



EE669: VLSI Technology

Thermal Oxidation of Silicon

Anil Kottantharayil,
Associate Professor,
Department of EE, IIT Bombay



Thermal oxidation

Success of Si in electronic applications is attributed to the availability of the native oxide SiO_2 , which:

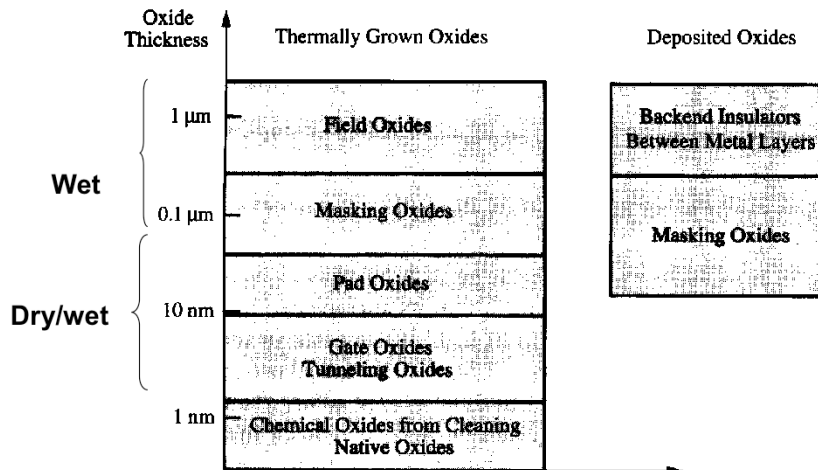
- Can be easily grown thermally on Si
- Has an excellent interface with Si: the best in terms of electrical and mechanical properties. Low density of defects and stable with time => interface passivation.
- Diffusion of most of the dopants in SiO_2 is slower than in Silicon. Stopping power of SiO_2 is better than Silicon => excellent mask for diffusion and ion implantation
- Si can be etched selective to the oxide and vice versa => etch mask
- SiO_2 is resistant to most chemicals used in fabrication but can be easily etched using HF

Some comparison may be in order to drive home the points. Si is not the best of semiconductor materials. It has an indirect band gap, the electron and hole mobilities are small compared to many other materials, the absorption coefficient is small etc. GaAs in comparison is a direct bandgap material, has better carrier mobility and has better optical absorption coefficient. However Si has a native oxide which is very robust in terms of its interface with Si, thermal stability, masking properties etc. This has made more than compensated for the short comings of silicon to the extent that Si is the choice of semiconductor for large scale electronic applications.

Ge, a close cousin of Si offers better carrier mobilities. However GeO_2 dissolves in water.



Applications of oxides in Silicon device technology



J. D. Plummer, M. D. Deal, P. G. Griffin, Silicon VLSI Technology, Pearson Education, 2001

2011 Monsoon

EE669: Thermal Oxidation of Silicon; Anil Kottantharayil

130

Typical applications of SiO₂ in silicon process technology and the thickness ranges are shown on this slide. There are broadly two methods to form SiO₂ films on Si wafers (1) thermal growth including by reaction of oxidizing species like water and H₂O₂ (2) deposition of films by chemical or physical vapor deposition methods. In the former case the Si from the wafer reacts and is oxidized to SiO₂. In the later case, the silicon from the wafer does not involve in the reaction.

Since the thermal oxide grows by oxidation of the silicon, it is generally not possible to grow thermal oxide once the wafer is covered with other materials (exception of course is poly-Si and a-Si). In such cases SiO₂ may be deposited by the CVD or PVD methods. CVD and PVD oxides can be deposited on substrates other than Si also.

Thermal oxidation can be carried out by exposing the wafer to an oxidizing ambient at high temperature, ~ 1000°C. The oxidizing ambient can be O₂, nitric oxide (NO) and nitrous oxide (N₂O). Oxidation process using these reactants is called dry oxidation. On the contrary oxidation carried out in H₂O ambient is called wet oxidation. H₂O can be in-situ generated by reacting H₂ and O₂ (pyrogenic oxidation) or vaporizing high purity H₂O using an inert carrier gas.

Thick oxides in the range of 200nm to ~ 1μm are used in electronic devices for isolation of devices and interconnects. Thermal oxides are used in isolation schemes like LOCOS (LOCAl Oxidation of Silicon – an isolation scheme used in older technologies) and box isolation (a laboratory scheme, not for products). When Si wafer surface is not exposed, isolation oxides have to be deposited by CVD or PVD.

Masking oxides are used mainly for masking select areas of the wafer surface from implants and diffusion. Implants may be also masked by photo resist. However diffusion is a high temperature process and hence requires a thermally stable material for masking. It should be noted that implantation into silicon is usually carried out through a "screening" oxides, especially for high dose and high energy implants. Screening oxide prevents or reduces damage to Si surface during implantation.

Oxide films are also used as masks for wet and dry etch of Silicon and other materials used in Silicon processing due to high selectivity of certain etch processes. For example, KOH would etch (100) silicon ~ 800 times faster than SiO₂. Hot phosphoric acid (~180°C) can be used to etch Si₃N₄ selective to SiO₂. SiO₂ layer can be used as etch stop layer in several etch processes. For example, typical reactive ion etching processes for Si₃N₄ are known to etch Si with the same rate or at a faster rate. A SiO₂ inserted between the nitride and Si can act as a good etch stop layer.

Pad oxides are typically used to isolate Si from noncompatible materials. For example, Si₃N₄ can introduce stress on the Si surface. A "pad" oxide on Si surface is usually grown or deposited prior to depositing nitride.

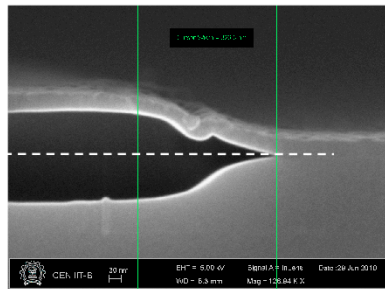
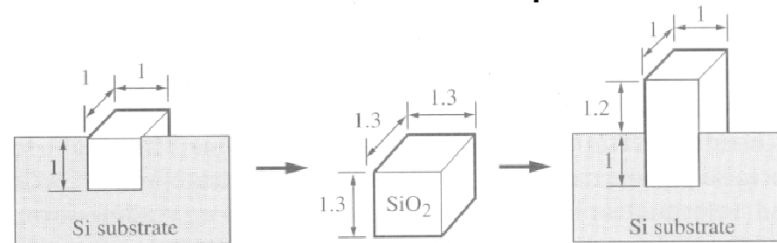
The applications discussed so far may be also satisfied with deposited oxides. However it should be noted that the thermal oxide is denser than deposited oxides and hence its etch rate is lower than deposited oxides. The electrical quality of thermal oxides like dielectric breakdown field, is better for thermal oxide.

The later criteria is very important for SiO₂ used for gate dielectric applications in MOS devices. Thermal oxide has excellent interface with Si and has high dielectric breakdown and low density of defects; both physical – for example pin holes and electrical – for example traps and fixed charge. Tunneling oxide used in flash memories is a typical example. Gate dielectric for CMOS logic applications have been scaled down to below 2 nm. For these range of thickness, nitridation of oxide is carried out to increase the physical thickness and relative permittivity => keeping the capacitance per unit area constant.

Recently high dielectric constant materials were introduced into CMOS technology. However the interface between these materials and Si is poor and contains large density of defects. More over during the deposition of such materials, a thin SiO₂ layer is formed between Si and the high-k material. Since the interface of Si and the gate dielectric decides several parameters of MOS devices, it is desirable to grow such films in a controlled manner. It should be noted that pristine Si surface would oxidize upon exposure to air or water to form 1-2nm of SiO₂ on the surface. The interfacial oxides in high-k technology are sometimes formed by chemical methods.



Thermal oxidation: volume expansion



J. D. Plummer, M. D. Deal, P. G. Griffin, *Silicon VLSI Technology*, Pearson Education, 2001

CMOS team, CEN, IITB

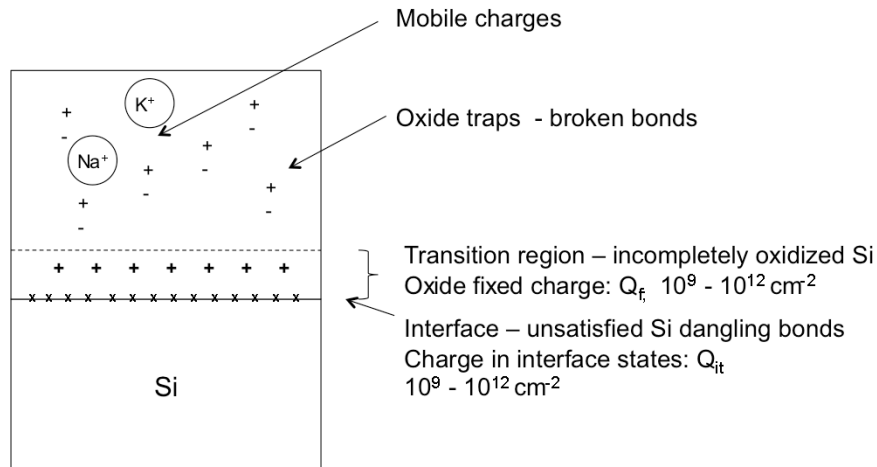
2011 Monsoon

EE669: Thermal Oxidation of Silicon; Anil Kottantharayil

131



Charges and traps in thermal oxides



J. D. Plummer, M. D. Deal, P. G. Griffin, Silicon VLSI Technology, Pearson Education, 2001

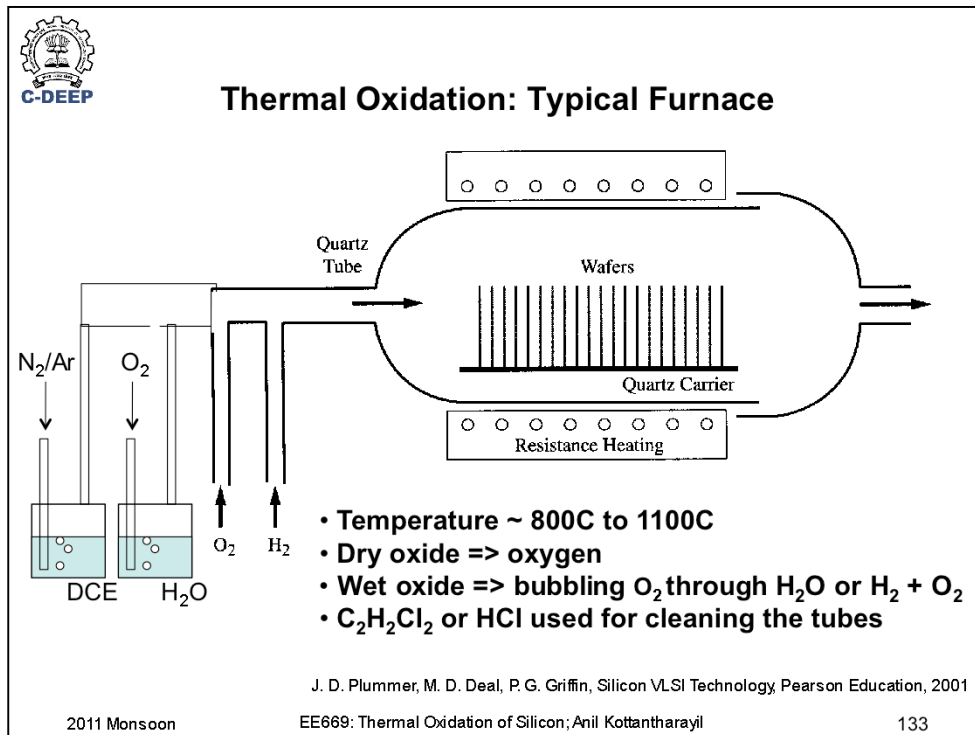
2011 Monsoon

EE669: Thermal Oxidation of Silicon; Anil Kottantharayil

132

We referred to this topic on the previous slide. The charges and traps within the the dielectric can be classified into 4 types as shown in the figure.

Except for the mobile charges, all other types can also be created during the operation of the device and can cause eventual device failure. However the density and number of these traps should be well controlled even at the fabrication stage. Unpredictable charge densities can cause large variations in characteristics of devices.



A typical furnace arrangement for furnace oxidation is shown. The furnace used for oxidation of silicon is made of a fused quartz cylindrical tube. The tube is heated by resistance heaters. Heating using lamps is another popular way to heat the furnace. Process gases are fed into the furnace through one end.

For dry oxidation only oxygen is required.

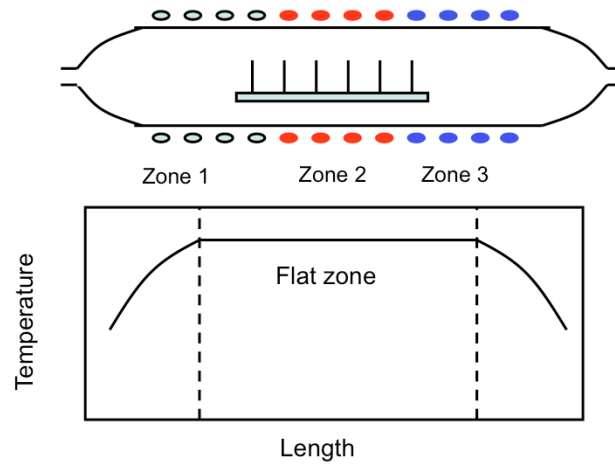
Water vapor can be generated for wet oxidation in two ways => bubbling & in-situ generation.

In-situ is preferred today due to the clean water vapor you get. But on the other hand there is a potential safety issue now.

The DCE and HCl are required for cleaning the tubes of various contaminants.



Atmospheric furnace for oxidation





Atmospheric furnace for oxidation (2)

- The flow in the furnace should be such a way as to avoid turbulence and have a laminar gas flow (smooth layer by layer flow) .
- Turbulence leads to nonuniform film growth
- Flow regime determined by Reynolds number

$$R_e = \frac{\rho u d}{\mu}$$

ρ : Volume density of the gas in kg/m³

u : Mean fluid velocity in m/sec

d : Tube diameter in m

μ : Viscosity in Pa. sec

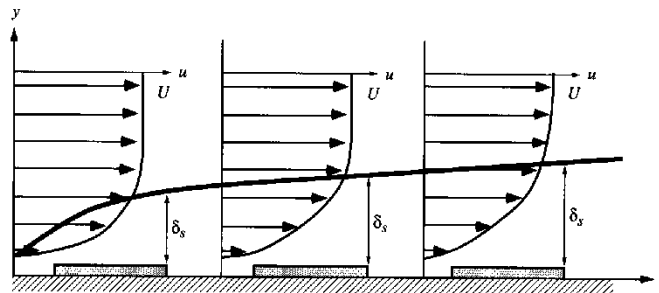
$R_e > 2000$: turbulent flow

$R_e < 1000$: laminar flow

J. D. Plummer, M. D. Deal, P. G. Griffin, Silicon VLSI Technology, Pearson Education, 2001



Boundary Layer in Horizontal Furnaces



$$\delta_s(x) = \left(\frac{\mu x}{\rho U} \right)^{1/2}$$

Where all the parameters are as defined on the previous slide. U is the velocity of the gas in the main flow above the boundary layer.

J. D. Plummer, M. D. Deal, P. G. Griffin, Silicon VLSI Technology, Pearson Education, 2001

2011 Monsoon

EE669: Thermal Oxidation of Silicon; Anil Kottantharayil

136

The development of a boundary layer can be thought of as the balance between two forces. The gas has a velocity along “ x ”. There is a retarding force acting on the gas due to friction. The friction is caused by the walls of the reactor and propagates through the viscosity of the gas. The boundary layer is a consequence of a sort of balance of these two counter acting forces.

The friction due to the walls of the reactor or wafer holders is fixed for a given reactor design and the viscosity of the gas is fixed for a given gas mixture and temperature. Consequently, a higher gas velocity in the main stream would result in a thinner boundary layer.

A good analogy may be wind. As you go up in the air, it is always windy. However on the ground, we may not feel the wind, unless the overall wind speed increases.

Explanation courtesy, Prof. Amit Aggrawal, Mechanical Engineering, IIT Bombay.



Boundary Layer in Thermal Oxidation

- In laminar flow regime, the reactants are transported from the main flow by diffusion to the wafer surface through the boundary layer
- Reaction byproducts, if any are transported from the wafer surface to the flow through the boundary layer. For example, the H_2 released in the reaction between Si and H_2O .
- Typically in thermal oxidation, the diffusion of the reactants through the boundary layer is assumed to be fast and hence would not limit the oxidation rate.



Atmospheric furnace for oxidation (3)

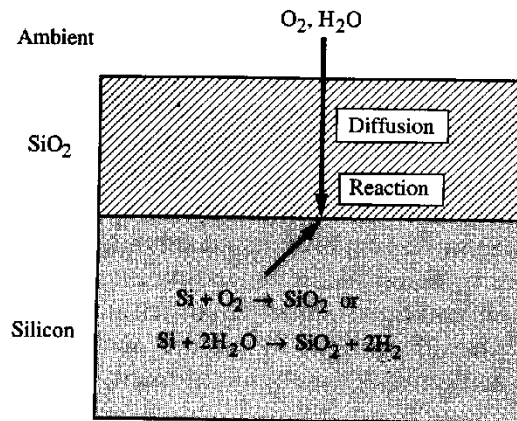
A typical furnace oxidation recipe:

Step	Name	Time	Temp	Gas
1.	load	11 min.	775 C	3% O ₂ /N ₂
2.	stab	10 min.	775 C	3% O ₂ /N ₂
3.	ramp up	10 C/min.	To 1000 C	3% O ₂ /N ₂
4.	stab	20 min.	1000 C	3% O ₂ /N ₂
5.	dryox	to XX nm	1000 C	O ₂
6.	purge	5 min.	1000 C	N ₂
7.	ramp down	4 C/min.	to 775 C	N ₂
8.	unload	16 min.	775 C	N ₂

A significant amount of time is taken for ramp up, stabilization and ramp down => high and extra thermal budget.



Thermal oxidation – model



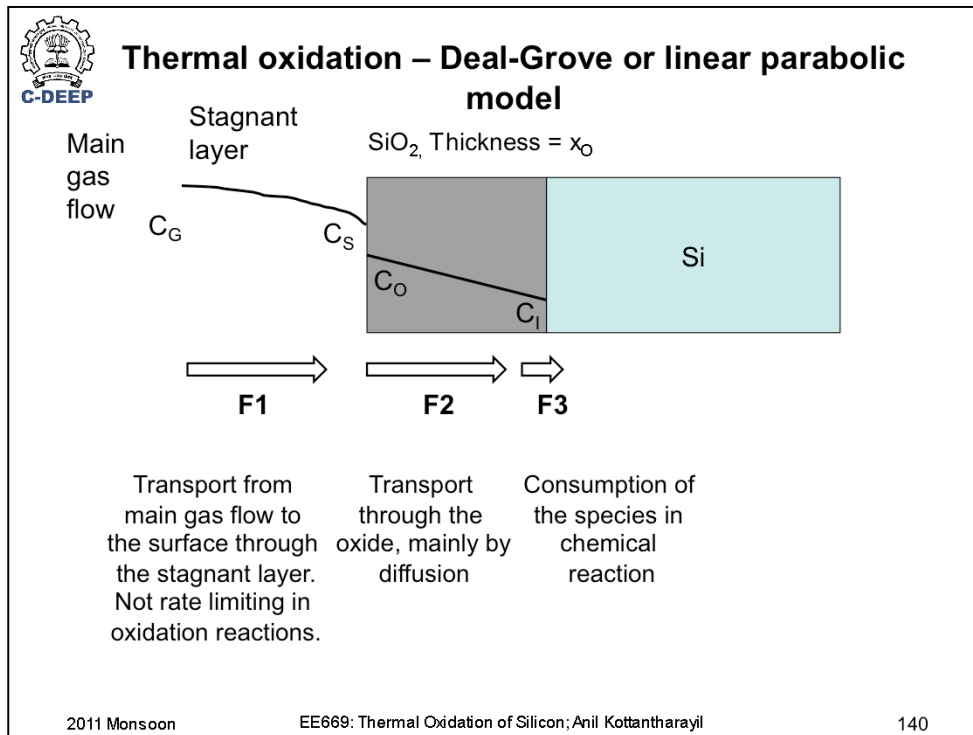
2011 Monsoon

EE669: Thermal Oxidation of Silicon; Anil Kottantharayil

139

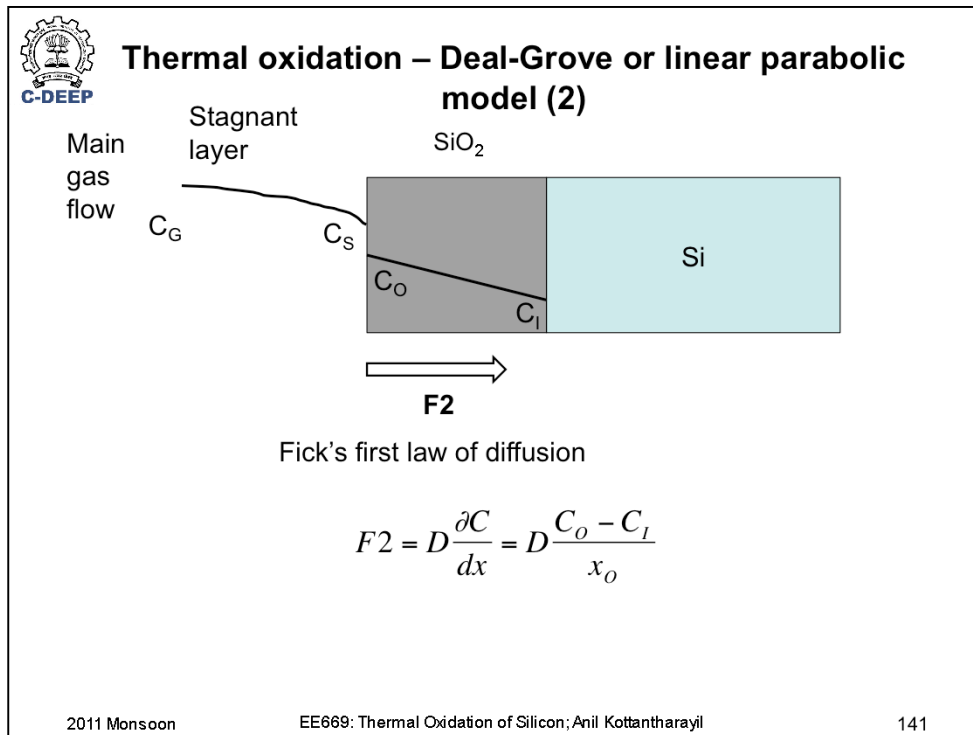
In a reaction between a solid and another material of any phase where the reaction by product is another solid, there are two ways in which the reaction can happen. The reaction would have produced a product through which either of the reactants have to diffuse. For example in the reaction between Si and oxygen, oxygen can diffuse through the already grown oxide and react with the Si at the Si-SiO₂ interface. The other possibility is that the silicon diffuses through the oxide and react with the oxygen at the surface.

By using isotopes of oxygen, it has been established that the oxygen is the diffusing species in the reaction between Si and O₂ at the normal oxidation temperatures used in silicon device processing.



C_G is the concentration of the oxidant species in the main gas flow in cm^{-3} , C_s is the concentration of the oxidant at the surface of the oxide and C_i is the concentration of the oxidant at the oxide – Si interface.

$F1$, $F2$ and $F3$ are the oxidant fluxes in $\text{cm}^{-2}\text{s}^{-1}$ through the stagnant layer, the oxide and at the SiO_2 -Si interface. The oxidation rate is found to be insensitive to gas flow rates for a large range of flow rates and hence $F1$ is not a rate limiting factor.

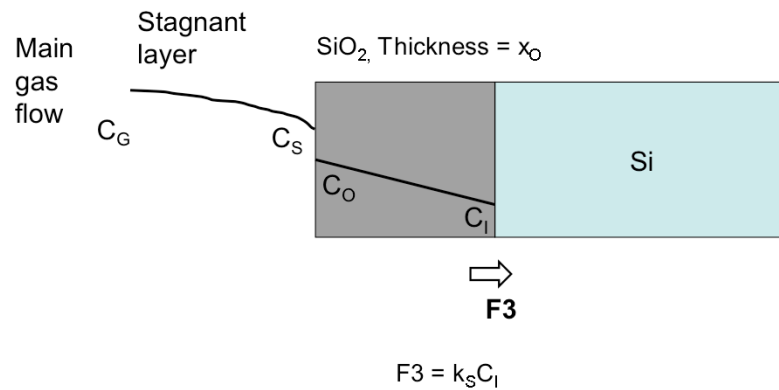


F_2 is the flux of the oxidant through the growing oxide. The oxidants are typically transported by diffusion. In the case of dry oxidation, i.e. using O_2 as oxidant, the diffusing species is molecular species. In the case of H_2O , there is less clarity on the diffusing species. It is believed that the H_2O weakens the SiO_2 structure by forming Si-OH. The oxidizing species can be OH groups or H_2O itself.

The diffusion flux can be written by considering the Fick's first law of diffusion, which states that the diffusion flux is proportional to the gradient of the concentration. If the oxidant is not lost during the diffusion in the oxide, which is a fairly good approximation, under steady state the flux through the oxide would be a constant. So the derivative can be approximated by the slope considering the concentration of the diffusing species at the two interfaces. D is the diffusion constant and has the units of $cm^2 s^{-1}$. The effective diffusivities of O_2 and H_2O are similar and $\sim 1.3 \times 10^{-8} cm^2 s^{-1}$ at 1100 C.



Thermal oxidation – Deal-Grove or linear parabolic model (3)



F3 is the rate of the oxidizing reaction at the interface between Si and SiO₂. It should be noted that the oxide formation can involve several reactions including, breaking of the oxidizing species (O₂ to 2O, for example), breaking of the Si-Si bonds, formation of the Si-O bonds. k_s captures the overall reaction rate. k_s has units of cm⁻¹ s⁻¹.



Thermal oxidation – Deal-Grove or linear parabolic model (4)

Under steady state,

$$F_2 = F_3 = F$$

$$F = D \frac{C_o - C_l}{x_o} = k_s C_l$$

$$F = \frac{C_o k_s}{1 + \frac{k_s x_o}{D}}$$

If “N” is the concentration of oxidant consumed per unit volume of the film grown

$$\frac{dx_o}{dt} = \frac{F}{N} = \frac{k_s C_o}{N(1 + \frac{k_s x_o}{D})}$$

$$N \int_{x_i}^{x_o} (1 + \frac{k_s x_o}{D}) dx_o = k_s C_o \int_0^t dt$$

2011 Monsoon

EE669: Thermal Oxidation of Silicon; Anil Kottantharayil

143

Under steady state the fluxes should all be the same and the overall reaction rate would be decided by the slowest reaction.

N for O₂ is 2.2 x 10²² per cm³ of oxide grown and H₂O is 4.4 x 10²² per cm³ of the oxide grown.

X_i is the thickness of the oxide already present on the wafer surface at the beginning of the process.



Thermal oxidation – Deal-Grove or linear parabolic model (5)

$$\frac{x_o^2 - x_i^2}{B} + \frac{x_o - x_i}{B/A} = t$$

where

$$B = \frac{2DC_o}{N}$$

$$\frac{B}{A} = \frac{C_o k_s}{N}$$

The growth equation can be rewritten as:

$$\frac{x_o^2}{B} + \frac{x_o}{B/A} = t + \tau$$

$$\tau = \frac{x_i^2 + Ax_i}{B}$$

2011 Monsoon

EE669: Thermal Oxidation of Silicon; Anil Kottantharayil

144

The growth equation is a parabolic – linear equation and hence the name. It should be noted that when x_o is small, the parabolic term can be neglected. Hence the initial growth proceeds linearly and the growth rate is essentially constrained by the reaction at the interface. B/A is called the linear rate constant.

On the other hand, when x_o is sufficiently large, the linear term can be neglected. The growth now is limited by the diffusion of the oxidant species through the already grown oxide. B is called the parabolic rate constant.

The linear parabolic model describes the oxide growth on lowly doped substrates with planar substrates for oxides thicker than 20nm.

The concentration of the oxidant at the surface of the oxide depends on the solubility in the oxide. The solubility of H_2O in oxide ($3 \times 10^{19} \text{ cm}^{-3}$ at 1100 C) is much higher than that of O_2 ($5 \times 10^{16} \text{ cm}^{-3}$ at 1100 C). As a consequence the oxidation rate is higher for wet oxidation than for dry oxidation.