Study Diary and discuss them with your Tutor at the next Study Support Meeting.

SAQ 6.1

Using the dispersion relation for the monoatomic linear lattice of N atoms with nearest neighbour interactions, show that the density of vibrational modes is given by

$$D(\omega) = \frac{2N}{\pi} \frac{1}{\sqrt{\omega_m^2 - \omega^2}}, \text{ where } \omega_m \text{ is the maximum frequency.}$$

SAQ 6.2

Consider a dielectric crystal made up of layers of atoms with rigid coupling between layers so that the motion of atoms is restricted to the plane of the layer (i.e. 2D solid). Using the Debye approximation, obtain the expression for the thermal energy and show that the phonon heat capacity in the low temperature limit is proportional to T^2 .

SAQ 6.3

In the Debye approximation, show that the mean square displacement of an atom at absolute zero is $\langle R^2 \rangle = \frac{3\hbar\omega_D^2}{4\pi^2\rho v^3}$, where v is velocity of sound. Estimate this value for Cu ($\theta_D = \hbar\omega_D/k_B = 343^0K$, $\rho = 8920 \text{ kg/m}^3$, $v = 3570 \text{ m/s}$).

References/ Further Readings

Ashcroft N. W. and Mermin N.D., (1976), Solid state Physics, Holt, Rinehart, Winston.

Study Session 7 Free-Electron Theory of Metals

Introduction

A free electron model is the simplest way to represent the electronic structure of metals. Although the free electron model is a great oversimplification of the reality, surprisingly in many cases it works pretty well, so that it is able to describe many important properties of metals.

According to this model, the valence electrons of the constituent atoms of the crystal become conduction electrons and travel freely throughout the crystal. Therefore, within this model we neglect the interaction of conduction electrons with ions of the lattice and the interaction between the conduction electrons. In this sense we are talking about a *free electron gas*. However, there is a principal difference between the free electron gas and ordinary gas of molecules. First, electrons are charged particles. Therefore, in order to maintain the *charge neutrality* of the whole crystal, we need to include positive ions. This is done within the *jelly model*, according to which the positive charge of ions is smeared out uniformly throughout the crystal. This positive background maintains the charge neutrality but does not exert any field on the electrons. Ions form a uniform jelly in which electrons move.

The second important property of the free electron gas is that it should meet the Pauli exclusion principle, which leads to important consequences.

Learning Outcome for Study Session 7

When you have studied this session, you should be able to:

- 7.1 Distinguish between the classical and quantum approaches to the free electron theory.
- 7.2 Describe the treatment of electron dynamics as a free electron gas.

7.1 Main Contents

7.1.1 One dimension

We consider first a free electron gas in one dimension. We assume that an electron of mass m is confined to a length L by infinite potential barriers. The wave function $\Psi_n(x)$ of the electron is a

solution of the Schrödinger equation $\Psi_n(x) = E_n \Psi_n(x)$, where E_n is the energy of electron orbital. Since we can assume that the potential lies at zero, the Hamiltonian H includes only the kinetic energy so that

$$H\Psi_n(x) = \frac{p^2}{2m} \Psi_n(x) = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \Psi_n(x) = E_n \Psi_n(x). \quad (7.1)$$

Note that this is a one-electron equation, which means that we neglect the electron-electron interactions. We use the term *orbital* to describe the solution of this equation.

Since the $\Psi_n(x)$ is a continuous function and is equal to zero beyond the length L , the boundary conditions for the wave function are $\Psi_n(0) = \Psi_n(L) = 0$. The solution of Eq.(7.1) is therefore

$$\Psi_n(x) = A \sin\left(\frac{\pi n}{L} x\right)$$

where A is a constant and n is an integer. Substituting (7.2) into (7.1) we obtain for the Eigen values

$$E_n = \frac{\hbar^2}{2m} \left(\frac{\pi n}{L}\right)^2$$

These solutions correspond to standing waves with different number of nodes within the potential well as is shown in Fig.7.1.

The energy levels are labeled according to the quantum number n which gives the number of half-wavelengths in the wave function. The wavelengths are indicated on the wave functions.

Now we need to accommodate N valence electrons in these quantum states. According to the Pauli Exclusion Principle, no two electrons can have their quantum number identical. That is, each electronic quantum state can be occupied by at most one electron. The electronic state in a one-dimensional solid is characterized by two quantum numbers n and m_s , where n describes the orbital $\Psi_n(x)$, and m_s describes the projection of the spin momentum on a quantization axis. Electron spin is equal to $s = 1/2$, so that there $(2s+1) = 2$ possible spin states with $m_s = \pm 1/2$. Therefore, each orbital labelled by the quantum number n can accommodate two electrons, one with spin up and the other with spin down orientation. Let n_F denote the highest filled energy

level, where we start filling the levels from the bottom ($n = 1$) and continue filling higher levels with electrons until all N electrons are accommodated.

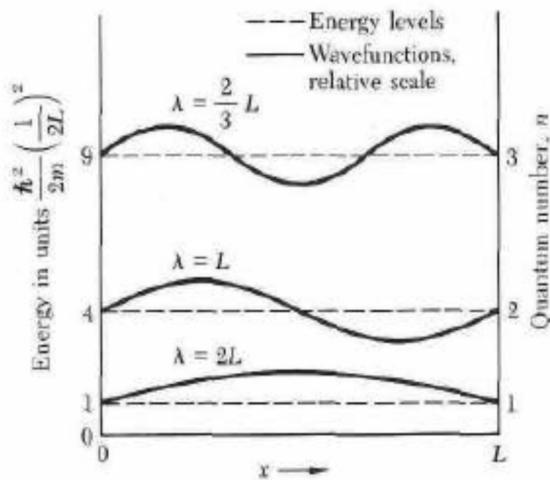


Fig.7.1 First three energy levels and wave-functions of a free electron of mass m confined to a line of length L

It is convenient to suppose that N is an even number. The condition $2n_F = N$ determines n_F , the value of n for the uppermost filled level. The energy of the highest occupied level is called the *Fermi energy* E_F . For the one-dimensional system of N electrons we find, using Eq. (7.3),

$$E_F = \frac{\hbar^2}{2m} \left(\frac{\pi N}{2L} \right)^2. \quad (7.4)$$

In metals the value of the Fermi energy is of the order of 5 eV. The ground state of the N electron system is illustrated in Fig.7.2a: All the electronic levels are filled up to the Fermi energy. All the levels above are empty.

In-text question

1. Q: Highlight the assumptions made in the free electron gas theory and deduce the solution to the Schrodinger's equation.

A: See pages 96-99.

7.1.2. The Fermi distribution

This is defined as the ground state of the N electron system at absolute zero. What happens if the temperature is increased? The kinetic energy of the electron gas increases with temperature. Therefore, some energy levels become occupied which were vacant at zero temperature, and some levels become vacant which were occupied at absolute zero. The distribution of electrons among the levels is usually described by the *distribution function*, $f(E)$, which is defined as the probability that the level E is occupied by an electron. Thus if the level is certainly empty, then, $f(E) = 0$, while if it is certainly full, then $f(E) = 1$. In general, $f(E)$ has a value between zero and unity.

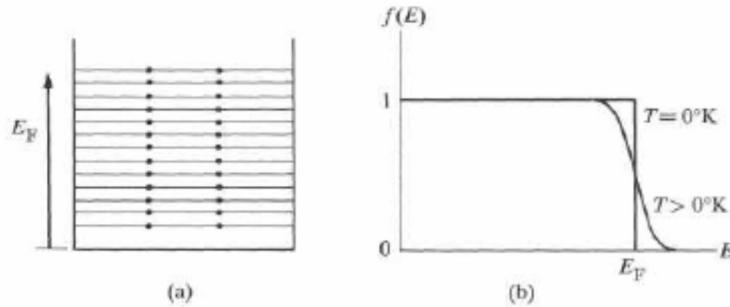


Fig. 7.2 (a) Occupation of energy levels according to the Pauli exclusion principle, (b) The distribution function $f(E)$, at $T = 0^{\circ}\text{K}$ and $T > 0^{\circ}\text{K}$.

It follows from the preceding discussion that the distribution functions for electrons at $T = 0^{\circ}\text{K}$ has the form

$$f(E) = \begin{cases} 1, & E < E_F \\ 0, & E > E_F \end{cases} \quad (7.5)$$

That is, all levels below E_F are completely filled, and all those above E_F are completely empty. This function is plotted in Fig. 7.2(b), which shows a discontinuity at the Fermi energy. When the system is heated ($T > 0^{\circ}\text{K}$), thermal energy excites the electrons. However, all the electrons do not share this energy equally, as would be the case in the classical treatment, because the electrons lying well below the Fermi level E_F cannot absorb energy. If they did, they would move to a higher level, which would be already occupied, and hence the exclusion principle would be violated.

Recall in this context that the energy which an electron may absorb thermally is of the order $K_B T$ ($= 0.025$ eV at room temperature), which is much smaller than E_F , this being of the order of 5 eV.

Therefore, only those electrons close to the Fermi level can be excited, because the levels above E_F are empty, and hence when those electrons move to a higher level there is no violation of the exclusion principle. Thus, only these electrons, which are a small fraction of the total number, are capable of being thermally excited.

The distribution function at non-zero temperature is given by the *Fermi distribution function*. The Fermi distribution function determines the probability that an orbital of energy E is occupied at thermal equilibrium

$$f(E) = \frac{1}{e^{(E-\mu)/k_B T} + 1} \quad (7.6)$$

This function is also plotted in Fig. 7.2(b), which shows that it is substantially the same as the distribution at $T = 0^{\circ}\text{K}$, except very close to the Fermi level, where some of the electrons are excited from below E_F to above Fermi level.

The quantity μ is called the chemical potential. The chemical potential can be determined in a way that the total number of electrons in the system is equal to N . At absolute zero, $\mu = E_F$.

In-text question

- Q: Explain briefly what happens to the physical properties of a metal when the temperature of the metal is increased.

A: See pages 99-100.

- Q: Write out the Fermi distribution function and explain what each of the term in the equation represents.

A: See pages 99-100.

7.1.3. Three dimensions

The Schrödinger equation in the three dimensions takes the form

$$H\psi(\mathbf{r}) = \frac{p^2}{2m}\psi(\mathbf{r}) = -\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{r}) = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\psi(\mathbf{r}) = E\psi(\mathbf{r}). \quad (7.7)$$

If the electrons are confined to a cube of edge L , the solution is the standing wave

$$\psi(\mathbf{r}) = A \sin\left(\frac{\pi n_x}{L}x\right) \sin\left(\frac{\pi n_y}{L}y\right) \sin\left(\frac{\pi n_z}{L}z\right), \quad (7.8)$$

where n_x , n_y , and n_z are positive integers.

In many cases, however, it is convenient to introduce periodic boundary conditions, as we did for lattice vibrations. The advantage of this description is that we assume that our crystal is infinite and disregard the influence of the outer boundaries of the crystal on the solution. We require then that our wave function is periodic in x , y , and z directions with period L , so that

$$\psi(x+L, y, z) = \psi(x, y, z), \quad (7.9)$$

and similarly for the y and z coordinates. The solution of the Schrödinger equation (7.7) which satisfies these boundary conditions has the form of the traveling plane wave:

$$\psi_{\mathbf{k}}(\mathbf{r}) = A \exp(i\mathbf{k} \cdot \mathbf{r}), \quad (7.10)$$

provided that the component of the wave vector \mathbf{k} are determined from

$$k_x = \frac{2\pi n_x}{L}; \quad k_y = \frac{2\pi n_y}{L}; \quad k_z = \frac{2\pi n_z}{L}, \quad (7.11)$$

where n_x , n_y , and n_z are positive or negative integers.

If we now substitute eq. (7.10) into Eq.(7.7), we will obtain for the energies of the orbital with the wave vector \mathbf{k}

$$E_{\mathbf{k}} = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2}{2m} (k_x^2 + k_y^2 + k_z^2). \quad (7.12)$$

The wavefunctions (7.10) are the eigenfunctions of the momentum $\mathbf{p} = -i\hbar\nabla$, which can be easily seen by differentiating (7.10) :

$$\mathbf{p}\psi_{\mathbf{k}}(\mathbf{r}) = -i\hbar\nabla\psi_{\mathbf{k}}(\mathbf{r}) = \hbar\mathbf{k}\psi_{\mathbf{k}}(\mathbf{r}). \quad (7.13)$$

The eigenvalue of the momentum is $\hbar\mathbf{k}$. The velocity of the electron is defined by
 $\mathbf{v} = \mathbf{p}/m = \hbar\mathbf{k}/m$.

In the ground state, a system of N electrons occupies states with lowest possible energies. Therefore, all the occupied states lie inside the sphere of radius K_F . The energy at the surface of this sphere is the Fermi energy E_F . The magnitude of the wave vector K_F and the Fermi energy are related by the following equation:

$$E_F = \frac{\hbar^2 k_F^2}{2m}. \quad (7.14)$$

The Fermi energy and the Fermi wave vector (momentum) are determined by the number of valence electrons in the system. In order to find the relationship between N and K_F , we need to count the total number of orbitals in a sphere of radius K_F which should be equal to N . There are two available spin states for a given set of k_x , k_y , and k_z . The volume in the \mathbf{k} space which is occupied by this state is equal to $(2p/L)^3$. Thus, in the sphere of $(4\pi k_F^3/3)$, the total number of states is

$$2 \frac{4\pi k_F^3 / 3}{(2\pi/L)^3} = \frac{V}{3\pi^2} k_F^3 = N, \quad (7.15)$$

where the factor 2 comes from the spin degeneracy. Then

$$k_F = \left(\frac{3\pi^2 N}{V} \right)^{1/3}, \quad (7.16)$$

which depends only on the electron concentration. We obtain, for the Fermi energy;

$$E_F = \frac{\hbar^2}{2m} \left(\frac{3\pi^2 N}{V} \right)^{2/3}. \quad (7.17)$$

and the Fermi velocity

$$v_F = \frac{\hbar}{m} \left(\frac{3\pi^2 N}{V} \right)^{1/3}. \quad (7.18)$$

A few estimates for Na: Na has bcc structure with cubic lattice parameter $a=4.2\text{\AA}$ and one valence electron per atom. Since there are 2 atoms in a unit cell, the electron concentration is $N/V = 2/(4.2\text{\AA}^3) = 3.10^{22}\text{cm}^{-3}$. Then, the Fermi momentum $K_F \approx (3.10 \cdot 3.10^{22}\text{cm}^3)^{1/3} \approx 10^8\text{cm}^{-1} = 1\text{\AA}^{-1}$.

The Fermi energy is given by $E_F = \frac{\hbar^2 K_F^2}{2m} = \frac{\hbar^2 K_F^2 a_0^2}{2m a_0^2} \approx 13.6eV \cdot 0.25 \approx 3.5eV$. The Fermi temperature is defined as $T_F = E_F/k_B \approx 40000K$ (1eV corresponds to $1.16 \cdot 10^4\text{K}$). The Fermi velocity can be found from $\frac{V_F}{c} = \frac{\hbar}{mc} k_F a_0 = \frac{\hbar}{mc} \frac{me^2}{\hbar^2} k_F a_0 \approx \frac{1}{137} 0.5 \approx 4 \cdot 10^{-3}$, which results in $V_F \approx 10^8\text{cms}^{-1}$. We see that the velocity of electrons is relatively large.

An important quantity which characterizes electronic properties of a solid is the *density of states*, which is the *number of electronic states per unit energy range*. To find it we use Eq.(7.17) and write the total number of orbitals of energy $< E$:

$$N(E) = \frac{V}{3\pi^2} \left(\frac{2mE}{\hbar^2} \right)^{3/2} \quad (7.19)$$

The density of states is then

$$D(E) = \frac{dN}{dE} = \frac{V}{2\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} E^{1/2} \quad (7.20)$$

or equivalently

$$D(E) = \frac{3N}{2E} \quad (7.21)$$

So within a factor of the order of unity, the number of states per unit energy interval at the Fermi energy, $D(E_F)$, is the total number of conduction electrons divided by the Fermi energy, just as we would expect.

The density of states is normalized in such a way that the integral

$$N = \int_0^{E_F} D(E)dE \quad (7.22)$$

gives the total number of electrons in the system. At non-zero temperature we should take into account the Fermi distribution function so that

$$N = \int_0^{\infty} D(E)f(E)dE, \quad (7.23)$$

This expression also determines the chemical potential.

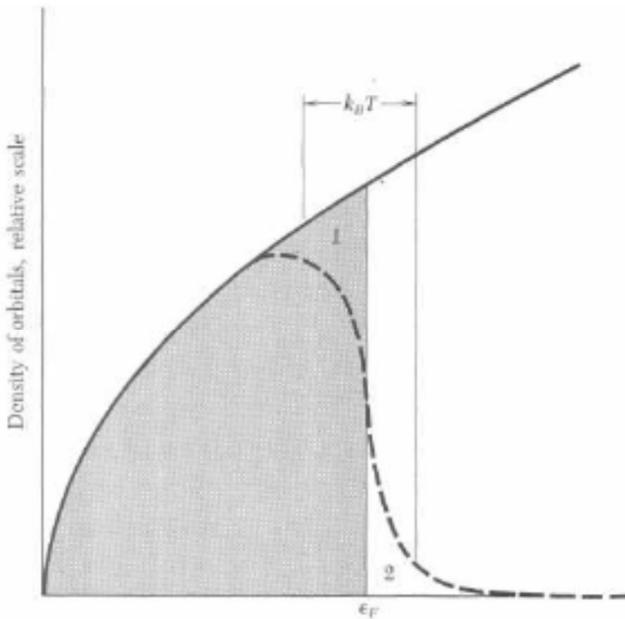


Figure 7.3 Density of single-particle states as a function of energy, for a free electron gas in three dimensions

The dashed curve represents the density $f(E, T)D(E)$ of filled orbitals at a finite temperature, but such that kT is small in comparison with E_F . The shaded area represents the filled orbitals at absolute zero. The average energy is increased when the temperature is increased from 0 to T , for electrons are thermally excited from region 1 to region 2.

In-text question

1. Q: What is the degree of degeneracy of the energy level $\left(\frac{38\hbar^2}{8ma^2}\right)$ of the particle in a cubical potential box of side ‘a’?

A: (6 1 1), (1 6 1), (1 1 6), (5 3 2), (5 2 3), (3 5 2), (3 2 5), (2 5 3), and (2 3 5) give the same energy level; therefore, the level is nine-fold degenerate.

2. Q: Find the energies of the six lowest energy levels of a particle in cubical box. Which of the levels are degenerate?

A: (1 1 1), (1 1 2), (1 2 2), (1 1 3), (2 2 2), (1 2 3)

The first and the fifth are non-degenerate; 2nd, 3rd and 4th are three-fold degenerate and the six-fold degenerate. Their energy values are:

$$\frac{3h^2}{8ma^2}, \frac{6h^2}{8ma^2}, \frac{9h^2}{8ma^2}, \frac{11h^2}{8ma^2}, \frac{12h^2}{8ma^2}, \frac{14h^2}{8ma^2}$$

Summary of Study Session 7

In this study session, you have learnt:

1. The consideration of free electron in 1D and solutions to its Eigen functions and energies were obtained.
2. About the changes observed in the electron gas when the temperature of the metal is increased and its Fermi distribution.
3. treating the free electron in 3D and its eigenfunctions and eigenvalues were obtained using Schrödinger equation.

Self Assessment Question for Study Session 7

Now that you have completed this study session, you can assess how well you have achieved its learning outcomes by answering the following questions. Write your answers in your Study Diary and discuss them with your Tutor at the next Study Support Meeting.

SAQ 7.1

Consider the free electron energy bands of an fcc crystal lattice in the reduced zone scheme in which all \mathbf{k} 's are transformed to lie in the first Brillouin zone. Plot roughly in the [111] direction the energies of all bands up to six times the lowest band energy at the zone boundary at $\mathbf{k} = (2\pi/a)(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. Explain what happens with these bands in the presence of a weak crystal potential.

SAQ 7.2

Suppose the crystal potential in a one-dimensional lattice of lattice constant a is composed of a series of rectangular wells which surround the atom. Suppose the depth of each well is U_0 and its width $a/5$.

- a. Calculate the values of the first three energy gaps. Compare the magnitudes of these gaps.

b. Evaluate these gaps for the case of $U_0 = 5$ eV and $a = 4\text{\AA}$.

SAQ 7.3

Using the solution for the energy bands near the zone boundary in the presence of a weak crystal potential, show that the electron velocity is parallel to the Bragg plane. Since the gradient is perpendicular to the surfaces on which a function is constant, this fact allows us to conclude that the constant energy surfaces (such as the Fermi surface) at the Bragg plane are perpendicular to that plane.

References/ Further Readings

Ashcroft N. W. and Mermin N.D., (1976), Solid state Physics, Holt, Rinehart, Winston.

Study Session 8 Metals: Electron Dynamics and Fermi Surfaces

Introduction

The next important subject we address is electron dynamics in metals. Our consideration will be based on a semi-classical model. The term “semi-classical” comes from the fact that within this model the electronic structure is described quantum-mechanically but electron dynamics itself is considered in a classical way, i.e., using classical equations of motion. Within the semi-classical model we assume that we know the electronic structure of metal, which determines the energy band as a function of the wave vector. The aim of the model is to relate the band structure to the transport properties as a response to the applied electric field.

Learning Outcomes for Study Session 8

When you have studied this session, you should be able to:

- 8.1 Determine the energy band as a function of wave vector using the semiclassical model.
- 8.2 Relate the band structure to the transport properties due to an applied electric field.

8.1 Main Content

8.1.1 Electron Dynamics

Given the functions $E_n(k)$, the semi-classical model associates with each electron, a position, a wave vector and a band index n . In the presence of applied fields, the position, the wave vector, and the index are taken to evolve according to the following rules:

- (1) The band index is a constant of the motion. The semiclassical model ignores the possibility of interband transitions. This implies that within this model, it is assumed that the applied electric field is small.
- (2) The time evolution of the position and the wave vector of an electron with band index n are determined by the equations of motion:

$$\frac{dr}{dt} v_n(k) = \frac{1}{\leftarrow} \frac{dE_n(k)}{dK} \quad (8.1)$$

$$\leftarrow \frac{dk}{dt} = F(r, t) = -eE(r, t) \quad (8.2)$$

Clearly, Eq.(8.2) has to be proved. It is identical to the Newton's second law if we assume that the electron momentum is equal to $\leftarrow k$. The fact that electrons belong to particular bands makes their movement in the applied electric field different from that of free electrons.

For example, if the applied electric field is independent of time, according to Eq. (8.2) the wave vector of the electron increases uniformly with time.

$$k(t) = k(0) - \frac{eEt}{\leftarrow} \quad (8.3)$$

Since velocity and energy are periodic in the reciprocal lattice, the will exhibit oscillation. This is in striking contrast to the free electron case, where v is proportional to k and grows linearly with time.

The k dependence (and, to within a scale factor, the t dependence) of the velocity is illustrated in Fig.8.1, where both $E(k)$ and $v(k)$ are plotted in one dimension. Although the velocity is linear in k near the band minimum, it reaches a maximum as the zone boundary is approached, and then drops back down, going to zero at the zone edge. In the region between the maximum of v and the zone edge, the velocity actually decreases with increasing k , so that the acceleration of the electron is opposite the externally applied electric force.

This extraordinary behavior is a consequence of the additional force exerted by the periodic potential, which is included in the functional form of $E(k)$. As an electron approaches a Bragg plane, the external electric field moves it in the opposite direction due to the Bragg-reflection.

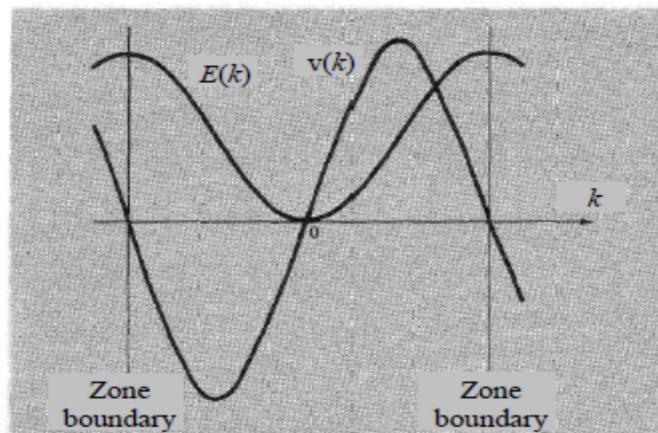


Fig.8.1. $E(k)$ and $v(k)$ vs. k in one dimension (or three dimensions, in a direction parallel to a reciprocal lattice vector that determines one of the first-zone faces)

Effective Mass

When discussing electron dynamics in solids, it is often convenient to introduce the concept of effective mass. If we differentiate Eq. (8.1) with respect to time, we find that

$$\frac{dv}{dt} = \frac{1}{\leftarrow} \frac{d^2E}{dk dt} = \frac{1}{\leftarrow} \frac{d^2E}{dk^2} \frac{dk}{dt}, \quad (8.4)$$

where the second derivative with respect to a vector should be understood as a tensor. Using Eq. (8.2) we find that

$$\frac{dv}{dt} = \frac{1}{\leftarrow} \frac{d^2E}{dk^2} \mathbf{F}, \quad (8.5)$$

In the one dimensional case, this reduces to

$$\frac{dv}{dt} = \frac{1}{\leftarrow} \frac{d^2E}{dk^2} \mathbf{F}. \quad (8.6)$$

This has the same form as the Newton's second law, provided that we define an effective mass by the relation:

$$\frac{1}{m^*} = \frac{1}{\leftarrow} \frac{d^2E}{dk^2}. \quad (8.7)$$

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 (Fig. 8.2).

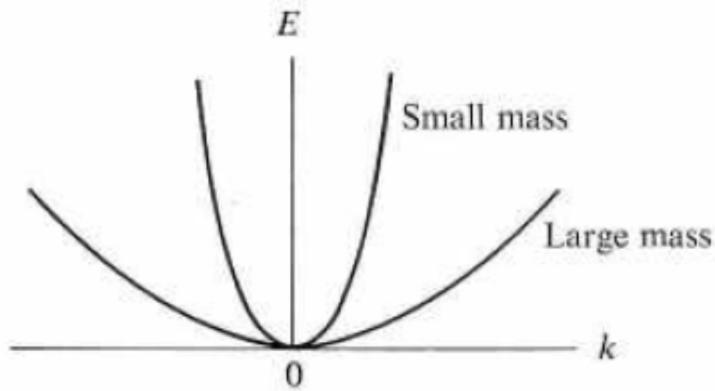


Fig. 8.2 The inverse relationship between the mass and the curvature of the energy band

In a general case, the effective mass is a tensor which is defined by

$$\left(\frac{1}{m^*}\right)_{\mu\nu} = \frac{1}{\leftarrow^2} \frac{d^2E}{dk_\mu dk_\nu}, \quad (8.8)$$

Where k_μ and k_ν are Cartesian coordinates.

The effective mass can be different depending on the directions in the crystal.

Current Density

The current density within a free electron model is defined as $\mathbf{j} = -en\mathbf{v}$, where n is the number of valence electrons per unit volume, and \mathbf{v} is the velocity of electrons. This expression can be generalized to the case of Bloch electrons. In this case, the velocity depends on the wave vector and we need to sum up over \mathbf{k} vectors for which there are occupied states available:

$$\mathbf{j} = \frac{-e}{V} \sum_{k, \text{occupied}} V(k) \quad (8.9)$$

Here the sum is performed within the extended zone scheme and V is the volume of the solid. It is often convenient to replace the summation by an integral. The volume of \mathbf{k} -space per allowed \mathbf{k} value is $\Delta k = 8\pi^3/V$, we can therefore, write the sum over \mathbf{k} as

$$\sum_k = \frac{V}{8\pi^3} \int dk. \quad (8.10)$$

Taking into account the spin degeneracy, we obtain for the current density:

$$\mathbf{j} = -e \int_{\text{occupied}} \frac{dk}{4\pi^3} V(k). \quad (8.11)$$

Using this expression we show now that *completely filled bands do not contribute to the current*. For the filled bands Eq. (8.11) should be replaced by

$$\mathbf{j} = -e \int_{\text{zone}} \frac{dk}{4\pi^3} V(k) \frac{dE(k)}{dk}, \quad (8.12)$$

This vanishes as a consequence of the theorem that the integral over any primitive cell of the gradient of a periodic function must vanish.

Proof:

Let $f(\mathbf{r})$ be any function with the periodicity of the lattice. The integral over the primitive cell

$$I(\mathbf{r}') = \int_{\text{cell}} dr f(\mathbf{r} + \mathbf{r}'), \quad (8.13)$$

is independent of \mathbf{r}' . Therefore,

$$0 = \frac{dI(r')}{dr'} = \int_{cell} dr \frac{df(r+r')}{dr'} = \int_{cell} dr \frac{df(r+r')}{dr}. \quad (8.14)$$

At point $r' = 0$ this results in

$$\int_{cell} dr \frac{df(r)}{dr} = 0. \quad (8.15)$$

Since the Brillouin zone is a primitive cell in the reciprocal space, the integral (8.12) vanishes. This implies that filled bands do not contribute to the current. Only partially filled bands need be considered in calculating the electronic properties of a solid. This explains why the Drude's theory assumption is often successful: in many cases, those bands derived from the atomic valence electrons are the only ones that are partially filled.

In-text question

1. Q: State the rules that the position wave vector and index evolve through in the presence of an applied field.
A: See pages 108-109.
2. Q: Show that completely filled bands do not contribute to the current in the presence of an applied field.
A: See pages 111-112.

8.1.2 Hole

One of the most impressive achievements of the semiclassical model is its explanation for the phenomena on that free electron theory can account for only if the carriers have a positive charge. We now introduce the concept of a hole.

The contribution of all the electrons in a given band to the current density is given by Eq. (8.11), where the integral is over all occupied levels in the band. By exploiting the fact that a completely filled band carries no current,

$$0 = \int_{zone} \frac{dk}{4\pi^3} V(k) = \int_{occupied} \frac{dk}{4\pi^3} V(k) + \int_{unoccupied} \frac{dk}{4\pi^3} V(k). \quad (8.16)$$

we can equally write Eq. (8.11), in the form:

$$\mathbf{j} = +e \int_{unoccupied} \frac{dk}{4\pi^3} V(k). \quad (8.17)$$

Thus the current produced by electrons occupying a specified set of levels in a band is precisely the same as the current that would be produced if the specified levels were unoccupied and all other levels in the band were occupied but with particles of charge $+e$ (opposite to the electronic charge).

Therefore, even though the only charge carriers are electrons, we may, whenever it is convenient, consider the current to be carried entirely by fictitious particles of positive charge that fill all those levels in the band that are unoccupied by electrons. The fictitious particles are called *holes*. It must be emphasized that particles cannot be mixed within a given band. If one wishes to regard electrons as carrying the current, then the unoccupied levels make no contribution; if one wishes to regard the holes as carrying the current, then the electrons make no contribution. One may, however, regard some bands using the electron picture and other bands using the hole picture, at one's convenience.

Normally, it is convenient to consider transport of the holes for the bands which are almost occupied, so that only a few electrons are missing. This happens in semiconductors in which a few electrons are excited from the valence to the conduction bands. Similar to electrons, we can introduce the effective mass for the holes. It has a negative sign.

Fermi surface

The ground state of N Bloch electrons is constructed in a similar fashion as that for free electrons, i.e., by occupying all one-electron energy levels with band energies $E_n K$ less than E_F , where E_F is determined by requiring the total number of levels with energies less than E_F to be equal to the total number of electrons. The wave vector \mathbf{k} must be confined to a single primitive cell of the reciprocal lattice. When the lowest of these levels are filled by a specified number of electrons, two quite distinct types of configuration can result:

1. A certain number of bands may be completely filled, all others remaining empty. Because the number of levels in a band is equal to the number of primitive cells in the crystal (and because each level can accommodate two electrons (one of each spin), a configuration with a band gap can arise only if the number of electrons per primitive cell is even.

2. A number of bands may be partially filled. When this occurs, the energy of the highest occupied level, the Fermi energy E_F , lies within the energy range of one or more bands. For each partially filled band, there will be a surface in \mathbf{k} -space separating the occupied from the unoccupied levels. The set of all such surfaces is known as the Fermi surface, and is the generalization to Bloch electrons of the free electron Fermi sphere. The parts of the Fermi surface arising from individual partially filled bands are known as branches of the Fermi surface. Analytically, the branch of the Fermi surface in the n -th band is that surface in \mathbf{k} -space determined by

$$E_n K = E_F \quad (8.18)$$

Thus, the Fermi surface is a constant energy surface (surfaces) in \mathbf{k} -space.

Since the $E_n K$ are periodic in the reciprocal lattice, the complete solution to Eq.(8.18) for each n is a \mathbf{k} -space surface with the periodicity of the reciprocal lattice. When a branch of the Fermi surface is represented by the full periodic structure, it is said to be described in a *repeated zone scheme*. Often, however, it is preferable to take just enough of each branch of the Fermi surface so that every physically distinct level is represented by just one point of the surface. This is achieved by representing each branch by that portion of the full periodic surface contained within a single primitive cell of the reciprocal lattice. Such a representation is described as a *reduced zone scheme*. The primitive cell chosen is often, but not always, the first Brillouin zone.

Brillouin zones

We consider now an example of building of a Fermi surface. We start from considering the Fermi surface for free electrons and then investigating the influence of the crystal potential. The Fermi surface for free electrons is a sphere centered at $\mathbf{k} = 0$. To construct the Fermi surface in the reduced-zone scheme, one can translate all the pieces of the sphere into the first zone through reciprocal lattice vectors. This procedure is made systematic through the geometrical notion of the higher Brillouin zones.

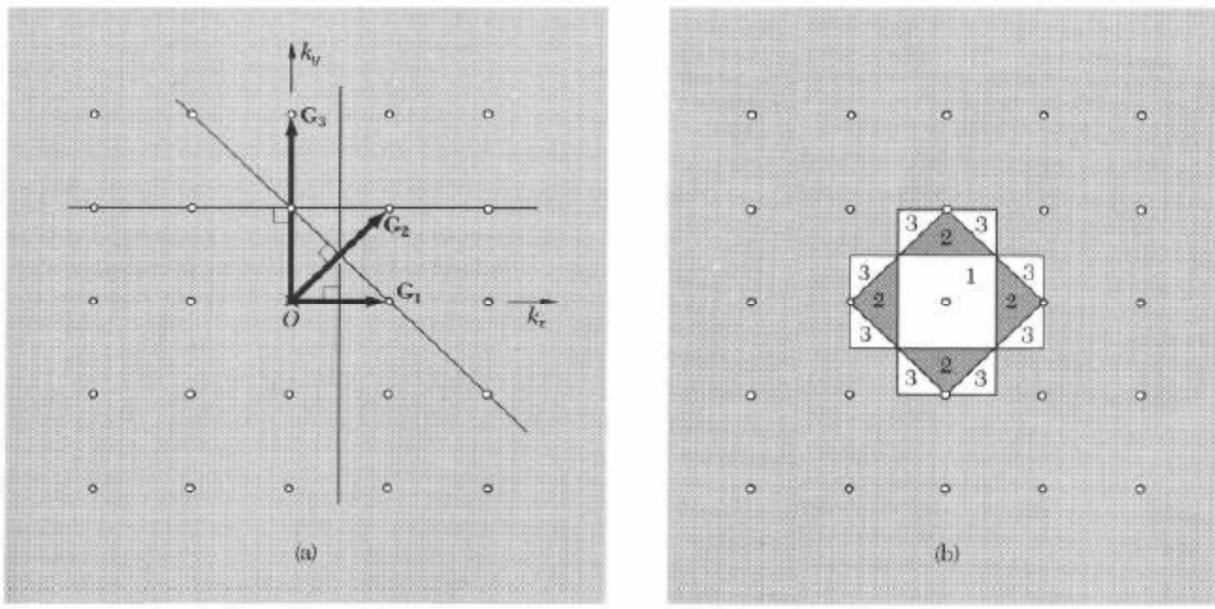


Figure 8.3 (a) Construction in k space of the first three Brillouin zones of a square lattice. The three shortest forms of the reciprocal lattice vectors are indicated as \mathbf{G}_1 , \mathbf{G}_2 , and \mathbf{G}_3 . The lines drawn are perpendicular bisectors of these \mathbf{G} 's.

(b) On constructing all lines equivalent by symmetry to the three lines in (a) we obtain the regions in k space which form the first three Brillouin zones. The numbers denote the zone to which the regions belong; the numbers here are ordered according to the length of vector \mathbf{G} involved in the construction of the outer boundary of the region.

We illustrate this construction for the two dimensional cubic lattice shown in Fig.8.3. Recall that the boundaries of the Brillouin zones are planes normal to \mathbf{G} at the midpoint of \mathbf{G} . The first Brillouin zone of the square lattice is the area enclosed by the perpendicular bisectors of \mathbf{G}_1 and of the three reciprocal lattice vectors equivalent by symmetry to \mathbf{G}_1 in Fig. 8.3a. These four reciprocal lattice vectors are $\pm(2\pi/a)\hat{k}_x$ and $\pm(2\pi/a)\hat{k}_y$.

The second zone is constructed from \mathbf{G}_2 and the three vectors equivalent to it by symmetry, and similarly for the third zone. The pieces of the second and third zones are drawn in Fig. 8.3b.

In general, the *first Brillouin zone* is the set of points in k -space that can be reached from the origin without crossing *any* Bragg plane. The *second Brillouin zone* is the set of points that can be reached from the first zone by crossing only one Bragg plane. The $(n+1)$ -th *Brillouin zone* is the set of points not in the $(n-1)$ -th zone that can be reached from the n -th zone by crossing only one Bragg plane.

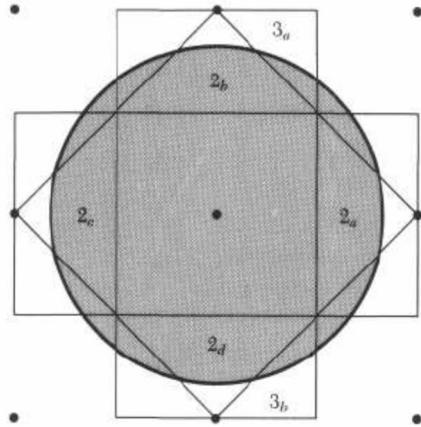


Figure 8.4 Brillouin zones of a square lattice in two dimensions.

The circle shown is a surface of constant energy for free electrons; it will be the Fermi surface for some particular value of the electron concentration. The total area of the filled region in K space depends only on the electron concentration and is independent of the interaction of the electrons with the lattice. The shape of the Fermi surface depends on the lattice interaction and the shape will not be an exact circle in an actual lattice.

The free electron Fermi surface for an arbitrary electron concentration is shown in Fig. 8.4. Now we perform a transformation to the reduced zone scheme as is shown in Figs. 8.5 and 8.6. We take the triangle labeled 2a and move it by a reciprocal lattice vector $G = -(2\pi/a)\hat{k}_x$ such that the triangle reappears in the area of the first Brillouin zone (Figure 8.5). Other reciprocal lattice vectors will shift the triangles 2b, 2c, 2d to other parts of the first zone, completing the mapping of the second zone into the reduced zone scheme. The parts of the Fermi surface falling in the second zone are now connected, as shown in Fig. 8.6.

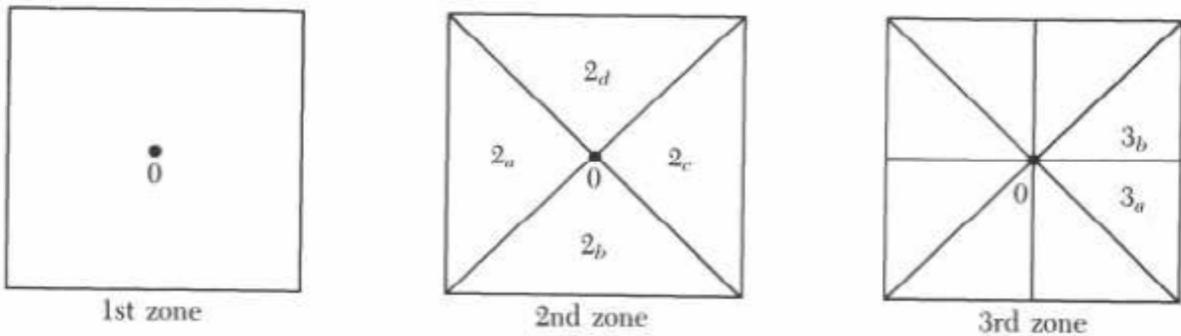


Figure 8.5 Mapping of the first, second, and third Brillouin zones in the reduced zone scheme.

The sections of the second zone in Fig. 8.5 are put together into a square by translation through an appropriate reciprocal lattice vector.

A different \mathbf{G} is needed for each piece of a zone.

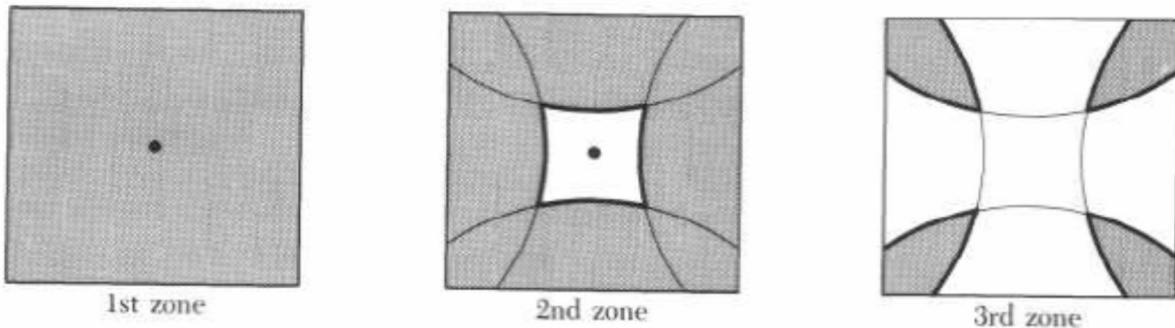


Fig.8.6 The free electron Fermi surfaces of Fig. 8.4, as viewed in the reduced zone scheme.

The shaded areas represent occupied electron states. Parts of the Fermi surface fall into the second, third and fourth zones. The fourth zone is not shown. The first zone is entirely occupied. The construction of Brillouin zones and Fermi surfaces in three-dimensions is more complicated. Fig.8.7 shows the first three Brillouin zones for bcc and fcc structures.

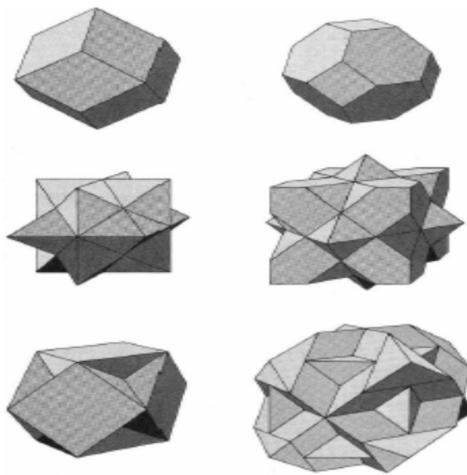


Figure 8.7 Surfaces of the first, second and third Brillouin zones for (a) body-centered cubic and (b) face-centered cubic crystals. (Only the exterior surfaces are shown). The interior surface of the nth zone is identical to the exterior surface of the (n-1) th zone. Evidently the surfaces bounding the zones become increasingly complex as the zone number increases.

The free electron Fermi surfaces for fcc cubic metals of valence 2 and 3 are shown in Fig.8.8.

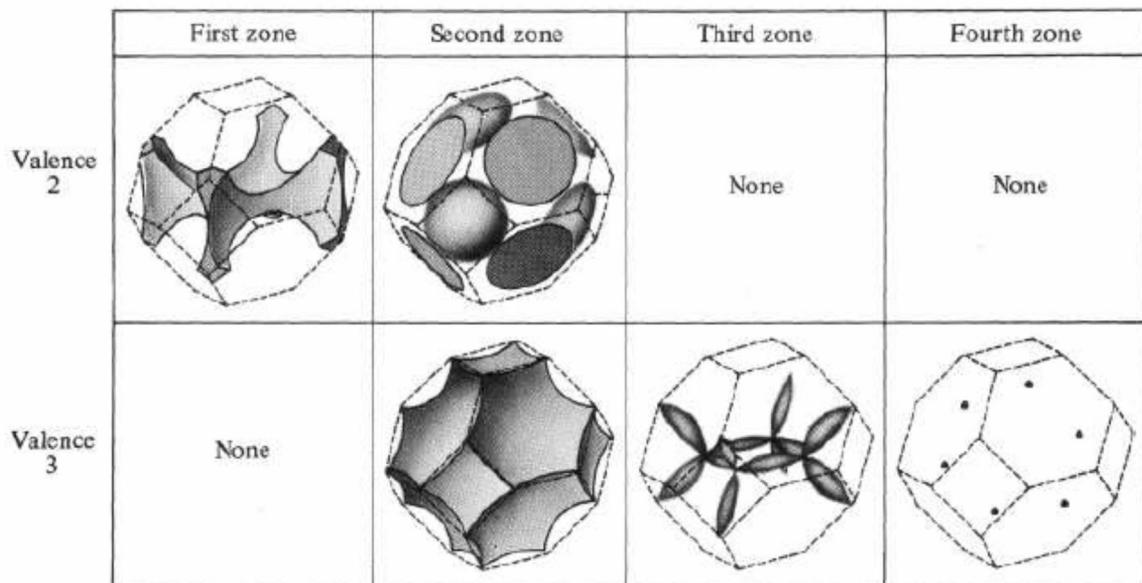


Fig.8.8 The free electron Fermi surfaces for face-centered cubic metals of valence 2 and 3.

For valence 1 the surface lies entirely within the interior of the first zone and therefore remains a sphere to lowest order. All branches of the Fermi surface are shown. The primitive cells in which they are displayed have the shape and orientation of the first Brillouin zone.

Effect of a crystal potential

How do we go from Fermi surfaces for free electrons to Fermi surfaces in the presence of a weak crystal potential? We can make approximate constructions freehand by the use of four facts:

- (i) The interaction of the electron with the periodic potential of the crystal causes energy gaps at the zone boundaries.
- (ii) Almost always, the Fermi surface will intersect zone boundaries perpendicularly. Using the equation for the energy near the zone boundary, it is easy to show that $\frac{dE}{dk} = \frac{-^2}{m}(k - \frac{1}{2}G)$, which implies that on the Bragg plane the gradient of energy is parallel to the Bragg plane. Since the gradient is perpendicular to the surfaces on which function $E_n(k)$ is constant, the constant energy surfaces at the Bragg plane are perpendicular to the plane.
- (iii) The crystal potential will round out sharp corners in the Fermi surfaces.
- (iv) The total volume enclosed by the Fermi surface depends only on the electron concentration and is independent of the details of the lattice interaction.

If a branch of the Fermi surface consists of very small pieces of surface (surrounding either occupied or unoccupied levels, known as "pockets of electrons" or "pockets of holes"), then a weak periodic potential may cause these to disappear. In addition, if the free electron Fermi surface has parts with a very narrow cross section, a weak periodic potential may cause it to become disconnected at such points.

In-text question:

1. Q: State the assumption made as regards the holes in conduction through metals.

A: See pages 112-113.

2. Q: State the distinct type of configuration that can result when the lowest of the energy levels are filled by a specified number of electrons on a Fermi surface.

A: See page 113-115

3. Q: How do you construct a Fermi surface in the reduced zone scheme?

A: See page 116-118.

Below are few examples for real metals.

8.1.3. Alkali metals

The radius of the Fermi sphere in bcc alkali metals is less than the shortest distance from the center of the zone to a zone face and therefore the Fermi sphere lies entirely within the first Brillouin zone. The crystal potential does not distort much the free electron Fermi surface and it remains very similar to a sphere.

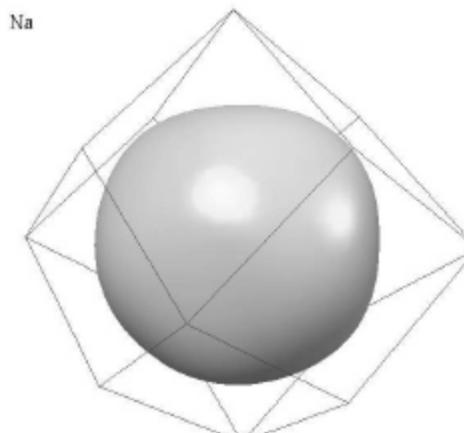


Figure 8.9 Fermi Surface of sodium

8.1.4. The noble metals

The Fermi surface for a single half-filled free electron band in an fcc Bravais lattice is a sphere entirely contained within the first Brillouin zone, approaching the surface of the zone most closely in the [111] directions, where it reaches 0.903 of the distance from the origin to the center of the hexagonal face. For all three noble metals therefore, Fermi surfaces are closely related to the free electron sphere. However, in the [111] direction contact is actually made with the zone faces, and the measured Fermi surfaces have the shape shown in Fig.8.10. Eight "necks" reach out to touch the eight hexagonal faces of the zone, but otherwise the surface is not grossly distorted from spherical.

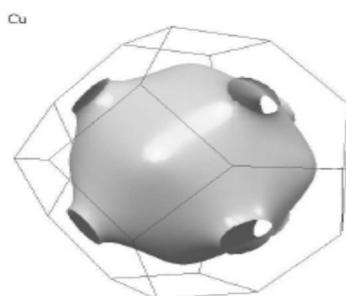


Figure 8.10 In the three noble metals the free electron sphere bulges out in the [111] directions to make contact with the hexagonal zone faces

The Fermi surface of aluminium is close to that of the free electron surface for fcc cubic monoatomic lattice with three conduction electrons per atom. The first Brillouin zone is filled and the Fermi surface of free electrons is entirely contained in the second, third and fourth Brillouin zones. When displayed in a reduced-zone scheme the second-zone surface is a closed structure containing unoccupied levels, while the third-zone surface is a complex structure of narrow tubes (Fig.8.8). The amount of surface in the fourth zone is very small, enclosing tiny pockets of occupied levels.

The effect of a weak periodic potential is to eliminate the fourth-zone pockets of electrons, and reduce the third-zone surface to a set of disconnected "rings" (Fig.8.11).

Aluminum provides a striking illustration of the theory of Hall coefficients. The high-field Hall coefficient should be $R_H = -1/(n_e - n_h) e$, where n_e and n_h are the number of levels per unit volume enclosed by the electron-like and hole-like branches of the Fermi surface. Since the first zone of aluminum is completely filled and accommodates two electrons per atom, one of the three valence electrons per atom remains to occupy second- and third-zone levels. Thus,

$$n_e^{II} + n_e^{III} = \frac{n}{3} \quad (8.19)$$

where n is the free electron carrier density appropriate to valence 3. On the other hand, since the total number of levels in any zone is enough to hold two electrons per atom, we also have

$$n_e^{II} + n_h^{II} = 2 \frac{n}{3} \quad (8.20)$$

Subtracting (8.20) from (8.19) gives

$$n_e^{III} - n_h^{II} = -\frac{n}{3} \quad (8.21)$$

Thus, the high-field Hall coefficient should have a positive sign and yield an effective density of carriers a third of the free electron value. This is precisely what is observed.

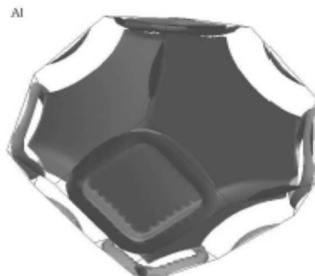


Figure 8.11 Fermi surface of Aluminum

In-text question

1. Q: Briefly explain the Fermi sphere for Alkali metals.

A: See page 120

Summary of Study Session 8

In this study session 8, you have learnt:

1. that conduction takes place through holes aside from electrons.
2. the formation and effect of brillouin zones.
3. the effect of crystal potentials on the Fermi surface.

Self-Assessment Questions for Study Session 8

Now that you have completed this study session, you can assess how well you have achieved its learning outcomes by answering the following questions. Write your answers in your Study Diary and discuss them with your Tutor at the next Study Support Meeting.

SAQ 8.1

What are Brillouin zones? How are they related to the energy of an electron in a metal?

SAQ 8.2

Explain the significance of Brillouin zones with particular reference to any cubic lattice.

SAQ 8.3

Illustrate the three Brillouin zones for a two dimensional square lattice.

References/ Further Readings

Ashcroft N. W. and Mermin N.D., (1976), Solid state Physics, Holt, Rinehart, Winston.

Study Session 9 Energy Bands

Introduction

The free electron model gives us a good insight into many properties of metals, such as the heat capacity, thermal conductivity and electrical conductivity. However, this model fails to help us with other important properties. For example, it does not predict the difference between metals, semiconductors and insulators. It does not explain the occurrence of positive values of the Hall coefficient. Also, the relation between conduction electrons in the metal and the number of valence electrons in free atoms is not always correct. We need a more accurate theory, which would be able to answer these questions.

Learning Outcomes for Study Session 9

When you have studied this session, you should be able to:

- 9.1 Differentiate metals, semiconductors and insulators from one another.
- 9.2 Distinguish between the free electron model and the many electron problem.
- 9.3 Calculate current density and life-time of holes and electrons.

9.1 Main Content

9.1.1 Electron in Periodic Field of Crystal

The problem of electrons in a solid is in general a many-electron problem. The full Hamiltonian of the solid contains not only the one-electron potentials describing the interactions of the electrons with atomic nuclei, but also pair potentials describing the electron-electron interactions. The many-electron problem is impossible to solve exactly and therefore we need simplified assumptions. The simplest approach we have already considered is a free electron model. The next step in building the complexity is to consider an independent electron approximation, assuming that all the interactions are described by an effective potential. One of the most important properties of this potential is that it is periodic on a lattice

$$U(r) = U(r + \mathbf{T}), \quad (9.1)$$

where \mathbf{T} is a lattice vector. Qualitatively, a typical crystalline potential might be expected to have a form shown in Fig.9.1, resembling the individual atomic potentials as the ion is approached closely and flattening off in the region between ions.

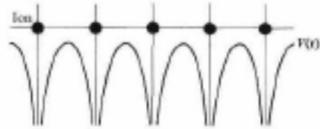


Fig. 9.1 The crystal potential seen by the electron.

Within the approximation of non-interacting electrons the electronic properties of a solid can be examined with the Schrödinger equation,

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + U(r) \right] \psi(r) = E \psi(r), \quad (9.2)$$

in which $\psi(r)$ is a wave function for one electron. Independent electrons, which obey a one-electron Schrödinger equation (9.2) with a periodic potential, are known as *Bloch electrons*, in contrast to "free electrons," to which Bloch electrons reduce when the periodic potential is identically zero.

Now we discuss general properties of the solution of the Schrödinger equation (9.2) taking into account the periodicity of the effective potential (9.1) and discuss main properties of Bloch electrons, which follow from this solution.

We represent the solution as an expansion over plane waves:

$$\psi(r) = \sum_k c_k e^{ikr}. \quad (9.3)$$

This expansion in a Fourier series is a natural generalization of the free-electron solution for a zero potential. The summation in (9.3) is performed over all \mathbf{k} vectors, which are permitted by the periodic boundary conditions. According to these conditions the wave function (9.3) should satisfy

$$\psi(x, y, z) = \psi(x + L, y, z) = \psi(x, y + L, z) = \psi(x, y, z + L), \quad (9.4)$$

so that

$$k_x = \frac{2\pi n_x}{L}; \quad k_y = \frac{2\pi n_y}{L}; \quad k_z = \frac{2\pi n_z}{L}, \quad (9.5)$$

where n_x , n_y , and n_z are positive or negative integers. Note that in general, $\psi(r)$ is *not* periodic in the lattice translation vectors. On the other hand, according to Eq.(9.1), the potential energy is periodic, i.e., it is invariant under a crystal lattice translation. Therefore, its plane wave expansion will only contain plane waves with the periodicity of the lattice. Therefore, only reciprocal lattice vectors are left in the Fourier expansion for the potential:

$$U(r) = \sum_G U_G e^{iGr}, \quad (9.6)$$

where the Fourier coefficients U_G are related to $U(r)$ by

$$U_G = \frac{1}{V_c} \int_{cell} e^{-iGr} U(r) dr, \quad (9.7)$$

where V_c is the volume of the unit cell. It is easy to see that, indeed, the potential energy represented by (9.6) is periodic in the lattice:

$$U(r + T) = \sum_G U_G e^{iG(r+T)} = e^{iGT} \sum_G U_G e^{iGr} = U(r), \quad (9.8)$$

where the last equation comes from the definition of the reciprocal lattice vectors $e^{iGT} = 1$. The values of Fourier components U_G for actual crystal potentials tend to decrease rapidly with increasing magnitude of \mathbf{G} . For example, for a Coulomb potential, U_G decreases as $1/G^2$. Note that since the potential energy is real, the Fourier components should satisfy $U_{-G} = U_G^*$.

We now substitute (9.3) and (9.6) in Eq.(9.2) and obtain:

$$\frac{\hbar}{2m} \sum_k k^2 c_k e^{ikr} + \sum_k \sum_G U_G c_k e^{i(k+G)r} = E \sum_k c_k e^{ikr}. \quad (9.9)$$

Changing the summation index in the second sum on the left from \mathbf{k} to $\mathbf{k}+\mathbf{G}$, this equation can be rewritten in form:

$$\sum_k e^{ikr} \left\{ \left(\frac{\hbar^2}{2m} k^2 - E \right) c_k + \sum_G U_G c_{(k-G)} \right\} = 0. \quad (9.10)$$

Since this equation must be satisfied for any \mathbf{r} the Fourier coefficients in each separate term of (9.10) must vanish and therefore

$$\left\{ \left(\frac{\hbar^2}{2m} k^2 - E \right) c_k + \sum_G U_G c_{(k-G)} \right\} = 0. \quad (9.11)$$

This is a set of linear equations for the coefficients c_k . These equations are nothing but a restatement of the original Schrödinger equation in the momentum space, simplified by the fact that the potential is periodic. This set of equations does not look very pleasant because, in principle, an infinite number of coefficients should be determined. However, a careful examination of Eq.(9.11) leads to important consequences.

First, we see that for a fixed value of \mathbf{k} the set of equations (9.11) couples only those coefficients, whose wave vectors differ from \mathbf{k} by a reciprocal lattice vector. In the one-dimensional case these are k , $k \pm 2\pi/a$, $k \pm 4\pi/a$, and so on. We can therefore assume that the \mathbf{k} vector belongs to the first Brillouin zone. The original problem is decoupled to N independent problems (N is the total number of atoms in a lattice) for each allowed value of \mathbf{k} in the first Brillouin zone. Each such problem has solutions that are a superposition of plane waves containing only the wave vector \mathbf{k} and wave vectors differing from \mathbf{k} by the reciprocal lattice vector.

Putting this information back into the expansion (9.3) of the wave function $\psi(r)$, we see that the wave function will be of the form

$$\psi_k(r) = \sum_G c_{k-G} e^{i(k-G)r}, \quad (9.12)$$

where the summation is performed over the reciprocal lattice vectors and we introduce index \mathbf{k} for the wave function. We can rearrange this so that

$$\psi_k(r) = e^{ikr} \sum_G c_{k-G} e^{-iGr}, \quad (9.13)$$

or

$$\psi_k(r) = e^{ikr} U_k(r), \quad (9.14)$$

where $U_k(r) = U_k(r + T)$ is a periodic function which is defined by

$$U_k(r) = \sum_G c_{k-G} e^{-iGr}. \quad (9.15)$$

Equation (9.14) is known as Bloch theorem, which plays an important role in electronic band structure theory. Now we discuss a number of important conclusions which follow from the Bloch theorem.

1. Bloch's theorem introduces a wave vector \mathbf{k} , which plays the same fundamental role in the general problem of motion in a periodic potential that the free electron wave vector \mathbf{k} plays in the free-electron theory. Note, however, that although the free electron wave vector is simply \mathbf{p}/\hbar , where \mathbf{p} is the momentum of the electron, in the Bloch case \mathbf{k} is not proportional to the electronic momentum. This is clear on general grounds, since the Hamiltonian does not have complete translational invariance in the presence of a non-constant potential, and therefore its eigenstates will not be simultaneous eigenstates of the momentum operator. This conclusion is confirmed by the fact that the momentum operator, $\mathbf{p} = -i\hbar\nabla$, when acting on $\psi_k(r)$ gives

$$-i\hbar\nabla\psi_k(r) = -i\hbar\nabla[e^{ikr} U_k(r)] = \hbar k \psi_k(r) - i\hbar e^{ikr} \nabla U_k(r), \quad (9.16)$$

which is not, in general, just a constant times $\psi_k(r)$; i.e $\psi_k(r)$ is not a momentum eigenstate.

Nevertheless, in many ways $\hbar k$ is a natural extension of \mathbf{p} to the case of a periodic potential. It is known as the *crystal momentum* or *quasi momentum* of the electron, to emphasize this similarity, but one should not be misled by the name into thinking that $\hbar k$ is a momentum.

2. The wave vector \mathbf{k} appearing in Bloch's theorem can always be confined to the first Brillouin zone (or to any other convenient primitive cell of the reciprocal lattice). This is because any \mathbf{k}' not in the first Brillouin zone can be written as

$$k' = k + G, \quad (9.17)$$

where \mathbf{G} is a reciprocal lattice vector and \mathbf{k} does lie in the first zone. Since $e^{i\mathbf{G}T} = 1$ for any reciprocal lattice vector, if the Bloch form (9.14) holds for \mathbf{k}' , it will also hold for \mathbf{k} . An example is given below for a nearly free electron model.

The energy E of free electrons which is plotted versus k in Fig.9.2a exhibits a curve in the familiar parabolic shape. Figure 9.2b shows the result of translations. Segments of the parabola of Fig.9.2a are cut at the edges of the various zones, and are translated by multiples of $G = 2\pi/a$ in order to ensure that the energy is the same at any two equivalent points. Fig.9.2c displays the shape of the energy spectrum when we confine our consideration to the first Brillouin zone only.

The type of representation used in Fig.9.2c is referred to as the *reduced-zone scheme*. Because it specifies all the needed information, it is the one we shall find most convenient. The representation of Fig.9.1, known as the *extended-zone scheme*, is convenient when we wish to emphasize the close connection between a crystalline and a free electron. Fig.9.2b employs the *periodic-zone* scheme, and is sometimes useful in topological considerations involving the \mathbf{k} -space. All these representations are strictly equivalent; the use of any particular one is dictated by convenience, and not by any intrinsic advantages it has over the others.

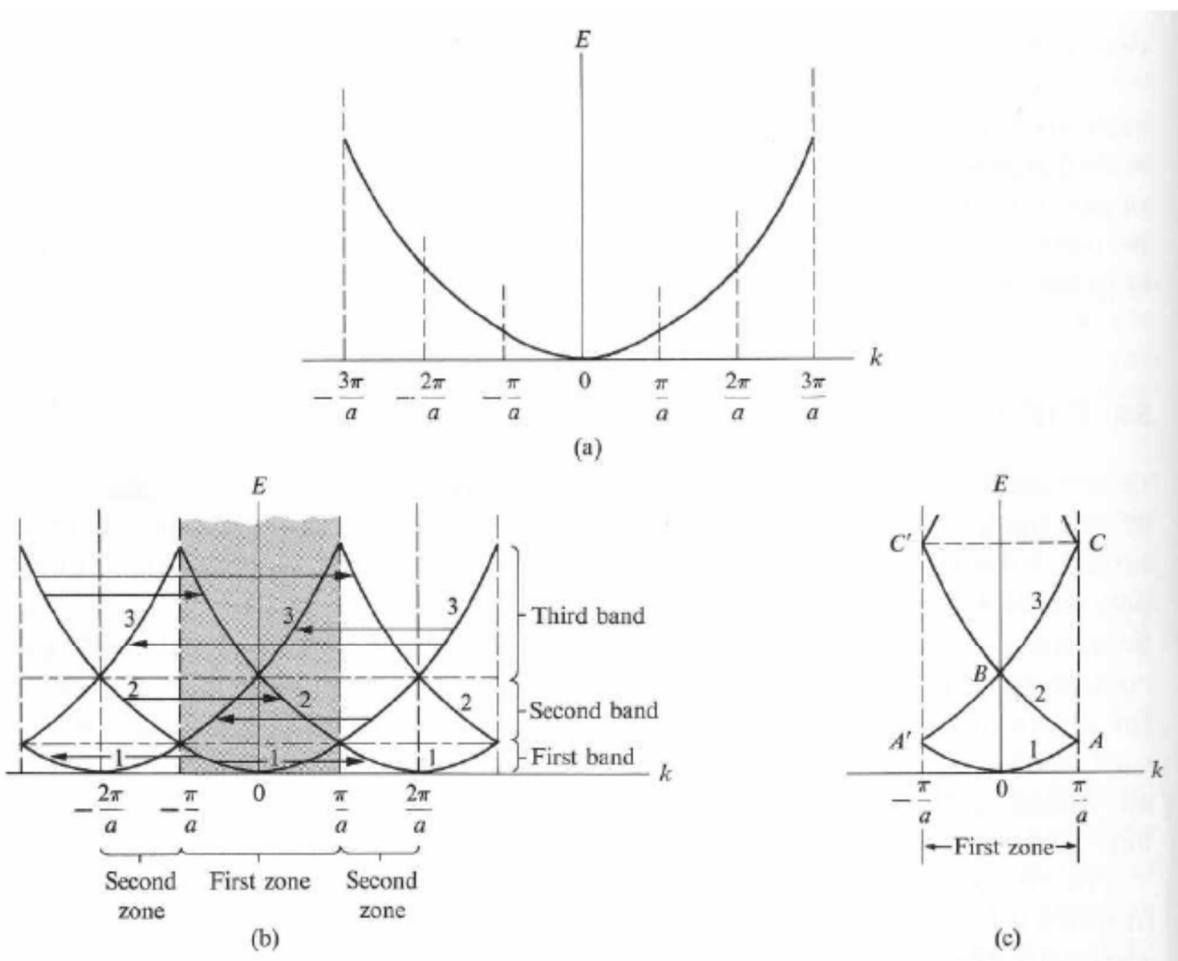


Fig.9.2 Free electron bands within *reduced-* (a), *extended-* (b) and *periodic-zone* (c) scheme.

3. An important consequence of the Bloch theorem is the appearance of the energy bands.

All solutions to the Schrodinger equation (9.2) have the Bloch form $\psi_k(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}U_k(\mathbf{r})$ where \mathbf{k} is fixed and $U_k(\mathbf{r})$ has the periodicity of the Bravais lattice. Substituting this into the Schrodinger equation, we find that $U_k(\mathbf{r})$ is determined by the eigenvalue problem

$$H(\mathbf{k})U_k(\mathbf{r}) = \left[-\frac{\hbar^2}{2m}(i + \mathbf{k}\nabla)^2 + U(\mathbf{r}) \right] U_k(\mathbf{r}) = E(\mathbf{k}) U_k(\mathbf{r}), \quad (9.18)$$

with boundary condition

$$U_k(\mathbf{r}) = U_k(\mathbf{r} + \mathbf{T}), \quad (9.19)$$

Due to the periodic boundary condition we can regard (9.18) as an eigenvalue problem restricted to a single primitive cell of the crystal. The eigenvalue problem is set in a fixed finite volume. We expect on general grounds to find an infinite family of solutions with *discretely* spaced eigenvalues, which we label with the band index n . The Bloch function can therefore be denoted by $\psi_{nk}(\mathbf{r})$ which indicates that each value of the band index n and the vector \mathbf{k} specifies an electron state, or orbital with energy $E_n(\mathbf{k})$.

Note that in terms of the eigenvalue problem specified by (9.18) and (9.19), the wave vector \mathbf{k} appears only as a parameter in the Hamiltonian $H(\mathbf{k})$. We therefore expect each of the energy levels, for given \mathbf{k} , to vary continuously as \mathbf{k} varies. In this way we arrive at a description of the levels of an electron in a periodic potential in terms of a family of continuous functions $E_n(\mathbf{k})$. For each n , the set of electronic levels specified by $E_n(\mathbf{k})$ is called an *energy band*. The information contained in these functions for different n and \mathbf{k} is referred to as the *band structure* of the solid.

1. Number of states in a band.

The number of orbitals in a band within the first Brillouin zone is equal to the number of unit cells N in the crystal. This is much the same as the statement made in connection with the number of lattice vibrational modes, and is proved in a like manner, by appealing to the boundary conditions. Consider first the one-dimensional case. The allowed values of k form a uniform mesh whose unit spacing is $2\pi/L$. The number of states inside the first zone, whose length is $2\pi/a$, is therefore equal to $(2\pi/a)/(2\pi/L) = L/a = N$, where N is the number of unit cells, in agreement with the assertion made earlier.

A similar argument may be used to establish the validity of the statement in two- and three-dimensional lattices. It has been shown that each band has N states inside the first zone. Since each such state can accommodate at most two electrons of opposite spins in accordance with Pauli Exclusion principle, it follows that the maximum number of electrons that may occupy a single band is $2N$. This result is significant, as it will be used in a later section to establish the criterion for predicting whether a solid is going to behave as a metal or as an insulator.

5. Now we show that an electron in a level specified by band index n and wave vector \mathbf{k} has a non-vanishing mean velocity, given by

$$V_n(\mathbf{k}) = \frac{dE_n(\mathbf{k})}{d\mathbf{k}}. \quad (9.20)$$

To show this we calculate the expectation value of the derivative of the Hamiltonian $H(\mathbf{k})$ in Eq.(9.18) with respect to \mathbf{k} :

$$\left\langle U_n \left| \frac{dH(\mathbf{k})}{d\mathbf{k}} \right| U_n \right\rangle = \left\langle U_n \left| -i \frac{\hbar^2}{m} (i\mathbf{k} + \nabla) \right| U_n \right\rangle = \left\langle \psi_n \left| \hbar \left(-\frac{i\hbar}{m} \nabla \right) \right| \psi_n \right\rangle. \quad (9.21)$$

Since $\mathbf{v} = (-i\hbar/m)\nabla$ is the velocity operator, this establishes (9.20).

This is a remarkable fact. It asserts that there are stationary levels for an electron in a periodic potential in which, in spite of the interaction of the electron with the fixed lattice of ions, it moves forever without any degradation of its mean velocity. This is in striking contrast to the idea of Drude, that collisions were simply encounters between the electron and a static ion.

In-text questions

1. Q: Describe how the electronic properties of non-interacting electrons can be examined.
A: See pages 124-125.
2. Q: What are the main conclusions resulting from the Bloch theorem.
A: See page 129
3. Q: Differentiate the reduced, extended and periodic zone scheme in band structure formation.
A: See page 128

9.1.2 Weak Potential

When the potential is zero the solutions of the Schrödinger equation (9.11) are plane waves:

$$E^0(\mathbf{k}) = \frac{\hbar^2 k^2}{2m}, \quad (9.22)$$

$$\psi_{\mathbf{k}}^0 = \frac{1}{\sqrt{V_c}} e^{ikr}, \quad (9.23)$$

where the wave function is normalized to the volume of unit cell V_c . In the reduced-zone representation shown in Fig.9.3, for each \mathbf{k} there is an infinite number of solutions which correspond to different \mathbf{G} (and can be labeled by index n), as we have already discussed. Each band in Fig.9.3 corresponds to a different value of \mathbf{G} in the extended scheme.

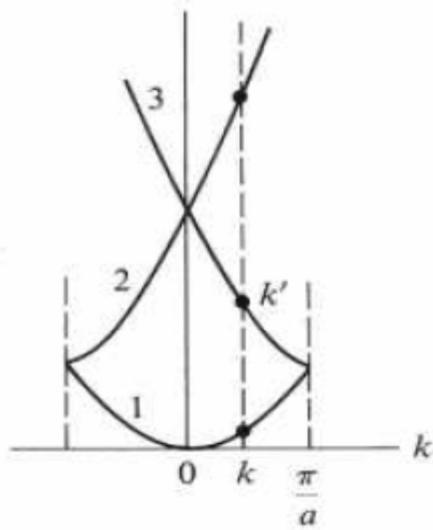


Fig.9.3 Only those states which have the same \mathbf{k} in the First Brillouin zone are coupled by perturbation

Suppose now that a weak potential is switched on. According to the Schrödinger equation (9.11) only those states, which differ by \mathbf{G} , are coupled by a perturbation. In the reduced zone scheme those states have same \mathbf{k} and different n (see Fig.9.3). As you know from quantum mechanics, if the perturbation is small compared to the energy difference between the states, which are coupled by the perturbation, we can use the perturbation theory to calculate wave functions and energy levels. Assuming for simplicity that we are looking for the correction to the energy of the lowest band $E^0(\mathbf{k})$, the condition for using the perturbation theory is

$$|E^0(\mathbf{k}) - E^0(\mathbf{k} - \mathbf{G})| \gg U, \quad (9.24)$$

for any $\mathbf{G} \neq 0$. According to the perturbation theory the energy is given by

$$E(\mathbf{k}) = E^0(\mathbf{k}) + \langle \psi_{\mathbf{k}}^0 | U | \psi_{\mathbf{k}}^0 \rangle + \sum_{\mathbf{G} \neq 0} \frac{\left| \langle \psi_{\mathbf{k}}^0 | U | \psi_{\mathbf{k}-\mathbf{G}}^0 \rangle \right|^2}{E^0(\mathbf{k}) - E^0(\mathbf{k}-\mathbf{G})}, \quad (9.25)$$

The first term in Eq.(9.25) is the undisturbed free-electron value for the energy. The second term is the mean value of the potential in the state $\psi_{\mathbf{k}}^0(\mathbf{r})$:

$$\langle \psi_{\mathbf{k}}^0 | U | \psi_{\mathbf{k}}^0 \rangle = \frac{1}{V_c} \int_{cell} U(\mathbf{r}) d\mathbf{r} = U_0. \quad (9.26)$$

This term gives us a constant independent of \mathbf{k} . Its effect on the spectrum is a rigid shift by a constant value without causing any change in the shape of the energy spectrum. This term can be set equal to zero. The third term can be rewritten as

$$\langle \psi_{\mathbf{k}}^0 | U | \psi_{\mathbf{k}-\mathbf{G}}^0 \rangle = \frac{1}{V_c} \int_{cell} e^{-ikr} U(\mathbf{r}) e^{-i(k-\mathbf{G})r} dr = \frac{1}{V_c} \int_{cell} e^{-iGr} dr = U_G. \quad (9.27)$$

Finally, we obtain for the energy:

$$E(\mathbf{k}) = E^0(\mathbf{k}) + \sum_{\mathbf{G} \neq 0} \frac{|U_G|^2}{E^0(\mathbf{k}) - E^0(\mathbf{k}-\mathbf{G})}. \quad (9.28)$$

The perturbation theory breaks down, however, in those cases when the potential cannot be considered as a small perturbation. This happens when the magnitude of the potential becomes comparable with the energy separation between the bands, i.e.,

$$|E^0(\mathbf{k}) - E^0(\mathbf{k}-\mathbf{G})| \leq U, \quad (9.29)$$

In this case we have to include these levels in the Schrödinger equation and solve explicitly.

There are special \mathbf{k} points for which the energy levels become degenerate and the relationship (9.29) holds for any non-zero value of the potential. For these \mathbf{k} points

$$E^0(\mathbf{k}) = E^0(\mathbf{k}-\mathbf{G}), \quad (9.30)$$

and consequently

$$|\mathbf{k}| = |\mathbf{k} - \mathbf{G}| . \quad (9.31)$$

The latter condition implies that \mathbf{k} must lie on a Bragg plane bisecting the line joining the origin of \mathbf{k} space and the reciprocal lattice point \mathbf{G} , as is shown in Fig.9.4.

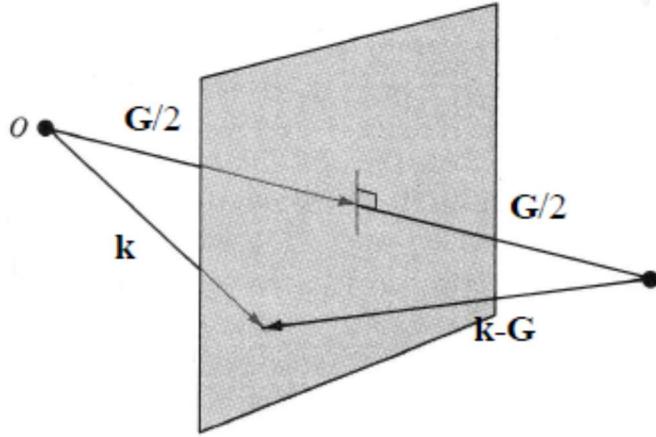


Fig. 9.4 $|\mathbf{k}| = |\mathbf{k} - \mathbf{G}|$, then the point \mathbf{k} must lie in the Bragg plane determined by \mathbf{G} .

Therefore, a weak periodic potential has its major effect on those free electron levels whose wave vectors are close to ones at which the Bragg reflection can occur. In order to find the energy levels and the wave functions near these points we will include in equation (9.11) only two levels: one which corresponds to \mathbf{k} and the other which corresponds to $\mathbf{k}-\mathbf{G}$, assuming that \mathbf{k} lies near the Bragg plane:

$$\begin{aligned} (E^0(\mathbf{k}) - E)c_k + U_G c_{k-G} &= 0 , \\ (E^0(\mathbf{k} - \mathbf{G}) - E)c_{k-G} + U_{-G} c_k &= 0 . \end{aligned} \quad (9.32)$$

These equations have the solution when the determinant is equal to zero, i.e.

$$\begin{vmatrix} E^0(\mathbf{k}) - E & U_G \\ U_G & E^0(\mathbf{k} - \mathbf{G}) - E \end{vmatrix} = 0 , \quad (9.33)$$

which leads to the quadratic equation

$$(E^0(\mathbf{k}) - E)(E^0(\mathbf{k} - \mathbf{G}) - E) - |U_G|^2 = 0 . \quad (9.34)$$

The two roots are

$$E = \frac{1}{2}(E^0(\mathbf{k}) + E^0(\mathbf{k} - \mathbf{G})) \pm \left[\frac{1}{4}(E^0(\mathbf{k}) - E^0(\mathbf{k} - \mathbf{G}))^2 + |U_G|^2 \right]^{1/2} . \quad (9.35)$$

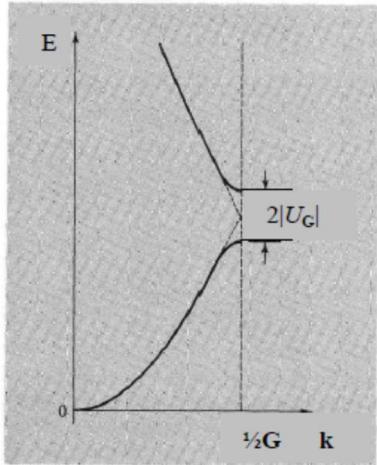


Figure 9.5 Plot of the energy bands given by eq.(9.35) for \mathbf{k} parallel to \mathbf{G} .

the lower band corresponds to the choice of the minus sign in eq. (9.35) and the upper band to the plus sign. When $k = \frac{1}{2}G$, the two bands are separated by a band gap of magnitude $2|U_G|$. When k is far removed from the Bragg plane, the levels (to leading order) are indistinguishable from their free electron values (denoted by dotted lines).

This results is particularly simple for points lying on the Bragg plane, since in this case $E^0(\mathbf{k}) = E^0(\mathbf{k} - \mathbf{G})$. We find from (9.35) then that

$$E = E^0(\mathbf{k}) \pm |U_G|. \quad (9.36)$$

Thus, at all points in the Bragg plane, one level is uniformly raised by $|U_G|$ and the other is uniformly lowered by the same amount. This means that there are no states in the energy interval between $E_1 = E^0(\mathbf{k}) - |U_G|$ and $E_2 = E^0(\mathbf{k}) + |U_G|$, which implies the creation of the band gap. The magnitude of the band gap is equal to twice the Fourier component of the crystal potential. We illustrate this behavior using a one-dimensional lattice shown in Fig.9.6. We see the splitting of the bands at each Bragg plane in the extended-zone scheme (Fig.9.6b). This results in the splitting of the bands both at the boundaries and at the centre of the first Brillouin zone (Fig.9.6a).

There are two important points to note. First, since the energy there increases as k^2 , the higher the band, the greater its width. Second, the higher the energy, the narrower the gap; this follows from the fact that the gap is proportional to a Fourier component of the crystal potential and that the

order of the component increases as the energy rises. Since the Fourier components of the potential decrease rapidly as the order increases, this leads to a decrease in the energy gap. It follows therefore that, as we move up the energy scale, the bands become wider and the gaps narrower; i.e., the electron behaves more and more like a free particle.

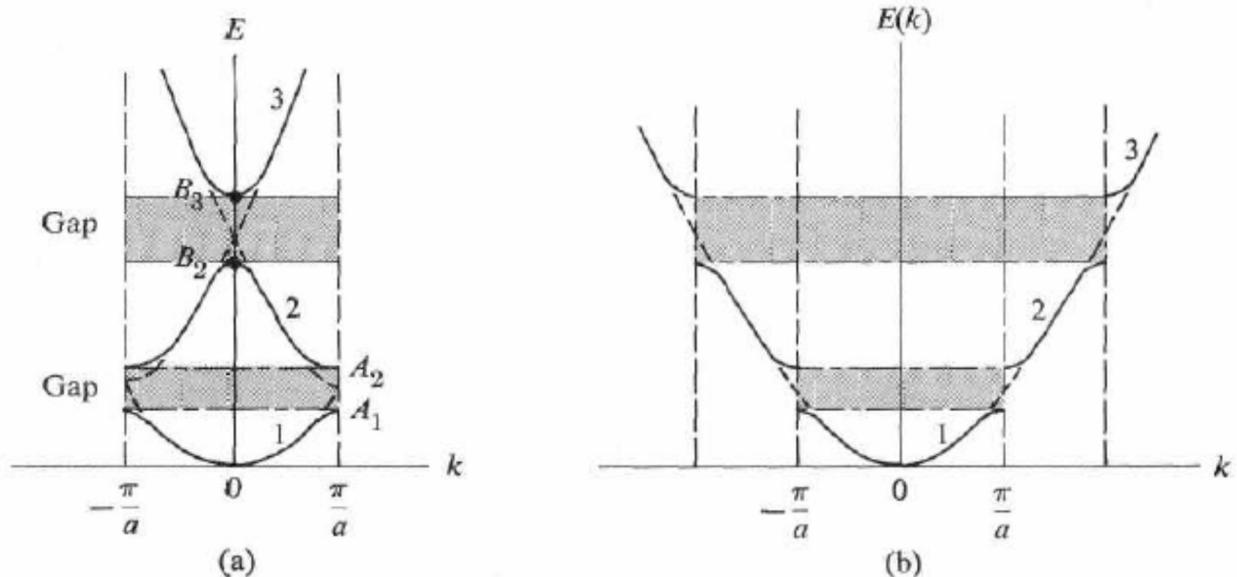


Fig. 9.6 (a) Dispersion curves in the nearly-free-electron model, in the reduced-zone scheme;
(b) The same dispersion curves in the extended-zone scheme.

Now we discuss the origin of the appearance of the band gaps at the Bragg planes. When \mathbf{k} lies on a Bragg plane we can easily find the form of the wave function corresponding to the two solutions (9.36). Assuming for simplicity that the potential is real, we obtain from Eq.(9.32),

$$c_k = \pm c_{k-G} . \quad (9.37)$$

For simplicity we consider a one-dimensional lattice, for which the Bragg reflection occurs at $\mathbf{k} = 1/2 \mathbf{G}$. We have then

$$\psi_{\pm} = \frac{1}{\sqrt{2V_c}} [e^{iG^r/2} \pm e^{iG^r/2}] . \quad (9.38)$$

We see that at the zone edge, the scattering is so strong that the reflected wave has the same amplitude as the incident wave. The electron is represented there by a standing wave, very unlike a free particle.

The distribution of the charge density is proportional to $|\psi|^2$, so that

$$\begin{aligned} |\psi_+|^2 &\propto \cos(G \cdot r/2), \\ |\psi_-|^2 &\propto \sin(G \cdot r/2), \end{aligned} \quad (9.39)$$

Since the origin lies at the ion, the ψ_- state distributes the electron so that it is piled predominantly at the nuclei (see Fig.9.7). Since the potential is most negative there, this distribution has a low energy. The function ψ_- therefore corresponds to the energy at the top of band 1, that is, point A1 in Fig. 9.6a.

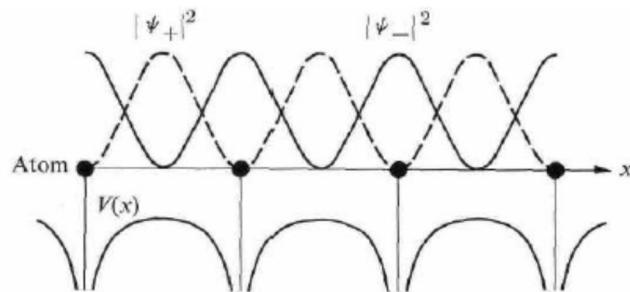


Figure 9.7 Spatial distributions of the charge density described by the functions Ψ_+ and Ψ_-

By contrast, the function ψ_+ deposits its electrons mostly between the ions (as shown in Fig.9.7), and corresponds to the bottom of band 2 in Fig.9.6a, that is, point A2. The gap arises, therefore, because of the two different distributions for the same value of k , the distributions having different energies.

In-text question

1. Q: Highlight how the wave functions and energy levels can be calculated when there is a weak potential.
A: See pages 132-137.

9.1.3 Metals and Insulators

Solids are divided into two major classes: metals and insulators. A metal – or a conductor – is a solid in which an electric current flows under the application of electric field. By contrast, application of an electric field produces no electric current in an insulator. There is a simple

criterion for distinguishing between the two classes on the basis of the band structure. If the valence electrons exactly fill one or more bands, leaving others empty, the crystal will be an insulator. An external electric field will not cause current flow in an insulator. Provided that a filled band is separated by an energy gap from the next higher band, there is no continuous way to change the total momentum of the electrons if every accessible state is filled. Nothing changes when the field is applied.

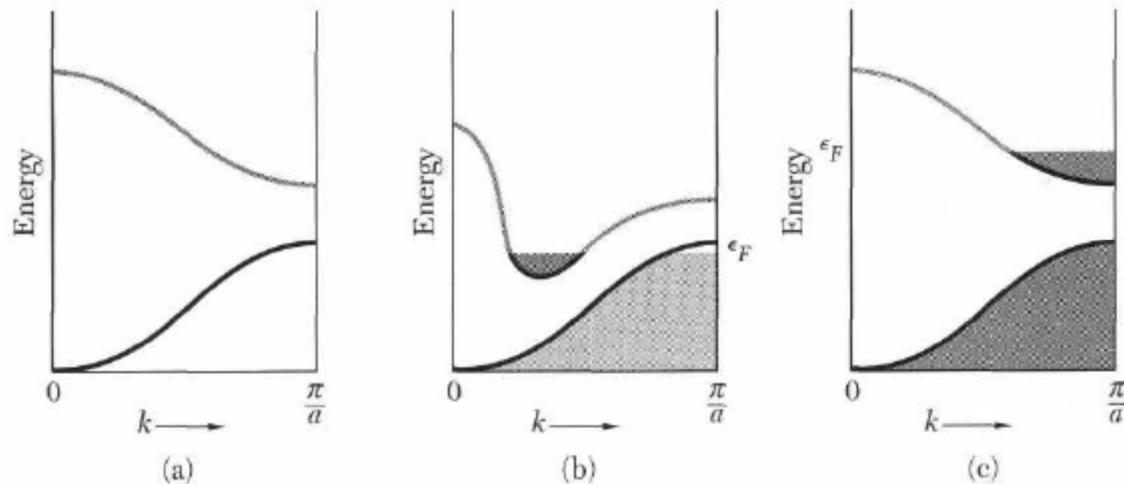


Fig. 9.8 Occupied states and band structures giving (a) an insulator, (b) a metal or a semimetal because of band overlap, and (c) a metal because of electron concentration.

In (b) the overlap need not occur along the same directions in the Brillouin zone. If the overlap is small, with relatively few states involved, we speak of a semimetal. On the contrary, if the valence band is not completely filled, the solid is a metal. In a metal there are empty states available above the Fermi level as in a free electron gas. An application of an external electric field results in current flow. It is possible to determine whether a solid is a metal or an insulator by considering the number of valence electrons. A crystal can be an insulator only if the number of valence electrons in a primitive cell of the crystal is an even integer. This is because each band can accommodate only two electrons per primitive cell. For example, diamond has two atoms of valence four, so that there are eight valence electrons per primitive cell. The band gap in diamond is 7eV and this crystal is a good insulator.

However, if a crystal has an even number of valence electrons per primitive cell, it is not necessarily an insulator. It may happen that the bands overlap in energy. If the bands overlap in

energy, then instead of one filled band giving an insulator, we can have two partly filled bands giving a metal (Fig.9.8b). For example, the divalent metals, such as Mg or Zn, have two valence electrons per cell. However, they are metals, although poor ones – their conductivity is small. If this overlap is very small, we deal with semimetals. The best known example of a semimetal is bismuth (Bi).

If the number of valence electrons per cell is odd, the solid is a metal. For example, the alkali metals and the noble metals have one valence electron per primitive cell, so that they have to be metals. The alkaline earth metals have two valence electrons per primitive cell; they could be insulators, but the bands overlap in energy to give metals, but not very good metals. Diamond, silicon, and germanium each have two atoms of valence four, so that there are eight valence electrons per primitive cell; the bands do not overlap, and the pure crystals are insulators at absolute zero. There are substances which fall in an intermediate position between metals and insulators. If the gap between the valence band and the band immediately above it is small, then electrons are readily excitable thermally from the former to the latter band. Both bands become only partially filled and both contribute to the electric condition. Such a substance is known as a semiconductor. Examples are Si and Ge, in which the gaps are about 1 and 0.7 eV, respectively.

Roughly speaking, a substance behaves as a semiconductor at room temperature whenever the gap is less than 2 eV. The conductivity of a typical semiconductor is very small compared to that of a metal, but it is still many orders of magnitude larger than that of an insulator. It is justifiable, therefore, to classify semiconductors as a new class of substance, although they are, strictly speaking, insulators at very low temperatures.

In-text questions

1. Q: Differentiate between Metals and Insulators in terms of energy band gap and occupied states.
A: See pages 138-139.

2. Q: Briefly describe a semiconductor.
A: See page 139.

Summary of Study Session 9

In this study session, you have learnt the:

1. Calculation of wave functions and energy levels in both strong and weak potential.
2. Distinguishing features of metals, semiconductors and insulators.

Self-Assessment Questions for Study Session 9

Now that you have completed this study session, you can assess how well you have achieved its learning outcomes by answering the following questions. Write your answers in your Study Diary and discuss them with your Tutor at the next Study Support Meeting.

SAQ 9.1

Compute the concentration of electrons and holes in an intrinsic semiconductor InSb at room temperature ($E_g = 0.2\text{eV}$, $m_e = 0.01m$ and $m_h = 0.018 m$). Determine the position of the Fermi energy.

SAQ 9.2

Indium antimonide has $E_g = 0.23\text{ eV}$; dielectric constant $\epsilon = 18$; electron effective mass $m_e = 0.015 m$. Calculate the donor ionization energy and the radius of the ground state orbit. At what minimum donor concentration will appreciable overlap effects between the orbits of adjacent impurity atoms occur? This overlap tends to produce an impurity band - a band of energy levels which permit conductivity presumably by a hopping mechanism in which electrons move from one impurity site to a neighboring ionized impurity site.

SAQ 9.3

Given the data for Si: $\mu_e = 1350 \text{ cm}^2/\text{V}\cdot\text{s}$, $\mu_h = 475 \text{ cm}^2/\text{V}\cdot\text{s}$, $m_e = 0.19m$, $m_h = 0.16m$ and $E_g = 1.1 \text{ eV}$, calculate

- a) The lifetimes of electrons and holes.
- b) The intrinsic conductivity, σ , at room temperature.

References/ Further Readings

Ashcroft N. W. and Mermin N.D., (1976), Solid state Physics, Holt, Rine hart, Winston.

Study Session 10 Semiconductors I

Introduction

Semiconductors are materials whose electronic properties are intermediate between those of metal and insulators. These intermediate properties are determined by the crystal structure, bonding

characteristics, and electronic energy bands and also by the fact that, unlike metals, a semiconductor has both positive (hole) and negative (electron) carriers of electricity whose densities can be controlled by doping the pure semiconductor with chemical impurities during crystal growth. The climax of investigations into these materials was the invention and development of the transistor by J. Bardeen, W. Brattain and W. Shockley in 1948 for which they received the Nobel Prize in 1956.

Learning Outcomes for Study Session 10

When you have studied this session, you should be able to explain:

- 10.1 The crystal structure, band structure and crystal bonding in semiconductors.
- 10.2 The different classes of semiconductors.
- 10.3 Statistical distribution of carriers in semiconductors.
- 10.4 Electron conductivity and mobility in semiconductors.

10.1 Main Content

10.1.1 Crystal Structure and Bonding

Semiconductors include a large number of substances of widely different chemical and physical properties. These materials are grouped into several classes of similar behavior, the classification being based on the position in the periodic table of the elements.

The best-known class is the Group IV semiconductors - C (diamond), Si, Ge, - all of which lie in the fourth column of the periodic table. They have been studied intensively, particularly Si and Ge, which have found many applications in electronic devices. The elemental semiconductors all crystallize in the diamond structure. The diamond structure has an fcc lattice with a basis composed of two identical atoms, and is such that each atom is surrounded by four neighboring atoms, forming a regular tetrahedron. Group IV semiconductors are covalent crystals, i.e., the atoms are held together by covalent bonds. These bonds consist of two electrons of opposite spins distributed along the line joining the two atoms. The covalent electrons forming the bonds are hybrid sp^3 atomic orbitals.

Another important group of semiconductors is the Group III-V compounds, so named because each contains two elements, one from the third and the other from the fifth column of the periodic

table. The best-known members of this group are GaAs and InSb (indium antimonite), but the list also contains compounds such as GaP, InAs, GaSb, and many others.

These substances crystallize in the zincblende structure, which is the same as the diamond structure, except that the two atoms forming the basis of the lattice are now different. Thus, in GaAs, the basis of the fcc lattice consists of two atoms Ga and As. Due to this structure, each atom is surrounded by four others of the opposite kind, and these latter atoms form a regular tetrahedron, just as in the diamond structure.

The bonding in the III-V compounds is also primarily covalent. The eight electrons required for the four tetrahedral covalent bonds are supplied by the two types of atoms, the trivalent atom contributing its three valence electrons, and the pentavalent atom five electrons. The bonding in this group is not entirely covalent. The two elements in the compound are different; the distribution of the electrons along the bond is not symmetric, but displaced toward one of the atoms. As a result, one of the atoms acquires a net electric charge. Such a bond is called heteropolar, in contrast to the purely covalent bond in the elemental semiconductors, which is called homopolar.

The distribution of electrons in the bond is displaced toward the atom of higher electronegativity. In GaAs, for instance, the As atom has a higher electronegativity than the Ga, and consequently the As atom acquires a net negative charge, whose value is -0.46e per atom (a typical value in Group III-V compounds). The Ga atom correspondingly acquires a net positive charge of 0.46e. Charge transfer leads to an ionic contribution to the bonding in Group III-V compounds. Their bonding is therefore actually a mixture of covalent and ionic components, although covalent ones predominate in most of these substances.

10.1.2. Band Structure

A semiconductor is a solid in which the highest occupied energy band, the valence band, is completely full at $T = 0^{\circ}k$, but in which the gap above this band is also small, so that electrons may be excited thermally at room temperature from the valence band to the next-higher band, which is known as the conduction band. Generally, the number of excited electrons is appreciable (at room temperature) whenever the energy gap E_F is less than 2 eV. The substance may then be classified as a semiconductor. When the gap is larger, the number of electrons is negligible, and the substance is an insulator.

When electrons are excited across the gap, the bottom of the conduction band (CB) is populated by electrons, and the top of the valence band (VB) by holes. As a result, both bands are only partially full, and would carry a current if an electric field were applied. The conductivity of the semiconductor is small compared with the conductivities of metals because of the small number of electrons and holes involved, but this conductivity is nonetheless sufficiently large for practical purposes.

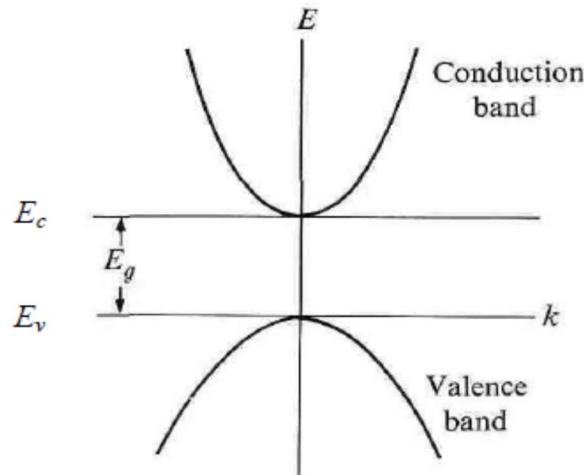


Fig. 10.1 Band structure in a semiconductor.

The simplest band structure of a semiconductor is indicated in Fig. 10.1. Since we are interested only in the regions which lie close to the band gap, where electrons and holes lie, we can ignore a more complex variation of the energy bands far away from the gap.

The energy of the conduction band has the form

$$E_c(k) = E_c + \frac{\hbar^2 k^2}{2m_e} \quad (10.1)$$

where k is the wave vector and m_e the effective mass of the electron. The energy E_g represents the energy gap. The zero-energy level is chosen to lie at the top of the valence band.

The energy of the VB (Fig. 10.1) may be written as

$$E_v(k) = E_v - \frac{\hbar^2 k^2}{2m_h} \quad (10.2)$$

where m_h is the effective mass of the hole, which is positive. (Because of the inverted shape of the VB, the mass of an electron at the top of the VB is negative, but the mass of a hole is positive.) Within this simple picture of the semiconductor, the primary band-structure parameters are thus the electron and hole masses m_e and m_h and the band gap E_g . Table 10.1 gives the parameters for various semiconductors. Note that the masses differ considerably from the free-electron mass. In many cases they are much smaller than the free-electron mass. The energy gaps range from 0.18 eV in InSb to 3.7 eV in ZnS. The table also shows that the wider the gap, the greater the mass of the electron.

Table 10.1. Band Structure parameters of Semiconductors

Group	Crystal	E_g (eV)	m_e/m	m_h/m
IV	C	5.3		
IV	Si	1.1	0.19	0.16
IV	Ge	0.7	0.08	0.04
III-V	GaAs	1.4	0.07	0.09
III-V	GaP	2.3	0.12	0.50
III-V	InSb	0.2	0.01	0.18
II-VI	ZnS	3.6	0.40	5.41
II-VI	ZnSe	2.7	0.10	0.60
II-VI	CdSe	1.7	0.13	0.45

The energy gap for a semiconductor varies with temperature, but the variation is usually slight. That variation with temperature should exist at all can be appreciated from the fact that the crystal, when it is heated, experiences a volume expansion, and hence a change in its lattice constant. This, in turn, affects the band structure, which is a sensitive function of the lattice constant.

The band structure in Fig.10.1 is the simplest possible structure. Band structures of real semiconductors are somewhat more complicated, as we shall see, but for the present the simple structure will suffice for our purposes.

In-text questions

1. Q: Differentiate between homopolar and heteropolar bonding in semiconductors.

A: See page 143.

2. Q: Briefly explain how semiconductors are formed.

A: See pages 142-143.

3. Q: How many free electrons does one donor atom contribute?

A: See page 143.

4. Q: Give the graphical representation of the energy band model of silicon.

A: See page 144 (Figure 10.1)

10.1.3. Types of Semiconductors

Intrinsic Semiconductors

In the field of semiconductors, electrons and holes are usually referred to as free carriers, or simply carriers, because it is these particles which are responsible for carrying the electric current. The number of carriers is an important property of a semiconductor, as this determines its electrical conductivity. Intrinsic semiconductors are semiconductors in which the number of carriers and the conductivity are not influenced by impurities. Intrinsic conductivity is typical at relatively high temperatures in highly purified specimens.

In order to determine the number of carriers, we need some of the basic results of statistical mechanics. The most important result in this regard is the Fermi-Dirac (FD) distribution function

$$f(E) = \frac{1}{e^{(E-\mu)/k_B T} + 1} \quad (10.3)$$

This function gives the probability that an energy level E is occupied by an electron when the system is at temperature T.

The function is plotted versus E in Fig.10.2. Here, we see that as the temperature rises, the unoccupied region below the Fermi level E_F becomes longer, which implies that the occupation of high energy states increases as the temperature is raised, a conclusion which is most plausible, since increasing the temperature raises the overall energy of the system.

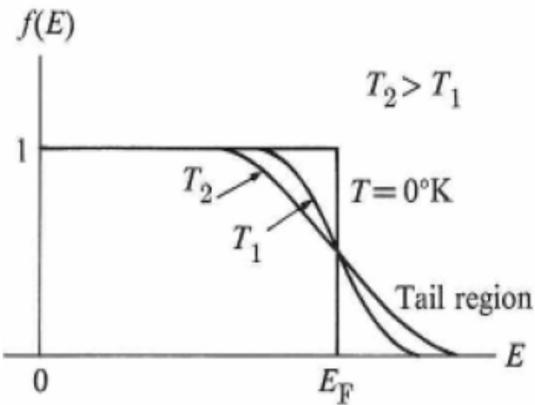


Fig. 10.2 The Fermi-Dirac distribution function

We shall see later that the Fermi level in intrinsic semiconductors lies close to the middle of the band gap. Therefore, we can represent the distribution function and the conduction and valence bands of the semiconductor by Fig.10.3.

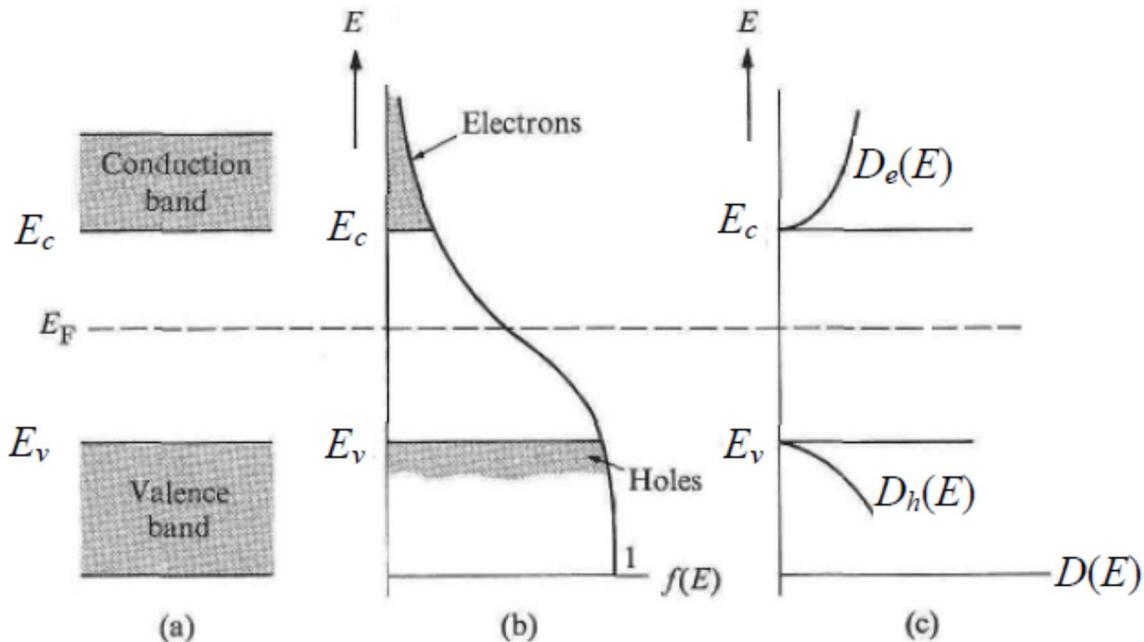


Fig. 10.3 (a) Conduction and valence bands, (b) The distribution function, (c) Density of states for electrons and holes.

First we calculate the concentration of electrons in the CB. The number of states in the energy range (E , $E + dE$) is equal to $D_e(E)dE$, where $D_e(E)$ is the density of electron states. Since each of these states has an occupation probability $f(E)$, the number of electrons actually found in this energy range is equal to $f(E) D_e(E)dE$. The concentration of electrons throughout the CB is thus given by the integral over the conduction band

$$n = \int_{E_c}^{\infty} f_e(E) D_e(E) dE \quad (10.4)$$

Where E_c is the bottom, the conduction band, as shown in Fig.10.3. The band gap in semiconductors is of the order of 1eV, which is much larger than kT . Therefore, $(E - \mu) \gg k_B T$ and we can neglect the unity term in the denominator of the distribution function (10.3), so that

$$f_e(E) \approx e^{-(E-\mu)/k_B T}. \quad (10.5)$$

The density of states for the conduction band is given by

$$D_e(E) = \frac{1}{2\pi^2} \left(\frac{2m_e}{\hbar^2} \right)^{3/2} (E - E_c)^{1/2}. \quad (10.6)$$

Note that $D_e(E)$ vanishes for $E < E_c$, and is finite only for $> E_c$, as shown in Fig. 10.3.

When we substitute equations for $f_e(E)$ and $D_e(E)$ into Eq. (10.4), we obtain

$$n = \frac{1}{2\pi^2} \left(\frac{2m_e}{\hbar^2} \right)^{3/2} e^{\mu/k_B T} \int_{E_c}^{\infty} (E - E_c)^{1/2} e^{-E/k_B T} dE. \quad (10.7)$$

By changing the variable, and using the result

$$\int_0^{\infty} x^{1/2} e^{-x} dx = \frac{\sqrt{\pi}}{2} \quad (10.8)$$

one can readily evaluate the integral in (10.7). The electron concentration then reduces to the expression

$$n = 2 \left(\frac{m_e k_B T}{2\pi \hbar^2} \right)^{3/2} e^{(\mu - E_c)/k_B T} \quad (10.9)$$

The electron concentration is still not known explicitly because the Fermi energy μ is so far unknown.

Essentially the same ideas employed above may also be used to evaluate the number of holes in the VB. The probability that a hole occupies a level E in this band is equal to $1 - f_e(E)$, since $f_e(E)$ is the probability of electron occupation. Assuming that the Fermi level lies close to the middle of the band gap, i.e. $(\mu - E_c) \gg k_B T$ for the valence band, we find for the distribution function of holes

$$f_e(E) = 1 - \frac{1}{e^{(E-\mu)/k_B T} + 1} = \frac{1}{e^{(\mu-E)/k_B T} + 1} \approx e^{-(\mu-E)/k_B T}. \quad (10.10)$$

The density of states for the holes is

$$D_h(E) = \frac{1}{2\pi^2} \left(\frac{2m_h}{\hbar^2} \right)^{3/2} (E_v - E)^{1/2}, \quad (10.11)$$

where E_v is the energy of the valence band edge. Proceeding in a similar fashion as we did for electrons, we find the concentration of holes in the valence band

$$p = \int_{-\infty}^{E_v} f_h(E) D_h(E) dE = 2 \left(\frac{m_h k_B T}{2\pi\hbar^2} \right)^{3/2} e^{(E_v-\mu)/k_B T}. \quad (10.12)$$

The electron and hole concentrations have thus far been treated as independent quantities. For intrinsic semiconductors, the two concentrations are, in fact, equal, because the electrons in the CB are due to excitations from the VB across the energy gap, and for each electron thus excited, a hole is created in the VB. Therefore,

$$n = p \quad (10.13)$$

and

$$(m_e)^{3/2} e^{(\mu-E_c)/k_B T} = (m_h)^{3/2} e^{(E_v-\mu)/k_B T}. \quad (10.14)$$

We obtain, then, for the Fermi energy

$$\mu = \frac{E_v + E_c}{2} + \frac{3}{4} k_B T \ln \frac{m_h}{m_e} \quad (10.15)$$

The second term on the right of (10.15) is very small compared with the first, and the energy level is close to the middle of the energy gap. This is consistent with earlier assertions that both the bottom of the CB and the top of the VB are far from the Fermi level.

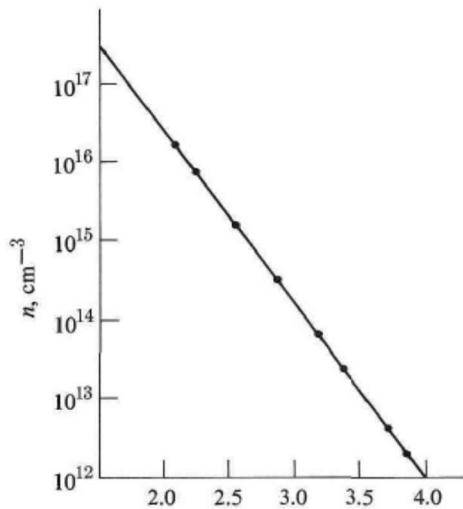


Fig. 10.4 Electron concentration n versus $1/T$ in Ge (in units $10^{-3}K^{-1}$)

The concentration of electrons may now be evaluated explicitly by using the above value of μ . Substitution of Eq. (10.15) into Eq. (10.9) yields

$$n = 2 \left(\frac{k_B T}{2\pi\hbar^2} \right)^{3/2} (m_e m_h)^{3/4} e^{-E_g/2k_B T} \quad (10.16)$$

where $E_g = E_c - E_v$ is the band gap. The important feature of this expression is that n increases very rapidly - exponentially - with temperature, particularly by virtue of the exponential factor. Thus, as the temperature is raised, a vastly greater number of electrons is excited across the gap.

Figure 10.4. is a plot of $\log n$ versus $1/T$. The curve is a straight line of slope equal to $(-E_g/2k_B)$. [The $T^{3/2}$ - dependence in (10.16) is so weak in comparison with the exponential dependence that the former may be disregarded for the purpose of this discussion.)

For Ge at room temperature $\sim 10^{14} \text{ electrons/cm}^3$, a typical value of carrier concentration in semiconductors.

Note that the expression (10.16) also gives the hole concentration, since $n = p$. Our discussion of carrier concentration in this section is based on the premise of a pure semiconductor. When the substance is impure, additional electrons or holes are provided by the impurities. In that case, the

concentrations of electrons and holes may no longer be equal, and the amount of each depends on the concentration and type of impurity present. When the substance is sufficiently pure so that the concentrations of electrons and holes are equal, we speak of an intrinsic semiconductor. That is, the concentrations are determined by the intrinsic properties of the semiconductor itself. On the other hand, when a substance contains a large number of impurities which supply most of the carriers, it is referred to as an extrinsic semiconductor.

Impurity States (extrinsic semiconductor)

A pure semiconductor has equal numbers of both types of carriers, electrons and holes. In most applications, however, one needs specimens which have one type of carrier only, and none of the other. By doping the semiconductor with appropriate impurities, one can obtain samples which contain either electrons only or holes only.

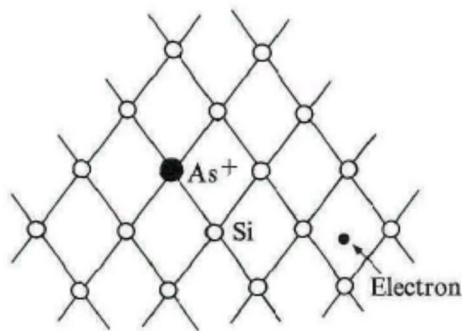


Fig 10.5 An As impurity in a Si crystal.

The extra electron migrates through the crystal. Consider, for instance, a specimen of Si which has been doped by As. The As atoms (the impurities) occupy some of the lattice sites formerly occupied by the Si host atoms. The distribution of the impurities is random throughout the lattice. But their presence affects the solid in one very important respect: The As atom has valence 5 while Si has valence 4. Of the five electrons of As, four participate in the tetrahedral bond of Si, as shown in Fig. 10.5. The fifth electron cannot enter the bond, which is now saturated, and hence this electron detaches from the impurity and is free to migrate through the crystal as a conduction electron, i.e., the electron enters the CB. The impurity is now actually a positive ion, As^+ (since it has lost one of its electrons), and it tends to capture the free electron. But we shall show shortly that the attraction force is very weak, and not enough to capture the electron in most circumstances. The net result is that the As impurities contribute electrons to the CB of the

semiconductors, and for this reason these impurities are called donors. Note that the electrons have been created without the generation of holes.

When an electron is captured by an ionized donor, it orbits around the donor much like the situation in hydrogen. We can calculate the binding energy by using the familiar Bohr model. However, we must take into account the fact that the coulomb interaction here is weakened by the screening due to the presence of the semiconductor crystal, which serves as a medium in which both the donor and ion reside. Thus, the coulomb potential is now given by

$$V(r) = -\frac{e^2}{\epsilon r}, \quad (10.17)$$

where ϵ is the reduced dielectric constant of the medium. The dielectric constant $\epsilon = 11.7$ in Si, for example, shows a substantial decrease in the interaction force. It is this screening which is responsible for the small binding energy of the electron at the donor site.

Using this potential in the Bohr model, we find the binding energy corresponding to the ground state of the donor to be

$$E_d = -\frac{e^4 m_e}{2\epsilon^2 \hbar^2}. \quad (10.18)$$

Note that the effective mass m_e has been used rather than the free mass m . The binding energy of the hydrogen atom is equal to 13.6 eV. The binding energy of the donor is reduced by the factor $1/\epsilon^2$, and also by the mass factor m_e/m , which is usually smaller than unity. Using the typical values $\epsilon \sim 10$ and $m_e/m \sim 0.1$ we find that the binding energy of the donor is about 10^{-3} of the hydrogen energy, i.e., about 0.01 eV. This is indeed the order of the observed values.

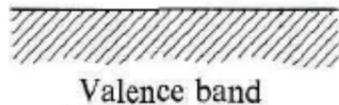
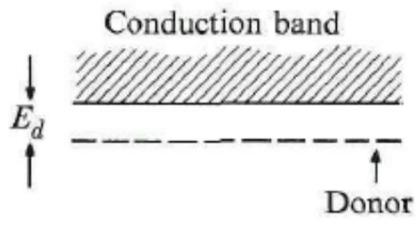


Fig. 10.6 The donor level in a semiconductor.

The donor level lies in the energy gap, very slightly below the conduction band, as shown in Fig.10.6. The level is so close to the CB, almost all the donors are ionized at room temperature, their electrons having been excited into the CB.

It is instructive to evaluate the Bohr radius of the donor electron. Straightforward adaptation of the Bohr result leads to

$$r_d = \epsilon \frac{m}{m_e} a_0 . \quad (10.19)$$

where a_0 is the Bohr radius, equal to 0.53 Å. The radius of the orbit is thus much larger than a_0 , by a factor of 100, if we use the previous values for ϵ and m_e . A typical radius is thus of the order of 50 Å. Since this is much greater than the interatomic spacing, the orbit of the electron encloses a great many host atoms, and our picture of the lattice acting as a continuous, polarisable dielectric is thus a plausible one.

Since the donors are almost all ionized, the concentration of electrons is nearly equal to that of the donors. Typical concentrations are about 10^{15} cm^3 . But sometimes much higher concentrations are obtained by heavy doping of the sample, for example, 10^{18} cm^3 or even more.

Acceptors

An appropriate choice of impurity may produce holes instead of electrons. Suppose that the Si crystal is doped with Ga impurity atoms. The Ga impurity resides at a site previously occupied by

a Si atom, but since Ga is trivalent, one of the electron bonds remains vacant (Fig. 10.7). This vacancy may be filled by an electron moving in from another bond, resulting in a vacancy (or hole) at this latter bond. The hole is then free to migrate throughout the crystal. In this manner, by introducing a large number of trivalent impurities, one creates an appreciable concentration of holes, which lack electrons.

The trivalent impurity is called an acceptor, because it accepts an electron to complete its tetrahedral bond. The acceptor is negatively charged, by virtue of the additional electron it has entrapped. Since the resulting hole has a positive charge, it is attracted by the acceptor. We can evaluate the binding energy of the hole at the acceptor in the same manner followed above in the case of the donor. Again, this energy is very small, of the order of 0.01 eV. Thus essentially all the acceptors are ionized at room temperature.

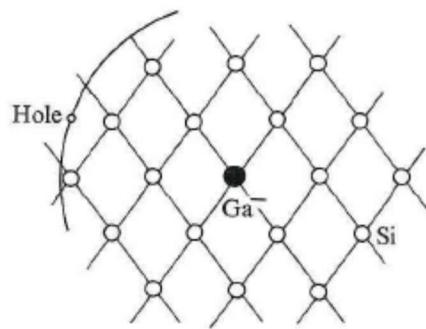


Fig. 10.7 A Ga impurity in a Si crystal (The extra hole migrates through the crystal).

The acceptor level lies in the energy gap, slightly above the edge of the VB, as shown in Fig. 10.8.

This level corresponds to the hole being captured by the acceptor. When an acceptor is ionized (an electron excited from the top of the VB to fill this hole), the hole falls to the top of the VB, and is now a free carrier. Thus, the ionization process, indicated by upward transition of the electron on the energy scale, may be represented by a downward transition of the hole on this scale.

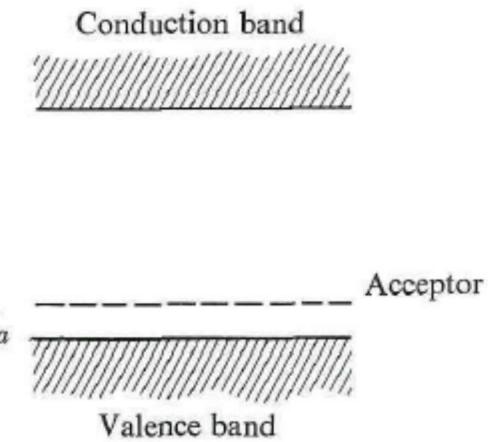


Fig. 10.8 The acceptor level in a semiconductor.

In-text questions

1. Q: How are n-type and p-type semiconductor produced?

A: See pages 146-152.

2. Q: Write down as many as you can of the ways in which the extrinsic semiconductors differ from an intrinsic semiconductor.

A: See pages 151-152.

3. Q: Mark the Fermi-level for (i) an intrinsic semiconductor (ii) n-type semiconductor and (iii) p-type semiconductor.

A: See pages 147-148.

Summary of Study Session 10

In this study session, you have learnt:

1. The definition of semiconductors and understand band structures of semiconductors.
2. The different types of semiconductors and their properties.
3. The distribution of carriers in semiconductor using Fermi-Dirac statistics.

Self –Assessment Questions for Study Session 10

Now that you have completed this study session, you can assess how well you have achieved its learning outcomes by answering the following questions. Write your answers in your Study Diary and discuss them with your Tutor at the next Study Support Meeting.

SAQ 10.1

Enumerate some of the properties of semiconductors.

SAQ 10.2

Explain with suitable diagrams the conduction band, valence band and the forbidden band and hence explain the contribution of electrons and holes to electrical conduction.

SAQ 10.3

What is a hole in a semiconductor and how is it formed?

SAQ 10.4

Distinguish between intrinsic and impurity semiconductors. Give an example for each class of semiconductors.

References/ Further Readings

Ashcroft N. W. and Mermin N.D., (1976), Solid state Physics, Holt, Rine hart, Winston.

Study Session 11 Semiconductors II

Introduction

As previously stated, at infinite temperature some electrons will acquire sufficient thermal energy to raise them from the valence band to the conduction band. The actual number depends on the number of permissible electron energy levels and the probability of these levels being occupied.

Learning Outcomes for Study Session 11

When you have studied this session, you should be able to explain:

- 11.1 The various electronic processes in semiconductors.
- 11.2 The Statistical distribution of carriers in semiconductors.
- 11.3 The characteristics of excitons.

11.1 Main Content

11.1.1. Semiconductor Statistics

Semiconductors usually contain both donors and acceptors. Electrons in the CB can be created either by thermal excitation or by thermal ionization of the donors. Holes in the VB may be generated by interband excitation, or by thermal excitation of electrons from the VB into the acceptor level. In addition, electrons may fall from the donor levels to the acceptor level. Figure 11.1 indicates these various processes.

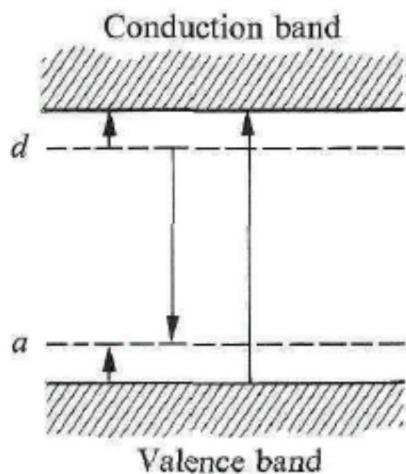


Fig.11.1 The various electronic processes in a semiconductor.

Finding the concentrations of carriers, both electrons and holes, taking all these processes into account, is quite complicated. We shall treat a few special cases, which are often encountered in practice. Two regions may be distinguished, depending on the physical parameters involved: The intrinsic and the extrinsic regions.

The intrinsic region

The concentration of carriers in the intrinsic region is determined primarily by thermally induced interband transitions. In this region $n = p$. The intrinsic region obtains when the impurity doping is small. When we denote the concentrations of donors and acceptors by N_d and N_a , the requirement for the validity of the intrinsic condition is

$$n \gg N_d, N_a \quad (11.1)$$

Since n increases rapidly with temperature, the intrinsic condition becomes more favorable at higher temperatures. All semiconductors, in fact, become intrinsic at sufficiently high temperatures (unless the doping is unusually high).

The extrinsic region

Quite often, the intrinsic condition is not satisfied. For the common dopings encountered, about 10^{15} cm^{-3} , the number of carriers supplied by the impurities is large enough to change the intrinsic concentration appreciably at room temperature. The contribution of impurities, in fact, frequently exceeds those carriers that are supplied by interband excitation. When this is so, the sample is in the extrinsic region. Two different types of extrinsic regions may be distinguished. The first occurs when the donor concentration greatly exceeds the acceptor concentration, that is, when $N_d \gg N_a$. In this case, the concentration of electrons may be evaluated quite readily. Since the donor's ionization energy (i.e. the binding energy) is quite small, all the donors are essentially ionized, their electrons going into the conduction band. Therefore, to a good approximation,

$$n = N_d \quad (11.2)$$

A semiconductor in which $n \gg p$ is called an n-type semiconductor (n for negative). Such a sample is characterized, as we have seen, by a great concentration of electrons. The other type of extrinsic region occurs when $N_a \gg N_d$, that is, the doping is primarily by acceptors. Using an argument similar to the above, one then has

$$p = N_a \quad (11.3)$$

i.e., all the acceptors are ionized. Such a material is called a p-type semiconductor. It is characterized by a preponderance of holes.

In discussing ionization of donors (and acceptors), we assumed that the temperature is sufficiently high so that all of these are ionized. This is certainly true at room temperature. But if the temperature is progressively lowered, a point is reached at which the thermal energy becomes too small to cause electron excitation. In that case, the electrons fall from the conduction band into the donor level, and the conductivity of the sample diminishes dramatically. This is referred to as freeze-out, in that the electrons are now "frozen" at their impurity sites. The temperature at which freeze-out takes place is $E_d \sim kT$, which gives a temperature of about 100°K.

The variation of the electron concentration with temperature in an n-type sample is indicated schematically in Fig. 11.2.

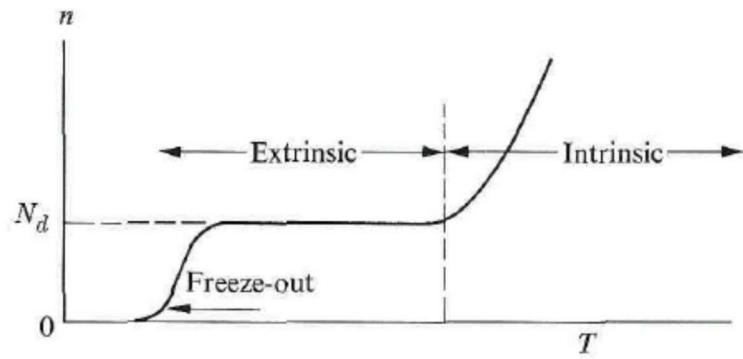


Fig. 11.2 Variation of electron concentration n with temperature in an n-type semiconductor.

11.1.2. Electrical conductivity; Mobility

Electrical conductivity is, of course, the quantity of primary interest in semiconductors. Both electrons and holes contribute to electric current. Assume first that a sample is strongly n-type and contains only one type of carrier: electrons. The conductivity can be treated according to the free-electron model:

$$\sigma_e = \frac{ne^2\tau_e}{m_e} \quad (11.4)$$

where m_e is an effective mass and τ_e is the lifetime of the electron. To estimate the value for σ_e , we substitute $n = 10^{14} \text{ cm}^{-3}$, which is eight orders in magnitude less than that in metals, and $m_e = 0.1m$. This leads to $\sigma_e \sim 10^{-7} (\mu\text{ohm.cm})^{-1}$, which is a typical figure in semiconductors. Although this is many orders of magnitude smaller than the value in a typical metal, where $\sigma_e \sim 1 (\mu\text{ohm.cm})^{-1}$, the conductivity in a semiconductor is still sufficiently large for practical applications.

Semiconductor physicists often use another transport coefficient: mobility. The mobility μ_e is defined as the proportionality coefficient between the electron drift velocity and the applied electric field, i.e.

$$|v_{e_e}| = \mu_e E, \quad (11.5)$$

where $|v_e|$ is the absolute value of the velocity. Taking into account that $\mathbf{j}_e = -en_e v_e$ and $\mathbf{j}_e = \sigma_e E$, we find that

$$\mu_e = \frac{e\tau_e}{m_e}. \quad (11.6)$$

As defined, mobility is a measure of the rapidity of the motion of the electron in the field. The longer the lifetime of the electron and the smaller its mass, the higher the mobility.

We can now express electrical conductivity in terms of mobility. We can write

$$\sigma_e = ne\mu_e \quad (11.7)$$

indicating that σ_e is proportional to μ_e . A typical value for μ_e may be obtained by substituting $\sigma_e \sim 10^{-7} (\mu\text{ohm.cm})^{-1}$ and $n = 10^{14} \text{ cm}^{-3}$ in Eq.(11.7). This yields

$$\mu_e = 10^3 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1} \quad (11.8)$$

What we have said about electrons in a strongly n-type substance can be carried over to a discussion of holes in a strongly p-type substance. The conductivity of the holes is given by

$$\sigma_h = \frac{pe^2\tau_h}{m_h} = pe\mu_h \quad (11.9)$$

where μ_h is the hole mobility.

Let us now treat the general case in which both electrons and holes are present. When a field is applied, electrons drift opposite to the field and holes drift in the same direction as the field. The currents of the two carriers are additive, however, and consequently the conductivities are too. Therefore,

$$\sigma = \sigma_e + \sigma_h \quad (11.10)$$

i.e., both electrons and holes contribute to the current. In terms of the mobilities, one may write

$$\sigma = ne\mu_e + pe\mu_h \quad (11.11)$$

The carriers' concentrations n and p may be different if the sample is doped, as discussed before. Also, one or the other of the carriers may dominate, depending on whether the semiconductor is n- or p-type. When the substance is in the intrinsic region, however, $n = p$, and Eq. (10.30) becomes

$$\sigma = ne(\mu_e + \mu_h) \quad (11.12)$$

where n is the intrinsic concentration. Even now the two carriers do not contribute equally to the current. The carrier with the greater mobility - usually the electron - contributes the larger share.

Dependence on Temperature

Conductivity depends on temperature, and this dependence is often pronounced. Consider a semiconductor in the intrinsic region. Its conductivity is expressed by (11.12). But in this situation, the concentration n increases exponentially with temperature, as may be recalled from (10.16). We may write the conductivity in the form

$$\sigma = F(T)e^{-E_g/2kT} \quad (11.13)$$

where $F(T)$ is a function which depends only weakly on the temperature. (This function depends on the mobilities and effective masses of the carriers.) Thus, conductivity increases exponentially with temperature from 0.0022 K^{-1} as shown in Fig.11.3.

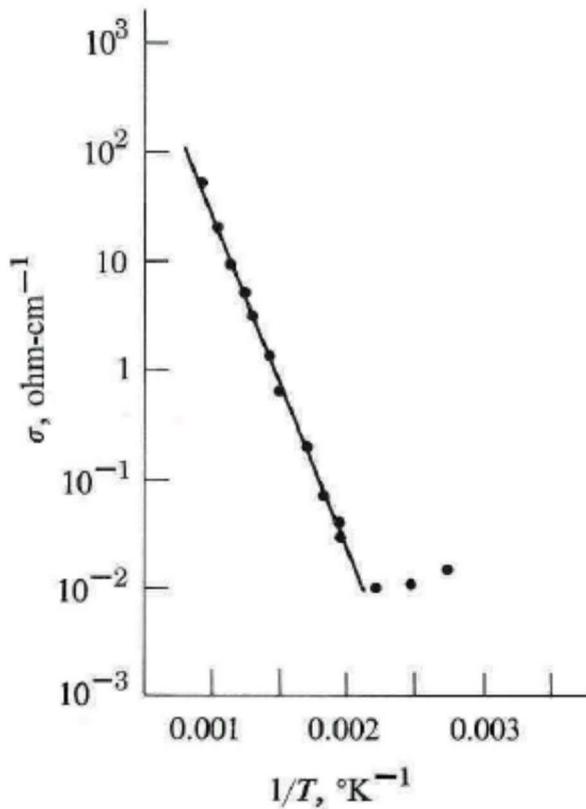


Fig. 11.3 Conductivity of Si versus 1/T in the intrinsic range.

This result can be used to determine the energy gaps in semiconductors. In the early days of semiconductors this was the standard procedure for finding the energy gap. Nowadays, however, the gap is often measured by optical methods. When the substance is not in the intrinsic region, its conductivity is given by the general expression (11.11). In that case, the temperature dependence of the conductivity on T is not usually as strong as indicated above. To see the reason for this, suppose that the substance is extrinsic and strongly n-type. The conductivity is

$$\sigma = ne\mu_e \quad (11.14)$$

But the electron concentration n is now a constant equal to N_d , the donor (hole) concentration. Also, any temperature dependence present must be due to the mobility of electrons or holes.

Mobility versus Temperature: scattering mechanisms

Mobility of electrons (or holes) varies with temperature. In n-type semiconductor, it is expressed as :

$$\mu_e = \frac{e\tau_e}{m_e} = \frac{el_e}{m_e V_e}. \quad (11.15)$$

Since the lifetime of the electron, or its collision time, varies with temperature, its mobility also varies with temperature. Normally, both lifetime and mobility diminish as the temperature rises.

The relaxation time is given by $\tau_e = l_e/V_e$, where l_e is the mean free path of the electron and V_e is the drift velocity. The velocity of electrons is different depending on their location in the conduction band. Electrons at the bottom of the conduction band in a semiconductor obey the classical statistics and not the highly degenerate Fermi statistics prevailing in metals. The higher electrons are in the band, the greater their velocity.

We can evaluate the conductivity by assuming that V_e is the average velocity. The average velocity can be estimated using the procedure of the kinetic theory of gases:

$$\frac{1}{2} m_e V_e^2 = \frac{3}{2} kT \quad (11.16)$$

This introduces a factor of $T^{-1/2}$ dependence in the mobility:

$$\mu_e = \frac{el_e}{m_e^{1/2}(3kT)^{1/2}}. \quad (11.17)$$

The mean free path l_e also depends on the temperature, and in much the same way as it does in metals. l_e is determined by the various collision mechanisms acting on the electrons. These mechanisms are the collisions of electrons with thermally excited phonons and collisions with impurities. At high temperatures, at which collisions with phonons is the dominant factor, l_e is inversely proportional to temperature, that is, $l_e \sim T^{-1}$. In that case, mobility varies as $\mu_e \sim T^{-3/2}$. Figure 11.4. shows this for Ge.

Another important scattering mechanism in semiconductors is that of ionized impurities. When a substance is doped, the donors (or acceptors) lose their electrons (or holes) to the conduction band.

The impurities are thus ionized and are quite effective in scattering the electrons (holes). At high temperatures, this scattering is masked by the much stronger phonon mechanism, but at low temperatures, this latter mechanism becomes weak and the ionized-impurity scattering gradually takes over.

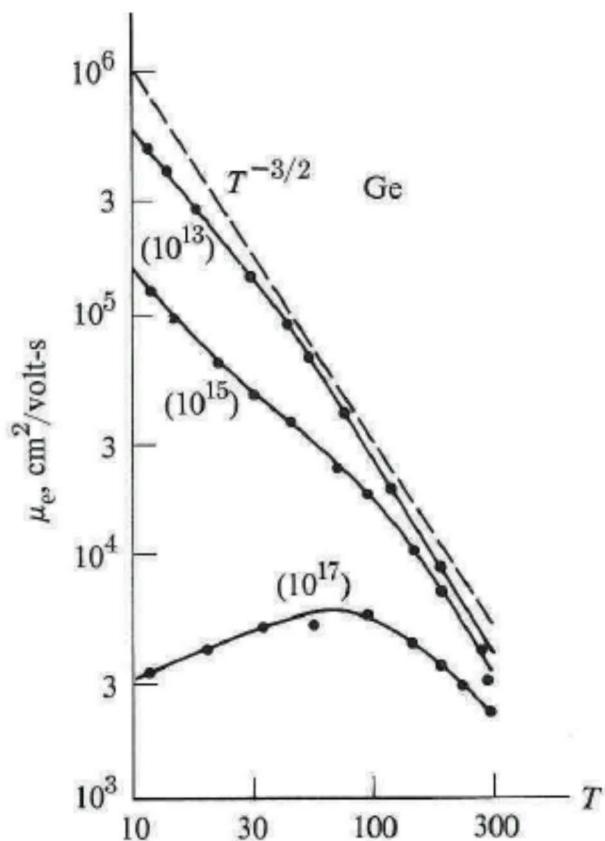


Fig. 11.4 Electron mobility versus T in Ge. The dashed curve represents pure phonon scattering; numbers in parentheses refer to donor concentrations.

11.1.3 Band Structure of Real Semiconductors

So far, we have assumed the simplest possible band structure, namely, a conduction band of a standard form, centered at the origin, $k = 0$, and a valence band of a standard inverted form, also centered at the origin. Such a simple structure is applicable for elucidating many observed phenomena, but it does not represent the actual band structures of many common semiconductors.

Only when one uses the actual band structure is it possible to obtain a quantitative agreement between experiments and theoretical analysis.

A material whose band structure comes close to the ideal structure is GaAs (Fig. 11.5). The conduction band has a minimum at the origin $k = 0$, and the region close to the origin is well represented by a quadratic energy dependence, $(k) = \frac{\hbar^2 k^2}{2m_e}$, where $m_e = 0.072m$. Since the electrons are most likely to populate this region, one can represent this band by a single effective

mass. Note however that, as k increases, the energy $E(k)$ is no longer quadratic in k , and those states may no longer be represented by a single, unique effective mass. The dependence of energy on k in the neighborhood of this *secondary* minimum is quadratic, and hence an effective mass may be defined locally, but its value is much greater than that of the primary minimum (at the center). The actual value is $0.36m$. Due to cubic symmetry, there are six equivalent secondary minima, or *valleys*, in all, along the [100] directions.

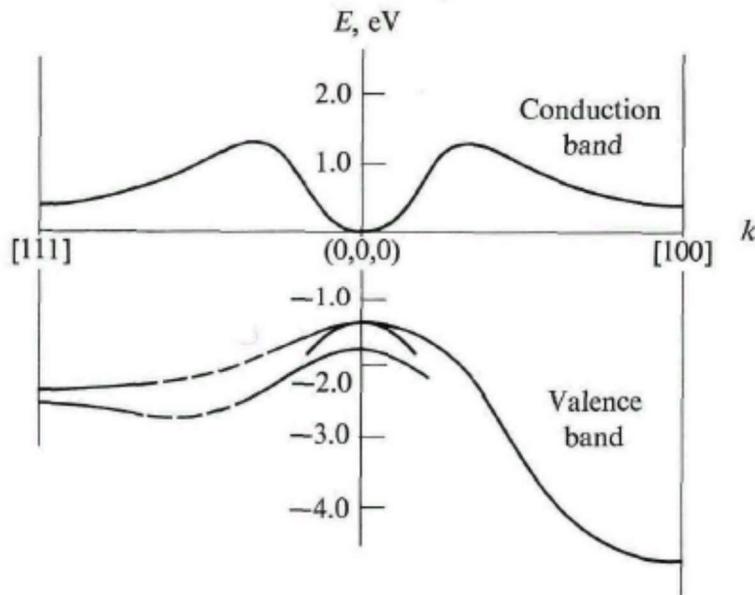


Fig. 11.5 Band structure of GaAs plotted along the [100] and [111] directions.

These secondary valleys do not play any role under most circumstances, since the electrons usually occupy only the central or primary valley. In such situations, these secondary valleys may be disregarded altogether. There are also other secondary valleys in the [111] directions, as shown in Fig. 11.5. These are higher than the [100] valleys, and hence are even less likely to be populated by electrons.

The valence band is also illustrated in Fig. 11.5. Here, it is composed of three closely spaced subbands. The curvatures of the bands are different, so are the effective masses of the corresponding holes. One speaks of light holes and heavy holes.

Other III-V semiconductors have band structures quite similar to that of GaAs.

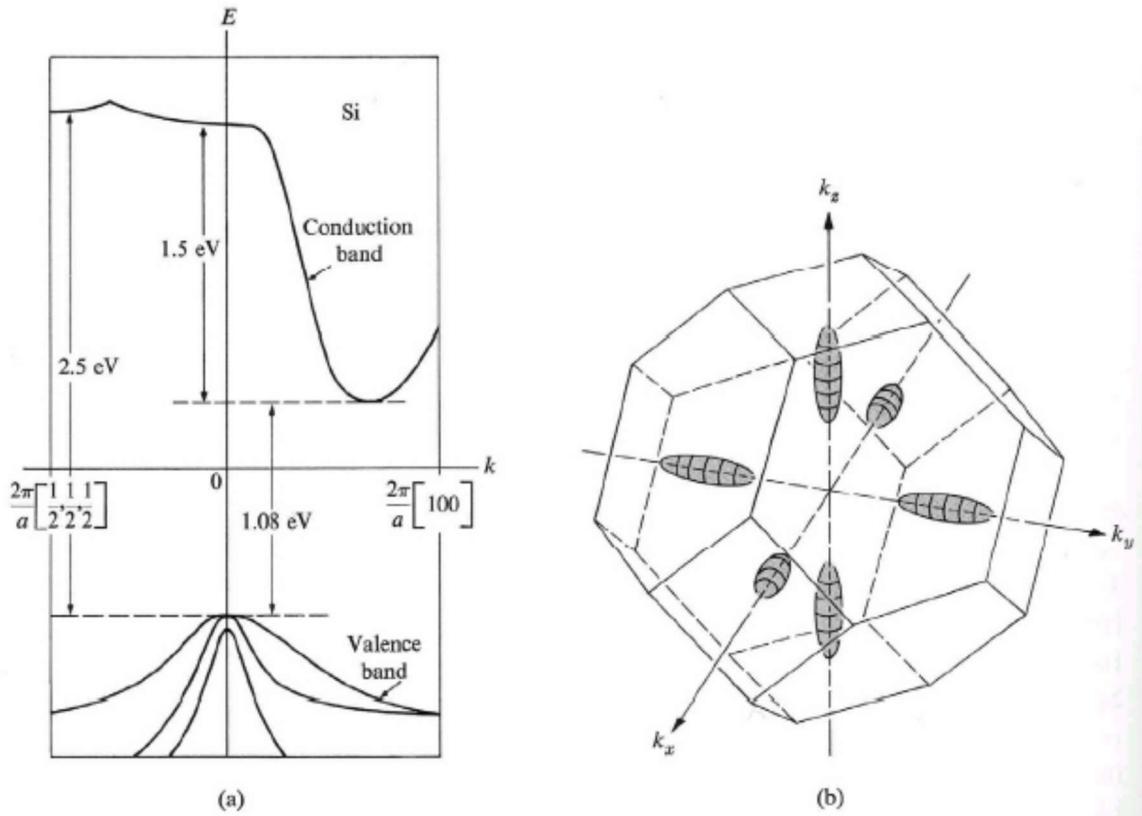


Fig.11.6 (a) Band structure of Si plotted along the [100] and [111] directions (b) Ellipsoidal energy surfaces corresponding to primary valleys along the [100] directions.

Figure 11.6(a) shows the band structure of Si. An interesting feature is that the conduction band has its lowest (primary) minimum not at $k=0$. The minimum lies along the [100] direction, at about 0.85 the distance from the center to the edge of the zone. Note that the bottom of the conduction does not lie directly above the top of the valence band. This type of semiconductors is known as indirect gap semiconductors. These should be distinguished from direct gap semiconductors such as GaAs.

Due to the nature of the cubic symmetry, there are actually six equivalent primary valleys located along the [100] directions. These are illustrated in Fig. 11.6(b). The energy surfaces at these valleys are composed of elongated ellipsoidal surfaces of revolution, whose axes of symmetry are along the [100] directions. There are two different effective masses which correspond to these surfaces: the longitudinal and the transverse effective masses. The longitudinal mass is $m_l =$

0.97m, while the two identical transverse masses are $m_t = 0.19m$. The mass anisotropy ratio is about 5.

The valence band in silicon is represented by three different holes (Fig.11.6a). One of the holes is heavy ($m_h = 0.5m$), and the other two are light. The energy gap in Si, from the top of the valence band to the bottom of the conduction band, is equal to 1.08 eV. The fact that the bottom of the conduction does not lie directly above the top of the valence band, is irrelevant to the definition of the band gap.

11.1.4. Excitons

An electron and a hole may be bound together by their attractive coulomb interaction, just as an electron is bound to a proton to form a neutral hydrogen atom. The bound electron-hole pair is called an exciton, Fig.11.7. An exciton can move through the crystal and transport energy; it does not transport charge because it is electrically neutral. It is similar to positronium, which is formed from an electron and a positron.

Excitons can be formed in every insulating crystal. All excitons are unstable with respect to the ultimate recombination process in which the electron drops into the hole.

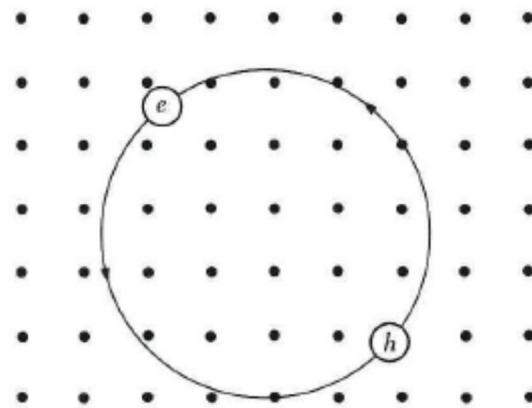


Fig.11.7 The exciton shown is weakly bound, with an average electron-hole distance large in comparison with a lattice constant.

An exciton is a bound electron-hole pair, usually free to move together through the crystal.

The binding energy of the exciton can be measured by optical transitions from the valence band, by the difference between the energy required to create an exciton and the energy to create a free electron and free hole, Fig.11.8.

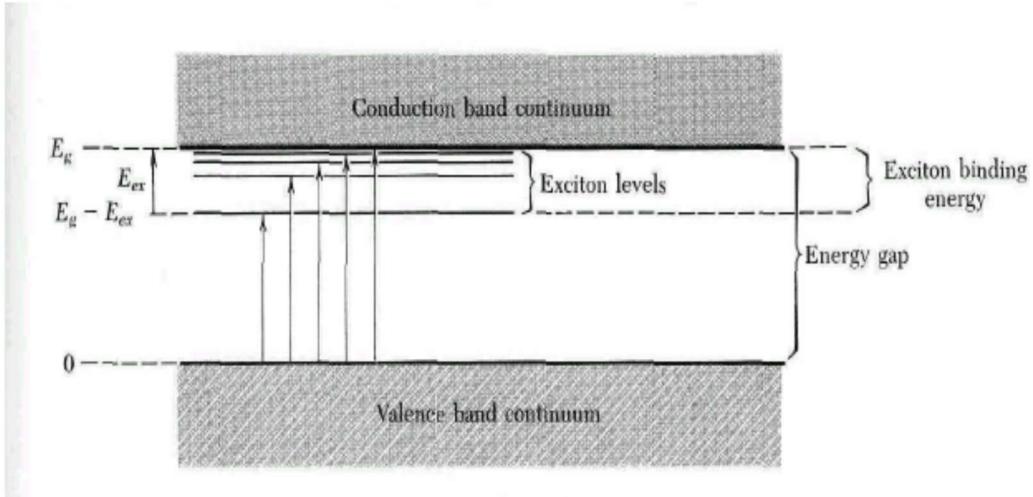


Fig.11.8 Energy levels of an exciton.

Optical transitions from the top of the valence band are shown by the arrows; the longest arrow corresponds to the energy gap. The binding energy of the exciton is E_{ex} , referred to a free electron and free hole.

Energy levels of an exciton can be calculated as follows. Consider an electron in the conduction band and a hole in the valence band. The electron and hole attract each other by the Coulomb potential

$$V(r) = -\frac{e^2}{\epsilon r}, \quad (11.18)$$

where r is the distance between the particles and ϵ is the appropriate dielectric constant. There will be bound states of the exciton system having total energies lower than the bottom of the conduction band.

The problem is the hydrogen atom problem if the energy surfaces for the electron and hole are spherical and non-degenerate. The energy levels are given by

$$E_n = E_c - \frac{e^4 \mu}{2\mu^2 \hbar^2 n^2}. \quad (11.19)$$

Here n is the principal quantum number and μ the reduced mass:

$$\frac{1}{\mu} = \frac{1}{m_e} + \frac{1}{m_h} \quad (11.20)$$

formed from the effective masses of the electron and hole.

The exciton ground state energy is obtained on setting $n = 1$. This is the ionization energy of the exciton.

In-text questions:

1. Q: What are Excitons? Explain briefly how the energy levels of excitons can be calculated.

A: See page 166-167.

2. Q: Calculate the position of Fermi level E_f and the conductivity at 300k for a germanium crystal containing 5×10^{22} arsenic atoms/m³. Also calculate the conductivity if the mobility of the electron is $0.39 \text{ m}^2\text{V}^{-1}\text{s}^{-1}$.

A: $E_c - E_f = 0.16 \text{ eV}$; $\sigma_x = 3120 \Omega^{-1} \text{ m}^{-1}$

3. Q: Obtain the equation for the conductivity of an intrinsic semiconductor in terms of carrier concentration and carrier mobilities. Suggest a method for evaluating the energy gap of semiconductors.

A: See pages 163-165.

Summary of Study Session 11

In this study session, you have learnt:

1. The distribution of carriers in semiconductors using the Fermi-Dirac statistics.
2. About excitons and its characteristics.

Self –Assessment Questions for Study Session 11

Now that you have completed this study session, you can assess how well you have achieved its learning outcomes by answering the following questions. Write your answers in your Study Diary and discuss them with your Tutor at the next Study Support Meeting.

SAQ 11.1

Consider a sample of n-type silicon with $N_d = 10^{21}/\text{m}^3$. Find n and p at 300K. The number of intrinsic carriers at 300K is 9.8×10^{15} .

SAQ 11.2

In a semiconductor the effective mass of the electron is $0.07m_0$ and that of a hole is $0.4m_0$, where m_0 is the free electron mass. Assuming that the average relaxation time for the holes is half that for electrons, calculate the mobility of the holes when the mobility of the electrons is $0.8\text{m}^2\text{V}^{-1}\text{s}^{-1}$.

SAQ 11.3

Compare the density of charge carriers in a pure silicon crystal at the two temperatures 27°C and 57°C . E_g for Si is 1.1eV.

SAQ 11.4

Intrinsic semiconductor material A has an energy gap 0.36eV while material B has an energy gap of 0.72 eV . Compare the intrinsic density of carriers in these two semiconductors at 300K . Assume that the effective masses of all the electrons and holes are equal to the free electron mass.

References/Further Readings

Richard Turton (2000), The Physics of Solids, Oxford University Press.

Study Session 12 Hall Effect

Introduction

It is often necessary to determine whether a material is n-type or p-type. Measurement of the conductivity of a specimen will not give this information since it cannot distinguish between positive hole and electron conduction. The Hall Effect can be utilized in distinguishing between the two types of carriers, and it also allows the density of the charge carriers to be determined.

Learning Outcomes for Study Session 12

When you have studied this session, you should be able to:

- 12.1 Understand the Hall Effect.
- 12.2 Distinguish between the two types of carriers using Hall Effect.
- 12.3 Determine the density of the charge carriers using Hall Effect.

12.1 Main Content

12.1.1 Hall Voltage

If a piece of conductor (metal or semiconductor) carrying a current is placed in a transverse magnetic field, an electric field is produced inside the conductor in a direction normal to both the current and magnetic field. This phenomenon is known as the Hall Effect, and the generated voltage is called the *Hall voltage*. See Fig. 12.1 below

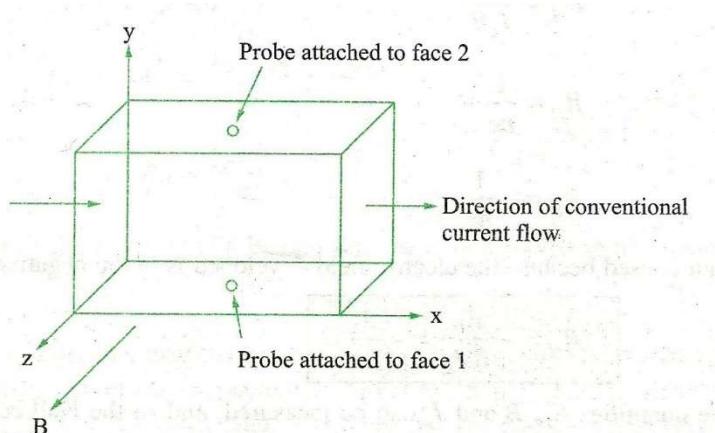


Fig.12.1. Illustration of the Hall Effect

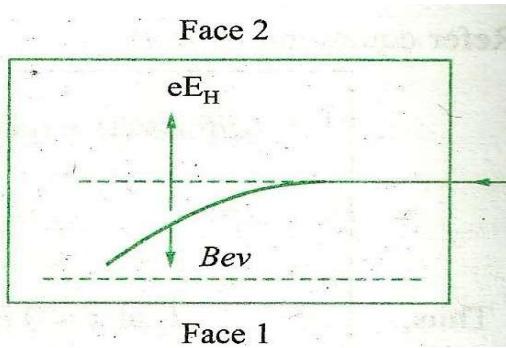


Fig.12.2. Motion of electrons in a n-type semiconductor

Hall Effect may be explained by reference to Fig. 12.2 which shows the front face of the slab only.

If v is the velocity of electrons at right angles to the magnetic field, there is a downward force on each electron of magnitude Bev . This causes the electron current to be deflected in a downward direction and causes a negative charge to accumulate on the bottom face of the slab (face 1). A potential difference is therefore established from top to bottom of the specimen with bottom face negative. This potential difference causes a field E_H in the negative y -direction, and so there is a force of eE_H acting in the upward direction on the electron. Equilibrium occurs when

$$eE_H = eBv \quad (12.1)$$

$$E_H = Bv \quad (12.2)$$

If J_x is the current density in the x -direction, then

$$J_x = nev$$

Where n is the concentration of current carriers.

$$\text{Thus, } E_H = \frac{BJ_x}{ne} \quad (12.3)$$

The Hall Effect is described by means of the Hall coefficient R_H , defined in terms of the current density J_x by the relation

$$E_H = R_H J_x B$$

$$\text{Or } R_H = \frac{E_H}{J_x B} \quad (12.4)$$

$$\text{i.e. } R_H = \frac{1}{ne} \quad (12.5)$$

$$\text{In this case } R_H = -\frac{1}{ne} \quad (12.6)$$

Negative sign is used because the electric field developed is in the negative y-direction

$$R_H = -\frac{E_H}{J_x B} = -\frac{1}{ne} \quad (12.7)$$

All the three quantities E_H , B and J_x can be measured, and so the Hall coefficient and carrier density n can be found,

$$R_H = \frac{E_H}{J_x B} = \frac{1}{pe} \quad (12.8)$$

Where p is the positive hole density.

In-text questions:

1. Q: Explain the term Hall Effect.

A: See pages 174-176.

2. Q: Show that for a p-type semiconductor the Hall co-efficient R_H is given by

$$R_H = \frac{1}{pe}.$$

A: See page 176.

3. Q: Briefly discuss the physical origin of Hall Effect.

A: See page 176.

11.1.2. Determination of the Hall Co-efficient

The Hall coefficient is determined by measuring the Hall voltage that generates the *Hall field*. If V_H is the Hall voltage across the sample of thickness t , then

$$V_H = E_H t$$

Substituting for E_H from equation (12.4), one gets

$$V_H = B t R_H J_x \quad (12.9)$$

If b is the width of the sample, then its cross-section will be bt and the current density

$$J_x = I_x/bt$$

$$\text{Thus } V_H = \frac{BtR_HI_x}{bt} = \frac{BR_HI_x}{b}$$

$$\text{Hence, } R_H = \frac{V_H b}{I_x B} \quad (12.10)$$

Note that the polarity of V_H will be opposite for n -type and p -type semiconductors.

In-text question:

1. Q: Briefly explain how the Hall co-efficient can be determined.

A: See pages 175-176.

12.1.3 Applications of Hall Effect

(i) *Determination of semiconductor type:*

For an n -type semiconductor the Hall coefficient is negative whereas for a p -type semiconductor it is positive. Thus, the sign of the Hall coefficient can be used to determine whether a given semiconductor is n or p -type.

(ii) *Calculation of Carrier Concentration:*

The Hall voltage V_H is measured as usual by placing the two probes at the centres of the top and bottom faces of the sample. If the magnetic flux density is B Wb/m², then

$$n = \frac{1}{eR_H}$$

$$\text{where } R_H = \frac{1}{ne} = \frac{V_H b}{I_x B}$$

(iii) *Determination of Mobility:*

If the conduction is due to one type of carriers, e.g., electrons, we have

$$\sigma = ne\mu_n$$

$$\mu_n = \sigma/ne = \sigma R_H$$

$$\mu_n = \sigma(V_H B/I_x B)$$

i.e., knowing σ , the mobility μ_n can be determined.

(iv) *Measurement of Magnetic Flux Density:*

Since Hall voltage V_H is proportional to the magnetic flux density B for a given current I_x through a sample, the Hall effect can be used as the basis for the design of a magnetic flux density meter.

(v) *Measurement of Power in an Electromagnetic Wave:*

In an electromagnetic wave in free space the magnetic field H and the electric field E are at right angles. Thus, if a semiconductor sample is placed parallel to E it will derive a current I in the semiconductor. The semiconductor is subjected simultaneously to transverse magnetic field H producing a Hall voltage across the sample. The Hall voltage will be proportional to the product E and H , i.e. to the magnitude of the *poynting vector* of the electromagnetic wave. Thus, the Hall effect can be used to determine the power flow in an electromagnetic wave.

(vi) *Hall Effect Multiplier:*

If the magnetic flux density B is produced by passing a current I' through an air-core coil, B will be proportional to I' . The Hall voltage is thus proportional to the product of I and I' . this forms the basis of the mutiplier.

In-text questions:

1. Q: Explain how Hall Effect is used to determine the mobility of Charge carriers.

A: See page 176.

2. Q: Give the principle of Hall Effect multiplier.

A: See page 177.

3. Q: List the applications of Hall Effect.

A: See page 176-177.

Summary of Study Session 12

In this study session, you have learnt:

1. The definition of Hall Effect and how to determine Hall voltage and Hall coefficient.
2. That the Hall coefficient is related to charge mobility and conductivity of semiconductors.
3. The different applications of the Hall Effect.

Self-Assessment Questions for Study Session 12

Now that you have completed this study session, you can assess how well you have achieved its learning outcomes by answering the following questions. Write your answers in your Study Diary and discuss them with your Tutor at the next Study Support Meeting.

SAQ 12.1

A semiconducting crystal of length 0.012m, width 0.0005m, and thickness 0.001m is placed in magnetic field, $B=1\text{Wb.m}^2$ and the current passed through it is 20mA and $E_y = 7.4\text{mV}$. Calculate the Hall Co-efficient.

SAQ 12.2

A semiconducting crystal in the form of a thin rectangular wafer is placed with its plane normal to a magnetic induction, B while a current is made to flow along its length under e.m.f, E_x . Show that an e.m.f. E_y develops across its width.

SAQ 12.3

The Hall coefficient of a certain specimen of silicon was found to be $-7.35 \times 10^{-5} \text{ m}^3\text{C}^{-1}$ from 100 to 400K. Is this semiconductor intrinsic or extrinsic at room temperature, and is it *n-type* or *p-type*? The electrical conductivity at room temperature was found to be $200\Omega^{-1}\text{m}^{-1}$. Calculate the density and mobility of charge carriers at room temperature.

SAQ 12.4

The resistivity of a doped silicon sample is $8.9 \times 10^{-3}\Omega\text{-m}$. The Hall coefficient was measured to be $3.6 \times 10^{-4}\text{m}^3/\text{C}$. Assuming single carrier conduction, find the mobility and density of the charge carriers.

References/Further Readings

Richard Turton, (2000), The Physics of Solids, Oxford University Press.

Study Session 13 Superconductivity

Introduction

The electrical resistivity of all metals and alloys decreases when they are cooled. When the temperature is lowered, the thermal vibrations of the atoms decrease and the conduction electrons are less frequently scattered. The decrease of resistance is linear down to a temperature equal to one-third of the characteristic Debye temperature of the material, but below this, the resistance decreases less rapidly as the temperature falls.

Learning Outcomes for Study Session 13

When you have studied this session, you should be able to:

- 13.1 Explain Superconductivity.
- 13.2 Explain Meissner effect and the BCS theory.
- 13.3 Explain the effect of a magnetic field on superconductors.

13.1 Main Content

13.1.1 A survey of superconductivity

To understand anything about superconductivity, we have to know something about ordinary conductivity. All ordinary materials resist the flow of electricity to a lesser or greater extent. Electricity or electrical current consists of the movement of electrons through a material. In order to make electrons flow, we need to supply energy. Electrons in a metal behave like water in a hose. The water would not flow unless we supply some pressure. The amount of pressure it takes depends on the kind of hose we use.

Superconductors are extraordinary because they alone are immune to the effects of joule heating. It takes no energy at all to make current flow in conductor, and no energy is lost to friction to sustain the current either. Its electrical resistance is precisely zero. Superconductivity was first discovered by the Dutch physicist Heike Kamerlingh Onnes in 1911. Not all pure metals have been found to be superconductors; for example; copper, iron, sodium have not shown superconductivity down to the lowest temperature to which they have been so far cooled. Superconductivity can be exhibited in a large number of alloys e.g Bi-Pd. Also, superconductivity can be shown by conductors which are not metals in the ordinary sense; e.g the semiconducting mixed oxide of barium, lead, bismuth, and conducting polymer.

13.2 An account of the mechanism of superconductivity

The theory of superconduction has been developed over many years, but Bardeen, Cooper, and Schrieffer (BCS) theory in 1957, published the first really satisfactory account of the mechanism. They show that superconductivity occurs when a special state of affairs exists between the conduction electrons.

Two electrons in free space will be mutually repelled by the Coulomb force between, but, in the solid state, the force between the two electrons will be modified by the interaction of the electrons within the crystal lattice. In certain substances, the lattice interaction is so great that simple repulsive force become modified into an attractive force binding certain electrons together into what are called *Cooper pairs*. For two electrons to become bound into a Cooper pair, they must be in equilibrium condition, i.e., with no current flowing, have equal and opposite momenta.

Some of the important properties of superconductors are the following:

1. The current in the superconductors persists for a very long time.
2. Superconductivity has normally been observed only for those metallic substances for which the number of valence electrons Z lies between 2 and 8.
3. In all cases involving transition metals, the variation of T_c with the number of valence electrons shows sharp maxima for $Z = 3, 5, 7$.
4. For a given value of Z , certain crystal structures seem more favourable than others. For example, β -tungsten and α -manganese structure are conductive to the phenomenon of superconductivity.
5. The magnetic field does not penetrate into the body of the superconductor. The property known as the *Meissner effect*, is the fundamental characterization of superconductivity. However, when the magnetic field B is greater than critical value $B_c(T)$, the superconductor becomes a normal conductor [$B_c(T)$ is zero at $T = T_c$] and has the largest value at $T = 0$.
6. Superconductivity occurs in materials having high normal resistivities.
7. When the current through the superconductor is increased beyond a critical value $I_c(T)$, the superconductor again becomes a normal conductor.

8. The specific heat of the material shows an abrupt change at $T = T_c$, jumping to a large value for $T < T_c$.

If one observes the total list of superconducting materials, the general features to be noted are:

- Monovalent metals are not generally superconductors.
- Ferromagnetic and antiferromagnetic materials are not superconductors.
- Good conductors at room temperature are not superconductors, and superconducting materials are not very good conductors at room temperature as the normal metals.
- Amorphous thin films of Be, Bi, and Fe show superconductivity.
- Bismuth, antimony and tellurium become superconducting under high pressure. However, there are exceptions.

In-text questions:

1. Q: Define Superconductivity.

A: See page 179-180.

2. Q: List some of the important properties of superconductors.

A: See pages 180-181.

3. Q: What are Cooper pairs?

A: See page 180.

13. 3 Effects of Magnetic Field

The superconducting state of a metal exists only in a particular range of temperature and field strength. The condition for superconducting state to exist in the metal is that some combination of temperature and field strength should be less than a critical value. Superconductivity will disappear if the temperature of the specimen is raised above its T_c , or if a sufficiently strong magnetic field is employed. The curves are nearly parabolic and can be represented by the relation.

$$H_c = H_0 \left[1 - \left(\frac{T}{T_c} \right)^2 \right] \quad (13.1)$$

where H_c is the maximum critical field strength at the temperature, T , H_0 is the maximum *critical field strength* occurring at absolute zero, and T_c is the *critical temperature*, the highest temperature for superconductivity. Thus the above equation defines a curve which divides the normal region of the field temperature diagram of the metal from the superconducting region.

13.4 Critical Current

The magnetic field which causes a superconductor to become normal from a superconducting state need not necessarily be an external applied field. It may arise as a result of electric current flow in the conductor. The minimum current that can be passed in a sample without destroying its superconductivity is called *critical current* I_c . If a wire of radius r of a type I superconductor carries a current I , there is a surface magnetic field, $H_I = 1/2\pi r$ associated with the current. If H_I exceeds H_c , the material will go normal. If in addition, a transverse magnetic field H is applied to the wire, the condition for the transition to the normal state at the surface is that the sum of the applied field and the field due to the current should be equal to the critical field. Thus, we have

$$H_c = H_I + 2H$$

$$H_I = \frac{I_c}{2\pi r} = H_c - 2H$$

$$I_c = 2\pi r(H_c - 2H) \quad (13.2)$$

This is called Silsbee's rule. The critical current I_c will decrease linearly with increase of applied field until it reaches zero at $H = H_c/2$. If the applied field is zero, $I_c = 2\pi rH_c$.

In-text questions:

1. Q: Briefly explain the effect of magnetic field on superconductors.

A: See page 181.

2. Q: What is meant by the term critical current?

A: See page 182.

3. Q: Describe the Silsbee rule.

A: See page 180.

13.5 Flux Exclusion: The Meissner Effect

It was assumed that the effect of a magnetic field on a superconductor would be as that in a metal. However, in 1933 Meissner and Ochsenfeld measured the flux distribution outside tin and lead specimens which had been cooled below their transition temperatures while in a magnetic field. They found that at their transition temperatures, the specimens spontaneously became perfectly diamagnetic, cancelling all flux inside even though they had been cooled in a magnetic field.

This experiment was the first to demonstrate that superconductors are something more than materials which are perfectly conducting; they have an additional property that a merely resistanceless metal would not possess: a metal in the superconducting state never allows a magnetic flux density to exist in its interior. That is to say, inside a superconducting metal we always have

$$\mathbf{B} = \mathbf{0}$$

whereas inside a merely resistanceless metal there may or may not be a flux density, depending on the circumstances. When a superconductor is cooled in a weak magnetic field, at the transition temperature, persistent currents arise on the surface and circulate so as to cancel the flux density inside, the same way as a magnetic field is applied after the metal has been cooled. This effect, whereby a superconductor never has a flux density even when in applied magnetic field, is called *Meissner effect*.

Summary of Study Session 13

In this study session, you have learnt:

1. The definition and characteristics of superconductors.
2. The Meissner effect.
3. The BCS theory.

Self-Assessment Questions for Study Session 13

Now that you have completed this study session, you can assess how well you have achieved its learning outcomes by answering the following questions. Write your answers in your Study Diary and discuss them with your Tutor at the next Study Support Meeting.

SAQ 13.1

What are superconductors? Mention the important property changes that occur in materials when they change from normal to superconducting state. Give some examples of practical uses that exploit the above property change.

SAQ 13.2

Name two superconducting materials. Define the critical magnetic field and derive the thermodynamic relation connecting the critical magnetic field and the transition temperature.

SAQ 13.3

What is BCS theory of superconductivity? How is superconductivity affected by magnetic field?

SAQ 13.4

Calculate the critical current density for 1mm diameter wire of lead at

- (a) 4.2K
- (b) 7K.

A parabolic dependence of H_c upon T may be assumed. Given: T_c for lead is 7.18 K and H_0 for lead is 6.5×10^4 ampere/metre.

References/Further Readings

Richard Turton, (2000), The Physics of Solids, Oxford University Press.

SOLUTIONS TO SELF ASSESSMENT QUESTIONS

STUDY SESSION 1

SELF ASSESMENT QUESTIONS

1.1. Please see text.

1.2. Please see text.

1.3. Please see text.

STUDY SESSION 2

SELF ASSESMENT QUESTIONS

2.1. Please see text.

2.2. Please see text.

2.3. Please see text.

2.4. 6.9eV or $667 \times 10^3 \text{ kJ/kmol}$.

2.5. Please see text.

STUDY SESSION 3

SELF ASSESMENT QUESTIONS

3.3 Please see text.

3.4 Please see text.

3.5 Please see text.

STUDY SESSION 4

SELF ASSESMENT QUESTIONS

4.1 Please see text.

4.2 Please see text.

STUDY SESSION 5

SELF ASSESSMENT QUESTIONS

5.1 The total energy of the wave is the sum of the kinetic energy E_{kin} and the potential energy E_{pot} . The kinetic energy is the sum of the kinetic energies of all atoms, i.e.

$$E_{kin} = \frac{1}{2}M \sum_n \left(\frac{du_n}{dt} \right)^2$$

where M is the mass of atoms and $\frac{du_n}{dt}$ is the velocity of n -th atom. The potential energy is the potential energy of all the “springs” connecting atoms. For two atoms this energy is the same as that for a harmonic oscillator, i.e. $\frac{1}{2}Cx^2$, where C is interatomic force constant and x is the change in distance between the atoms from the equilibrium distance. Therefore, for atoms n and $n+1$ having displacements u_n and u_{n+1} respectively this energy is $\frac{1}{2}C(u_{n+1} - u_n)^2$. The total potential energy is given by

$$E_{pot} = \frac{1}{2}C \sum_n (u_{n+1} - u_n)^2$$

The total energy of the wave is therefore

$$E = E_{kin} + E_{pot} = \frac{1}{2}M \sum_n \left(\frac{du_n}{dt} \right)^2 + \frac{1}{2}C \sum_n (u_{n+1} - u_n)^2.$$

5.2 Please see text.

5.3 Please see text.

STUDY SESSION 6

SELF ASSESSMENT QUESTIONS

6.1 Please see text.

6.2 Please see text.

6.3. Please see text.

STUDY SESSION 7

SELF ASSESSMENT QUESTIONS

7.1 Please see text.

7.2.(a) Please see text.

(b) Please see text.

7.3. Please see text.

STUDY SESSION 8

SELF ASSESSMENT QUESTIONS

8.1. Please see text

8.2. Please see text

8.3 Please see text

STUDY SESSION 9

SELF ASSESSMENT QUESTIONS

9.1. Please see text

9.2. Please see text

9.3. Please see text

STUDY SESSION 10

SELF ASSESSMENT QUESTIONS

1.(i) Please see text

(ii) Please see text

(iii) Please see text

(iv) Please see text

2. (i) Please see text

(ii) Please see text

(iii) Please see text

3. (i) Please see text

(ii) $E_c - E_f = 0.16\text{eV}$; $\sigma_x = 3120\Omega^{-1}\text{m}^{-1}$

(iii) Please see text

STUDY SESSION 11

SELF ASSESSMENT QUESTIONS

11.1 $n = 10^{21}/\text{m}^3$, $p = 9.6 \times 10^{18}/\text{m}^3$

11.2 $0.07\text{m}^2\text{V}^{-1}\text{s}^{-1}$

11.3 0.128

11.4 1015

STUDY SESSION 12

SELF ASSESSMENT QUESTIONS

12.2 Please see text

12.3 $8.5 \times 10^{22}/\text{m}^3$, $1.47 \times 10^{-2}\text{m}^2\text{V}^{-1}\text{s}^{-1}$

12.4 $0.0344\text{ m}^2\text{V}^{-1}\text{s}^{-1}$, $2.04 \times 10^{22}/\text{m}^3$

STUDY SESSION 13

SELF ASSESSMENT QUESTIONS

13.1 Please see text

13.2 Please see text

13.3 Please see text

13.4 $1.71 \times 10^8\text{A}/\text{m}^2$