Chapter II

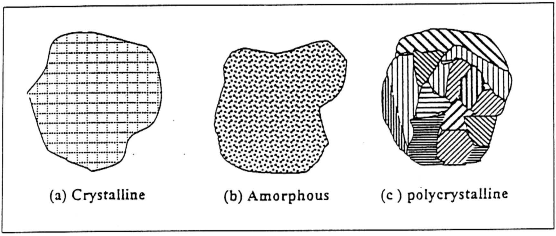
Theory

Chapter II

Crystalline and Amorphous Semiconductors

2.1 Introduction

Solids can be classified into three categories: viz. crystalline, amorphous and polycrystalline materials. The basic arrangement of atoms in these materials differ. In crystalline materials the atoms are arranged periodically and such a basic arrangement is repeated throughout the entire solid. The crystalline solids have, therefore, both short-range and long-range orders of atomic arrangements. In amorphous solids, on the other hand, there is a lack of long range atomic arrangement and there is no regular periodic structure. The polycrystalline solids, are composed of many small granules of the single crystal material.



(a) Crystalline (b) Amorphous (c) polycrystalline

Figure 2.1 : The three types of solids classified according to atomic arrangements. (a)crystalline (b) amorphous is illustrated by microscopic views of the atoms whereas (c)polycrystalline structure is illustrated by a more macroscopic view of adjacent single crystalline regious. [1]

In this chapter, we describe the basic theories of crystalline and amorphous semiconducting materials which we shall be using in analysing our experimental results. The semiconducting materials can be divided into the following types:

i. Elemental semiconductors,

ii. Compound semiconductors,

iii. Amorphous semiconductors, and

iv. Organic semiconductors

i) Elemental Semiconductors

They may be of intrinsic and extrinsic types. Intrinsic semiconduction is a property of the pure material such as pure Silicon(Si) or pure Germanium(Ge). The basic charge carriers in them are electrons and holes. Extrinsic semiconduction results due to the impurity additions in the form of dopants and the process of adding these components has been called doping. There are two different types of extrinsic semiconductions : (a) n-type in which negative charge carriers(electrons) are dominant and (b) p-type , where the positive charge carriers (holes) have the upper hand.

ii) Compound Semiconductors

These semiconductors are formed of elements existing on both sides of group IV of the periodic table. The III-V compounds have formulas of the type MX where M is a +3 valence element while X is a +5 valence element. Similarly II - VI compounds combine a +2 valence element with a +6 valence element. Pure III-V and II-VI compounds are intrinsic semiconductors. They can be made extrinsic by suitable dopings in a fashion similar to those used for elemental semiconductors.

iii) Amorphous Semiconductors

This semiconductor is noncrystalline in nature. The various amorphous semiconductors which are available are : Si , Ge , Se , Te, GeSe etc. The amorphous silicon is often prepared by the decomposition of silane(SiH_4). The process is often incomplete and "amorphous silicon" in reality is a silicon-hydrogen alloy. [2]

iv) Organic Semiconductor

This kind of a semiconductor is neither a covalently bonded nor ionically bonded but a molecular crystal. The central problem in the study of organic semiconductors is to understand how electric charge enters and is transported through the molecular crystal which are often highly disordered on atomic and molecular scales. They may also have applications analogous to other types of semiconductors.

2.2 Crystalline Semiconductors

In extremely "pure" form, Germanium and Silicon exhibits intrinsic conductivity over a wide temperature range. By intrinsic conductivity , we mean the conductivity generated by electrons and holes which are present inside the solid. To understand the nature of electrical conductivity of such a pure material , it is important to realise that the crystal structures of Germanium and Silicon are the same as that of the diamond (crystalline carbon). These belongs to the group IV of the periodic table. These are also referred to as valence crystals, for reasons which will soon become clear. The essential features of the so called diamond structure crystals appear in Fig. 2.2 . Each of the atom is surrounded by four nearest neighbours as shown in the figure. A covalent bond is formed by pairing an electron of each atom with each one of its four neighbours. These covalent bonds are the important cement that holds together the atoms of the diamond crystal structure. These bonds are also responsible for electrical behaviour of the crystal [3].

Semiconductors in their natural crystalline forms are known as intrinsic. If traces of certain chemical impurities are introduced into them, they give arise to energy levels lying within the energy gaps. Such impurities are able to control their conductivities over wide limits. Their temperature dependence of conductivities distinguishes them from metals and insulators.

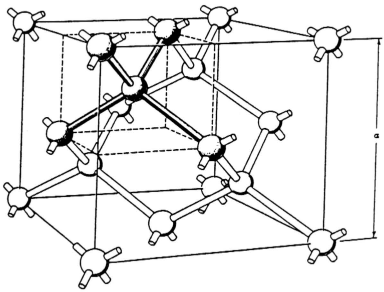


Figure 2.2 : Diamond Crystal Structure).

2.2.1 Band Structure of Crystalline Semiconductors

The electronic energy levels of individual atoms no longer represents the energy levels of electrons inside a crystal. Once a crystal is formed, it is found that the energy levels of the inner-shell electrons are not affected appreciably but those of the outer-shell electrons are, because they are shared by more than one atom in the crystal. The new energy levels of the outer-shell electrons may however be determined by making use of wave mechanics. As a result of such an attempt, it has been found out that the coupling amongst outer-shell electrons of atoms results in a band of closely spaced electronic levels. Before undertaking such an analysis, the qualitative discussion of the formation of energy bands in a semiconductor crystal has to be done. By considering Si or Ge crystal, the outer sub-shell of such an element is known to contain 2s and 2p electrons. To form a crystal of N atoms, it should be possible to vary the spacing between any two of them. When the atoms are situated very far apart, the electrostatic interaction among them is negligible. In such a case the electronic energy levels of isolated atoms will merge with one another. The electronic energy levels of the crystal will then be the same as those of the isolated atoms. If the inner-shell are completely ignored there will be 2N electrons completely filling the 6N p-levels above them. This is illustrated in fig. 2.3 [3]. If the interatomic spacing is gradually decreased, the forces exerted by atoms on their neighbours are no longer negligible, the electronic wave functions of individual atoms overlap and the crystal becomes a single electronic unit which must obey Pauli's exclusion principle. As a result, 2N electronic states belonging to the same energy levels will split into a band. The difference between any two energy levels in a band is very small. Such a group of closely spaced energy levels has been illustrated in fig. 2.3. The 2N states in the lower band are completely filled with 2N s-electrons. The upper band of 6N states are however filled partially with 2N p-electrons. When the atomic distance becomes too small, the bands may overlap as shown in fig. 2.4. This will lead to the 6N upper states merging with 2N lower ones giving rising to a total of 8N allowed states. Each atom of the crystal may then be presumed to contribute four electrons to the lower band. These electrons are then capable of moving freely inside the crystal and the band in which these electrons move is termed the valence band. When the atomic distance in the crystal is reduced beyond the stage, the band overlapping will be enhanced and the interaction amongst atoms becomes extremely strong. At the lattice spacing, shown in fig. 2.4, the valence band filled with 4N electrons is separated by a forbidden band gap of width E_g from an empty band consisting of another 4N vacant state at $T = 0^0K$. The upper vacant band is termed the conduction band and the lower one is the valence band [3]. The energy bands of solids have been studied theoretically by a variety of methods. For semiconductors, the three methods most frequently used are:

i. Orthogonalized Plane Wave (OPW) Method,
ii. Pseudopotential Method, and
iii. k.P. Method

The fig. 2.5 shows the results of studies of the energy band structures of Ge, Si and GaAs. Notice that for any semiconductor, there is forbidden energy region in which allowed states cannot exist under any circumstances. Energy bands are only permitted above and below this energy gap. The upper bands are called conduction bands and the lower bands are the valence bands. The separation between them is the energy gap, which is the most important parameter in semiconductor physics. Fig. 2.6 depicts the simplified band structures of semiconductor.

At room temperature and normal atmosphere, the values of band gaps are 0.66 eV for Ge , 1.12 eV for Si and 1.42 eV for GaAs. These values are applicable to high purity materials. For heavily doped materials, band gaps become smaller. Near room temperatures, the band gaps of Ge and GaAs increase with pressure while that of Si decreases with pressure. In ideal semiconductors, there is no state within the energy gaps. But in real problems, the situation is much different. This is due to several reasons such as;

i. A semiconductor crystal obtained from industry is not a perfect crystalline and has a number of point defects, different impurities and crystal defects.

ii. The semiconductor crystal will contain surface states due to incomplete covalent bonds and other defects.

The presence of these defects is taken care of by placing appropriate energy states within the energy gaps of an otherwise perfect crystal [4].

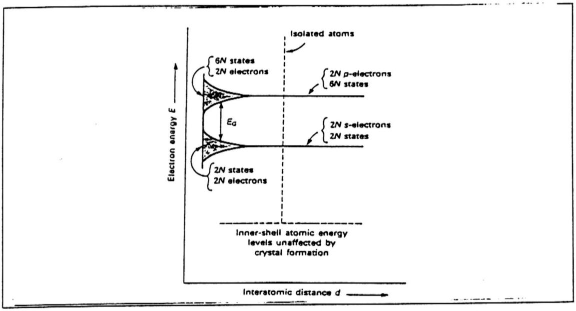


Figure 2.3 : Energy band in a semiconductor crystal for interatomic spacing (d) > lattice spacing

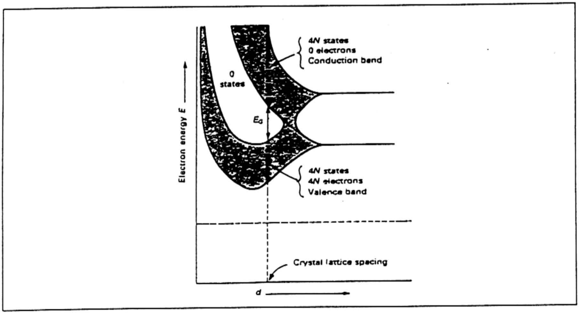


Figure 2.4 : Energy band in a semiconductor crystal for interatomic spacing (d) ~ lattice spacing

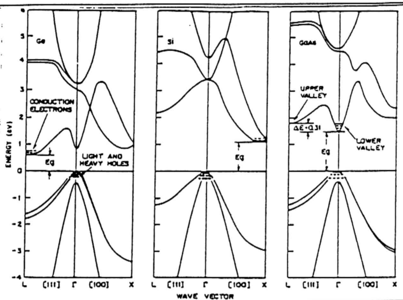


Figure 2.5: Energy-band structures of Ge , Si and GaAs , where E_g is the energy band gap. Plus(+) signs indicate holes in valence bands and minus(-) signs indicate electrons in the conduction bands.

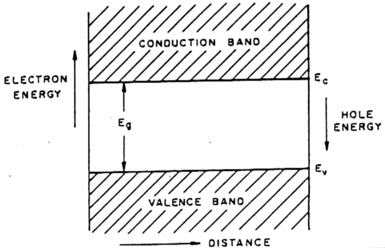


Figure 2.6: Typical simplified band structure of a semiconductor

2.2.2 Electrical Properties

The operation of semiconductor devices depends upon electrons and holes. To discuss their electronic properties, the Si crystal has been taken as an example. Like most other semiconductors, it crystallises in the diamond structure. The bonds in the crystal are of the shared-electron or covalent type. The electrons involved in a particular bond have their spins oppositely directed. The situation prevailing in semiconductors is different from that in a metal where the outer-shell electrons form a cloud that permeates the entire solid. At room temperatures some energy is always associated with the lattice in the form of atomic vibrations which are capable of breaking some electronic bonds. An electronic bond may also be broken by a light quantum. The electrons so liberated by heat or light are free to wander about in the conduction band like free electrons in metals. The second consequence of bond breaking of electrons is the creation of a vacancy or a hole in the valence structure. A valence electron from any neighbouring site may then jump over to fill the original vacancy by virtue of its finite thermal energy, leaving another vacancy or a hole in its place during the process. As this may occur repeatedly, the holes may appear to wander about in the valence band like the electronic motion in the conduction band. The single event of bond breaking then gives arise to both electron and hole conduction in a semiconductor.

The important electrical parameters of crystalline semiconductors are resistivity, mobility and carrier-lifetime. Since the electrical properties directly affect the device characteristics , great attention has been paid to their evaluations. The resistivity type (n or p) is determined by the thermoelectric probe method in which the sign of the thermo-emf determines whether the material is n or p type. In the semiconductor industry the most widely used technique for resistivity measurement is the four-point probe. The usual geometry of the probes is linear as shown in fig. 2.7 , with the probes spaced typically 1mm apart. If the probes are uniformly spaced ($S = S_1 = S_2 = S_3$) , as usually the case , and for probes resting on a semiconductor-infinite medium , the resistivity(ρ) is given by ;

$$\rho = \frac{2\pi SV}{I} \qquad (2.1)$$

Equation 2.1 must be corrected for finite geometry of the semiconductor samples. For an arbitrarily shaped sample the resistivity is given by;

$$\rho = \frac{2\pi SFV}{I} \qquad (2.2)$$

where F is the correction factor appropriate to the sample geometry. The main advantage of this technique is its simplicity and for this reason , it is also employed for resistivity mapping of a semiconductor wafer. The damages introduced when the probes touch the surface is undesirable. Contact less resistivity measurements have also been explored in recent years. Among the various optical and electrical methods of resistivity measurement, only the eddy current method has found practical applications[5].

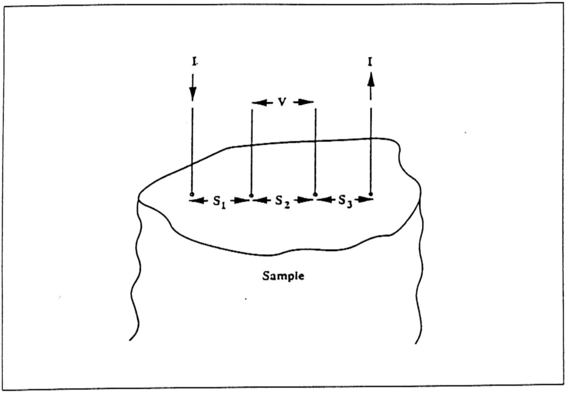


Figure 2.7 : Linear four - point resistivity probe.

The temperature dependence of semiconductor conductivity distinguishes semiconductors from metals and insulators. To analyse the variation of semiconductor conductivity as a function of temperature , the semiconductor conductivity can be expressed as ;

$$\sigma_n(T) = q\mu_n(T)n(T) \qquad (2.3)$$

The mobility $\mu_n(T)$ to be determined by the lattice scattering may be expressed as ;

$$\mu_n(T) = \frac{A_1}{T^{\frac{3}{2}}} \qquad (2.4)$$

The carrier density in "n" type of semiconductor can be written as ;

$$n(T) = A_2\, T^{3/2} \exp\left[\frac{-(E_c - E_f)}{kT}\right] \qquad (2.5)$$

On combining (2.3) , (2.4) and (2.5) we get,

$$\sigma(T) = A\exp\left[\frac{-(E_c - E_f)}{kT}\right] \qquad (2.6)$$

But in doped semiconductors the Fermi level position is not independent of temperature. If we ignore the variations of E_f with T , σ according to equation (2.6) essentially exhibits an exponential variation up to about 60^0K. With further rise in temperature , the supply of electrons from the donor states decreases. Around room temperatures, therefore , n deviates from its exponential dependence and levels of at $n \sim N_d$. In that case , the semiconductor conductivity according to eqn.(2.3) is given by ;

$$\sigma(T) = q[\frac{A_1}{T^{\frac{3}{2}}}]N_d \qquad (2.7)$$

Thus when the carrier density has stopped increasing , σ(T) according to eqn.(2.7) must drop off with an increse in T. If the temperature of the semiconductor is increased sufficiently so that the direct excitation of carriers from the valence band to the conduction band takes place, the specimen tends to become intrinsic. It then starts displaying a characteristic exponential dependence as expressed by ;

$$\sigma_i = q(\mu_n+\mu_p)n_i$$

$$\sigma_i = q[\mu_n(T)+\mu_p(T)](N_c\, N_v)^{1/2}\exp\left[\frac{-E_g}{kT}\right] \qquad (2.8)$$

The variation of lnσ as a function of 1/T is shown in fig. 2.8 [6]. The conductivity measurement cannot reveal which carrier is in majority. However , this information can be obtained from the Hall Effect measurement, which is also the basic tool for the determination of mobility. A static magnetic field applied directly perpendicular to the direction of the current flow produces a force on the flowing change in a direction perpendicular to both. When this happens electrons, and holes will be separated by opposite force acting on them. They in will turn produce an electric field (E_H) which depends on the cross product of the magnetic intensity, H, and the current density, J. This fact is demonstrated in fig. 2.9.

$$E_H = R\,J \times H \qquad (2.9)$$

where R is called the Hall coefficient. Now let us consider a bar of a semiconductor, having dimension, x, y and z. Let J is directed along x and H along z then E_H will be along y, as shown in fig. 2.10. We could then write,

$$R = \left[\frac{V_H / Y}{JH}\right] = \left[\frac{V_H Z}{IH}\right] \qquad (2.10)$$

where V_H is the Hall voltage appearing between the two surfaces perpendicular to y and z. In general , the Hall voltage is not a linear function of the magnetic field , i.e. the Hall coefficient is not generally a constant , but a function of the applied magnetic field. Consequently , the interpretation of the Hall voltage is not usually a simple matter. However , it is easy to calculate this Hall voltage if it is assumed that all carriers have the same drift velocity. We will do this in two steps (a) by assuming that carriers of only one type are present , and (b) by assuming that carriers of both types are present.

(a) Hall Effect for One Type of a Carrier :

Metals and degenerate semiconductors are the examples where only one type of a carrier dominates. The magnetic forces on the carrier is $E_m = q\,(v \times H)$ and is compensated by the Hall field $F_H = q\,E_H$, where v is the drift velocity of the carriers. Assuming the direction of various vectors as before, one finds,

$$v \times H = E_H \qquad (2.11)$$

From simple reasoning , the current density J is the charge q multiplied by the number of carriers traversing unit area in unit time , which is equivalent to the carrier density multiplied by the drift velocity i.e. $J = qnv$. By putting these values in equation (2.10) , we get ,

$$R = \frac{E_H}{JH} = \frac{VH}{qnvH} = \frac{1}{nq} \qquad (2.12)$$

From this equation, the sign of R would be positive for p-type and negative for n-type. When one carrier dominates , the conductivity of the material is

$$\sigma = nq\mu \qquad (2.13)$$

where μ is the mobility of the charge carriers. Thus ,

$$\mu = R\sigma \qquad (2.14)$$

Equation 2.14 provides a method for experimental measurement of mobility in unit cm^2 $volt^{-1}$ sec^{-1}.

(b) Hall Effect for Two Types of Carriers

Intrinsic and lightly doped semiconductors are the two examples of this type. In such cases , the quantitative interpretation of Hall Coefficient is more difficult since both types of carriers contributes to the Hall field. It is also clear that for the same electric and the magnetic field directions , the Hall voltage for p-carriers will be of opposite sign from that of n-carriers. As a result , both mobilities enter into the calculations of Hall coefficient and a weighted average is the result , which is given by ,

$$R = \frac{\mu_h^2 p - \mu_n^2 n}{2(\mu_h p - \mu_n n)} \qquad (2.15)$$

where μ_n and μ_p are the mobilities of electrons and holes respectively , p and n respectively are the densities of holes and electrons. Since the mobility μ_n and μ_p which are not constants but functions with temperature (T), the Hall coefficient given by the equation (2.15), is also a function of temperature. In general $\mu_n > \mu_p$, so that the inversion in the sign of the Hall coefficient may happen only if $p > n$. Thus : "Hall Coefficient inversion" is characteristic only of p-type semiconductors. At point of zero Hall coefficient , it is possible to determine the ratio of mobility and their relative concentrations. [7]

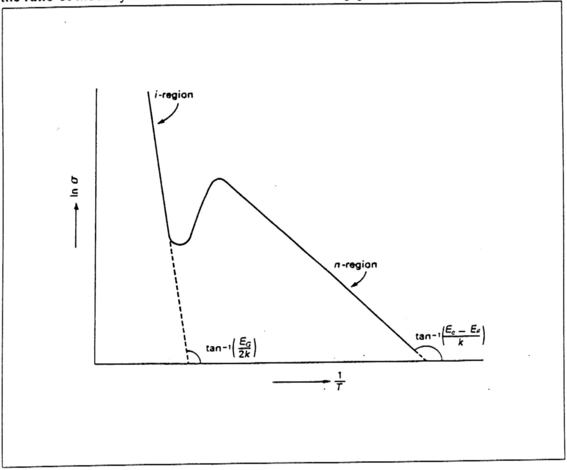


Figure 2.8 : Variation of $\ln\sigma$ versus T^1

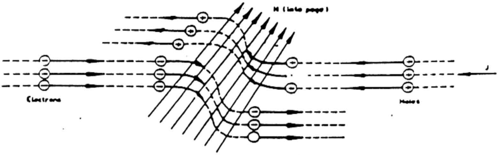


Figure 2.9 : Carrier Separation Due To A Magnetic Field

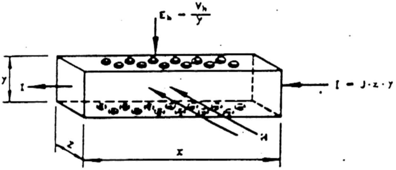


Figure 2.10 : Sample For Studying Hall Effect

2.2.3. Optical Properties

Optical measurements constitute one of the most important means of determining the band structure of semiconductors. Photon induced electronic transitions can occur between different bands, which lead to the determination of the energy gap, or within the same band in the form of the free carrier absorption. Optical measurements can also be used to study lattice vibrations. The transmission coefficient(T) and the reflection coefficient(R) are the two important quantities generally measured. For normal incidence they are given by;

$$\mathrm{T} = \frac{\left(1 - R^2\right)\exp\left(\frac{-4\pi k}{\lambda}\right)}{\left(1 - R^2\right)\exp\left(\frac{-8\pi k}{\lambda}\right)} \qquad (2.16)$$

$$\mathrm{R} = \frac{\left(1 - n^2\right) + K^2}{\left(1 + n^2\right) + K^2} \qquad (2.17)$$

where λ is the wavelength, n the refractive index, K the absorption constant and x the thickness of the sample. The absorption coefficient per unit length is given by;

$$\alpha \equiv \frac{4\pi K}{\lambda} \qquad (2.18)$$

By analysing the T-λ and/or R-λ data at normal incidence , or by making observations of R or T for different angles of incidences , both n and K can be obtained and related to the transition energy between bands. Near the absorption edge the absorption coefficient can be expressed as ,

$$\alpha \approx (h\nu - E_g)^{\gamma} \qquad (2.18)$$

where $h\nu$ is the photon energy , E_g is the band gap and γ is a constant. In the one electron approximation γ equals ½ and 1½ for allowed direct transitions and forbidden direct transitions , respectively (with $K_{min} = K_{max}$ as transition (a) and (b) shown in fig. 2.11.) ; the constant γ equals 2 for indirect transitions (transition (c) in fig. 2.11) , where phonons are involved. In addition there is a bound electron - hole pair (exciton) transition with the energy levels in the band gap which moves through the crystal lattice as a unit. Near the absorption edge , where the values of ($h\nu - E_g$) becomes comparable with the binding energy of an exciton , the Coulomb interaction between the free hole and the electron must be taken into account. When $h\nu \leq E_g$ the absorption merges continuously into the absorption caused by the higher excited states of the exciton. When $h\nu >> E_g$, higher energy bands participate in the transition process and complicated band structures are reflected in the absorption coefficient.[8]

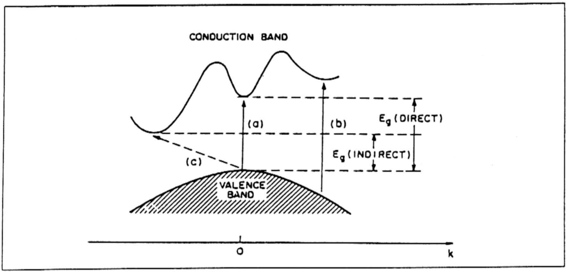


Figure 2.11: Optical Transitions : (a) and (b) direct transitions; (c) indirect transition

2.2.4. Structural Properties

The application of the X-ray diffraction to the surface characterisation requires some considerations of the definition of the material surface. The theory of X-ray scattering grew up in several directions. The diffraction law which has been studied in the X-ray diffraction is the Bragg's law. The glancing angle of the incidence for X-ray reflection, θ, is given by;

$$n\lambda = 2\, d \sin\theta \qquad (2.20)$$

where λ is the wavelength, n is the order of the beam, d is the spacing between lattice planes. It also has been pointed out there, that if the scattering material is distributed such that if $n = 1$, the relation simplifies to $\lambda = 2\, d \sin\theta$. For a single crystal, diffraction of extraneous radiation from tube contaminants can give peaks comparable in intensity to the thin films peaks. After obtaining the diffraction pattern of the film and separating the peaks due to the film from those due to a substrate, a peak-shift analysis will give information on the residual strain and a combination of peak shift and peak broadening data will give informations on microstrain, faulting probability and particle size.

2.3 Amorphous Semiconductors

Amorphous semiconductors are non-crystalline materials. They lack long-range periodic ordering of their constituent atoms. Unlike amorphous metals, amorphous semiconductors do not exist in closed-packed forms, but rather they contain covalently bonded atoms arranged in an open network with correlations in ordering up to the third or fourth nearest neighbours. The short range order is directly responsible for the observable semiconductor properties such as optical absorption edge and activated electrical conductivity[10]. Amorphous semiconductors can be broken down into ionic and covalent materials. The ionic materials which have been studied most are the halides and oxide glasses, particularly the transition metal oxide glasses. Pure materials have just positional disorders and when there is a presence of impurities the transition metal ions will possess some degree of electronic disorder. These types of materials are prepared by rf-sputtering technique. The covalent amorphous semiconductors can be classified into two types such as, purely elemental materials and binary materials. This pure elemental material has perfect covalent bonds and the binary materials are also covalently bonded but they are multicomponent borides, arsenides and chalcogenide glasses. The significance of these covalent amorphous semiconductors are that, they possess compositional as well as positional disorders. The amorphous semiconductors are surrounded by a tetrahedron of four other atoms with an average separation essentially equal to that in the corresponding crystal. Deposited amorphous Ge and Si are far from ideal, and contain many voids within the structure. These voids will lead to the interior dangling bonds and they can dominate the electrical properties of amorphous materials just as vacancies or impurities can control the transport properties of crystalline materials. These voids can be made to disappear however by annealing the amorphous films below the crystallisation temperature[11].

A distinction should be made between amorphous and polycrystalline materials. Polycrystalline semiconductors are composed of grains with each grain containing a periodic array of atoms surrounded by a layer of interconnective or boundary atoms. For smaller and smaller grains, i.e; microcrystallites, the surface layer of each grain contains a larger and larger numbers of atoms relative to the periodically arrayed interior atoms. Eventually for small enough grains, the distinction between the interior and surface is lost and the concept of microcrystallites with a definable periodic region loses its meaning. While attempts have been made to model amorphous semiconductors as microcrystalline , it is now generally accepted that network models are more applicable [12]. These amorphous film are commonly obtained by one of four main methods: such as vapour deposition, sputtering in argon, electrolysis of solutions of materials such as $GeCl_4$ and by means of a glow discharge from silane(SiH_4) or germane(GeH_4) gaseous.

2.3.1. Band Structures of Amorphous Semiconductors

Amorphous semiconductor materials are not only characterised by extended states as in crystalline materials but also by localised states due to slight different magnitude of the bond angle and the length in such materials. It is also not easily doped to change its conductivity while it is difficult to identify whether the conduction is due to the intrinsic carriers or extrinsic mobile carriers. Based on the Anderson's theory , Mott argued that the spatial fluctuation in the potential caused by the configurational disorder in amorphous materials may lead to the formation of localised states, which do not occupy all the different energies in the band, but form a tail above and below the normal band. According to Mott's postulate , there should be a sharp boundary between the energy ranges of extended and localised states. These states are called localised in the sense that an electron placed in such a region will not diffuse at zero temperature to other regions with corresponding potential fluctuations. The concept of localised states in the band tails is used to propose several band models of amorphous semiconductors. Fig. 2.12 illustrates schematically the main features of the various models.

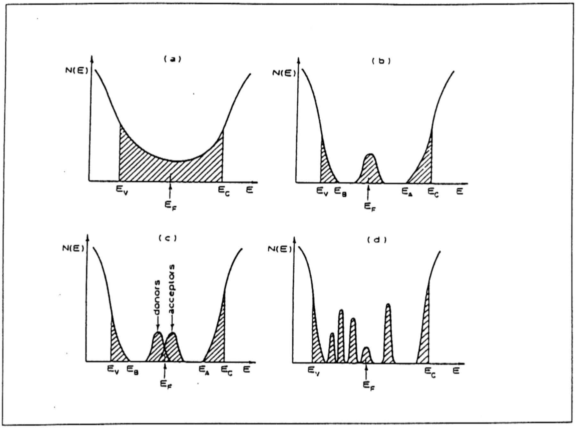


Figure 2.12 : Schematic density of states diagrams for amorphous semiconductors. (a) The Cohen-Fritzsche-Ovshinsky model, (b) The Davis- Mott model (c) Modified Davis-Mott model (d) a "real" glass with defect states.

(a) The Cohen-Fritzsche-Ovshinsky (CFO) model

The CFO model shown in fig. 2.12(a) , assumes that the tail states extend across the gap in a structureless distribution. This gradual decrease of the localised states destroys the sharpness of the conduction and valence band edges. In this model the tail, the conduction and the valence bands overlap , leading to an appreciable density of states in the middle of the gap.

(b) The Davis - Mott Model

According to Davis and Mott , the tails of localised states should be narrow and should extend a few tenths of an electron volt into the forbidden gap. The existence of a band of compensated levels near the middle of the gap , originating from defects in the random network , e.g. dangling bonds , vacancies , etc. Fig. 2.12(b) sketches the Davis-Mott model; here E_c and E_v represent the energies which separate the ranges where the states are localised and extended. The central band may split into a donor and a acceptor band, which will also pin at the fermi level (fig. 2.12 c). Mott suggested that at the transition from the extended to localised states the mobility drops by several orders of magnitude producing a mobility edge. The interval between the energies E_c and E_v acts as s pseudogap and is defined as the mobility gap. The presence of appreciable densities of different types of structural defects is not surprising if one realises that most of the amorphous semiconductors are prepared by fast deposition from the vapour phase or by freezing in the liquid state or by quenching. The position of the fermi-level is largely determined by the charge distribution in the gap states.

On the basis of the Davis-Mott model , there can be three processes leading to conduction in amorphous semiconductors. Their relative contributions to the total conductivity will predominate in different temperature regions. At very low temperatures conduction can occur by thermally assisted tunnelling between states at the fermi level. At relatively higher temperatures the carriers are excited to higher localised states of the band tails; carriers in these localised states take part in the electric charge transport only by hopping. At still higher temperatures carriers are excited across the mobility edges into the extended states. The mobility in the extended states are much higher then in the localised states[13]. Therefore the electrical conductivity measurements over a wide temperature range are needed to study the electronic structure of an amorphous semiconductor.

2.3.2. Electrical Transport Properties

To investigate and explain the electronic transport properties of amorphous semiconductors one may start from the Davis-Mott model.

(a) D.C. Electrical Conductivity

The essential features of the Davis-Mott model for the band structure of amorphous semiconductors are the existence of narrow tails of localised states at the extremities of the valence and conduction bands and furthermore a band of localised levels near the middle of the gap. These lead to three basically different channels for conduction.

i) Extended State Conduction

The conductivity of any semiconductor can be expressed as ,

$$\sigma = -e \int N(E)\ \mu(E) kT \left(\frac{\partial f(E)}{\partial E} \right) dE \qquad (2.21)$$

where f(E) is the fermi-Dirac distribution function ,

$$f(E) = \frac{1}{1 + \exp\left(\frac{E - E_f}{kT} \right)} \qquad (2.22)$$

Using relationship

$$\left(\frac{\partial f(E)}{\partial E} \right) = -f(E) \left[\frac{1 - f(E)}{kT} \right] \qquad (2.23)$$

σ can be written as ,

$$\sigma = e\int N(E)\mu(E)f(E)[1-f(E)]dE \qquad (2.24)$$

In the Davis-Mott model, the fermi level E_f is situated near the middle of the gap and thus sufficiently far from E_c, the energy which separate the extended from the localised states, so that Boltzman statistics can be used to describe the occupancy of states as,

$$f(E) = \exp\left[\frac{-(E-E_f)}{kT}\right] \qquad (2.25)$$

According to Mott's view[14] , the mobility drops sharply at the critical energy E_c (or E_v) but at the present it is not exactly known how the mobility depends on the energy in both the conduction regimes. In the non-degenerate case and under the assumption of a constant density of states and constant mobility the conductivity due to electrons excited beyond the mobility edge in to the extended states is given by ;

$$\sigma = eN(E)kT\mu_c\exp\left[\frac{-(E-E_f)}{kT}\right] \qquad (2.26)$$

where the μ_c is the average mobility. The number of electron is given by ;

$$n = \int_{E_c}^{\infty} N(E_c)\exp\left(\frac{-(E-E_f)}{kT}\right)dE$$

$$n = N(E_c)kT\ \exp\left[\frac{-(E_c-E_f)}{kT}\right] \qquad (2.27)$$

To get and idea of the order of magnitude of μ_c we shall follow Mott's treatment[15]. We define $\sigma(E_c) = e\ N(E_c)\ \mu_c\, kT$. If $N(E_c) \cong \langle N(E)\rangle/3$, where $\langle N(E)\rangle$ is the average density of states over the band , then ;

$$\sigma(E_c) = e\langle N(E_c)\rangle\mu_c\ \frac{kT}{3} \qquad (2.28)$$

Mott calculated the lowest value of the electrical conductivity before the start of the activated process , i.e. just at E_c. The quantity he called the "minimum metallic conductivity". He derived the equation ;

$$\sigma_{min} = \text{constant} \,.\, \left(\frac{e_2^2}{ha}\right) \quad (2.29)$$

where the constant lies in the range between 0.026 and 0.1 ; $\sigma_{min}$is usually of the order 200-300Ω^{-1}cm^{-1}. Taking constant = 0.026 , one finds the final mobility to be ,

$$\mu_c = 0.078\left(\frac{e}{ha}\right) <N(E)>kT \qquad (2.30)$$

In the nearly free electron model , N(E) is given by ;

$$N(E) = \frac{km}{2\pi^2 h^2} \qquad (2.31)$$

On the other hand , the maximum energy E_{max} of the band , which also yields the width B of the band , is expressed as ,

$$E_{max} = \frac{h^2\pi^2}{2ma^2} andk_{max} = \frac{\pi}{3} \qquad (2.32)$$

if $N(E) \approx \frac{1}{a^3}$ B

Introducing this result in the expression (2.30) yields ,

$$\mu_c = \frac{0.078ea^2B}{hkT} \qquad (2.33)$$

By taking a = 2Å , B = 5 eV , one finds at room temperature that $\mu_c \approx 10$ cm^{2}V^{-1}s^{-1}. This values correspond to a mean free path comparable to or less than the interatomic distance. Cohen [16] suggested that the conductivity in this case would be properly described by a diffusive Brownian-type motion. In this regime the mobility can be obtained with the help of the Einstein's relation,

$$\mu = \frac{eD}{kT} \qquad (2.34)$$

The diffusion coefficient D may be written as ,

$$D = \frac{1}{6} \nu\, a^2 \qquad (2.35)$$

where ν is the jump frequency and a , the interatomic separation. The mobility in the Brownian-motion regime is then given by ,

$$\mu_c = \frac{ea^2\nu}{6kT}$$

This expression yields the same temperature dependence as (2.33) derived by Mott. Since $\mu_c \propto 1 / kT$, one expects that the expression for conductivity is of the form ,

$$\sigma = \sigma_0 \exp\left[\frac{-(E_c - E_f)}{kT}\right] \qquad (2.37)$$

where σ_o is preexponential factor, E_c-E_f is the activation energy, k is the Boltzmann constant and T is the experimental temperature. Mott [17] , has made an estimation of the pre-exponential σ_0 which is found to lie between 10 and 10^3 $\Omega^{-1}cm^{-1}$ in most amorphous semiconductors.

ii) Conduction in Band Tails

If the wave function are localised , so that $\sigma(E) = 0$, conduction can only occur by thermally activated hopping. Every time an electron moves from one localised states to another , it will exchange energy with a phonon. It may be expected that the mobility will have a thermally activated nature ,

$$\mu_{hop} = \mu_0 \exp.\left[\frac{-W(E)}{kT}\right] \qquad (2.38)$$

The pre-exponential factor μ_0 has the form ,

$$\mu_0 = \frac{1}{6}\nu_{ph}\,\frac{eR^2}{kT} \qquad (2.39)$$

where ν_{ph} is the phonon frequency and R is the distance covered in one hopping. For typical phonon frequency $\nu_{ph} = 10^{13}\ s^{-1}$ and $W \cong kT$, (2.38) yields mobility of the order of $10^{-2}\ cm^{-2}V^{-1}s^{-1}$ at room temperature. Comparison of this value with the one calculated for conduction in the extended states suggests , as postulated by Mott , that the mobility may drop by a factor of at least 100 at the energy which separates the localised states from the non localised ones. The conductivity being an integral over all available energy states , will depend on the energy distribution of the density of localised states. The conductivity σ_{hop} due to electrons can be easily expressed by the following equation.

$$\sigma_{0\,hop} = \frac{1}{6}\nu_{ph}e^2\,R^2\,N(E_c) \qquad (2.40)$$

(b) A.C. Electrical Conductivity

For highly resistive amorphous films which has reactant and non-reactant components it is suggested that the alternating current (a.c.) measurement rather than the d.c. measurement be performed because of the polarisation occurring at the electrodes due to the inability of the mobile carriers to cross the thin film or electrode interface producing extremely small d.c. current. Common measurements by the four probe method normally does not solve the above mentioned problem. To reduce the error in the measurement, it is better to have an effective area to be as small as possible. Another problem while measuring the d.c. effect is that it does not provide overall information as compared to the a.c. counterpart. The a.c. analysis over a frequency range is a complex plane analysis which allows the determination of the individual component values from the reduced data obtained over a range of frequency.

The electrical conductivity $\sigma(\omega)$ is measured as a function of the frequency (ω) of an alternating electric field. The frequency range in the measurement extends from 10^2 to 10^{18}Hz. A difference in behaviour of $\sigma(\omega)$ is expected when conduction occurs by the motion of charge carriers in the extended states or by hopping in the localised states. The enhancement of a.c. conductivity is achieved through the contributions of three mechanism:

a) Transport by mobile carriers excited to the extended or non-localised states near conduction band edge E_c or valence band edge E_v. No frequency dependence of the conductivity due to the carriers in the extended states is expected up to 10^7 Hz.

b) Transport by carriers into the localised states at the edge of the valence and conduction band. Here the transport is by the "hopping" mechanism with the conductivity increasing with frequency and temperature.

c) Hopping transport by carriers with energies near the fermi level takes place provided the electrons density N_v is finite. The distinct different with the transport in (b) is that the conductivity increases with the frequency and temperature[20 , 21].

The angular frequency-dependent conductivity $\sigma(\omega)$ is modified according to the Austin-Mott formula[22] with a.c. conductivity having a frequency dependence proportional to 10^{s} , with exponent $s \approx 0.7 - 1.0$ for frequencies below 10^8 Hz. When the conduction occurs by the phonon-assisted hopping between localised states one, expects the conductivity to increase with frequency. Austin and Mott [23], have derived the following formula when hopping conduction takes place near E_f:

$$\sigma(\omega) = \frac{\pi}{3} e^2 k T [N(E_f)]^2 \alpha^{-5} w[\ln\left(\frac{\upsilon_{ph}}{\omega}\right)]^4 \tag{2.41}$$

where ν_{ph} is a phonon frequency (typically of the order of $10^{13}s^{-1}$) and α describes the decay of localised state wave function ($1/\alpha$ is effective range of wave function). Two important features can be drawn from this equation. (I) $\sigma(\omega)$ varies linearly with temperature ; (II) because of the presence of logarithmic term $[\ln\left(\frac{\upsilon_{ph}}{\omega}\right)]^4$, the slope of a plot $\ln\sigma(\omega)$ versus $\ln(\omega)$ in not constant , but decreases slightly with increasing frequency of the applied field. The above equation can be an approximated by an expression of the form ;

$$\sigma(\omega) \approx \text{const. } W^s$$

where the exponent s is defined as $d(\ln\sigma) / d(\ln\omega)$ and is given by ,

$$s = 1 - \frac{4}{\ln\left(\frac{\nu_{ph}}{\omega}\right)} \qquad (2.42)$$

Taking $\nu_{ph} \approx 10^{13}\ s^{-1}$ the exponent decreases from a value $s = 0.84$ at $\omega = 10^2\ s^{-1}$ to $s = 0.65$ at $\omega = 10^8 s^{-1}$. If one makes the proper choice of the parameters α^{-1}, (2.41) offers a straight forward way of evaluating the density of states at the fermi levels. For amorphous semiconductors the constant is estimated in the range 10^{18} - 10^{20} $ev^{-1}\ cm^{-3}$. The exponent power β and s are written as;

$$\beta = \frac{6k_BT}{W_m} \qquad (2.43)$$

$$s = 1 - \beta \qquad (2.44)$$

$$1 - s = \frac{6k_BT}{W_m} \qquad (2.45)$$

where k_B is the Boltzmann constant, T is the absolute temperature and W_m is the binding energy. Equation (2.44) is only valid at lower temperatures when β has smaller temperature dependence. Thus, the minimum hopping distance r_{min} is given by [24] ,

$$r_{min} = \frac{2e^2}{\pi\varepsilon\varepsilon_0 W_m} \qquad (2.46)$$

where ε_0 is the primittivity of vacuum.

2.3.3 Optical Properties

The transition process from the valence band to the conduction band is known as the fundamental absorption. In a fundamental absorption process, where an electron absorbs a photon, the energy gap must be equal to or smaller than the photon energy ($h\nu \geq E_g$). It was observed that there is a similarity between the amorphous semiconductors and crystalline semiconductors where the absorption spectra is concerned. On the other hand, the obvious difference was noted in the cut-off frequency, where it was found that it is higher and sharply defined in crystalline states compared to the amorphous ones. In the amorphous semiconductors the excitation of electrons take place from the localized states in the valence bands to the delocalized states in conduction bands. This causes the difference in the cut-off frequencies of the absorption spectra in the crystalline and the amorphous states. The cut-off (minimum) frequency is referred to as the absorption edge where the excitation of electrons from valence band to conduction band occurs when $\nu_0 = E_g/h$. The absorption spectra of amorphous semiconductors is shown in fig. 2.13. the absorption spectra of amorphous semiconductors has three absorption regions that is:

i. Region A : high absorption region ($\alpha > 10^4$ cm^{-1}), where the absorption coefficient α is detailed by the realtion $h\nu\alpha \sim (h\nu - E_g)^r$ (r is a constant of the order 1, E_g is the optical energy gap). At this region, the energy gap is obtained for all amorphous semiconductors.

ii. Region B : this is the exponential absorption region and is frequently observed in crystalline semiconductors, which extends over 4 orders. From this region, the width of the band tail (E_e) is obtained.

iii. Region C : this is the week absorption tail that emanates from the transition of defect states at the mid gap.

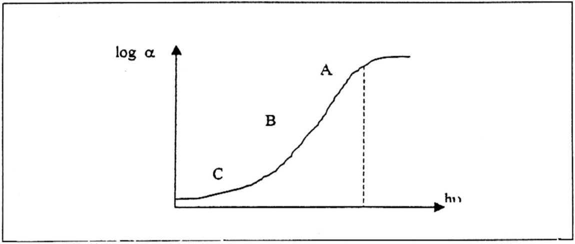


Figure 2.13 : Absorption spectrum of amorphous semiconductor.

The optical properties of thin films can be derived from the thin film interferometry. In principle, the determination of the amplitudes and intensities of the beam of light reflected or transmitted by single layer of a film or multilayer films are straight forward. Theoretical and experimental analyses on the optical behaviour of thin films deal primarily with the optical reflection coefficient (R), the transmission coefficient (T), and the absorption coefficient (A), such that;

$$R + T + A = 1 \qquad (2.47)$$

These three parameters are directly related to the optical constant of the films as well as the film thickness (t_f). The optical constants of solids may be determined in two ways:

i. Intensity measurement (non-polarised light) , which can be either for normal or oblique incidence.

ii. polarisation measurement with oblique incidence [25].

Conventional optical methods for determining the optical constants and thicknesses are , employed by several researches [26 , 27 , 28]. They can be summarised as follows,

$$T \equiv T(n, k, t, \lambda)$$

$$R \equiv R(n, k, t, \lambda) \qquad (2.48)$$

where T and R are the spectral transmittance and reflectance respectively of a film and λ is the wavelength of the incident light. It is to be noted that the conventional methods mentioned above always lead to multiple solutions [29 , 30] and to choose the correct solution, an efficient criterion has to be applied[31 , 32]. Fig. 2.14 represents a thin film with a complex refractive index , $n^*_2 = n - ik$, bounded by two transparent media with refractive indices n_1 and n_3.

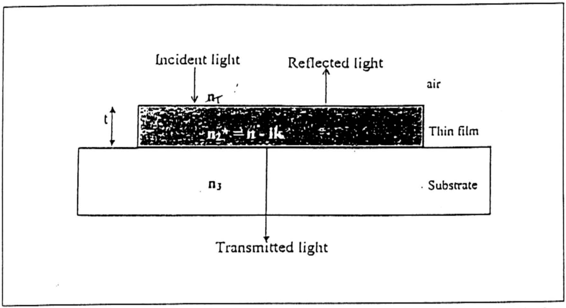


Figure 2.14 : Transmission and reflection of light by a single thin film. (t = thickness of thin film , n_1 = refractive index of air, n_3 = refractive index of substrate , n^*_2 = complex refractive index of thin film , k = extinction coefficient)

The transmission and reflectance normal to the incident film are given by ;

$$T = \frac{n_1}{n_3} \frac{|t_{32} t_{12} \exp.(i\beta)|^2}{|1 + r_{32} r_{12} \exp.(i2\beta)|^2} \qquad (2.49)$$

$$R = \frac{|r_{12} + r_{23} \exp.(i2\beta)|^2}{|1 + r_{12} r_{23} \exp.(i2\beta)|^2} \qquad (2.50)$$

The indices 1 , 2 and 3 correspond to air , film and substrate respectively ; r_{ij} and t_{ij} are Fresnel coefficients which for normal incidences are ,

$$t_{ij} = \frac{2N_i}{N_i + N_j} \qquad (2.51)$$

$$r_{ij} = \frac{N_i - N_j}{N_i + N_j} \qquad (2.52)$$

Taking $N_2 = n + ik$ with n and k real functions of λ, the complex angle β is given by ;

$$\beta = \frac{2\pi tn}{\lambda} = \frac{2\pi tn}{\lambda} + \frac{i2\pi tk}{\lambda} \qquad (2.53)$$

where t is the film thickness and λ is the vacuum wavelength of the incident radiation. Since the surrounding media (glass and air) is assumed to be transparent , N_1 and N_3 are real quantities which are represented by n_1 and n_3. By replacing Fresnel coefficient (equation 2.51) in to equation 2.49 , and considering a weak absorption i.e. $k^2 << (n_2 - n_1)^2$ and $k^2<<(n_2 - n_3)^2$, equation (2.49) can be approximated as:

$$T = \frac{16n_1n_3n^2A}{C_1^2 + C_2^2A_2 + 2C_1C_2ACos\left(\frac{4\pi tn}{\lambda}\right)} \qquad (2.54)$$

where $C_1 = (n_1 + n)(n + n_3)$; $C_2 = (n - n_1)(n_3 - n)$ and $A = \exp(-4\pi tk/\lambda) = \exp(-\alpha t)$. where the absorption coefficient is [33-36],

$$\alpha = \left(\frac{4\pi k}{\lambda}\right) \qquad (2.55)$$

2.3.4. Structural Properties

X-ray diffraction studies have played an important role in identification and characterisation of solids and by distinguishing crystalline from amorphous substances. The determination of the crystal structure, space groups, unit cell dimensions, atomic co-ordinates and actual electron density distribution around atoms can be determined by using the X-ray diffraction. Thin films consist of many single crystals known as grains or crystallites , it is then called a polycrystal whose grain orientations are random. The bulk crystal has grain sizes much larger than a thin film which is more disordered. In order to “see” the fine details of molecular structures of many amorphous thin films , it is necessary to use radiation of wavelengths in the neighbourhood of 0.1mm (10^{18}Hz) or smaller about the dimension of atoms. Such a radiation beam is readily available in the form of X-rays produced by bombarding a target composed of an element of intermediate atomic number between Cr and Mo in the periodic table. The formation of an image of the object under scrutiny is not possible when X-rays, neutrons and high energy electrons are used to provide atomic resolution.

The pattern of radiation which was scattered is known as the diffraction pattern and according the Bragg's equation it is;

$$n\lambda = 2\ d\ \sin\theta \qquad (2.56)$$

where λ is radiation wavelength, n is integer, d is the parallel plane lattice spacing and θ is the angle of incidence of the beam. The diffracted rays produced from the crystals are divided in two stages;

i. scattering by individual atoms, due to regularity of the atomic lattice.

ii. mutual interference between the scattered rays where the wavelength is of the same order as the interatomic distance[37, 38].

2.3.5. The Role of Hydrogen Content and Dangling Bonds in Hydrogenated Amorphous Silicon

In order to understand the role played by hydrogen during the film growth, it is helpful to examine the net reaction equation [43].

$$SiH_x(\text{plasma}) \leftrightarrow Si(\text{solid}) + xH(\text{plasma})$$

The forward reaction represents the film deposition, while the reverse one, represents the film erosion. The film formation is actually due to two types of reactions namely deposition and hydrogen "etching" of the growth surface. The exact rate of the film growth depends on the balance between the two. Under normal deposition conditions, the reaction is very far from equilibrium and moves predominantly forward, leading to film formation. But even in the case of pure SiH_4 plasma deposition the reverse reaction still takes place. The reverse reaction is termed as etching and this reaction does not lead to the Si network propagation. The reaction occurring on the growing surface leads to the reaction of hydrogen atom eroding the Si-surface and involves other more complex reaction forming volatile species which leaves the film surface. The success of a-Si:H originates from its low defect density and the key to achieving a low defect density material is the presence of hydrogen during the film growth. On the other hand, adding larger amounts of hydrogen to dilute reactant gasses in the plasma will produce microcrystalline silicon or even polycrystalline material. The formation of microcrystalline material will be dominant due to the chemical equilibrium in the above reaction. Hydrogen is present in abundance during the film growth because hydrogen is the by product of the silane dissociation reaction and this hydrogen also plays an important role in the plasma deposition of a-Si:H. The effectiveness of hydrogen in reducing defects depends strongly on the way it is incorporated , rather than the hydrogen content itself. For hydrogenated amorphous silicon (a-Si:H) films deposited at substrate temperature above $\approx 200^0C$, most of the hydrogen present is in the monohydride form, a-Si:H film deposited at lower temperature contains much more hydrogen but the hydrogen is bonded in the most defective modes of SiH_2, SiH_3 or $(SiH_2)_n$ [44]. The presence of hydrogen tends to saturate the dangling bonds on internal surface of microvoids and on point defects in the structure. It has by no means been established that bond saturation by hydrogen is the only reason for the much lower overall density of states in glow discharge a-Si:H compared with levels found in evaporated or sputtered materials [45]. It is very likely that different deposition methods leads to different defect structures. Comparatively large microvoids , containing numerous dangling bonds, appear to dominate the properties of evaporated a-Si:H [46], whereas in the glow discharge material, point defects possibly of the divacancy type predominate. The electron energy levels of these dangling bonds (unpaired electrons) lie in between the valence and conduction band states of fully paired bonded electrons.

When the amorphous material is heated or annealed at sufficiently high temperature (350°C - 600°C) , but below the crystallisation temperature, hydrogenated silicon loses most of its hydrogen. As expected, paramagnetic dangling bonds are left behind, and a large increase of the magnetic resonance signal is observed [47]. It is possible to produce a complete chemical reaction between the broken bonds of the pure material with external hydrogen, thus transforming the first material into low density of states hydrogenated silicon. Of course, this cannot be reached with the molecular gas, since the breaking of the H-H bond would require too much energy. The reaction is therefore performed in a hydrogen plasma containing a large propotion of atomic hydrogen. The size of the atom is quite small, at a tempereture of 500°C, it can easily diffuse into the bulk of the pure amorphous silicon and makes a stable covalent bond with the localised defects of the material. These types of annealing effects are due to the healing of the dangling bonds by the reconstruction and the rearrangement of the amorphous network. However, it was established that under certain conditions the effects of annealing in reducing dangling bonds are not much. Even large compensated microvoids may be created depending on the deposition conditions. Hydrogen content generally increases as the substrate temperature increases, the SiH_4 preassure decreases and as the rf power in the electrodeless discharge increases [48]. The key to achieve low defects density material is to fabricate a-Si:H with ideal hydrogen content suited for microcrystalline devices. The hydrogen incorporated during depositions may depend perhaps critically, on the detailed plasma conditions at the specimen surfaces, as has been established for some of the other properties. Schottky Barrier and pn junction diodes have been fabricated with amorphous silicon. When operated solar cells and these diodes behave almost like corresponding crystal devices. Hydrogen and other additivies might, beside passivating the gap states, also play active role of enlarging the band gap, changing the electron lattice-coupling etc. How much hydrogenation is needed to reveal the most interesting phenomena is still an open question.