CHAPTER 2 THEORY

This chapter remarks the introduction to titanium aluminium nitride, crystal structure and physical properties of titanium aluminum nitride, the related theories, documents, and glow discharge process, dc sputtering deposition system, dc magnetron sputtering deposition system, unbalanced magnetron sputtering deposition system, thin film deposition by reactive sputtering method and characterization of thin films.

2.1 TiAlN Crystal Structures

Titanium aluminum nitride was developed in the late 1980s which is the ternary nitride structure occured from three different kind of element consist of transition metals as titanium and aluminum atoms forming with nitrogen atom under optimum condition The TiAlN thin film coating wide spread deposited in industrial scale by physical vapor deposition (PVD) methods. This evident that the solid-state solution of one or more solutes in a solvent. Such a mixture is considered a solution rather than a compound when the crystal structure of the solvent remains unchanged by addition of the solutes atom then mixed together is called "solid solution" [37]. The process begin by add different amount of Al atom into face center cubic (FCC) TiN structure resulting the TiAlN with B1-NaCl structure. The structure of (TiAl)N films depends a lot on the quantity of aluminum that is added to the compound. As greater amounts of aluminum are added, the color of the coating changes, turning dark gray, while the coating takes on the golden color of the TiN when less aluminum is added [38]. Generally, aluminium contents higher 50 at.%, is called "Aluminium titanium nitride (AlTiN)"[39].

The Limited information is available on the ternary phase diagram of Ti-Al-N (Fig. 2.2). Two ternary phases exist at 1000°C with an approximate composition of Ti_3AlN and Ti_2AlN. Ti_3AlN has a cubic structure with a lattice parameter of a = 4.112 A°. It coexists with TiN_{1-x}, Ti_3Al and Ti_2AlN compounds. The lattice parameters of hexagonal Ti_2AlN phase are a = 2.991 A° and c = 13.621 A°. It was found to coexist with Ti_3AlN, TiN_{1-x}, AlN, TiAl and Ti_3AlN compounds at 1000°C. The a-Ti phase extends up to over 30% N [18].

Application The range of applications for TiAlN coatings are attractive for many application fields due to behaviour of TiAlN films is superior mechanical and tribology properties as follow hardness friction coefficient wear resistance to that of other nitride coatings as shown in table 2.1. For these reasons, Ti– Al–N is an excellent candidate for use as optical coatings in industry, diffusion barriers in microelectronics, electrodes in micro-electro-mechanical-devices, biomedical, sattlelight and hard wear-resistant coatings in machining applications [10, 11, 13, 40, 41].

Table 2.1 shows the properties of TiAlN thin films compare with the metal nitride hard coating, structures of TiAlN compare with TiN were presented in Figure 2.1. Ternary phase diagram of Ti-Al-N was showed in Figure 2.2

Table 2.1 Properties of hard coating thin films [42]

Coating	Hardness (HV)	Friction coefficient	Wear resistance	Maximun service temperature (°C)	Color
CrN	1500	Excellent	Good	700	Silver/Gray
TiN	2200	Good	Excellent	525	Gold
ZrN	2500	Good	Excellent	600	Light-Gold
TiCN	3000	Good	Excellent	400	Violet
TiAlN	3600	Good	Excellent	750	Violet Black

(http://www.northeastcoating.com/Coatings_1.html)

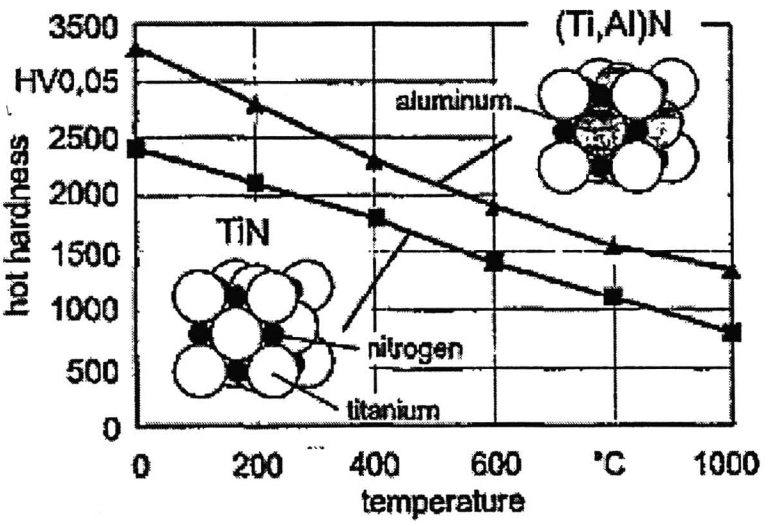

Figure 2.1 Hardness and crystal structure of titanium aluminium nitride and titanium nitride [43]

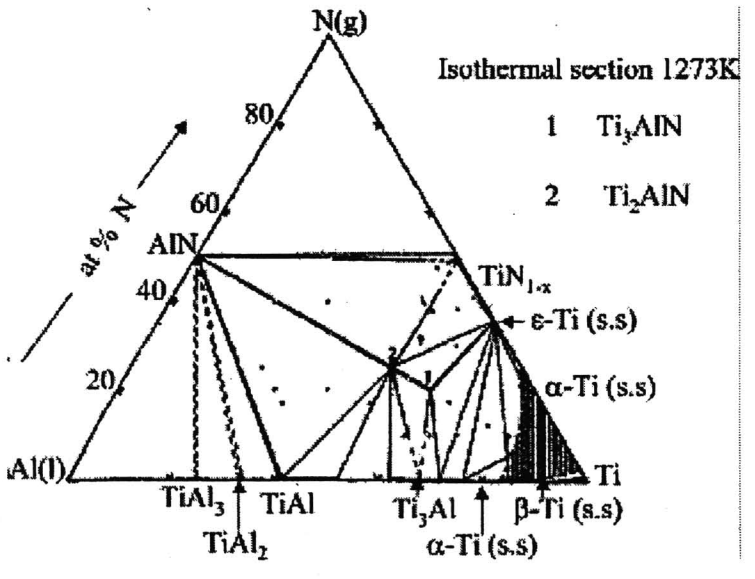

Figure 2.2 Ternary phase diagram of Ti-Al-N at 1273 K [44]

2.2 Sputtering Mechanisms [45]

Let us consider what happens when an ion approaches the surface of a solid (of the same or different material); the solid is usually called the *target.* One or all of the following phenomena may occur (Figure 2.2):

1. The ion may be reflected, probably being neutralized in the process. This reflection is the basis of an analytical technique known as *Ion Scattering* Spectroscopy, which enables us to characterize the surface layers of the material, and also tell us a lot of about the fundamental ion-surface interaction.
2. The impact of the ion may cause the target to eject an electron, usually referred to as a *secondary electron*
3. The ion may become buried in the target. This is the phenomenon of ion implantation, which is already used extensively in integrate circuit technology for selectively doping silicon wafer with precisely controlled amounts and depth profiles of specific impurities, and is likely to find many other applications such as surfaces treatment of steels.
4. The ion impact may also be responsible for some structural rearrangements in the target material. "Rearrangements" may vary from simple vacancies (missing atoms) and interstitials (atoms out of position) to more gross lattice defected such as changes of stochiometry (i.e. relative proportions) in alloy or compound targets, or to charge levels and distribution, and are usually collectively referred to as *radiation damage*, which is a subject of great importance, particularly with relation to nuclear energy. Radiation damage can often be removed by annealing (heat treatment) but it is not always unwanted, and perhaps the alternative name or *altered surface layers*, used mostly to describe the stoichiometry changes, is more apt.
5. The ion impact may set up a series of collisions between atoms of target, possibly leading to the ejection one of these atoms. This ejection process is known as *sputtering.*

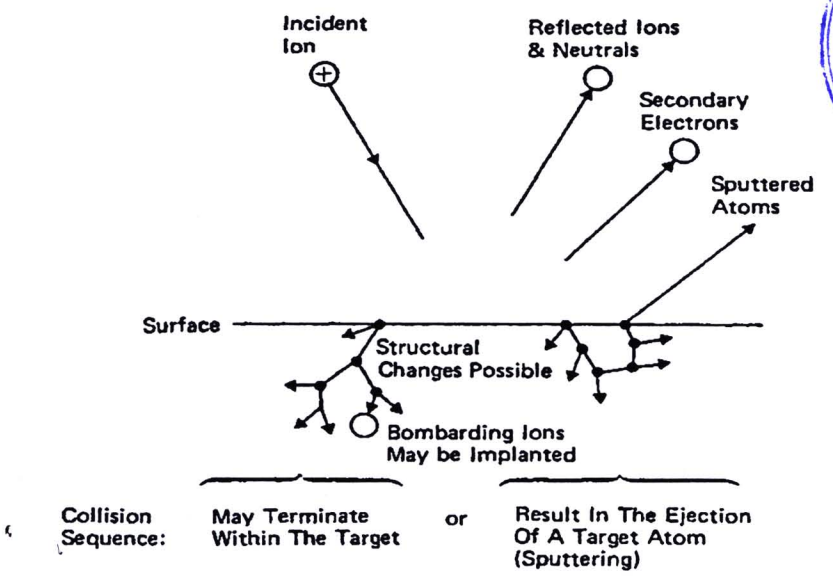

Figure 2.3 The interactions of ions with surfaces [47]

The sputtering process is considered in detail for the following three different energy region of the incident ions [46].

1. Threshold region (100eV)
2. Low energy region(0.1keV)
3. High energy region (10-60keV)

According to the elastic-collision theory, the maximum possible energy transferred in the first collision T_m is given by

$$T_m = \frac{4M_1M_2}{(M_1+M_2)^2}E \qquad \text{.....2.1}$$

where M_1 and M_2 are the masses of the incident ions and target atoms respectively, and E is the energy of incident ions. In the first order of approximation the sputter yield S is proportional to the T_m, the sputter yield of a given target material bombarded with different elements is given by

$$S = k\frac{1}{\lambda(E)\cos\theta}\frac{M_1M_2}{(M_1+M_2)^2}E \qquad \text{.....2.2}$$

where k is constant which includes different target material constants, λ is the mean free path for elastic collisions near the target surface, and θ is the angle between the normal on the target surface and the direction of incidence ions. The mean free path is given by

$$\lambda = \frac{1}{\pi R^2 n_0} \qquad \text{.....2.3}$$

where n_0 is the number of lattice atoms per unit volume, and R is the collision radius.

The collision radius R for the rigid sphere model can be calculate for a screened potential as

$$R = C\frac{a_0}{\left(z_1^{2/3}+z_2^{2/3}\right)^{1/2}}\ln\frac{z_1 z_2 e^2}{\varepsilon_0 RE'} \qquad \text{.....2.4}$$

where $E' = M_1E/(M_1+M_2)$, C is constant, a_0 is the radius of the hydrogen atom($=0.57\times10^{-8}$cm), e the elementary charge, ε_0 the dielectric constant in the vacuum, Z_1e and Z_2e are the nuclear charges for M_1 and M_2 respectively. The relation (2.2) gives qualitative information about the sputter yield.

2.3 Conventional DC Sputtering System [45, 55]

The material we wish to sputter is made into a sputtering target which becomes the cathode of an electrical circuit, and has a high negative voltage V(dc) applied to it as shown in Figure 2.3. The target is nearly always solid, although powders and even liquids are sometime used. The substrate which we wish to coat is placed on an electrically grounded anode a few inches away. These electrodes are housed in a chamber which is vacuum. Argon gas is introduced into the chamber to some specified pressure. The action of the electric field is to accelerate electrons which in turn collide with argon atoms, breaking some of them up into argon ions and more electrons to produce the glow discharge that we will discuss in the next section. The charged particles thus produced are accelerated by the field, the electrons tending towards the anode (causing more ionization on the way) and the ions towards the cathode, so that a current I flows.

When the ions strike the cathode, they may sputter some of the target atoms off. They may also liberate secondary electron from the target and it is these secondary electrons which are responsible for maintaining the electron supply and sustaining the glow discharge. The sputtered atoms from the target fly off in random directions, and some of them land on the substrate (on anode), condense there, and form a thin film.

The voltage V which is required to drive the current I through the system is a function of the system pressure p. The rate of thin film formation on the substrate will depend on the amount of sputtering at the target; this in turn will depend on the ion flux at the target and so (linearly) on the current. However, the amount of sputtering also depends on the sputtering yield S and so on the ion energy and hence on V, which determines the sheath voltage at the target. The choice of sputtering pressure p and the implied choice

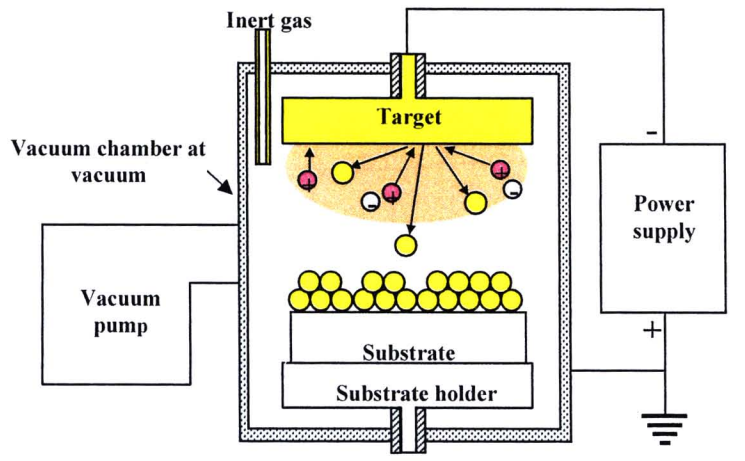

Figure 2.4 Vacuum coating by sputtering method [55]

of the V-I characteristic, are thus rather important and so we will now look at the criteria used in selecting operating parameters and in deciding on the nature of the sputtering gas.

2.4 DC Glow Discharges for Sputtering [47, 55]

Direct current (DC) plasma used in low-pressure deposition processes are based mainly on the glow discharge. A glow discharge is generated by applying a voltage of a few hundred to a few thousand volts, e.g., to planar electrodes (planar diode) immersed in a gas atmosphere as shown in Figure 2.4. The pressure in the chamber is typically in the range of about 30-130 Pa and the electrodes are separated for a few cm or more. At these conditions, when the voltage is applied, a breakdown takes place and the discharge is quickly self-sustaining. The self-sustained form of the discharge includes, at least, a cathode dark space and negative glow. If the electrodes are separated by more than a few centimeters, a thin, so-called Faraday dark space is seen behind the negative glow followed by a positive column that fills the rest of the space toward the anode. At the anode surface, another dark space is seen.

A distinguishing feature of a glow discharge is that almost the entire applied voltage is dropped across the cathode dark space ("cathode fall") creating conditions for a self- sustaining discharge. Ions from the negative glow enter the cathode dark space, are accelerated there toward the cathode, and bombard it causing emission of secondary electrons. The secondary electrons are accelerated across the cathode fall toward the negative glow and ionize gas molecule thus replacing ions lost at the cathode and in recombination in the volume and at the walls.

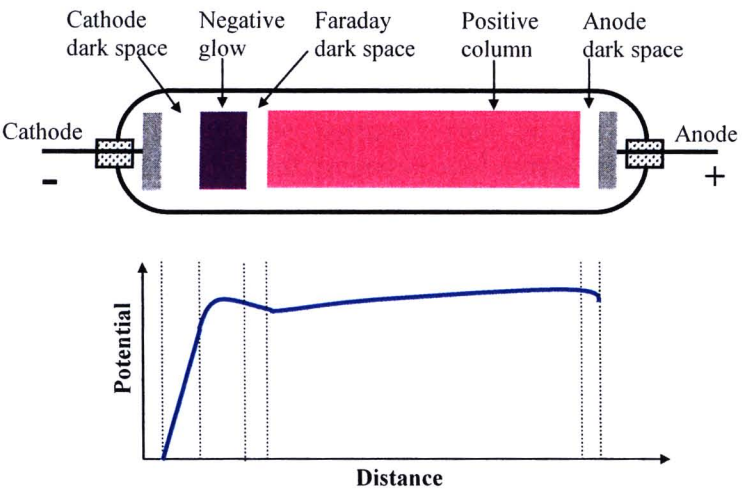

Figure 2.5 Schematic of the main parts of a glowing discharge (above) and the potential distribution in a normal glow discharge (below) [55]

This equilibrium is the basis for self-sustaining conditions. In an effort to keep the ionization rate by secondary electrons at the necessary level, the density of gas molecules, i.e. means the pressure should be not less than about 30 Pa.

Electron and ion densities in the positive column as well as in the negative glow are about equal creating a low-temperature non-equilibrium plasma (a term first used by Langmuir). The electron plasma density or just plasma density in negative glow region is usually of the order of magnitude of 10^{10} cm^{-3} or less, and being lower in the positive column.

Ion bombardment of a cathode in a glow discharge results in sputtering off the cathode material. The flux of sputtered material is mainly in the form of atoms and molecules and can be used for deposition of thin films. Unfortunately, at a pressure of about 30 Pa and above the scattering of sputtered atoms on their way to the anode (substrate) can be avoided only if the cathode-anode distance is much less than 5 cm which is a very impractical constraint. In addition, in a regular glow discharge the ion current density at the cathode surface in a conventional glow discharge is of at the order of 0.01-1.00mA/cm^2. In order to achieve the industrially desirable deposition rates of metals (about 5 nm/s) the current densities should be about be above 1.00 mA/cm^2. A current density of 1.00 mA/cm^2 is obtained running the glow discharge in the abnormal mode using a voltage of a few kV, but this creates an additional drawback, for example, by argon ions is most effective at energies of a few hundred volts. At energies of a few kV, the sputtering yield is significantly lower and the whole deposition process is not energy efficient.

To make the glow discharge plasma more dense, operate the discharge at lower pressures, and increase the ion current density at the cathode, confinement of the glow discharge by a magnetic field (magnetic confinement) or a special geometry (geometrical confinement) is employed.

2.5 DC Magnetron Sputtering Deposition System [47, 55]

Magnetic confinement of a glow discharge is achieved by creating crossed magnetic and electric fields near the cathode surface. This is realized in various magnetron configurations: plannar, cylindrical, rotatable, and others. As an example, let us consider a planar circular magnetron with an axisymmetric field as shown in Figure 2.6 The magnets are placed behind the target in such way that all magnetic lines that exit the magnets from the target side cross it twice returning back to the magnets and creating a magnetic field parallel to the target of about 200-500 Gauss. The component of the magnetic field parallel to the target surface forces the emitted electron, after acceleration in the cathode fall, to move parallel to the target surface and to produce more secondary electrons. As a result of such a magnetic trap effect, the negative glow plasma is intensified and shifts closer to the cathode surface (target). The negative glow in a circular and rectangular planar magnetron has a shape of a bagel and a race track, respectively as shown in Figure 2.7.

The plasma density in the negative glow of a magnetron discharge can be by tow orders of magnitude higher than in the glow discharge and reach values of the order of 10^{12} cm^{-3}. The high plasma density near the cathode allows a large number of ions to be extracted and accelerated into the direction of the cathode. The ion current density at the cathode can be as high as 100 mA/cm^2, and, therefore, provide a high sputtering rate. A further increase in the sputtering rate can also come from the substantially lower cathode voltages which in case of a magnetron discharge seldom exceed 600 V and, therefore, are closer to the ion energy where there is a maximum in the sputtering yield dependence on ion energy. Magnetrons are usually operated in the pressure range of 0.27-1.33Pa (2-10 mtorr).

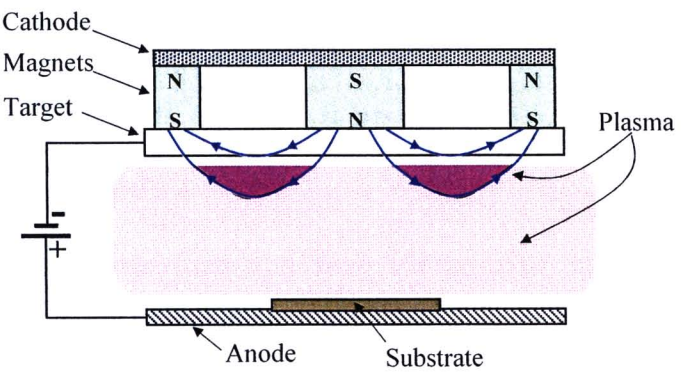

Figure 2.6 Schematic of a magnetron system [55]

At these pressures to avoid substantial scattering of sputtered off atoms, the cathode-to-substrate distance can be in the convenient range of about 10 cm without appreciable scattering of sputtered atoms.

Closing the cathode-anode electrical circuitry is achieved by those electrons that missed the magnetic field trap or escape from it. (The last ones escape in all directions except for the direction toward the cathode). As a result, weak plasma is created around the magnetron reaching densities of the order of 10^8 cm^{-3} or less at the substrate.

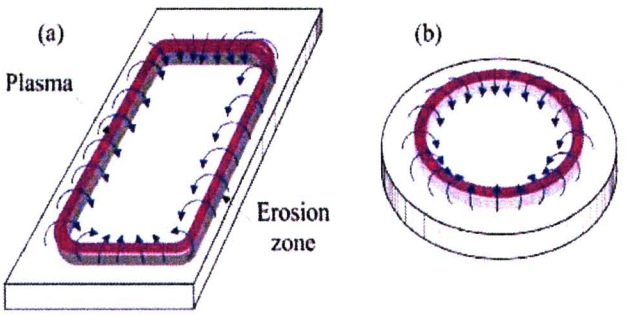

Figure 2.7 Schematic of planar magnetron: (a) rectangular and (b) circular

2.6 Thin Film Deposition by Reactive Sputtering Method [48, 55]

Reactive sputter deposition from an elemental target relies on: (a) the reaction of the depositing species with a gaseous species, such as oxygen or nitrogen, (b) reaction with an adsorbed species, or (c) reaction with a co-depositing species such as carbon to form a compound. The reactive gas may be in the molecular state (e.g.N_2, O_2) or may be "activated" to form a more chemically reactive or more easily adsorbed species. Typically, the reactive gases have a low atomic masses (N=14, O=16) and are thus not effective in sputtering. It is therefore desirable to have a heavier inert gas, such as argon, to aid in sputtering. Mixing argon with the reactive gas also aids in activating the reactive gas by the Penning ionization/excitation processes.

Typically, a problem in reactive sputter deposition is to prevent the "poisoning" of the sputtering target by the formation of a compound layer on its surface. Poisoning of a target surface greatly reduces the sputtering rate and sputtering efficiency. This problem is controlled by having a high sputtering rate (magnetron sputtering) and controlling the availability of the reactive gas, such that there will be enough reactive species to react with the film surface to deposit the desired compound, but not so much that it will unduly poison the target surface.

The appropriate gas composition and flow for reactive sputter deposition can be established by monitoring the partial pressure of the reactive gas as a function of reactive gas flow, or by impedance of the plasma discharge. Figure 2.7 shows the dynamics of this process as a function of reactive gas flow is depicted. At low flow rates all of the reactive gas is incorporated into the deposited film (metallic mode). As the gas flow rate is increased, a threshold is reached where the target surface experiences compound formation. When this compound formation exceeds the removal rate of material, a threshold is reached that is accompanied by a sharp decrease in sputtering rate and discharge voltage (compound mode). This decrease is due to the face that compounds generally have lower sputtering rates and higher secondary electron emissions. Additionally, a third cause for the drop in sputtering rate is due to less efficient sputtering by reactive gas ions than by inert ions. When the reactive gas flow is sufficiently reduced the system will revert back to metallic mode. However, metallic mode sputtering will not commence at the same flow rate, because the compounds remaining on the target need to be removed before normal sputtering can resume [20].

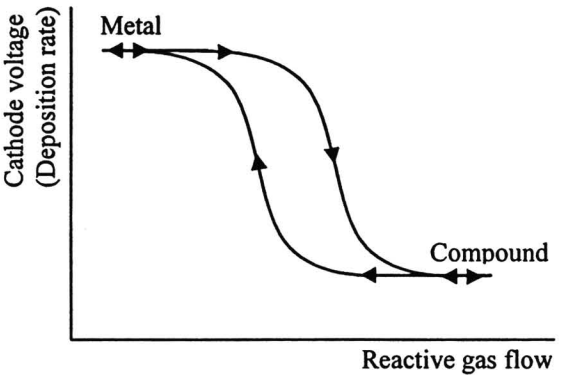

Figure 2.8 Generic hysteresis curve for voltage vs. reactive gas flow rate [55]

2.7 The Structure of Thin Film [33, 55]

Upon coalescence of the surface nuclei to form a continuous film, the nucleation step of film deposition is complete, and the fourth step begins: development of the bulk film structure. The form of this structure changes dramatically with the amount of thermal motion taking place during film growth, which scales with the ratio of the substrate T to the melting point of the film, T_s/T_m (in K, not °C), known as the "homologous" or "reduced" T. Structural form also changes with the amount of additional energy being delivered to the growth surface. Three structure zones (Z1, Z2 and Z3) were initially in an evaporative deposition study which included both metals and ceramics, and these zones have since been observed in a wide variety of film materials deposited by all of the vapor-phase processes. A fourth "transitional" zone (ZT) between Z1 and Z2 was identified in sputter deposition and has since been found prominent in other energy-enhanced processes. Occasionally, anomalous structural forms occur, in particular the whiskers, illustrated in Figure 2.9, that have been seen in Ti, for example. The occurrence of whiskers implies an extreme preference for growth along the vertical direction. The characteristic structures of the four basic zones are also illustrated in Figure 2.9. They are described briefly below and are analyzed in more detail in the next two subsections.

1. Z1 zone occurs at T_s/T_m so low that surface diffusion is negligible; that is, the diffusion length more than the hop distance. Z1 consists of columns typically tens of nm in diameter separated by voids a few nm across. The columns have poor crystallinity (many defects) or are amorphous. In thicker films, there become superimposed upon this structure an array of cones with wider voids between them. The cones terminate in domes at the surface, and the size of the domes increases with film thickness.
2. ZT zone also occurs when the diffusion length less than the hop distance. It contains defected columns similar to those of Z1, but the voids and domes are absent. ZT is usually associated with energy-enhanced processes.

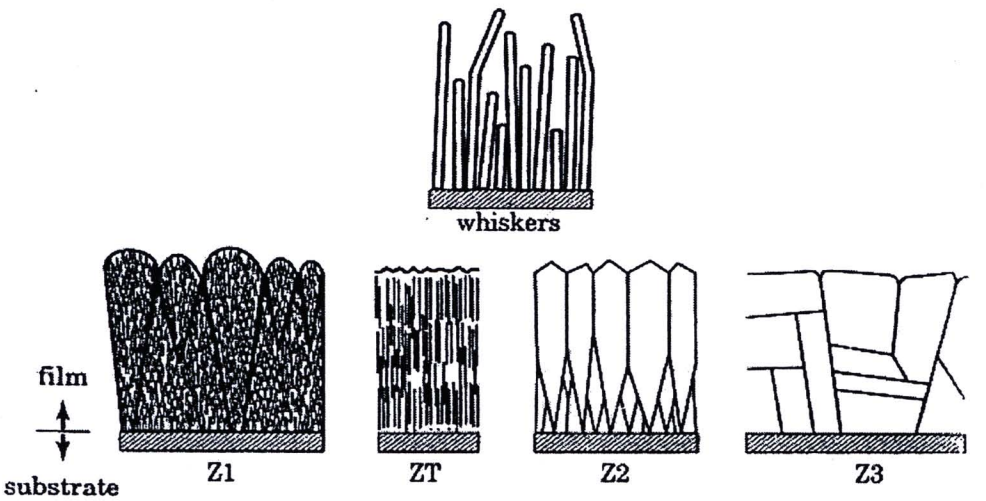

Figure 2.9 Characteristics of the four basic structure zones and of whiskers, in cross section. The ratio of substrate T_s to film melting T_m (T_s/T_m) increases in the direction Z1→ZT→Z2→Z3 [49]

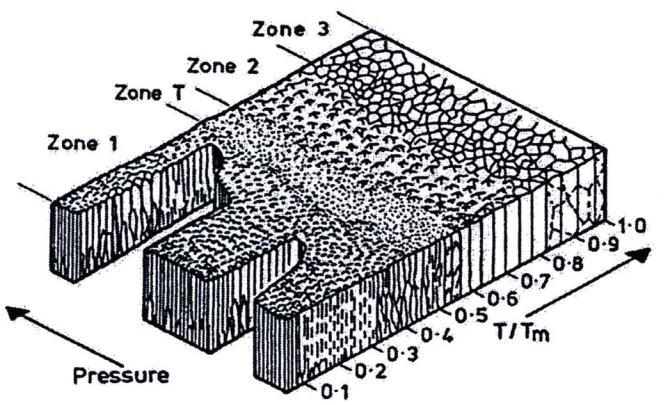

Figure 2.10 Conclusion of Thornton's structure-zone diagram [50]

3. Z2 zone occurs at $T_s/T_m > 0.3$ or so, high enough so that surface diffusion is becoming significant. It consists of columns having tight grain boundaries between them and having a characteristic diameter which increases with T_s/T_m. Crystalline columns are less defected than in Z1 and ZT and are often facetted at the surface. The Z2 structure can also occur in amorphous films; there, the column boundaries are planes of reduced bonding rather than planes of crystallographic discontinuity.
4. A transition to Z3 occurs in certain instances at $T_s/T_m > 0.5$ or so high enough so that considerable bulk annealing of the film is taking place during deposition. Z3 is characterized by more isotropic or equiaxed crystallite shapes. Film surfaces are often smoother at $T_s/T_m > 0.5$ for either Z2 or Z3; however, the grain boundaries can develop grooves. Figure 2.10 shows conclusion all four zones which has been called Thornton's structure-zone diagram.

The following several points about these zones need to be kept in mind.

1. All four zones cannot always be identified for a given material. In particular, Z3 is often not observed.
2. The transition from one zone to another is not always abrupt with T, and the transition T varies with deposition conditions and material. The growth mode can also change from Z3 or Z2 or from ZT to Z1 moving up through the thickness of film.
3. The surface topography shown in Figure 2.9 is typical, but it can vary considerably with factors such as surface-energy anisotropy and incident angle of depositing vapor.
4. Epitaxial films exhibit none of this bulk structure-at least when they are free of mosaic texture. However, they can still have highly facetted and therefore very rough surfaces when grown on planes of high surface energy, due to restructuring of the surface to minimize energy.
5. Amorphous films exhibit bulk structure only when they are inhomogeneous, because they have no crystallographic pattern by which to define the boundaries between grains.

Homologous T is the main determinant of structure. Z1 and ZT films result from "quenched" processes in which thermal migration of the adsorbed material is negligible, whereas Z2 and Z3 films result from thermally activate rearrangement on or within the films, respectively. We will discuss these two regimes of structural development separately below, and we will see how the growth dynamics affect both the bulk structure and the surface topography.

2.8 Characterization of Thin Film

The characters or properties of thin film can be analyzed or investigated by different techniques. This work will study thin film's characteristics, for examples, the crystal structure, phase and surface morphology by several techniques as follows.

2.8.1 The Study of Crystal Structure by X-Ray Diffractrometer (XRD) [51]

X-ray diffraction (XRD) is a very important experimental technique used to investigate the crystal structure of bulk solids, including the determination of lattice constants, identification of unknown materials, orientation of single crystals and preferred orientation of polycrystalline. However, extending the application of this method to thin films has been avoided for the reason that the penetration depth of x-rays with typical incident angles is large, so that the substrate, rather than the film, dominates the scattered x-rays signal. Nevertheless, the x-ray diffraction has the advantage over the TEM method that it doesn't require film removal from the substrate.

X-rays are electromagnetic waves. The phenomenon of the X-ray diffraction by crystals results from a scattering process in which X-rays are scattered by the electrons of the atoms without change in wavelength. A diffracted beam is produced by such scattering only when certain geometrical conditions are satisfied, which may be expressed in either of two forms, the Bragg law or the Laue equations. The resulting diffraction pattern of a crystal, comprising both the positions and intensities of the diffraction effects, is a fundamental physical property of the substance. Analysis of the positions of

the diffraction effect leads immediately to a knowledge of the size, shape and orientation of the unit cell. The equipment used in this work is a Rigaku Rint 2000 X-ray diffractometer with Cu K_α radiation with wavelength 1.5406 Å operated at 40 kV and 40 mA, as shown in Figure 2.11. The XRD spectra were collected in the 20 - 65° of 2θ range with a measurement step of 0.02°. Graphic diagrams for XRD measurement can also notify preferred orientation and phase of crystalline structures, compared to data from Joint Committee on Powder Diffraction Standard (JCPDS).

Figure 2.11 An illustration of the Rigaku Rint 2000 X-ray diffractrometer

2.8.2 Crystallite Size [51, 55]

A typical x-ray diffraction pattern together with the line spectrum of $AlTi_3N$ thin film is shown in Figure 2.12. In order to characterize the microstructure of the films, three parameters should be obtained from the diffractograms: the position of the peaks 2θ, the height of the peaks *h* and the width at the half height of the peaks β. The crystallite size (grain size) of the $AlTi_3N$ films was determined from the full width at half maximum (FWHM) of the X-ray diffraction peak using Scherrer's formula:

$$D = \frac{0.94\lambda}{\beta \cos\theta} \qquad \text{.....2.5}$$

where λ is the X-ray wavelength, β is the FWHM of a diffraction peak and θ is the diffraction angle.

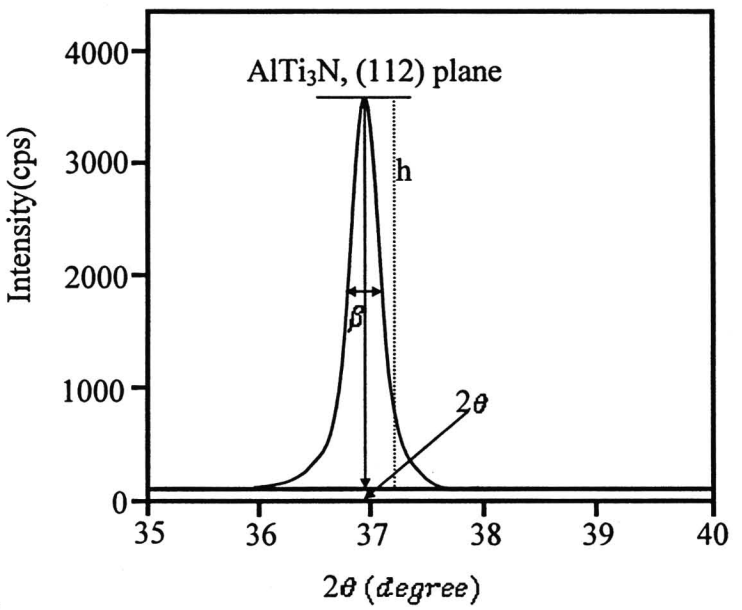

Figure 2.12 Reflection of (112) plane by XRD for typical $AlTi_3N$ thin films

2.8.3 Lattice Constant [51, 55]

The distance between the two adjacent lattice planes is derived from the experimental peak position by means of Bragg formula:

$$d_{exp} = \lambda_0 / 2\sin\theta \qquad \text{.....2.6}$$

whereas its theoretical value d_{hkl} can be calculated from the lattice constant (a = 1.0118 nm) by :

$$d_{hkl} = \frac{a}{\sqrt{h^2 + k^2 + l^2}} \qquad \text{.....2.7}$$

where h, k and l are Miller indices of the corresponding lattice planes (Figure 2.13).

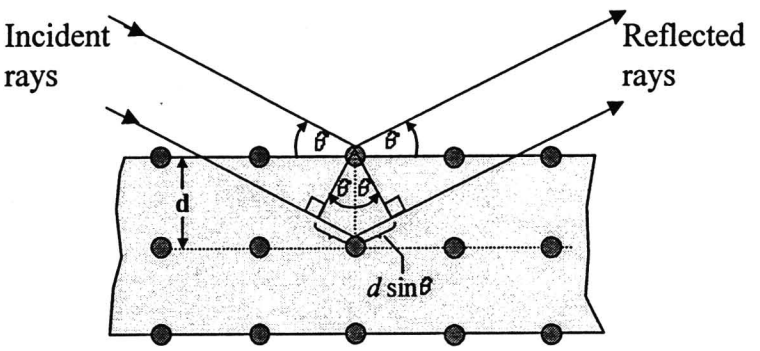

Figure 2.13 Reflection of x-rays from a crystal [55]

Figure 2.14 shows peak intensities which present the crystal plane of material at the different reflection ray angles. The position of diffraction pattern may shift by higher angle since residual stress occurs in deposited material. Moreover, x-ray diffraction technique is also give phase composition data and lattice parameter.

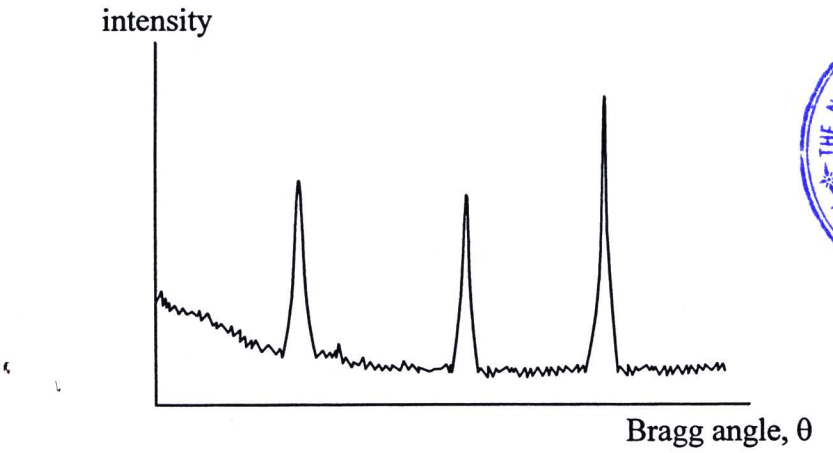

Figure 2.14 Peak intensities which present crystal planes of material at different reflection ray angles [55]

2.8.4 Atomic Force Microscopy (AFM) [52, 55]

The important character of thin film is surface morphology and thickness. The favorite technique that uses for study thickness and surface morphology of nanoscale thin film is atomic force microscope (AFM). The AFM consists of a cantilever with a sharp tip (probe) at its end that is used to scan the specimen surface. The cantilever is typically silicon or silicon nitride with a tip radius of curvature in the order of nanometers. When the tip is brought into proximity of a simple surface, forces between the tip and the sample lead to a deflection of the cantilever according to Hooke's law. Depending on the situation, forces that are measured in AFM include mechanical contact forces, van der Waals forces, etc. Typically, the deflection is measured using a laser spot reflected from the top surface of the cantilever into an array of photodiodes as shown in Figure 2.17. Other methods that are used include optical interferometry, capacitive sensing or piezoresistive AFM cantilevers. These cantilevers are fabricated with piezoresistive elements that act as a strain gauge. Using a Wheatstone bridge, strain in the AFM cantilever due to deflection can be measured, but this method is not as sensitive as laser deflection or interferometry.

If the tip was scanned at a constant height, a risk would exist that the tip collides with the surface, causing damage. Hence, in most cases a feedback mechanism is employed to adjust the tip-to-sample distance to maintain a constant force between the tip and the sample. Traditionally, the sample is mounted on a piezoelectric tube that can move the sample in the z direction for maintaining a constant force, and the x and y directions for scanning the sample. Alternatively a 'tripod' configuration of three piezo crystals may be employed, with each responsible for scanning in the x,y and z directions. This eliminates some of the distortion effects seen with a tube scanner. In newer designs, the

tip is mounted on a vertical piezo scanner while the sample is being scanned in x and y using another piezo block. The resulting map of the area s=f(x,y) represents the topography of the sample.

The AFM can be operated in a number of modes, depending on the application. In general, possible imaging modes are divided into static (also called contact) modes and a variety of dynamic (or non-contact) modes where the cantilever is vibrated as shown in Figure 2.15.

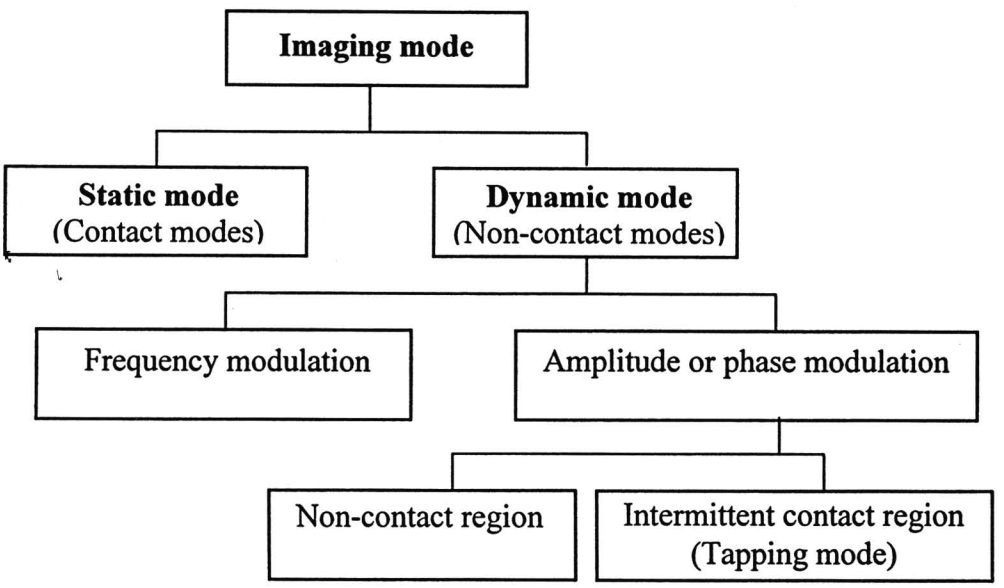

Figure 2.15 Diagram of imaging mode of AFM [55]

Static Mode

In the static mode operation, the static tip deflection is used as a feedback signal. Because the measurement of a static signal is prone to noise and drift, low stiffness cantilevers are used to boost the deflection signal. However, close to the surface of the sample, attractive forces can be quite strong, causing the tip to 'snap-in' to the surface. Thus static mode AFM is almost always done in contact where the overall force is repulsive as shown in Figure 2.16. Consequently, this technique is typically called 'contact mode'. In contact mode, the force between the tip and the surface is kept constant during scanning by maintaining a constant deflection.

Dynamic Mode

In the dynamic mode operation, the tip of the cantilever does not contact the sample surface. The cantilever is instead oscillated at a frequency slightly above its resonance frequency where the amplitude of oscillation is typically a few nanometers (<10nm). The van der Waals forces, which are strongest from 1 nm to 10 nm above the surface and are attractive region as shown in Figure 2.16, or any other long range force which

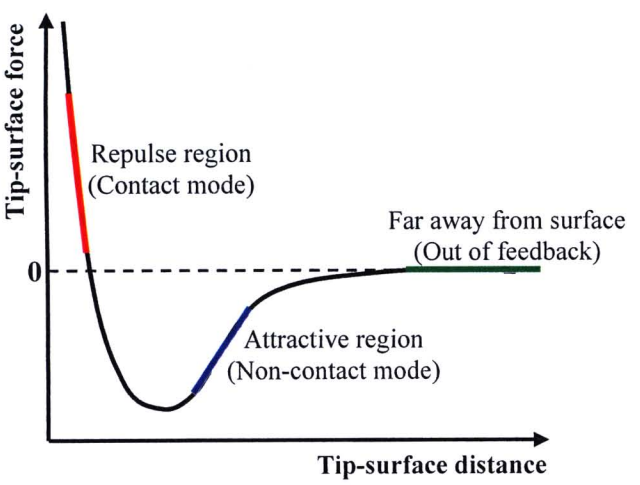

Figure 2.15 The interactive force between atoms at different material distances [53, 55]

extends above the surface acts to decrease the resonance frequency or change phase and oscillation amplitude of the cantilever. These changes are combined with the feedback loop system maintains a constant oscillation amplitude or frequency by adjusting the average tip-to-sample distance. Measuring the tip-to-sample distance at each (x,y) data point allows the scanning software to construct a topographic image of the sample surface.

Non-contact mode AFM does not suffer from tip or sample degradation effects that are sometimes observed after taking numerous scans with contact AFM. This makes non-contact AFM preferable to contact AFM for measuring soft samples. In the case of rigid samples, contact and non-contact images may look the same. However, if a few monolayers of adsorbed fluid are lying on the surface of a rigid sample, the images may look quite different. An AFM operating in contact mode will penetrate the liquid layer to image the underlying surface, whereas in non-contact mode an AFM will oscillate above the adsorbed fluid layer to image both the liquid and surface. In the dynamic mode operation can separate to frequency and amplitude modulations.

1. Frequency Modulation, changes in the oscillation frequency provide information about tip-sample interactions. Frequency can be measured with very high sensitivity and thus the frequency modulation mode allows for the use of very stiff cantilevers. Stiff cantilevers provide stability very close to the surface and, as a result, this technique was the first AFM technique to provide true atomic resolution in ultra-high vacuum conditions.
2. Amplitude modulation, changes in the oscillation amplitude or phase provide the feedback signal for imaging. In amplitude modulation, changes in the phase of oscillation can be used to discriminate between different types of materials on the surface. Amplitude modulation can be operated either in the non-contact or in the intermittent contact regime. In ambient conditions, most samples develop a liquid meniscus layer. Because of this, keeping the probe tip close enough to the sample for short-range forces to become detectable and preventing the tip from sticking to the surface, presents a major hurdle for the non-contact dynamic mode in ambient conditions. Dynamic contact mode (also called intermittent contact or tapping mode) was developed to bypass this problem. In dynamic contact mode, the cantilever is oscillated such that the separation distance between the cantilever tip and the sample surface is modulated.

In non-contact (tapping mode), the cantilever is driven to oscillate up and down at near its resonance frequency by a small piezoelectric element mounted in the AFM tip holder as shown in Figure 2.17. The amplitude of this oscillation is greater than 10 nm, typically 100 to 200 nm. Due to the interaction of forces acting on the cantilever when the tip comes to the surface, Van der Waals force or dipole-dipole interaction, electrostatic forces, etc cause the amplitude of this oscillation to decrease as the tip gets closer to the sample. An electronic servo uses the piezoelectric actuator to control the height of the cantilever above the sample. The servo adjusts the height to maintain a set cantilever oscillation amplitude as the cantilever is scanned over the sample. A tapping AFM image is therefore produced by imaging the force of the oscillating contacts of the tip with the sample surface. This is an improvement on conventional contact AFM, in which the cantilever just drags across the surface at constant force and can result in surface damage. Tapping mode is gentle enough even for the visualization of supported lipid bilayers or adsorbed single polymer molecules (for instance, 0.4 nm thick chains of synthetic polyelectrolytes) under liquid medium. At the application of proper scanning parameters, the conformation of single molecules remains unchanged for hours.

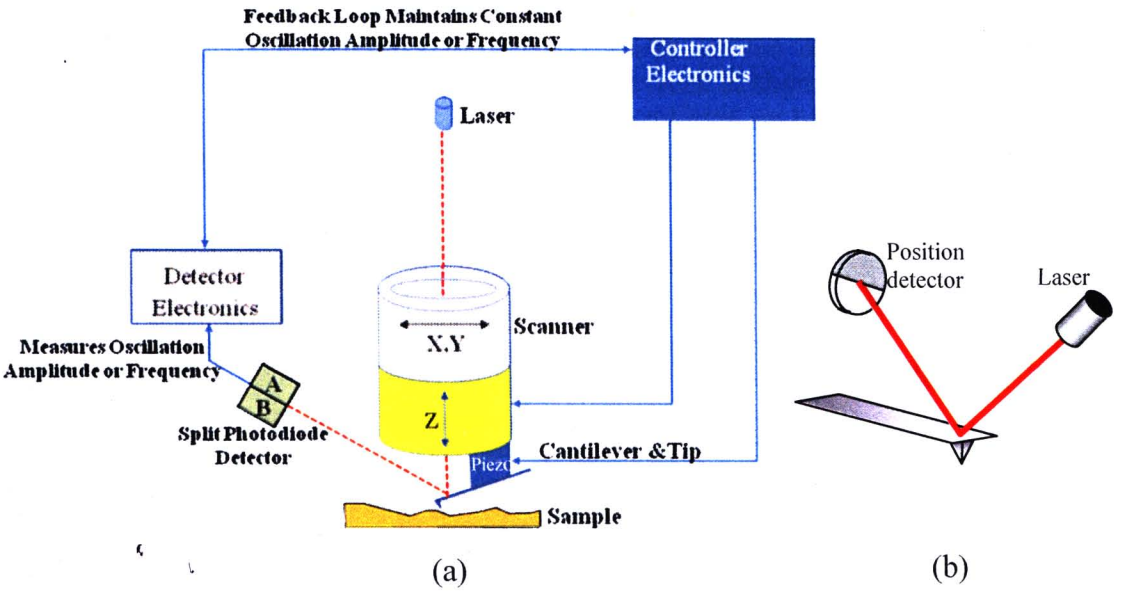

Figure 2.17 AFM: (a) Non-contact [52, 55] (b) Beam deflection system, using a laser and photodector to measure the beam position [53, 55]

The surface morphology of titanium dioxide thin film which obtained from AFM technique composes of 2D, 3D picture and roughness value. By the way, the roughness values are compose of average roughness ($\bar{z}$) and root-mean-square roughness (R_{rms}). The $\bar{z}$ value is calculated from equation 2.8.

$$\bar{Z} = \frac{1}{n}\sum_{i=1}^{n} | Z_i | \qquad \text{.....2.8}$$

The RMS roughness value is determined from equation 2.9

$$R_{rms} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (Z_i - \bar{Z})^2} \qquad \text{.....2.9}$$

z_i are maximum and minimum values of each position on surface, n is a number of maximum and minimum values which measured on surface.

2.8.5 Electron Microscopy [54, 55]

We have seen that microscopy limits us to the resolution of objects separated by a distance of not less than about 200nm. We could work at shorter optical wavelengths in the ultraviolet, although this will only push the resolution up by a factor of 3 or more. In addition, it is often difficult or impossible to find suitable materials to act as lenses in the ultraviolet. A radical approach is to abandon the use of light and to adopt instead electrons as the source of waves in the microscope. For an electron of momentum p, the particle-wave duality tells us that this behaves as a wave of wavelength λ given by $\lambda = h/p$, where h is the Planck constant. If the electrons have received this momentum by having been accelerated through some potential difference V, then their kinetic energy is $p^2/2m=eV$, so that $\lambda = h/(2meV)^{1/2}$. For an accelerating voltage of 100 keV, typical wavelength of electron microscopes is $\lambda \approx 0.004$ nm.

It can be seen that using electrons as the waves in a microscope will increase the resolution of a microscope. However, this is not the whole picture. To construct an electron microscope, we need some form of lens, a device which will alter the direction of motion of an electron as it passes through it. Electric and magnetic fields will accomplish this. Provided, that the electron beam dose not diverge much. Lenses based on electric and magnetic fields will not only alter the direction of the electron, but will focus them in much the same way as conventional lenses focus light. However, we must note the caveat on the electron beam divergence, and it is a very important point. Severe aberrations occur in electron lenses when the electrons travel at even moderate angles to the axis of the lens. These aberrations will give rise to blurring and distortion of the image. Of course, such effects occur in optical lenses, but the shape of an optical lens can be carefully designed to minimize them. This is not possible with electrostatic and magnetic lenses. For this reason, electron beams need to be tightly confined to the central region of electron lens. This makes the numerical aperture of electron lenses very small, typically around 0.01. the electron analogue of the optical microscope can therefore resolve points separated by a distance d = 0.61/(numerical aperture) = 0.61× 0.004/0.01=0.244nm .In addition, chromatic aberration is a problem with electron lenses, with the focal length depending on the wavelength of the electrons. In order to overcome this, highly monochromatic electron beams must be employed. There are several types of electron microscope in existence, but will discuss two types are transmission electron microscope (TEM) and scanning electron microscope (SEM).

2.8.5.1 The Transmission Electron Microscope (TEM) [55]

The transmission electron microscope (TEM) is electron analogue of the optical compound microscope. The sample is 'illuminated' uniformly with electrons which pass through it to form an image. The ultimate resolution of this instrument is determined by the wavelength of the electron and the numerical aperture of the lens, as described above.

A schematic diagram of the transmission electron microscope is shown in Figure 2.18(a). A heated tungsten filament is used as the source of electrons. This is biased to around 100 keV, which accelerates the electrons through the condenser lens on to the specimen. The condenser aperture restricts the size of the beam to reduce unnecessary illumination of the specimen, since electron beams do damage materials to some extent. The sample itself is very thin (around 100nm) so that most of the electrons pass through

it without significant inelastic scattering. Complex thinning procedures are often necessary for semiconductors, using chemical thinning or ion beam milling. Vibrations in the atomic composition or thickness of the sample produce changes in the electron scattering from the main beam. In addition, if the sample is crystalline, strong Bragg diffraction of the electron beam will occur in specific directions, forming sharp spots in the back focal plane of the objective lens. As with the optical microscope, the unscattered beam may be accepted into the image, contrast being obtained when features in the solid scatter electrons out of the beam so that they do not continue along the microscope (bright field imaging). If the contrast thereby obtained is small, then only the scattered electrons are

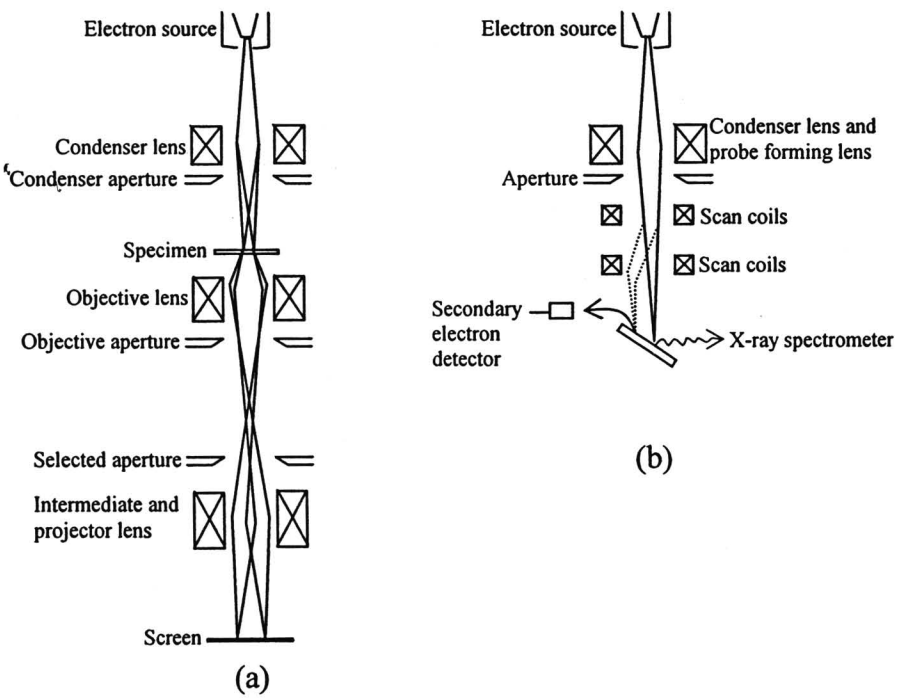

Figure 2.18 Schematic diagram of: (a) Transmission Electron Microscope (TEM) and (b) Scanning Electron Microscope (SEM) [54, 55]

collected (dark field imaging). The scattered electrons are collected by the objective lens forming a diffraction pattern of the object in the back focal plane of the lens and an intermediate image of the object in the image plane of the objective. The objective lens is the critical component for the microscope. As can be seen from the diagram, it is only lens for which the electrons have a large angle to the axis. It will therefore be a major contribution of aberration in the system.

The objective aperture provides contrast by selecting or intercepting scattered or unscattered electrons so that bright field or dark field imaging can be achieved. The intermediate and projector lenses magnify the image produced by the objective lens before projecting it on to a screen. These lenses do not contribute significantly to aberration since the angle the electron beam makes to their axes is small.

The TEM can therefore be used to produce an enlarged image of an object just as with an optical microscope, with magnifications of hundreds. In addition, It can be used to study the crystallographic structure of solids. This is accommodated by the intermediate lens. The focal length of a magnetic lens can be changed easily by changing its electric current. The focal length of the intermediate lens can be adjusted so that, instead of the image plane, the back focal plane of the objective lens is imaged at the object plane of the projector lens. The electron diffraction pattern of the crystalline solid occurs at the back focal plane of the objective lens, and it is this which is now amplified and projected on to the viewing screen. This image will contain information on the crystallographic structure of the solid. By employing a selected area aperture, the area of the specimen contributing to the diffraction pattern may be limited. The diffraction patterns obtained depend on the crystallographic orientation of the semiconductor. Defects present in the semiconductor also give characteristic diffraction pattern features which can be used to characterize them. Images of such defects in crystals can be produced by utilizing the local changes in the Bragg scattering they produce due to local distortions in the crystal planes. A diffraction beam can be selected (or blocked) by the objective aperture and monitored. The changes in Bragg scattering in the region of the defect will then produce contrast in the transmitted electron beam.

The diffracted beams may also be used to produce what are known as *lattice images*. To produce these, the undeflected beam and one of the deflected beams are allowed to reach the image plane. Since they have travelled different path lengths, they will interfere, producing fringes which represent the planes of the atoms in the solid. The angle of deflection of a diffracted beam will depend on the lattice plane spacing.

2.8.5.2 The Scanning Electron Microscope (SEM) [55]

The TEM forms an image of an object in essentially the same way as an optical microscope. The scanning electron microscope uses an entirely different technique. In the SEM, the scanning electron beam is focused to a small spot and scanned across the sample in a raster fashion. The microscope detects secondary electrons scattered back from the surface by the incident electron beam, and, using these synchronously with the beam scanning the surface of the solid, builds up a magnified object of the image. In addition, a small fraction of the incident electrons may also be backscattered. It does not look at transmitted electrons, and the sample can be relatively thick. The electron beam interacts with the solid via electrostatic forces, with the interaction taking place chiefly in a pear shaped region with an extent of about 1 μm below the surface, as shown in Figure 2.19. However, the secondary electrons produced within the solid and escaping from it can only come from the surface regions, the actual escape depth depending on the energy of the electrons. The SEM therefore is principally used to investigate the surface of solids. The backscattered secondary and primary electrons are collected by a detector, the output of which is used to modulate the intensity of a television screen. The process involved in SEM can result in the build-up of charge on the surface of relatively low-conductivity semiconductor, and this can distort the final image. To overcome this, the surface of the semiconductor is usually coated with a thin metallic layer (or the SEM is operated at lower voltages and hence resolution)

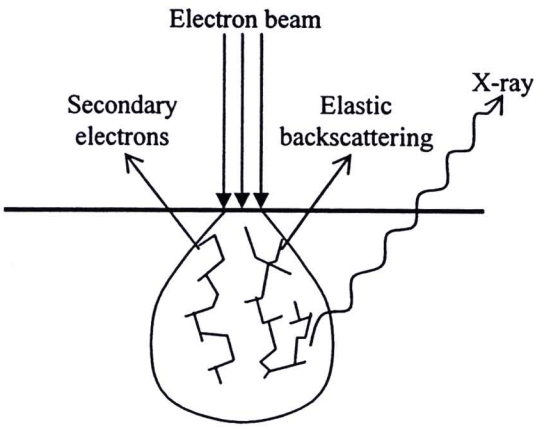

Figure 2.19 Interaction between incident electron beam and solid takes place in a pear-shaped volume [54, 55]

A schematic diagram of an SEM is shown in Figure 2.18(b). The condenser lenses and the probe forming lens are used to produce the electron spot. The can be scanned over the surface of a specimen by two set s of deflecting coils. Each set of coils can deflect the beam in two perpendicular directions, so that a square area on the sample may be scanned. The secondary electrons emitted from the surface are attracted to the detector by metal gauze in front of the detector which is held at some attracting electrostatic potential.

Contrast in the SEM arises chiefly from variations in the number of electrons reaching the detector from different regions of the spectrum. The simplest form of contrast comes from irregularities in the surface of the semiconductor. The number of secondary electrons reaching the detector depends sensitively on the angle between the incident electron beam and the surface at the point where the strikes. In addition, the brightness depends on the surface inclination in such a way that the image exhibits a three-dimensional effect. Variations in the secondary electron emission of the object will appear in the image if they extend over a distance of the order of or greater than the electron probe beam diameter. Ii is the sport size of the probe beam which determines the resolution of the SEM. Ultimately; the spot size is determined by the Airy pattern obtained from diffraction at the aperture of the final focusing lens. This depends on the wavelength of the electrons so that increasing the voltage accelerating the electron beam, hence decreasing the wavelength, would appear to be the factor which determines the resolution. However, it should be recognized that the secondary electron emission process is a weak process. Decreasing the spot size will reduce the signal due to secondary electron emission. Eventually, at sufficiently small spot size, the signal due the secondary electron emission will be smaller than the random signals (noise) in the electronics so that no image will be observed. It is this factor, rather than the wavelength, which determines the resolution of the SEM. Of course, we can increase secondary electron emission for a given spot size by increasing the probe beam current intensity, although eventually surface charging effects will dominate. The resolution of the SEM therefore depends on probe spot size and probe current intensity and not on wavelength as in the TEM.