Course Title: Processing & Fabrication Technology

Course No.: EEE 707

Class lecture -2

Introduction to Diffusion and Ion Implantation

Diffusion

Diffusion is the process by which impurity atoms move through the crystal lattice.

By this process a certain region of silicon substrate is doped to make a N+ or P+ region to form the source and drain regions in MOS device.

The idea of diffusing techniques to alter the type of conductivity in silicon or germanium was disclosed by Pfann in 1952.

Diffusion technique is used to form:

Bases, emitters, collectors and resistors in bipolar device fabrication and

Source and drain regions and to dope polysilicon in MOS device fabrication.

Diffusion…

The rate at which the dopants diffuse into the silicon is a strong function of temperature.

Diffusion is carried out at high temperature (1000C to 1200C) to obtain the desired doping profile.

When the substrate is cooled at room temperature, the impurities are essentially frozen to its position.

The diffusion process is performed in a furnace like the oxidation process.

The depth to which impurities diffuse depend on temperature and time of diffusion.

Common material used as dopants are boron (p-type dopants), and phosphorous and arsenic (n-type dopants). They have solubilities above $5x10^{20}$ cm^{-3} in the diffusion temperature range.

Diffusion…

- These dopants can be introduced in several ways: including
  - Solid sources (e.g. BN for boron, As_2O_3 for arsenic, and P_2O_5 for phosphorous),
  - Liquid sources (BBr_3, $AsCl_3$, and $POCl_3$), and
  - Gaseous sources (B_2H_6, AsH_3, and PH_3).
- However, liquid sources are most commonly used.
- An example of the chemical reaction for phosphorous diffusion using a liquid source is -

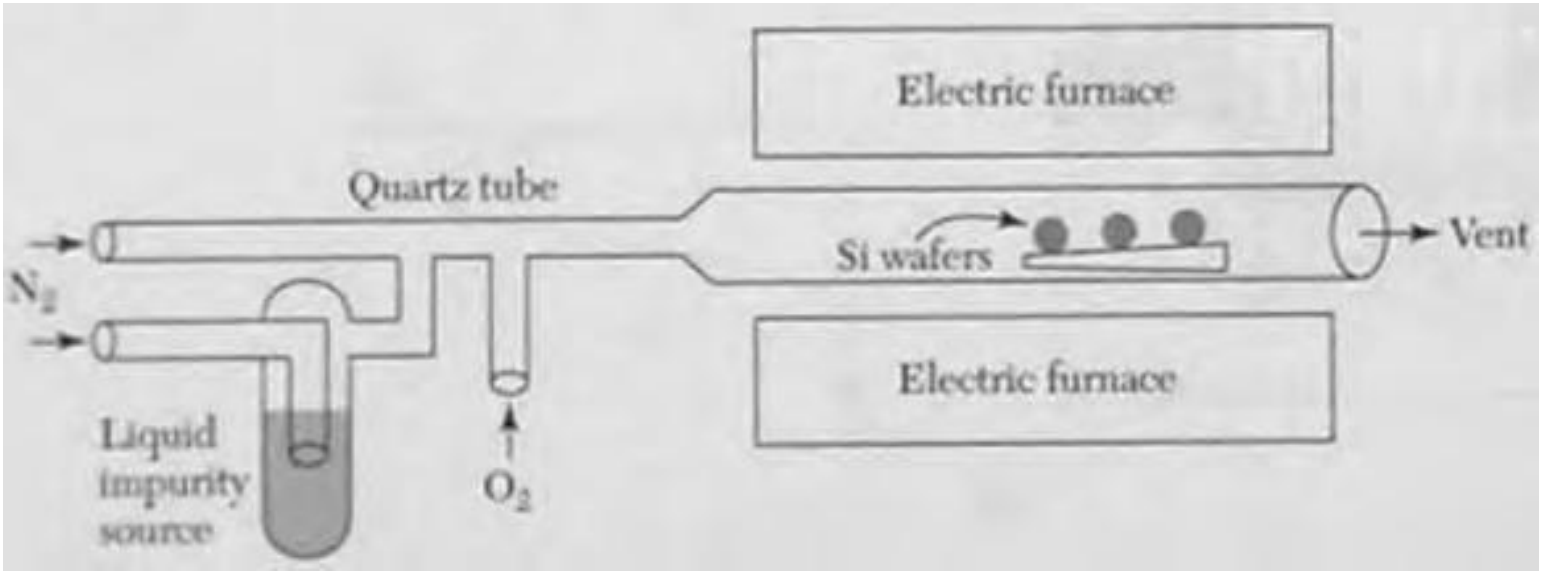
$$4POCl_3 + 3O_2 = 2P_2O_5 + 6Cl_2 \uparrow$$

- The P_2O_5 forms a glass-on-silicon wafer and is then reduced to phosphorus by silicon,

$$2P_2O_5 + 5Si = 4P + 5SiO_2$$

- The phosphorous is released and diffuse into the silicon and Cl_2 is vented.

Diffusion…



Schematic diagram of a typical open-tube diffusion system

- For diffusion in gallium arsenide, the high vapor pressure of arsenic requires special methods to prevent the loss of arsenic by decomposition or evaporation.

Ion Implantation

Ion implantation is another method used to introduce impurities into silicon.

An ion implanter produces ions of the desired impurity, accelerates them by an electric field, and allows them to strike the silicon surface.

The ions become embedded in the silicon.

The depth of penetration depends on the energy of ion beam, which is controlled by the accelerating field voltage.

The quantities of ions implanted can be controlled by varying the beam current (flow of ions). As both voltage and current can be accurately controlled, the ion implantation results in much more accurate and reproducible impurity profiles than can be obtained by diffusion.

Moreover, ion implantation can be performed at room temperature.

Annealing

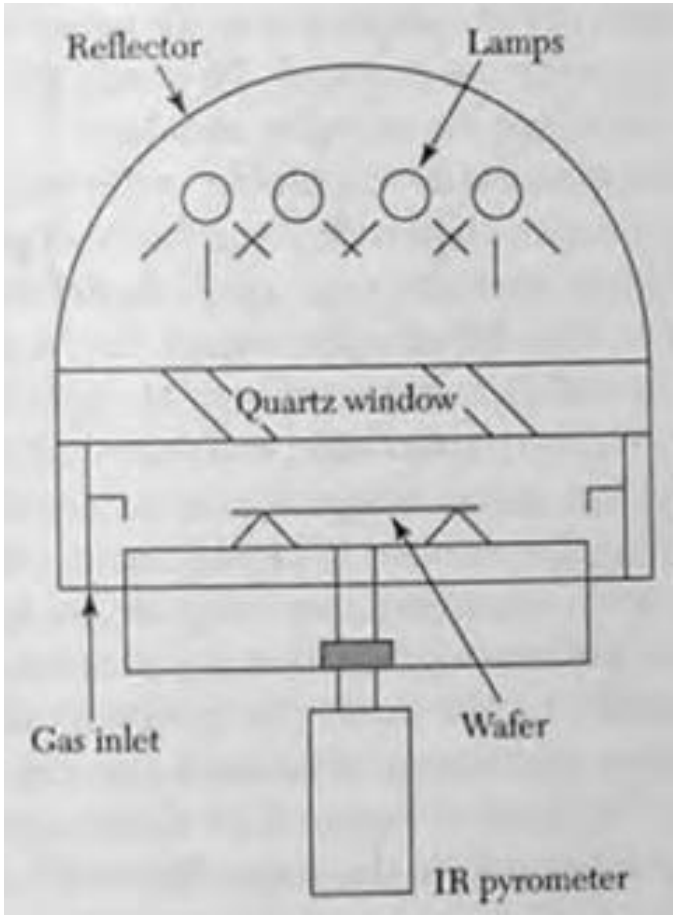
- Because of the damaged region and the disorder cluster that result from ion implantation, semiconductor parameters such as mobility and lifetime are severely degraded.
- In addition, most of the ions as implanted are not located in substantial site. To activate the implanted ions and to restore mobility and other material parameters, we must anneal the semiconductor at an appropriate combination of time and temperature.

Annealing…

- Conventional annealing uses an open-tube, batch furnace system similar to that used for thermal oxidation. This process requires a long time and high temperature to remove the implant damage. However, conventional annealing may cause substantial dopant diffusion and cannot meet requirements for shallow junctions and narrow doping profiles.
- Rapid thermal annealing (RTA) is an annealing process that employs a variety of energy sources with a wide range of times, from 100 seconds down to nanoseconds - all short compared with conventional annealing. RTA can activate dopants fully with minimal redistribution.

Rapid Thermal Annealing (RTA)

- The temperature measured from the heated wafer is usually from 600C to 1100C. A wafer is heated quickly under atmospheric conditions or at low pressure under isothermal conditions. Typical lamps in an RTA system are tungsten filaments or arc lamps. The processing chamber is made of either quartz, silicon carbide, stainless steel, or aluminum and has quartz windows through which the optical radiation passes to illuminate the wafer. The wafer holder is often made of quartz and contacts the wafer in a minimum number of places.



RTA system that is optically heated

Rapid Thermal Annealing (RTA)…

- A measurement system is placed in a control loop to set wafer temperature. The RTA system interfaces with a gas-handling system and a computer that controls system operation. Typically, wafer temperature in an RTA system is measured with a noncontact optical pyrometer that determines temperature from radiated infrared energy.

Comparison between Conventional annealing and RTA

- Rapid heating with temperature gradients in the wafers can cause wafer damage in the form of slip dislocations induced by thermal stress.
- On the other hand, conventional furnace processing causes significant problems, such as particle generation from the hot walls, limited ambient control in an open system and a large thermal mass that restricts controlled heating times to tens of minutes.
- In fact requirements on contamination, process control, and cost of manufacturing floor space have resulted in a paradigm shift to the RTA process.

Introduction to Oxidation

Oxidation

The oxidation of silicon is necessary throughout the modern IC fabrication process to form SiO_2.

Oxidation refers to the chemical process of silicon reacting with oxygen to form silicon dioxide (SiO_2) layer on the substrate.

SiO_2 has several applications:

- To serve as a mask (buffer material) against ion implantation or diffusion of dopants into silicon,
- To provide surface passivation,
- To isolate one device from another,
- To act as a component in MOS structures,
- To provide electrical isolation of multilevel metallization systems.

Oxidation Techniques

Techniques used to form Silicon dioxide layers:

1. Thermal oxidation (Dry oxidation and Wet oxidation)
   - Most important for Si devices and key process in modern Si IC technology.
2. Vapor phase technique (CVD)
3. Plasma oxidation

Growth Mechanism

Thermal Oxidation is performed at 700 - 1200 C.

Silicon surface has a high affinity to oxygen, SiO_2 layer forms rapidly when Si surface is exposed to an oxidizing ambient.

To speed up the reaction, the wafer is heated to 1000C to 1200C.

The heating is performed in a special high-temperature ultra-clean furnace.

All the fabrication steps are accomplished in an ultra-cleaned space. Because, a tiny dust particle may alter the electrical properties of the IC, even destroy the circuitry. Automatic wafer handling and robotics are used during the process steps. Even then, the wafers must be constantly cleaned to avoid contamination.

Growth Mechanism…

Oxygen used in the reaction is introduced as either (i) a high-purity gas (dry oxide) or (ii) water vapor (wet oxide).

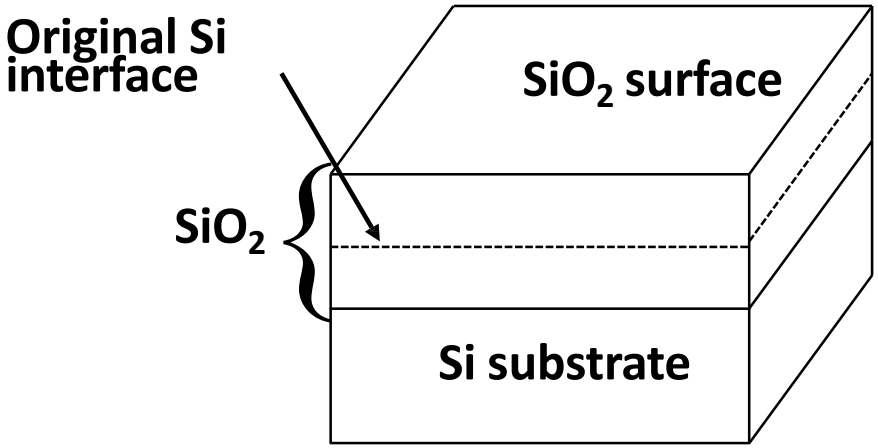
Chemical reactions describing the thermal oxidation of Si in oxygen or water vapor are:

$$Si\ (solid) + O_2\ (gas) \rightarrow SiO_2\ (solid) \qquad (i)$$

$$Si\ (solid) + 2H_2O \rightarrow SiO_2\ (solid) + 2H_2\ (gas) \qquad (ii)$$

The basic process involves the sharing of valence electrons between silicon and oxygen; the silicon-oxygen bond structure is covalent.

Growth Mechanism…



During the course of oxidation Si-SiO_2 interface moves into the Si. It is found that for a growth of an oxide thickness of d, a layer of Si with thickness of 0.44d is consumed.

Through experimental analysis it is established that oxidation proceeds by diffusion of the oxidizing species through the oxide to the Si-SiO_2 interface where the oxidation reaction occurs.

In literature, there are controversies regarding the fact whether charged or neutral species are transported through the oxide layer and the reaction at the Si-SiO_2 interface.

Kinetics of Thermal Oxidation

A silicon slice contacts the oxidizing species (oxygen or water vapor), resulting in a surface concentration of C_0 molecules/cm^3 for these species. The magnitude of C_0 equals the equilibrium bulk concentration of the species at the oxidation temperature.

The equilibrium concentration generally is proportional to the partial pressure of the oxidant adjacent to the oxide surface. At 1000C and a pressure of 1 atm, the concentration C_0 is 5.2 x 10^{16} molecules/cm^3 for dry oxygen and 3 x 10^{19} molecules/cm^3 for water vapor.

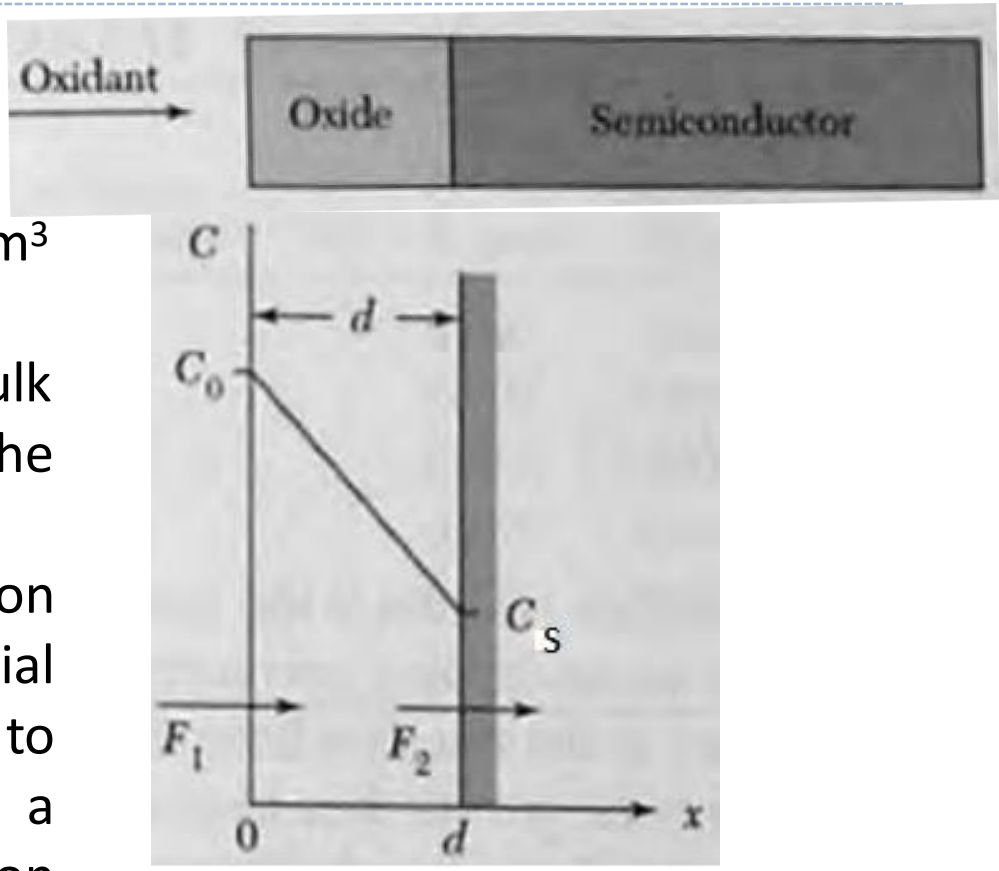


Fig. Basic model for the thermal oxidation of silicon

Kinetics of Thermal Oxidation…

The oxidizing species diffuse through the silicon dioxide layer, resulting in a concentration C_s at the surface of silicon. The flux F_1 can be written as

$$F_1 = D\frac{dC}{dx} \equiv \frac{D(C_0 - C_s)}{x} \quad \text{..... (1)}$$

Where, D is the diffusion coefficient of the oxidizing species, and x is the thickness of the oxide layer already present.

At the silicon surface, the oxidizing species reacts chemically with silicon. Assuming the rate of reaction is proportional to the concentration of the species at the silicon surface, the flux F_2 is given by

$$F_2 = kC_s \quad \text{..... (2)}$$

Where, k is the surface reaction rate constant for oxidation.

Kinetics of Thermal Oxidation…

At the steady state,

$F_1 = F_2 = F$

Combing equations (1) and (2) gives

$$F = \frac{DC_0}{x + (D/k)}$$

The reaction of the oxidizing species with silicon forms silicon dioxide. Let C_1 be the number of molecules of the oxidizing species in a unit volume of the oxide.

Thus the growth rate of the oxide layer thickness is given by

$$\frac{dx}{dt} = \frac{F}{C_1} = \frac{DC_0 / C_1}{x + (D/k)} \quad \text{..... (3)}$$

Kinetics of Thermal Oxidation…

We can solve this differential equation subject to the initial condition x (0) = d_0, where d_0 is the initial oxide thickness; d_0 can also be regarded as the thickness of oxide layer grown in an earlier oxidation step. Solving equation (3) yields the general relationship for the oxidation of silicon,

$$x^2 + \frac{2D}{k}x = \frac{2DC_0}{C_1}(t + \tau) \qquad \text{..... (4)}$$

Where $\tau \equiv (d_0^2 + 2Dd_0/k)C_1/2DC_0$, which represent a time coordinate shift to account for the initial oxide layer d_0.

The oxide thickness after an oxidizing time t is given by

$$x = \frac{D}{k}[\sqrt{1 + \frac{2C_0k^2(t+\tau)}{DC_1}} - 1] \qquad \text{..... (5)}$$

Kinetics of Thermal Oxidation…

For small values of t, equation (5) reduces to

$$x \cong \frac{C_0 k}{C_1}(t+\tau) \quad \text{..... (6)}$$

and for larger values of t, it reduces to

$$x \cong \sqrt{\frac{2DC_0}{C_1}(t+\tau)} \quad \text{..... (7)}$$

During the early stages of oxide growth, when surface reaction is the rate-limiting factor, the oxide thickness varies linearly with time. As the oxide layer becomes thicker, the oxidant must diffuse through the oxide layer to react at the silicon-silicon dioxide interface, and the reaction becomes diffusion limited. The oxide growth then becomes proportional to the square root of the oxidizing time, which results in a parabolic growth rate.

Kinetics of Thermal Oxidation…

equation (4) is often written in a more compact form:

$$x^2 = Ax = B(t + \tau)$$

Where, A = 2D/k, B = $2DC_0/C_1$ and B/A = kC_0/C_1. using this form, Eqs. (6) and (7) can be written as

$$x = \frac{B}{A}(t + \tau)$$

for the linear region and as

$$x^2 = B(t + \tau)$$

for the parabolic region.

Why is wet oxidation rate higher than dry oxidation rate?

Values of C_0 (Concentrations) for O_2 and H_2O species in SiO_2 at 1000°C are 5.2×10^{16} and 3.0×10^{19} respectively. Again flux of oxidant is proportional to C_0, which is almost three orders of magnitude greater for water than for oxygen. This is the reason of faster growth rate of wet oxidation than that of dry oxidation.

Though, wet oxidation has a faster growth rate but dry oxidation gives better electrical characteristics. The thermally grown oxide layer has excellent electrical insulation properties.

Factors affecting the oxide growth rate

1. Crystal orientation of Si affects the oxide growth rate, because reaction rate constant depends on crystal structure of Si surface.

2. Any unintentional moisture accelerates the dry oxidation rate.

3.High concentrations of sodium enhance the oxidation rate by changing the bond structure in the oxide (enhancing the diffusion and concentration of oxygen molecules in the oxide).

4. The common dopant elements of group III and V can enhance the oxidation behavior.

5. Typical 1-5% addition of dry HCl with dry oxygen increases the oxidation rate.

6. Oxide growth rate is enhanced with temperature and Pressure.

Thin Oxide

What is the importance of thin oxide?

For submicron range MOS VLSI, it is vital to grow oxide on crystalline Si with thickness of 50 to 200 Angstrom.

For example, dielectric material for MOS devices can be thin thermal oxide of composite structure of thermal oxide and silicon nitride. This dielectric is an active component of storage capacitor in DRAM, and its thickness determines the amount of charge that can be stored.

$$C = \frac{\varepsilon A}{d}$$

$$Q = CV$$

Properties of Oxide

The typical properties of oxides include:

- Susceptibility
- Leakage current
- Stress
- Dielectric breakdown
- Refractive index
- Oxide composition
- Density
- Etch rate
- Oxide charges
- Masking property of SiO2

Oxide charges

The Si-SiO_2 interface contains a transition region between crystalline Si and amorphous SiO_2. A charge trapped in the interface can induce a charge of opposite polarity in the underlying silicon, thereby affecting the ideal characteristics of MOS devices.

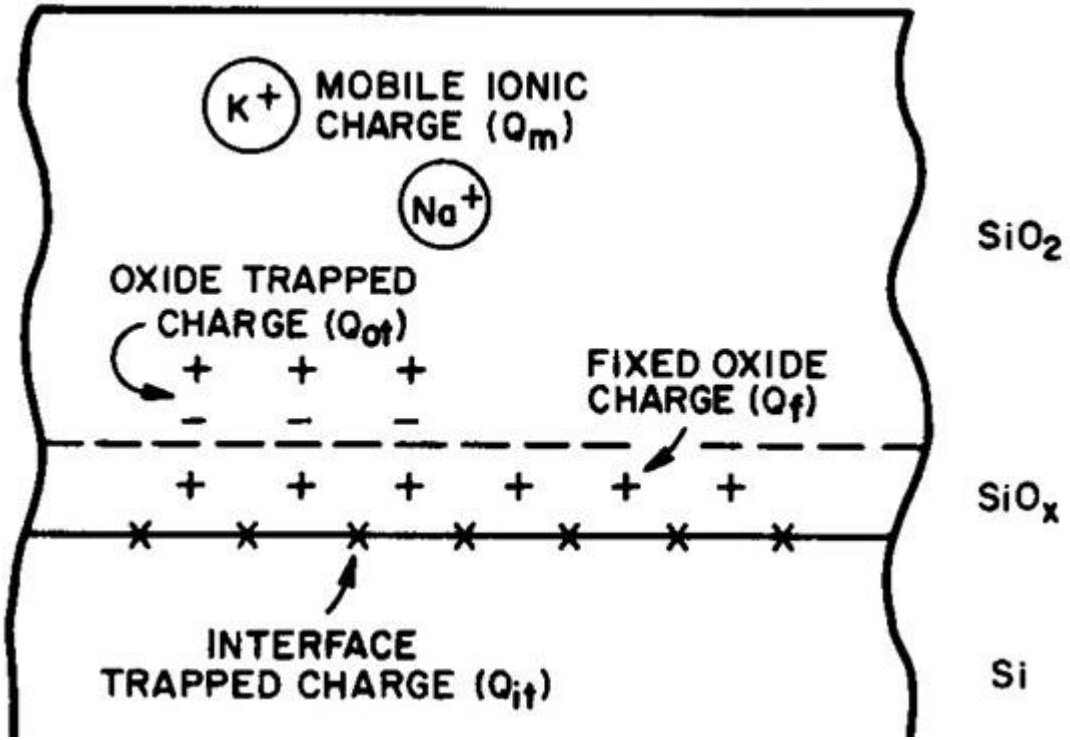


Fig.- Charges in thermally oxidized Silicon

Oxide charges…

Interface trapped charges (Q_{it}):

Results from structural defects related to oxidation process, metallic impurities, bond-breaking processes. A low-temperature hydrogen annealing (450 C) effectively neutralizes Q_{it}.

Fixed oxide charge (Q_f):

Q_f usually positive located approximately 30 angstrom of Si-SiO_2 interface and cannot be discharged. Rapid cooling after oxidation results low value of Q_f; inert ambient annealing also results low Q_f.

Oxide trapped charges (Q_{ot}):

Q_{ot} may be positive or negative, due to holes or electrons trapped in the bulk SiO_2. This charges may result ionizing radiation, avalanche injection or high currents in the oxide. It may be annealed out by low-temperature treatment.

Mobile ionic charge (Q_m):

Q_m is attributed to alkali ions such as sodium, potassium and lithium in oxide. Common techniques used to minimize this charge include cleaning the furnace tube in a chlorine ambient, and using masking layers such as silicon nitride.

Plasma Oxidation

Plasma oxidation is a low-temperature vacuum process, usually carried out at pure oxygen discharge.

Plasma oxidation offers:

- The possibility of growing high-quality oxides at lower temperatures (<600 C).
- High growth rate (1 μm/h).
- Reduced movement of previous diffusions and suppression of defect formation.
- Reasonably thick oxides (1 μm).

Plasma Oxidation…

- Plasma is generated either by a RF discharge or a DC electron source.
- Wafer is placed in uniform density plasma. Wafer is positively biased below plasma potential.
- The growth rate of oxide increases with increasing substrate temperature, plasma density and substrate dopant concentration.
- The quality of oxide grown by this method can be as good as that of thermally grown SiO2.

Growth of SiO_2 by CVD method

SiO_2 layer can be deposited also in CVD method (vaporized oxygen and silicon flow is maintained over a substrate).

Merits and Demerits of CVD:

- Oxide deposits at faster rate and at low temperature (below 500°C).
- But the quality is not as good as that obtained by thermal growth.