3.1 Introduction

The MOS (Metal Oxide Semiconductor) devices are practically a base part of all electronic circuits and the functionality of the final circuit depends on its electrical characteristics. Even though these devices are fabricated in a very clean environment of class 1000 or class 10,000, these devices are not assured to be free from impurities and defects as the fabrication process itself are likely to introduce certain impurities and defects which are undesirable causing more deviations from the ideal values. Electrical characterization of these devices to account for the charges and defects introduced during fabrication and also to find the exact mechanism of defect generation and annihilation becomes an important task as far as reliability concern exists. This chapter covers a wide range of topics starting from the physics of MOS structures to types of defects associated with fabrication and the treatment given to annihilate the defects.

The chapter is divided into twelve sections. Sections 3.2 and 3.3 brief the basic physics of MOS capacitors. Sections 3.3 to 3.7 deals with the charges and defects associated with oxide and Silicon bulk. A brief introduction to various types of deep defects in the energy gap of the Silicon bulk is also given. Section 3.8 discusses the experimental procedure followed to fabricate the MOS structure. Section 3.9 and 3.10 deals with the electrical characterization methods followed to extract the parameters of interest. Section 3.11 introduces to a new type of spectroscopic method, Deep Level Transient Spectroscopy (DLTS). Application of its principles in order to extract the defect energy levels, capture cross section and recombination lifetime are also discussed. The deep trap levels induced during processing of MOS structures were examined using DLTS. Isochronal treatment leading to annihilation of process induced defects is also explained in section 3.12.
3.2 Structure of MOS Capacitor

The MOS capacitor is an important test device in the semiconductor industry as the gate of the MOS Field Effect Transistor (MOSFET) is actually a MOS capacitor. The MOS capacitor is composed of three components: the substrate, the gate dielectric and the gate electrode. The Figure 3.1 below shows the structure of a typical MOS capacitor.

![Figure 3.1: Structure of MOS capacitor](image)

**Si Substrate**

The substrate forms the foundation of IC devices and currently, silicon wafers are used as substrates. Initially, the silicon is formed in a cylindrical shape called a boule. During the formation of the boule, dopants are added to the molten silicon. These dopants determine the type of wafer that will result from the boule. As the amount of dopant is increased the resistivity of the resulting wafer is decreased. At the present level of technology, dopant levels are in the $10^{18}$ cm$^{-3}$ range. With technological changes, the diameter of the wafer has also increased more than ~ 300 mm in diameter. This allows the number of dies per wafer to increase, resulting in higher product yield with the same number of processing steps.

**Gate Dielectric**

Presently, silicon dioxide (SiO$_2$) is the gate dielectric of choice. It is the naturally formed layer on silicon at room temperature in an air-ambient. But this thin layer of SiO$_2$ with undesirable characteristics is not for IC use. For device fabrication, a high quality oxide is needed. This is accomplished by growing an amorphous silicon-oxide on a silicon
substrate at high temperatures (i.e., > 950 °C). This type of oxide has many beneficial properties as shown in Table 3.1.

Table 3.1 Properties of silicon dioxide [1]

<table>
<thead>
<tr>
<th>Property (units)</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Melting Point (°C)</td>
<td>~1700</td>
</tr>
<tr>
<td>Density (g/cm³)</td>
<td>2.27</td>
</tr>
<tr>
<td>Bulk Refractive Index</td>
<td>1.46</td>
</tr>
<tr>
<td>Dielectric Constant</td>
<td>3.8-3.9</td>
</tr>
<tr>
<td>Dielectric Strength (V/cm)</td>
<td>~5 x 10⁸</td>
</tr>
<tr>
<td>Energy Gap (eV)</td>
<td>~8</td>
</tr>
<tr>
<td>Etch Rate in Buffered HF (Å/min)</td>
<td>1000</td>
</tr>
<tr>
<td>Molecules/cm²</td>
<td>2.3 x 10²²</td>
</tr>
</tbody>
</table>

**Gate Electrode**

Aluminum is the widely used gate metal. But due to the problems of its low melting point and tendency to spike into the oxide it was replaced by degenerately doped polysilicon. Polysilicon was chosen as the replacement for aluminum due to its higher melting point and compatibility with SiO₂. With the development of high-K dielectrics and problems with poly depletion, comes the need for new metal gate electrodes.

### 3.3 Band Bending in MOS Structure

Band diagrams are depictions of the electron energies and potentials in the device and are very useful in understanding the operation of a MOS capacitor. Each component of the capacitor can be represented with a band diagram. In Figure 3.2, the band diagram for an n-type MOS capacitor under ideal (i.e., \( V_{FB} = 0 \)) conditions is shown and the following discussion considers this ideal MOS capacitor. The band diagram of the semiconductor consists of the conduction band energy level, \( E_c \), the valence band energy level, \( E_v \), and the intrinsic energy level, \( E_i \), which is halfway between the conduction band and the valence band. These three bands, \( E_c, E_v, \) and \( E_i \), are specific to the intrinsic substrate material and always remain parallel to one another. The band gap energy, \( E_g \), is the energy difference between the two allowed bands of electron energy i.e. conduction band
energy and the valence band energy of the semiconductor. The Fermi energy level, $E_F$, is related to the substrate doping. The Fermi potential, $q\phi_F$, is the difference between the intrinsic energy level and the Fermi level of the semiconductor. The semiconductor work function, $\Phi_S$, is the difference between the Fermi level of the semiconductor and the vacuum level. Similarly, the metal work function, $\Phi_M$, is the difference between the vacuum level and the Fermi level of the metal.

![Energy Band Diagram of n-type Ideal MOS Capacitor](image)

**Figure 3.2: Energy band diagram of n-type ideal MOS capacitor**

When a bias is applied to the gate electrode, the bands bend according to the magnitude and polarity of the bias. When a positive bias is applied, the bands bend upward as shown in Figure 3.3 (a) and the Fermi levels of the metal and semiconductor differ by an amount proportional to the applied bias. The bending shows that a negative charge, due to electrons, has collected at the Si-SiO$_2$ interface to compensate the applied bias. This state is called accumulation since it is caused by the accumulation of electrons in the substrate at its interface with the oxide.

The application of a negative bias will cause the bands to bend downward. The band diagram for this state is shown in Figure 3.3 (b). The capacitor is now considered to be in depletion since the oxide / substrate interface has been depleted of electrons. Due to the absence of electrons there is a net positive charge in the substrate at the interface. As a larger negative bias is applied the bands bend even further as shown in Figure 3.3 (c). The capacitor goes into inversion when $E_i$ is bent below the Fermi level of the

28
semiconductor. There is an inversion layer in the substrate at the interface that acts as a p-type material.

![Diagram](image)

Figure 3.3: Energy band diagram of n-type ideal MOS capacitor when $V \neq 0$ for the following cases (a) accumulation; (b) depletion; (c) inversion [2]
Now consider a typical n-type MOS capacitor, where the flatband voltage is non-zero. In this case, the bands bend without an applied bias. For this non-ideal capacitor, the bands will be similar to what is shown in Figure 3.4 when the applied bias is zero. The band bending is caused by a difference between the work functions of the metal and the semiconductor as well as any charge located in the oxide or at the oxide / substrate interface. When a bias equal to the flatband voltage is applied, the MOS capacitor is in the flatband state and the bands of the semiconductor straighten or become flat. The metal-semiconductor work function is the difference between the work function of the metal and that of the semiconductor (i.e., $q\phi_{MS} = \Phi_M - \Phi_S$). Similar band diagrams for this non-ideal n-type MOS capacitor can be developed with accumulation occurring when the bias is larger than the flatband voltage, depletion when the bias is close to the flatband voltage, and inversion as the bias becomes greater than the flatband voltage.

![Energy band diagram of n-type non-ideal MOS capacitor (with metal-semiconductor workfunction difference)](image)

**Figure 3.4:** Energy band diagram of n-type non-ideal MOS capacitor (with metal-semiconductor workfunction difference)

### 3.4 Charges in SiO₂ and Si-SiO₂ Interface

A high quality oxide for device fabrication is always accomplished by growing an amorphous silicon-oxide on a silicon substrate at high temperatures which results in many beneficial properties of the oxide. However, silicon oxide is not without problems. The charges at the oxide and the Si-SiO₂ interface become the bane to reliability. In 1980, Deal proposed a standard terminology for oxide charges associated with thermally oxidized silicon based on the recommendation by a committee established by the
Electronics Division of the Electrochemical Society and the IEEE Semiconductor Interface Specialists Conference. As per the suggestions, the charges were categorized into four groups: fixed ($Q_f$), mobile ($Q_m$), oxide trapped ($Q_{ot}$) and interface trapped ($Q_n$) charges. In each case, $Q$ is the net effective charge per unit area at the SiO$_2$-Si interface. The basic classification of these traps and charges are shown in Figure 3.5.

![Figure 3.5: Charges associated with thermally oxidized silicon](image)

**Fixed Oxide Charge ($Q_f$)**

A positive charge, due primarily to structural defects (ionized silicon, i.e., trivalent silicon, and non-bridging oxygen) in the oxide layer within about 2.5 nm from the Si-SiO$_2$ interface. The density of this charge, whose origin is related to the oxidation process, depends on oxidation ambient and temperature, cooling conditions, and on silicon orientation and is also called slow states. $Q_f$ depends on the final oxidation temperature. The higher the oxidation temperature, the lower is $Q_f$. The fixed oxide charge is not in electrical communication with the underlying silicon [3].

**Mobile Ionic Charge ($Q_m$)**

Mobile charge comes from contaminants being incorporated into the oxide during processing. Sodium is the most common ionic contaminant in MOS processing. It leaches from glass containers into chemicals used in processing and form direct or indirect contact with body salts. Other possible contaminants are Lithium and Potassium ions [4]. These ions are mobile within the oxide under bias temperature aging conditions.
Negative ions and heavy metals may also contribute to this charge even though they are not mobile below 500 °C.

**Oxide Trapped Charge (Q_{ot})**

This positive or negative charge may be due to holes or electrons trapped in the oxide. Trapping may result from ionizing radiation, avalanche injection, Fowler-Nordheim tunneling, or other similar processes. The oxide trapped charges are generally positive and have generally been found to induce negative charges in the silicon. Additionally, it communicates with the underlying silicon through trapping-detrapping processes. Unlike fixed oxide charge, this type of charge is generally annealed out by low temperature (<500 °C) treatment, although neutral traps may remain.

**Interface Trapped Charge (Q_{in})**

A positive or negative charge, due to the interruption of periodic lattice structure such as (a) structural, oxidation induced defects, (b) metal impurities, and (c) other defects caused by ionizing radiation or similar bond breaking processes. In fact they are the results of dangling bonds. They are located within the silicon forbidden gap at the Si-SiO₂ interface. Unlike a fixed oxide charge, an interface trapped charge interacts strongly with the underlying silicon and can thus be charged or discharged, depending on the surface potential. These types of charges are also called surface states, fast states or interface states. An interface trap is considered a donor if it can become neutral or positive by donating an electron. An acceptor interface trap can become neutral or negative by accepting an electron. Rather than being a single discrete energy level, interface traps are continuum of energy levels and thus are quantified as interface trap density (D_{in}) expressed in number of states/(cm².eV).

### 3.5 Defects in Si Bulk

An ideal crystal lattice has its atom at its designated position and any deviations from this perfect structure are called imperfections or defects. These defects are likely to introduce electronic energy states into the semiconductor band gap, which can be placed into two categories: **shallow levels** and **deep levels**.

*Shallow levels* are located near their related band edges (valence band for acceptors and
conduction band for donors) i.e. \(-0.1\) eV from the band edge, thus these levels are thermally ionized at room temperature. Impurity elements which are used as dopants in semiconductors normally create these shallow levels which are ionized at room temperature and provide free carriers to form p-type or n-type semiconductor.

*Deep levels* are those defects positioned deeper in the band gap than the dopant levels and are found to bind the carriers much more strongly into highly compact, localized states. The deep levels have higher ionization energies; therefore contribute very little to the free charge carriers. Defects with deep levels in the band gap are often referred to as, ‘traps’, ‘recombination centers’, or ‘generation centers’. In case of a p-n junction, deep levels act as recombination centers when there are excess charge carriers in the semiconductors. Deep levels act as generation centers when the carrier density is below its equilibrium value as in the reverse biased space charge region (scr). The generation and recombination (G-R) mechanisms through a deep level are described below.

Consider a deep level impurity of energy \(E_T\) shown in Figure 3.6. Let \(N_T\) be the number of impurities/cm\(^3\) uniformly distributed throughout the semiconductor. The semiconductor has \(n\) electrons/cm\(^3\) and \(p\) holes/cm\(^3\) introduced by shallow level dopants.

![Figure 3.6: Electron energy band diagram for a semiconductor with deep level impurities](image)

In order to follow the various capture and emission processes, let the G-R center first capture an electron from the conduction band, shown in Figure 3.6 (a) and characterized by the capture co-efficient \(c_n\). After electron capture one of the two events takes place. The center can either emit the electron back to the conduction band from where it came
shown as $e_n$ in Figure 3.6 (b), or it can capture a hole from the valance band, shown as $e_p$ in Figure 3.6 (c). After either of these two events, the G-R center is occupied by a hole and again has two choices. Either it emits the hole back to the valance band characterized by $e_p$ as shown in Figure 3.6 (d), or capture an electron as shown in Figure 3.6 (a). These are the only four possible events between the conduction band, the impurity level and the valance band. A recombination event is represented by the Figure 3.6 (a) followed by 3.6 (c) and a generation event is represented by the Figure 3.6 (b) followed by 3.6 (d). The impurity level and both conduction and valance bands participate in recombination and generation. A third event, which is neither recombination nor generation, is the trapping event represented by the Figure 3.6 (a) followed by 3.6 (b) or the Figure 3.6 (c) followed by 3.6 (d). In either case a carrier is captured and subsequently emitted back to the band from which it came. Only one of the two bands and the impurity level participate in trapping event. Whether an impurity level acts as a trap or a G-R center depends on the location of the Fermi level in the band gap, the temperature and the capture cross sections of the impurity.

The study of deep levels is important in semiconductors since they modify the properties of the semiconductors and therefore, those of the devices fabricated using them. Even though Deep levels are desirable in some applications, they are also a nuisance if present in semiconductors that are used for photovoltaic and many other applications. Thus deep level study is of paramount importance in the semiconductor device industry.

The deep level defects in semiconductors can further be divided into two main categories; point defects and extended defects. Point defects are not extended in space in any dimension and this implies that the perturbation of the lattice is localized about a lattice site and involves only a few nearest neighbours. There are two kinds of point defects of great interest in semiconductor crystals, intrinsic (e.g. vacancies or self-interstitial) and extrinsic point defects (e.g. impurity atoms occupying substitutional or interstitial lattice sites. Small agglomerations of several point defects like divacancies, vacancy-impurity complexes, vacancy-donor etc are also generally considered as point defects. Extended defects are extended in nature (such as, grain boundaries, dislocations or stacking faults). The discussion in this section is focused more on point defects which are more relevant to the work covered in this thesis. The
Point defects are further classified as Primary and Secondary defects based on its nature and existence.

**Primary Defects**

The isolated lattice vacancy (V) and self-interstitial (I) are the primary defects produced after high-energy particle irradiation in semiconductors. The primary defects are mobile at room temperatures and therefore, deep level transient spectroscopy measurements of room temperature irradiated silicon will not reveal an isolated vacancy or silicon interstitials. Figure 3.7 gives a pictorial description of primary defects. The creation and repair process of primary defects are given below

Defect creation: \( \text{Particle} + \text{Si}_i \text{(Substitutional Si)} \rightarrow \text{Si}_i \text{(Interstitial Si)} + \text{V} \text{(Vacancy)} \)

Defect repair : \( \text{Si}_i + \text{V} \rightarrow \text{Si}_i \)

![Diagram representing the Vacancy, Self-Interstitial and Substitutional atom](image)

**Figure 3.7: Diagram representing the Vacancy, Self-Interstitial and Substitutional atom**

**Vacancy Defect:** If an atom is removed from its regular lattice site the empty lattice site is called a vacancy defect (V), and is shown in Figure 3.7. The vacancy in semiconductors can have up to five charge states, \( V^{++}, V^+, V^0, V^- \) and \( V^{--} \). The vacancy formation energy in Germanium (1.7 eV – 2.5 eV) is significantly smaller than in Silicon (~4.0 eV), for all charge states. When two neighbouring atoms are removed and also when two migrating vacancies meet and combine a divacancy is formed. A divacancy can also exist in four different charge states in Si [5].
**Interstitial Defect:** Interstitials are atoms, which occupy a site in the crystal structure, which is not a regular lattice site as shown in Figure 3.6. An interstitial defect can be of the same species as the atoms of the lattice (self-interstitial) or of a different nature (an interstitial impurity). Interstitials are generally high-energy configurations. Once again, the introduction of an interstitial induces a relaxation and distortion of the lattice, which surrounds it. The type of configuration the interstitial assumes, depends on its ability to make bonds with its neighbours and therefore can change with its charge state. A nearby interstitial defect and vacancy defect is called a Frenkel pair.

**Secondary Defect**

The vacancy and interstitials which survive the recombination of simple defects can diffuse into the semiconductor and interact with other intrinsic and extrinsic defects giving rise to complex room temperature stable defects. For example, in silicon when a vacancy becomes mobile, it can be trapped by an oxygen atom to form a V-O complex (A-center), or the doping impurity (e.g. P) to form a V-P complex (E'-center), or by another vacancy to form divacancies. Other complex defects are via the mobile interstitial (I). A schematic of A-center, E-center and Divacancy are shown in Figure 3.8.

![Figure 3.8: Representation of A-center, E-center and Divacancy](image-url)
**The Divacancy:** The divacancy is formed by the removal of two neighboring atoms. Generally divacancies can be created in semiconductors by particle irradiation either as a primary defect, when collision cascade is dense enough, or as a secondary defect by pairing of single vacancies diffusing randomly. The divacancies in silicon can appear in four charge states viz. $V_2^+$, $V_2^0$, $V_2^-$, $V_2^\text{m}$. 

![Equation](image)

**The E-center:** The $E$-center can be described as a vacancy trapped next to a substitutional donor atom. The $E$-center can be formed when the impurity atom captures a mobile vacancy. As to the formation of an $E$-center, although local atomic strain effect may need to be considered, the key role is played by the columbic interaction between a positively charged atom and a negatively charged vacancy [6]. Let’s consider the $E$-center in phosphorus-doped silicon, V-P. In the neutral charge state i.e. PV$^0$, two of the three silicon atoms surrounding the vacancy pull together to form an electron pair, leaving an unpaired electron in the orbital of the third silicon atom, while two electrons with antiparallel spins are accommodated by the phosphorus atom. When the Fermi level is above the $E$-center an extra electron is accepted and becomes negatively charged, PV$^-$. The formation of $E$-center in silicon removes two electrons in the conduction band by converting a positively charged P donor atom to a negatively charged V-P center.

![Equation](image)

**The A-center:** The $A$-center (V-O) may be regarded as a vacancy trapped next to an oxygen atom in an interstitial position. Similar to the $E$-center the $A$-center can be formed as a primary defect or when an oxygen impurity traps a mobile vacancy. The $A$-center competes for the vacancies with the $E$-center and its concentration is dependent on the relative O impurity concentration in the sample. The A-center has also been found to be an efficient recombination center [7], and therefore can be used to control minority carrier lifetime in silicon for fast switching device application.

![Equation](image)

Some of the other defect complexes are the Carbon interstitial (C$_i$), Boron interstitial...
(B\textsubscript{i}), interstitial boron – substitutional boron (B\textsubscript{i}-B\textsubscript{s}), interstitial boron – interstitial oxygen (B\textsubscript{i}-O\textsubscript{i}), interstitial boron – substitutional carbon (B\textsubscript{i}-C\textsubscript{s}), interstitial carbon – interstitial oxygen (C\textsubscript{i}-O\textsubscript{i}) and interstitial carbon – substitutional carbon (C\textsubscript{i}-C\textsubscript{s}). In case of a divacancy, when mobile or after dissociating, it can form trivacancies, quadravacancies, pentavacancies, and higher order defects.

Most of the primary and secondary defects discussed in this section are electrically active and introduce deep levels in the semiconductor band gap. A deep level may act as a minority carrier trap, majority carrier trap or recombination centre depending on its position in the band gap and on relative capture cross-section of minority and majority carriers. A majority carrier trap is an electron trap in n-type semiconductor or a hole trap in p-type semiconductor. Conversely a minority carrier is a hole trap in n-type semiconductor or an electron in p-type semiconductor. If a majority- or minority-carrier lives a mean lifetime in the captured state and is thermally ejected to the band from which it came, the center may be regarded as majority carrier trap or minority carrier trap respectively. From defect spectroscopy measurements such as DLTS it is possible to extract the defect properties such as the concentration, energy level and capture cross-section of defect level. The capture cross-sections can now be used to deduce whether the defect will act as minority carrier trap, majority carrier trap or a recombination centre.

Recombination centers are deep levels with approximately equal capture cross sections for both electrons and holes and these centers are normally located near the middle of the band gap. After capturing a majority carrier, if the majority carrier stays trapped at the center long enough for the trap to capture a minority carrier, then recombination takes place and the center is acting as a recombination center.

### 3.6 Defect Production by Ion Irradiation

High energy particles are decelerated when passing through matter and in the process transfer energy to the material. The transferred energy can modify the structure and properties of the materials. In case of crystalline semiconductors, particle induced materials modification will occur as long as the projectile particles can transfer energy larger than the displacement energy to the lattice atoms. The minimum energy necessary to displace an atom from its lattice position is called the threshold energy (T\textsubscript{d}).
Normally the threshold energy is assumed to be isotropic (i.e. independent of the direction in which the atom is displaced in the lattice.

**Energy Loss Mechanisms**

The capacity of a solid to slow down or stop a projectile is called the stopping power, and is defined as the amount of energy lost per unit length of trajectory in the solid. The stopping power depends on the type and energy of the projectile and on the mass of the target material. When a particle enters a target material there are two main mechanisms that cause an energy loss: (a) elastic collisions with the nuclei of the target material (nuclear stopping) and (b) inelastic collisions with bound or free electrons (electronic stopping). In electronic stopping, the term inelastic is used to signify that the collisions may result both in the excitations of bound electrons of the medium and in the excitations of the electron cloud of the ion. The relative effect of the two mechanisms depends on the mass of the target material as well as the energy, mass and charge of the incident particle. Hence the stopping power is given by

\[ S = \left( \frac{dE}{dx} \right) = \left( \frac{dE}{dx} \right)_{\text{nuclear}} + \left( \frac{dE}{dx} \right)_{\text{electronic}}. \]

These stopping contributions cause the energetic ion to stop at some predictable distance inside the medium. The range \( R \) of the projectile is related to its mean track length and for a projectile with initial energy \( E \) the range can be written as

\[ R = \int_U^E \frac{dE'}{S(E')} \]

where \( U \) the visibility threshold for the track and \( R \) is the range of the ion. If \( S(E') \) is the mean stopping power value, we obtain the mean range for the ions. The process of ion stopping is a statistical process and so some ions will undergo many collisions, stopping in a distance shorter than the average value. Hence, we do not obtain only one value for the range but a distribution. The range profile is often described by a mean projected range \( R' \). The range profile is an important parameter, because it describes the concentration distribution of implanted doping atoms. The concentration level affects the electrical properties of semiconductor materials.
Computer based simulation of implantation profiles are now possible using Monte Carlo-based techniques such as Transport of Ions in Matter (TRIM) [8] and Stopping Range of Ions in Matter (SRIM) [9]. These simulation codes should be treated with caution when used for crystalline material since they assume an amorphous material and ignore the effects of channeling, diffusion effects during and after ion implantation and also do not account for annihilation of vacancies and interstitials and thus overestimates the concentration of vacancies and interstitials produced by the implantation.

It should be pointed out that threshold energies for most materials is generally greater than the formation energy of Frenkel pairs, because defect formation is a complex multi-body collision process (e.g. a recoil atom can bounce back to its lattice position or kick back another atom to its lattice position). If an incident particle has energy much greater than the threshold energy ($T_d$), it will transfer energy to the target atom. The displaced atom may in turn collide and displace other atoms, creating cluster damage, until it eventually comes to rest, usually in an interstitial site and sometimes in a vacancy site. Light energetic particles (such as Si, He, Ar, neutrons, electrons and protons) tend to leave tracks of relative small defect concentrations. These ions initially slow down, mainly by electron stopping process with little displacement damage, until eventually nuclear stopping becomes dominant at the end of their range. Therefore, there is generally little lattice damage along the track except for the end of the range. Heavy ions by contrast may create damage clusters along their track. A comparison of lattice damage by light and heavy ions is depicted in Figure 3.9.

![Figure 3.9: Schematic of ion track in a solid, and associated damage for a light ion (top) and a heavy ion (bottom), redrawn from ref. 10.](image-url)
The heavy ions may undergo relatively higher degree of nuclear stopping than light ions even right from the surface. The volume of the crystal in which the ion energy is deposited is usually larger than the volume in which the lattice damage occurs. When the damaged areas start to overlap with increasing ion dose, an amorphous layer can result, which implies that all the nuclei have been displaced from their lattice position and the long range order which describes a crystal is no longer present.

3.7 Defect Annealing Mechanisms

The characteristic temperature at which a defect disappears is the defect’s annealing temperature. The defects annealing mechanism can be classified into two main categories:

(a) **Diffusion**, as the temperature is increased defects migrate to sinks (e.g. by moving to the surfaces or grain boundaries) or they are subsequently trapped by other defects or impurities (e.g. direct recombination of the interstitial with a vacancy, complex formation or hydrogen passivation) to form new defects. The ability of a defect to migrate through the crystal is determined by the thermal energy of the crystal and its charge state.

(b) **Dissociation**, which is the breaking-up of the complex defects.

It is interesting to note that most simple defects anneal out at between 200 °C and 400 °C in silicon, while higher order defects are introduced at higher temperatures between 350 °C and 500 °C.

3.8 Fabrication of MOS Structure

As the radiation hardness of a MOS device is strongly dependent on the type of the substrate and process history of the device, it is necessary to be very careful in the selection of the wafer and to decide about a proper process flow for device fabrication. For our purpose, Structures were fabricated on <100> oriented, 0.08 - 0.2 Ω-cm resistivity, n-type phosphorous doped, 1.5 inch diameter and 220 microns thick silicon wafers. This <100> orientation is always preferred over <111> orientation because of the fact that <100> orientation introduces about 10 times less interface trap density, which can cause charging at the Si-SiO₂ interface. This charging causes variations in the threshold voltage of the devices and is therefore undesirable.
Chemical Cleaning of Wafers

In order to remove the surface metallic impurities, organic contaminants etc the wafers were subjected to standard cleaning procedures like RCA-I, RCA-II and piranha. All chemicals used to process silicon were MOS grade with very high purity. The resistivity of the de-ionized water used to rinse wafers was 15-16 MΩ-cm.

To start with wafer cleaning, it is necessary to remove the grease which occurs during wafer cutting at the foundry or may be due to handling and the process is called degreasing which includes cleaning the wafers in Tri Chloro Ethylene (TCE), Acetone and Methanol. Once the wafer is thoroughly degreased, the wafers can now be processed with RCA and piranha. The detailed procedure used for cleaning the wafers before device processing are listed below.

RCA-I: This process removes organic residues and films on the silicon wafers.

- 5 parts of DI water was added to 1 part of Ammonium Hydroxide (NH₄OH) and heated to ~80 °C.
- 1 part of Hydrogen Peroxide (H₂O₂) was mixed to the solution.
- Wafers were kept immersed in the above mentioned solution for about 15 mins.
- Wafers were removed from the solution and thoroughly rinsed several times in DI water.

RCA-2: This process is used to remove metal ions from silicon wafers.

- A solution of DI water and Hydrogen Chloride (HCL) is prepared in the ratio 5:1 and heated to ~80 °C.
- 1 part of Hydrogen Peroxide (H₂O₂) is mixed to the heated solution.
- Wafers were kept immersed in the above mentioned solution for about 10 mins.
- Wafers were thoroughly rinsed several times in DI water.

Piranha cleaning:

- A solution of Hydrogen Peroxide (H₂O₂) and Sulfuric Acid (H₂SO₄) in the ratio 1:7 was prepared.
- Wafers were immersed in the solution for ~15 mins.
- Wafers were thoroughly rinsed in DI water.
• Wafers were given HF treatment and again rinsed in DI water.
• Degreasing process was again performed.
The wafers were then dried using nitrogen gun and taken for the next process i.e. Oxidation.

**Thermal Oxidation**

The ability of a silicon surface to form a silicon dioxide passivation layer is one of the key factors in silicon technology. Although silicon is a semiconducting material, silicon dioxide is a dielectric material. This combination, a dielectric layer formed on a semiconductor, along with other properties of silicon dioxide, makes it one of the most commonly used layers in silicon devices. Silicon dioxide layers find use in devices to pacify the silicon surface, to act as doping barriers and surface dielectrics, and to serve as dielectric parts of device structures.

*Atomic Structure of Silicon Dioxide*

- It is tetrahedron cell
- Strong adhesion to silicon and good electrical properties
- Native oxide grows at room temperature, non-uniform

*Figure 3.10: Atomic structure of SiO$_2$*
Thermal Oxidation Mechanisms

For the oxide layer to keep growing, the oxygen and silicon atoms must come in contact. However, the initially grown layer of silicon dioxide separates the oxygen in the chamber from the silicon atoms of the wafer surface. For oxide growth to continue, either the silicon in the wafer must migrate through the already grown oxide layer to the oxygen in the vapour, or the oxygen must migrate to the wafer surface. In the thermal growth of silicon dioxide, the oxygen migrates (the technical term is *diffuses*) through the existing oxide layer to the silicon wafer surface. Thus, the layer of silicon dioxide consumes silicon atoms from the wafer surface—the oxide layer grows into the silicon surface.

Consumption of Silicon during Oxidation

- For every 1000 Å, 450 Å of silicon is consumed
- Oxide growth occurs when the oxygen gas molecules move through the existing silicon oxide layer to the silicon wafer.

![Figure 3.11: Silicon consumption during oxidation](image)

Charge Buildup during Oxidation

- Positive charge is due to incomplete oxidation of Si
- Interface-trapped charge consisting of positive or negative charges that result from structural defects, oxidation induced defects, or metal impurities, and a mobile oxide charge
- Dangling bond can be annealed out using forming gas (H₂/N₂)
There are two types of oxidation. *Dry Oxidation* is done in O₂ gas environment. *Wet Oxidation* is done using H₂O vapour environment. The reaction that takes place during dry and wet oxidation is as given below:

**Dry oxidation**: \( \text{Si (solid) + O}_2 \, (\text{gas}) \xrightarrow{\text{Heat}} \text{SiO}_2 \, (\text{solid}) \)

**Wet Oxidation**: \( \text{Si (solid) + H}_2\text{O} \, (\text{gas}) \xrightarrow{\text{Heat}} \text{SiO}_2 \, (\text{solid}) + 2\text{H}_2 \, (\text{gas}) \)

Dry oxides are considered to be the best for gate dielectrics as compared to wet oxides. Hence in the present work dry oxide of different thicknesses were grown. The wafers were loaded into the oxidation furnace at 800 °C in N₂ ambient. The thermal oxide growth was done at 1050 °C in dry oxygen atmosphere for 20 mins and 40 mins to get an oxide layer of 720 Å and 820 Å respectively. The wafers were cooled at a rate of ~1 °C/s in N₂ ambient and unloaded from the furnace at 800 °C. The thickness of the oxide layer was measured using ellipsometer. The details of which are given in chapter 3.

**Metallization**

Metallization for gate and other contacts are normally done using evaporation method. In this method, the substrate is placed inside a vacuum chamber, in which a block (source) of the material to be deposited is also located. The source material is then heated to the point where it starts to boil and evaporate. The vacuum is required to allow the
molecules to evaporate freely in the chamber, and they subsequently condense on all surfaces. This principle is the same for all evaporation technologies, only the method used to the heat (evaporate) the source material differs. There are two popular evaporation technologies, which are, e-beam evaporation and resistive evaporation, each referring to the heating method. In e-beam evaporation, an electron beam is aimed at the source material causing local heating and evaporation. In resistive evaporation, a tungsten boat, containing the source material, is heated electrically with a high current to evaporate the material.

In the present work e-beam evaporation is used for both gate and back contact. Nickle was used for gate and for back contact titanium and gold was used. Vacuum was maintained at 2E-7 torr before evaporation. Evaporation was done for about 20-25 minutes in each case. For gate metallization metal masks with 2 mm dia were used.

![Figure 3.13: Photographs of Metal mask and metalized wafers](image)

**Back Oxide Etch**

While doing dry oxidation, oxide was grown on front as well as back side of the wafer. To form an ohmic contact at the substrate, the back oxide has to be etched off before back metallization. This was done using buffered oxide etchant. Then wafers were thoroughly rinsed in DI water and then dried using N\(_2\) gun. If the silicon dioxide has not been completely etched away, we observe a film of the DI water wetting or sheeting across the whole surface of the wafer. If the layer of SiO\(_2\) has been etched off, the DI water will not adhere to the back, except for a few isolated drops, and the wafer will appear dry and dull gray. This happens because SiO\(_2\) is hydrophilic as opposed to silicon which is hydrophobic.
The Buffered HF consists of 6 parts NH₄F (40%) + 1 part HF (49%). Etch rate of SiO₂ at room temperature in buffered HF is ~ 700 Å/min. The etch rate of SiO₂ increases with increasing weight % of HF in the etch solution, as well as higher ratios of NH₄F buffer in BHF solution. Etch rate also directly increases with increasing temperature.

After the completion of the fabrication of MOS structure, individual devices were diced and characterized using an automatic wafer level testing probe to extract the parameters of interest. The characterizations of the devices are given in the following sections.

Table 3.2: Device specifications for fabricated MOS capacitor

<table>
<thead>
<tr>
<th>Substrate</th>
<th>n-type silicon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orientation</td>
<td>&lt;100&gt;</td>
</tr>
<tr>
<td>Wafer diameter</td>
<td>1.5 inch</td>
</tr>
<tr>
<td>Wafer thickness</td>
<td>220 microns</td>
</tr>
<tr>
<td>Resistivity</td>
<td>0.08-0.2 Ω-cm</td>
</tr>
<tr>
<td>Dielectric material</td>
<td>SiO₂</td>
</tr>
<tr>
<td>Dielectric thickness</td>
<td>72 nm</td>
</tr>
<tr>
<td>Gate area</td>
<td>3.14 x 10⁻² cm²</td>
</tr>
<tr>
<td>Gate metal</td>
<td>Nickel</td>
</tr>
</tbody>
</table>

3.9 Parameter Extraction from C-V Curves

The single most important measurement in MOS work is a plot of the capacitance of a MOS structure as a function of its dc gate bias. This is known as C-V curve. Capacitance-voltage (C-V) measurements are commonly used in studying the gate-oxide quality and to account for the defects in detail. These measurements are made on a two-terminal device called a MOS capacitor (MOS cap), which is basically a MOSFET without a source and drain. C-V test results offer a wealth of device and process information, including bulk and interface charges. Many MOS device parameters, such as oxide thickness, flatband voltage, mid gap voltage, substrate doping etc. can be
extracted from the C-V data. With a high frequency C-V it is also possible to quantify
the oxide and interface charge densities. In this section the method of extracting the
parameters of interest from a high frequency C-V curve is explained. By definition,
capacitance is the change in charge (Q) in a device that occurs when there is a change in
voltage (V):
\[ C = \frac{\Delta Q}{\Delta V} \]

One general practical way to implement this is to apply a small AC voltage signal to the
device under test, and then measure the resulting current. Integrate the current over time
to derive Q and then calculate C from Q and V. C-V measurements in a semiconductor
device are made using two simultaneous voltage sources: an applied fixed AC voltage
signal and a DC voltage that is swept in time. The purpose of the DC voltage bias is to
allow sampling of the material at different depths in the device. The AC voltage bias
provides the small-signal bias so the capacitance measurement can be performed at a
given depth in the device.

**High Frequency C-V Curve**

The most important property of the MOS capacitor is that its capacitance changes with
an applied DC voltage. As a result, the modes of operation of the MOS capacitor change
as a function of the applied voltage. Figure 3.14 illustrates the high frequency (@ 1 MHz
ac signal) C-V curve (ideal) drawn as per device specifications (n-type substrate, 720 Å
thick SiO₂, 3.14x10⁻² cm² gate area) used for our experiments. As a DC sweep voltage is
applied to the gate, it causes the device to pass through accumulation, depletion, and
inversion regions.

**Normalization of Ideal C-V Curve**

The C-V curves can be normalized to its maximum value without loss of generality. This
removes the need to know the area of the MOS capacitor, also providing the other
necessary information. The normalized curve of the above said high frequency C-V
curve is as shown in Figure 3.15.
Relation between Oxide Thickness ($T_{OX}$) and Oxide Capacitance ($C_{OX}$)

For a relatively thick oxide (> 50Å), extracting the oxide thickness is fairly simple. The oxide capacitance ($C_{OX}$) is the high frequency capacitance when the device is biased for strong accumulation. In the strong accumulation region, the MOS-C acts like a parallel-plate capacitor and the oxide thickness ($T_{OX}$) may be calculated from $C_{OX}$ and the gate area using the following equation:
where:

\[ T_{OX} = \text{oxide thickness (nm)} \]
\[ A = \text{gate area (cm}^2\text{)} \]
\[ \varepsilon_{ox} = \text{permittivity of the oxide material (F/cm)} \]
\[ C_{OX} = \text{oxide capacitance (F)} \]
\[ 10^7 = \text{units conversion from cm to nm.} \]

**Series Resistance**

After generating a C-V curve, it may be necessary to compensate for series resistance in measurements. The series resistance \((R_s)\) can be attributed to either the substrate (well) or the backside of the wafer. For wafers typically produced in fabs, the substrate bulk resistance is fairly small (< 10Ω) and has negligible impact on C-V measurements. However, if the backside of the wafer is used as an electrical contact, the series resistance due to oxides can significantly distort a measured C-V curve. Without series compensation, the measured capacitance can be lower than the expected capacitance, and C-V curves can be distorted. The Figure 3.16 shows the 3 element and a simplified 2 element model of MOS capacitor in strong accumulation to determine series resistance.

In the figure below \(C_A\) is the accumulation capacitance, \(C_{OX}\) oxide capacitance and \(R_S\) is the series resistance.

![Diagram](image)

**Figure 3.16: Simplified MOS model in accumulation to determine Series Resistance**

50
The series resistance \((R_s)\) may be calculated from the capacitance and conductance values that are measured while biasing the MOS device in the accumulation region as follows:

\[
R_s = \frac{\left( \frac{G}{2\pi f C} \right)^2}{1 + \left( \frac{G}{2\pi f C} \right)^2} G
\]

where:

- \(R_s\) = series resistance (\(\Omega\))
- \(G\) = measured conductance (\(S\))
- \(C\) = measured parallel model capacitance (in strong accumulation) (\(\text{F}\))
- \(f\) = test frequency (\(\text{Hz}\)).

### Substrate Doping Concentration

The substrate doping concentration \((N_0)\) is related to the reciprocal of the slope of the \(1/C^2\) vs. \(V_G\) curve. The doping concentration is calculated using the equation [3]

\[
N_D = \frac{2}{q \varepsilon_s A \left( \frac{\Delta 1/C^2}{\Delta V_G} \right)}
\]

where:

- \(N_D\) = substrate doping concentration (\(\text{cm}^{-3}\)).
- \(q\) = electron charge \((1.60219 \times 10^{-19} \text{ C})\)
- \(A\) = gate area (\(\text{cm}^2\))
- \(\varepsilon_s\) = permittivity of the substrate material (\(\text{F/cm}\))
- \(V_G\) = gate voltage (\(\text{V}\))
- \(C\) = measured capacitance (\(\text{F}\)).
Flatband Capacitance and Flatband Voltage

The application of the flatband voltage ($V_{FB}$) results in the disappearance of band bending. At this point, known as the flatband condition, the semiconductor band is said to become flat. Because the band is flat, the surface potential is zero (with the reference potential being taken as the bulk potential deep in the semiconductor). Flatband voltage and its shift are widely used to extract other device parameters, such as oxide charges.

$V_{FB}$ can be identified from the C-V curve. One way is to use the flatband capacitance method. For this method, the ideal value of the flatband capacitance ($C_{FB}$) is calculated using the equation [3]

$$\frac{C_{FB}}{C_{OX}} = \frac{1}{1 + \left( \frac{136 \sqrt{T/300}}{T_{OX} \sqrt{N_D}} \right)}$$

where:

- $C_{OX}$ = Oxide capacitance (F)
- $T$ = Temperature (K),
- $T_{OX}$ = Oxide thickness (m)
- $N_D$ = Substrate doping concentration (cm$^{-3}$).

Metal-semiconductor Work Function Difference

The metal-semiconductor work function difference ($\phi_{ms}$) is commonly referred to as the workfunction. The work function represents the difference in work necessary to remove an electron from the gate and from the substrate. For an ideal MOS diode with n-type semiconductor, the work function difference is assumed to be zero and is derived as follows:

$$\phi_{ms} = \phi_n - \left( \chi + \frac{E_g}{2q} - \psi_S \right)$$

where:

- $\phi_n$ = work function difference (V)
\( \phi_m \) = metal work function (V)

\( \chi \) = electron affinity (V)

\( E_g \) = substrate material bandgap (eV)

\( \psi_B \) = bulk potential (V).

If the value of \( \phi_m \) is not zero, and if oxide charges exist, (assuming negligible interface traps) the experimental capacitance-voltage curve will be shifted from the ideal theoretical curve by an amount \( V_{FB} \) [11].

**Fixed Charge Density \((N_f)\)**

The shift in the flat band voltage is directly related to the work function difference and the density of fixed charges. The fixed charge density can be extracted from the C-V curve using the equation. Irradiation of devices causes the change in fixed charge density and the net fixed charge density \((\Delta N_f)\) is given by

\[
\Delta N_f = \frac{(\phi_m - \Delta V_{fb}) C_{OX}}{q}
\]

where:

\( \Delta N_f \) = net fixed charge density \((\text{cm}^2)\)

\( q \) = the elementary charge \((1.60219 \times 10^{-19} \text{C})\).

\( \phi_m \) = work function difference (V)

\( \Delta V_{fb} \) = Flat band voltage shift (V)

\( C_{OX} \) = oxide capacitance (F).

**Oxide Charge Density \((N_o)\)**

The accurate value Oxide charge density results from extracting the values at the mid-gap since the effect of interface charges at the mid-gap will be at its minimum. The net change in the oxide charge density in the devices due to irradiation is given by [12]

\[
\Delta N_{ox} = -\frac{C_{OX} \Delta V_{mg}}{qA}
\]
where:

$\Delta N_{ox} = \text{net oxide charge density (cm}^{-2})$

$q = \text{the elementary charge (1.60219} \times 10^{-19} \text{C)}.$

$\Delta V_{mg} = \text{mid-gap voltage shift (V)}$

$C_{ox} = \text{oxide capacitance (F)}$

$A = \text{Gate area (cm}^2).$

**Interface Charge Density (N$_{it}$)**

The interface trap-charge densities can be estimated from midgap-to-flatband stretch-out of high frequency C-V curves and is given by [12]

$$
\Delta N_{it} = \frac{C_{ox}(\Delta V_{fb} - \Delta V_{mg})}{qA}.
$$

where:

$\Delta N_{it} = \text{net interface charge density (cm}^{-2})$

$q = \text{the elementary charge (1.60219} \times 10^{-19} \text{C)}.$

$\Delta V_{fb} = \text{flat band voltage shift (V)}$

$\Delta V_{mg} = \text{mid-gap voltage shift (V)}$

$C_{ox} = \text{oxide capacitance (F)}$

$A = \text{gate area (cm}^2).$

### 3.10 Dielectric Loss Tangent (tan $\delta$) and Quality Factor (Q)

The *loss tangent* is a parameter of a dielectric material that quantifies its inherent dissipation of electromagnetic energy. The term refers to the angle in a complex plane between the resistive (lossy) component of an electromagnetic field and its reactive (lossless) component.

A capacitor is typically made of a dielectric placed between conductors. The lumped element model of a capacitor includes a lossless ideal capacitor in series with a resistor termed the equivalent series resistance (ESR), as shown in the Figure 3.17.
Figure 3.17: Lumped elemental model of a real capacitor

The ESR represents losses in the capacitor. In a low-loss capacitor the ESR is very small, and in a lossy capacitor the ESR can be large. The ESR is a derived quantity representing the loss due to both the dielectric's conduction electrons and the bound dipole relaxation phenomena. In a dielectric, only one of, either the conduction electrons or the dipole relaxation, typically dominates loss. For the case of the conduction electrons being the dominant loss, then

$$ESR = \frac{\sigma}{\epsilon'\omega^2 C},$$

where:

$C$ = lossless capacitance

$\sigma$ = free charge conduction

$\epsilon'$ = $\epsilon_0 \epsilon_r$ (Permittivity of free space x relative permittivity)

$\omega$ = angular frequency.

A capacitor's loss tangent is equal to the tangent of the angle between the capacitor's impedance vector and the negative reactive axis, as shown in the Figure 3.18.

Figure 3.18: Graphical representation of angle $\delta$
The loss tangent is then given by

\[ \tan \delta = \frac{ESR}{|X_c|} = \omega C \cdot ESR = \frac{\sigma}{\varepsilon \omega}. \]

The dielectric loss tangent is also called the Dissipation Factor (DF) which is equal to the reciprocal of its Quality Factor (Q). The quality factor is a dimensionless quantity.

\[ \tan \delta = DF = \frac{1}{Q}. \]

3.11 Deep Level Transient Spectroscopy (DLTS)

In contrast to the steady-state high frequency C-V methods described so far, transient capacitance spectroscopy gives information by measuring how the nonsteady-state high-frequency capacitance changes with time t. DLTS was invented by Lang [13] and is now widely used to detect traps of so-called "deep levels" in the Si band gap. DLTS is a capacitance transient thermal scanning technique and it overcomes the drawbacks of other conventional techniques like Thermally Stimulated Current (TSC) and Thermally Stimulated Capacitance (TSCAP), in view of its better immunity to noise and surface channel leakage current. It is sensitive (detects trap concentrations as low as $10^4$ in order), spectroscopic (exhibits a peak for each trap detected) and allows to obtain parameters from either minority (positive peak in the spectrum) or majority (negative peak) carrier traps. Easy and direct interpretation of experimental results obtained from DLTS study makes it a valuable tool for defect analysis [14 - 15].

Initially, the method utilized measurements of transient capacitance following the pulsed bias in a p-n junction or Schottky barrier diode to monitor changes in the charge state of defect centers. Schulz and Johnson [16] extended applying DLTS to study the charge emission from interface states in MOS structures. Therefore, for MOS structure, if the charge emission from interface states of Si is detectable, it is worthwhile to use DLTS for characterization.

The DLTS signal is the difference of capacitances at two different times after a filling pulse. It shows peaks for different trap levels in the sample at the respective temperatures $T$. Sign of the DLTS peak indicates whether the trap is near valance band or conduction
Height of DLTS peak is proportional to the trap concentration. Position of the peak, on temperature scale is uniquely determined by the thermal emission properties of the trap. If traps are filled by a filling pulse and the reverse bias is switched on again, the sample is in thermal non-equilibrium and relaxes into equilibrium by detrapping the charges back. This relaxation process is related with a capacitance transient. Its time constant is governed by the thermal emission rate $e_m$ which depends on the trap energy $E_t$ and the temperature $T$. By analyzing a DLTS spectrum, one can measure the activation energy, concentration profile and capture cross section for each trap.

Lang introduced the rate window concept to deep level characterization. If the C-t curve from the transient capacitance experiment is processed so that a selected decay rate produces a maximum output, then a signal whose decay time changes monotonically with temperature, reaches a peak when the signal passes through the rate window of a boxcar average. When observing a repetitive C-t transient through such a rate window while varying the decay time constant by varying the sample temperature, a peak appears in the temperature versus output plot. This plot is named DLTS spectrum. One such DLTS spectrum corresponding to the temperature dependent capacitance transient, obtained by the rate window set by $t_1$ & $t_2$ is shown in Figure 3.19.

![Figure 3.19: Temperature dependent capacitance transient and corresponding DLTS signal](image-url)
If $C(t_1)$ is the capacitance at sampling time $t_1$ and $C(t_2)$ is the capacitance at sampling time $t_2$ then the difference signal $\Delta C = C(t_1) - C(t_2)$ gives the DLTS signal. There is no difference between the capacitance at the two sampling times for very slow and very fast transients, corresponding to low and high temperatures respectively. A difference signal is generated when the time constant is of the order of $t_2 - t_1$ and goes to maximum as temperature varies. Thus the values of $t_1$ & $t_2$ determine the rate window for a DLTS thermal scan. Figure 3.20 shows the DLTS spectra of as processed MOS capacitor corresponding to the 7 rate windows employed in the present study.

![Figure 3.20: DLTS spectra corresponding to seven rate windows](image)

**Analysis of DLTS Spectrum**

The mechanism of recombination and generation of charge carriers through a deep level is described by a mathematical model. According to this model, a G-R center can be in one of the two charge states. When occupied by an electron, it is said to be in the $n_T$ state, and when occupied by a hole, it is said to be in the $p_T$ state. The concentration of G-R centers occupied by electrons $n_T$, and holes $p_T$ must be equal to the total concentration $N_T$, i.e., $N_T = n_T + p_T$. In other words, a center is either occupied by an electron or a hole. During the recombination or generation of electrons and holes, the electron concentration
in the conduction band \( n \), the hole concentration in the valance band \( p \), \( n_T \) and \( p_T \) change with time. The rate of change of \( n \) due to G-R mechanisms is given by,

\[
\frac{dn}{dt} = e_n n_T - c_n n p_T.
\]

Here \( c_n \) represents the capture rate which is the number of electrons captured by the G-R center from the conduction band. The capture co-efficient \( c_n \) is given by \( c_n = \sigma_n v_{th} \) where \( \sigma_n \) is the capture cross section of the G-R center and \( v_{th} \) is the thermal velocity of the electron. The emission rate \( e_n \) represents the number of electrons emitted per second from electron occupied G-R centers. \( e_n \) is the reciprocal of emission time constant \( \tau_e \) which is a function of temperature (T). \( \tau_e \) is given by,

\[
\tau_e = \frac{\exp\left(\frac{E_c - E_T}{kT}\right)}{\gamma_n \sigma_n T^2}.
\]

Here \( (E_c - E_T) \) represents the activation energy of the G-R center, \( k \) is the Boltzmann’s constant and \( \gamma_n \) is the material co-efficient. For Silicon, \( \gamma_n \sim 10^{21} \text{ cm}^{-2}\text{s}^{-1}\text{K}^{-2} \).

- A graph of \( I_n(T, T^2) \) versus \( \frac{1}{T} \) is called Arrhenius plot.
- Slope of the Arrhenius plot gives the activation energy of the G-R center.
- Y-intercept results in the calculation of capture cross section (\( \sigma_n \)) of the G-R center.
- Recombination of carriers through deep levels is called Shockley Read Hall (SRH) recombination [3]. SRH recombination life time is calculated using the equation,

\[
\tau_{SRH} = \frac{1}{\sigma_n v_{th} N_T}
\]

- The effective carrier life time (\( \tau_{eff} \)) is calculated using the relation,

\[
\frac{1}{\tau_{eff}} = \frac{1}{\tau_1} + \frac{1}{\tau_2} + \frac{1}{\tau_3} + \ldots
\]
where each term in the right hand side of the above equation represents the carrier life time corresponding to different trap levels.

- The identification of the defect type is made on the basis of their fingerprint such as activation energy, annealing temperature and capture cross section by comparing with those reported in the literature.

### 3.12 Process Induced Deep Level Defects

As purchased silicon wafer has a defect content which is below the measurable limit and can be essentially considered to be defect free. But during device processing which includes a number of steps viz., oxidation, diffusion, ion implantation, metallization etc., the introduction of defects in the substrate is most likely. The behaviour of defects in semiconductor structures induced during fabrication process is a major concern for microelectronics device technologies. As the technology is leading to a very large scale integration, millions of components are to be fabricated on a single chip. To ensure proper functioning of the chip, the fraction of the defective components must be smaller than $10^{-6}$ components per chip [17]. This implies that even a small number of defects may potentially cause a serious threat to the functionality and reliability of the semiconductor devices.

An important process in the silicon technology is the growth of an oxide on the substrate by means of its thermal oxidation. Conversion of silicon to silicon dioxide in this manner results in a volume change by a factor 2.23. This fact, combined with the difference in the coefficients of thermal expansion of these materials, often results in the formation of dislocations in the silicon during the cool down process [18]. The mismatch in the thermal expansion coefficients results in the stress of the order of $10^9$ dyn/cm$^2$ and could easily lead to interface shear stresses that are higher than the critical stress for shear flow ($\sim 5 \times 10^7$ dyn/cm$^2$) in silicon. This indicates that generation of defects in the Silicon crystal lattice is possible during oxidation or post oxidation cooling [19]. The lattice defects are usually characterized by highly localized states situated deep in the band gap as well as even close to the valence or conduction band edges which are commonly called deep level defects. Impurities, vacancies, interstitials, or their clusters are some of the examples of such deep level defects. Deep level defects have larger capture cross
sections than hydrogenic shallow defects and in most cases determine the minority carrier lifetime. Therefore, these defects play an important role in manufacturing high-speed electronic and optoelectronic devices. With technological scaling of semiconductor devices, the processing steps tend to become more complicated and are likely to result in more process induced defects. Hence defect investigation and characterization are interesting from the cognitive and practical point of view, as they govern the effective production of perfect semiconductor devices [20].

**DLTS Measurements**

The DLTS system (IMS-2000) employed for the present study consists of a boxcar averager, a pulse generator, a thousand point digitizer, a voltage generator and a high speed capacitance meter. The pulse generator is capable of generating pulses with widths ranging from 100 ns to 10 s. The pulse height could be programmed from -12 V to +12 V. The boxcar averager is capable of generating seven rate windows. The time constants can be varied from 1 ms to 2 s. In the present study, DLTS spectra are recorded with a reverse bias of 5 V and pulse width of 20 ms applied between the gate and the substrate. DLTS spectra were recorded for ‘as processed’ MOS capacitor. The device was subjected to isochronal annealing for 30 min at various temperatures (200 °C, 250 °C, 300 °C, and 350 °C) and the DLTS spectra were recorded after each annealing temperature. The trap concentration, activation energy and capture cross section of different deep levels were determined by DLTS data.

Figure 3.21 exhibits the DLTS spectra of MOS capacitor before and after annealing at four different annealing temperatures. The decrease in height of the peaks after annealing is a clear indication of reducing trap concentration after annealing. Figure 3.22 (a to d) exhibits the Arrhenius plots of deep level defects for ‘as processed’ and annealed MOS samples. As one can see from the Figure 3.21, the DLTS spectrum of the ‘as processed’ sample exhibits one peak at a temperature of 302 K which corresponds to a deep level (labeled E0) with activation energy of $E_C - 0.49$ eV in the forbidden gap of silicon. This defect can be attributed to interstitial silicon or the self interstitial Si(i). This defect is most likely generated due to oxidation of silicon samples [21, 22]. Thermal annealing of the device at various temperature results in a shift in the temperature as well as peak height. The defect acquires a different energy level with reduction in the peak height as...
the annealing temperature is increased. Table 3.3 exhibits the assignment of defect and its characteristics at different annealing temperatures. The trap concentration and capture cross section of all the deep level defects are calculated from the DLTS spectra. The activation energy of all the defects has been measured to an accuracy of 0.001 eV.

![DLTS spectra of MOS capacitor before and after annealing](image)

**Figure 3.21:** DLTS spectra of MOS capacitor before and after annealing

![Figure 3.22(a) 'as processed' device](image)  ![Figure 3.22(b) device annealed at 200 °C](image)

**Figure 3.22(a)** ‘as processed’ device  **Figure 3.22(b)** device annealed at 200 °C

62
Figure 3.22(c) device annealed at 250 °C  Figure 3.22(d) device annealed at 300 °C

Figure 3.22: Arrhenius plots corresponding to deep level defects observed for ‘as processed’ and annealed MOS samples

Table 3.3: Characteristics of defects in ‘as processed’ and annealed MOS capacitor

<table>
<thead>
<tr>
<th>Annealing temperature (°C)</th>
<th>Defect type</th>
<th>Activation energy (eV)</th>
<th>Trap concentration (cm⁻³)</th>
<th>Capture cross section (cm²)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>As processed (before annealing)</td>
<td>Si(i)</td>
<td>Eₐ - 0.49</td>
<td>1.10 x 10¹⁵</td>
<td>6.2 x 10⁻¹⁷</td>
<td>[21], [22]</td>
</tr>
<tr>
<td>200</td>
<td>Phosphorous vacancy pair (V-P)</td>
<td>Eₐ - 0.43</td>
<td>7.15 x 10¹⁴</td>
<td>6.96 x 10⁻¹⁸</td>
<td>[23]</td>
</tr>
<tr>
<td>250</td>
<td>Vacancy related complex</td>
<td>Eₐ - 0.46</td>
<td>3.65 x 10¹⁴</td>
<td>4.49 x 10⁻¹⁷</td>
<td>[24], [25]</td>
</tr>
<tr>
<td>300</td>
<td>Multi-vacancy-oxygen (V₂O₂ or V₃O)</td>
<td>Eₐ - 0.34</td>
<td>1.62 x 10¹⁴</td>
<td>4.87 x 10⁻¹⁹</td>
<td>[23], [26]</td>
</tr>
</tbody>
</table>

The identification of defect type is made on the basis of their finger prints such as activation energy and capture cross section by comparing with those reported in the literature. Carrier lifetime has been calculated from the DLTS data. Figure 3.23 shows that the carrier lifetime increases with annealing temperature. This is in conformity with the decrease in trap density with increase in annealing temperature as shown in Table 3.3.
Figure 3.23: Carrier lifetimes as a function of annealing temperature

The variation in substrate doping concentration with annealing temperature is shown in Figure 3.24.

Figure 3.24: Substrate doping measured after each annealing temperature

In silicon integrated circuit fabrication, the silicon is doped with dopants during crystalline growth or during device processing. When doped silicon is oxidized, a redistribution of dopant between the silicon and the growing oxide occurs. The dopant could preferentially segregate in the silicon or in the oxide, and the dopant distribution
will be affected by the relative dopant diffusivities in silicon and SiO$_2$. In case of phosphorus, the dopants piles up in the silicon at the interface. The variation in the dopant concentration in the silicon affects the electrical properties of the devices and hence a clear understanding of the dopant redistribution is very important [19].

From the above observations, it is well established that the interstitial silicon defect introduced during oxidation completely anneals at 350 °C. The present study thus brings about the importance of thermal annealing of devices to remove process induced defects in semiconductor fabrication.

References

4. DeWitt G. Ong, *Modern MOS Technology*


