CHAPTER 4: Oxidation

Oxidation of silicon is an important process in VLSI. The typical roles of SiO_2 are:

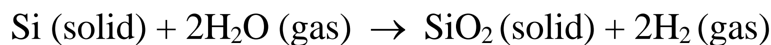
1. mask against implant or diffusion of dopant into silicon
2. surface passivation
3. device isolation
4. component in MOS structures (gate oxide)
5. electrical isolation of multi-level metallization systems

There are several techniques to form oxide layers, namely thermal oxidation, wet anodization, chemical vapor deposition, and plasma oxidation. Of the four techniques, thermal oxidation tends to yield the cleanest oxide layer with the least amount of interfacial defects.

4.1 Theory of Oxide Growth

The chemical reactions describing thermal oxidation of silicon in dry oxygen or water vapor are:

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



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

During the course of the oxidation process, oxygen or water molecules diffuse through the surface oxide into the silicon substrate, and the Si-SiO_2 interface migrates into the silicon (***Figure 4.1***). Thermal oxidation of silicon results in a random three-dimensional network of silicon dioxide constructed from tetrahedral cells. Since the volume expands, the external SiO_2 surface is not coplanar with the original silicon surface. For the growth of an oxide of thickness d, a layer of silicon equal to a thickness of $0.44d$ is consumed.

CityU 香港城市大學 City University of Hong Kong

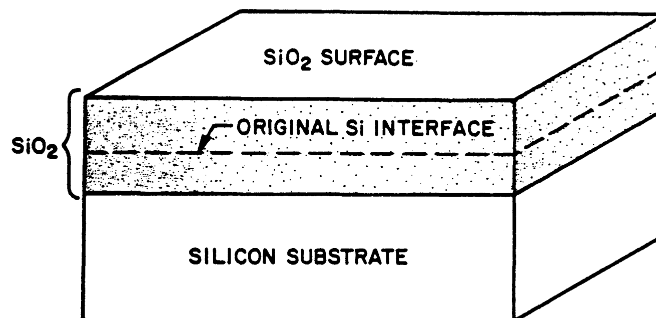


Figure 4.1: Growth of SiO_2.

A model elucidating the kinetics of oxide growth has been developed by Deal and Grove. The model, which is generally valid for temperatures between 700ºC and 1300ºC, partial pressure between 0.2 and 1.0 atmosphere, and oxide thickness between 30 nm and 2000 nm for oxygen and water ambients, is schematically illustrated in ***Figure 4.2***.

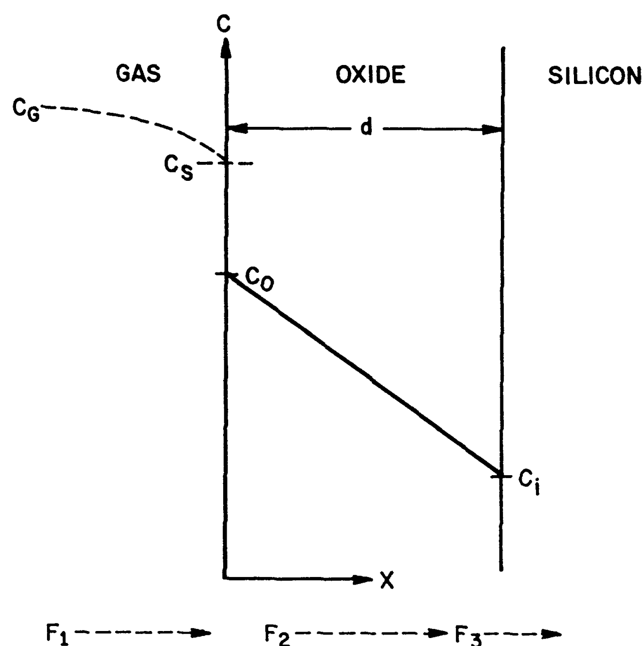


Figure 4.2: Basic model for thermal oxidation of silicon.

In steady state, the three fluxes, F_1 (flux of oxidizing species transported from the gas phase to the gas-oxide interface), F_2 (flux across the existing oxide toward the silicon substrate), and F_3 (flux reacting at the Si-SiO_2 interface) must be equal. F_1 can be approximated to be proportional to the difference in concentration of the oxidizing species in the gas phase and on the oxide surface.

$$F_1 = h_G (C_G - C_S) \qquad \textit{(Equation 4.1)}$$

where h_G is the gas-phase mass-transfer coefficient, C_G is the oxidant concentration in the gas phase, and C_S is the oxidant concentration adjacent to the oxide surface. Substituting $C = P/kT$ into *Equation 4.1*,

$$F_1 = (h_G/kT)(P_G - P_S) \qquad \textit{(Equation 4.2)}$$

Henry's Law states that, in equilibrium, the concentration of a species within a solid is proportional to the partial pressure of that species in the surrounding gas. Thus,

$$C_o = HP_S \qquad \textit{(Equation 4.3)}$$

where C_o is the equilibrium concentration of the oxidant in the oxide on the outer surface, H is the Henry's Law constant, and P_S is the partial pressure of oxidant in the gas phase adjacent to the oxide surface. Furthermore, we denote the equilibrium concentration in the oxide, that is, the concentration which would be in equilibrium with the partial pressure in the bulk of the gas P_G by the symbol C^*, and

$$C^* = HP_G \qquad \textit{(Equation 4.4)}$$

Hence,

$$C^* - C_o = H (P_G - P_S), \text{ and}$$

$$F_1 = (h_G/HkT)(C^* - C_o) = h (C^* - C_o) \qquad \textit{(Equation 4.5)}$$

where $h = h_G/HkT$ is the gas-phase mass-transfer coefficient in terms of concentration in the solid.

Oxidation is thus a non-equilibrium process with the driving force being the deviation of concentration from equilibrium. Henry's Law is valid only in the

absence of dissociation effects at the gas-oxide interface, thereby implying that the species diffusing through the oxide is molecular.

The flux of the oxidizing species across the oxide is taken to follow Fick's Law at any point d in the oxide layer. Hence,

$$F_2 = D(C_o - C_i)/d_o \qquad \textit{(Equation 4.6)}$$

where D is the diffusion coefficient, C_i is the oxidizer concentration in the oxide adjacent to the SiO_2/Si interface, and d_o is the oxide thickness.

The chemical reaction rate at the SiO_2/Si interface is assumed to be proportional to the reactant concentration. Therefore,

$$F_3 = k_S C_i \qquad \textit{(Equation 4.7)}$$

where k_S is the rate constant.

Under steady-state conditions, $F_1 = F_2 = F_3$. Thus,

$$h(C^* - C_o) = D(C_o - C_i)/d_o = k_S C_i$$

$$C_i = DC_o/(k_S d_o + D) \qquad \textit{(Equation 4.8)}$$

$$C_i = C^*/[1 + k_S/h + k_S d_o/D]\text{, and} \qquad \textit{(Equation 4.9)}$$

$$C_o = [(1 + k_S d_o/D)C^*]/[1 + k_S/h + k_S d_o/D] \qquad \textit{(Equation 4.10)}$$

When D is large, *Equation 4.8* becomes $C_i = C_o$, implying that the oxidation rate is controlled by the reaction rate constant k_S and by C_i (= C_o), that is, a reaction-controlled case. When D is very small, $h(C^* - C_o) = 0 = k_S C_i$. Therefore, $C^* = C_o$ and $C_i = 0$. The latter case is called diffusion-controlled case, as the oxidation rate depends on the supply of oxidant to the interface.

In order to calculate the oxide growth rate, we define N_1 as the number of oxidant molecules incorporated into a unit volume of the oxide layer. If oxygen is the reactant, $N_1 = 2.2 \times 10^{22}$ atoms/cm^3 because the density of SiO_2 is 2.2×10^{22} cm^{-3}. If water is used, N_1 becomes 4.4×10^{22} cm^{-3} as two H_2O molecules are incorporated into each SiO_2 molecule. The differential equation for oxide growth is given by

$$N_1\left[\frac{d(d_o)}{dt}\right]=k_sC_i=\frac{k_sC^*}{1+\frac{k_s}{h}+\frac{k_sd_o}{D}} \qquad \textit{(Equation 4.11)}$$

With an initial condition of $d_o(t = 0) = d_i$, the solution of *Equation 4.11* may be written as:

$$d_o^2 + Ad_o = B\,(t + \tau) \qquad \textit{(Equation 4.12)}$$

where $A \equiv 2D\,[1/k_S + 1/h]$ *(Equation 4.13)*

$$B \equiv 2DC^* / N_1 \qquad \textit{(Equation 4.14)}$$

$$\tau \equiv (d_i^2 + Ad_i) / B \qquad \textit{(Equation 4.15)}$$

The quantity τ represents a shift in the time coordinate to account for the presence of the initial oxide layer d_i. Solving *Equation 4.12* for d_o as a function of time gives

$$\frac{d_o}{A/2}=\left[1+\frac{t+\tau}{A^2/4B}\right]^{1/2}-1 \qquad \textit{(Equation 4.16)}$$

For long oxidation times, i.e., $t >> \tau$ and $t >> A^2/4B$, $d_o^2 \cong Bt$. B is therefore called the parabolic rate constant. For short times, i.e., $(t + \tau) << A^2/4B$, $d_o \cong [B/A](t + \tau)$, and B/A is referred to as the linear rate constant.

Example 4.1

In wet oxidation of silicon at 950°C, the following data are obtained:

t (hour)	0.11	0.30	0.40	0.50	0.60
d_o (oxide thickness in μm)	0.041	0.100	0.128	0.153	0.177

Show how to graphically determine the linear and parabolic rate constants from these experimental data. Assume that $\tau = 0$ for wet oxidation.

Solution

$d_o^2 + Ad_o = B(t + \tau)$ and for wet oxidation, $\tau = 0$

Rearranging the equation, $d_o = \frac{Bt}{d_o} - A$

Thus, a plot of d_o versus t/d_o will give B as the slope and A as the intercept.

t (hour)	0.11	0.30	0.40	0.50	0.60
d_o (oxide thickness in μm)	0.041	0.100	0.128	0.153	0.177

d_o (μm)	0. 041	0.100	0.128	0.153	0.177
t/d_o (h/μm)	2.683	3.000	3.125	3.268	3.390

From the plot, the slope of the line yields B = 0.2 μm^2/h

The intercept of the line yields A = 0.50 μm.

4.2 Experimental Fits

Rate constants are usually obtained graphically based on experimental results. ***Table 4.1*** and ***Table 4.2*** list the experimental rate constants for wet and dry oxidation of silicon, respectively. The Deal-Grove model provides excellent agreement with experimental data. The only exception is for SiO_2 films less than about 300Å thick grown in dry oxygen. In this case, an anomalously high oxidation rate is observed with respect to the model. However, a proposed modification to the Deal-Grove model, assuming that although diffusion through the oxide is still by molecular oxygen or water, the oxidation of silicon occurs by the reaction of a small concentration of atomic oxygen dissociated near the Si-SiO_2 interface, provides an excellent fit to the experimental data.

Table 4.1: Rate constants for wet oxidation of silicon.

Oxidation temperature (°C)	A (μm)	Parabolic rate constant B (μm^2/h)	Linear rate constant B/A (μm/h)	τ (h)
1200	0.05	0.720	14.40	0
1100	0.11	0.510	4.64	0
1000	0.226	0.287	1.27	0
920	0.50	0.203	0.406	0

Table 4.2: Rate constants for dry oxidation of silicon.

Oxidation temperature (°C)	A (μm)	Parabolic rate constant B (μm^2/h)	Linear rate constant B/A (μm/h)	τ (h)
1200	0.040	0.045	1.12	0.027
1100	0.090	0.027	0.30	0.076
1000	0.165	0.0117	0.071	0.37
920	0.235	0.0049	0.0208	1.40
800	0.370	0.0011	0.0030	9.0

4.3 Orientation Dependence

The rate of oxidation depends on the availability of reaction sites on the silicon substrates. Hence, as the surface areal density of atoms is dependent on crystal orientation, oxidation rates are expected to be orientation dependent. Oxidation on the <111> crystal plane has a higher rate because there are a higher number of surface atoms, i.e. reaction sites or chemical bonds, when compared to a <100> plane. ***Table 4.3*** depicts some of the properties of different silicon crystal planes, whereas ***Table 4.4*** and ***Figure 4.3*** compare the oxidation rates obtained from the <100> and <111> planes. Note that since the parabolic rate constant, *B*, is diffusion limited, it has very little dependence on the crystal orientation.

Table 4.3: Calculated properties of silicon crystal planes.

Orien-tation	Area of unit cell (cm^2)	Si atoms in area	Si bonds in area	Bonds available	Bonds (10^{14} cm^{-2})	Available bonds, N (10^{14} cm^{-2})	N relative to <110>
<110>	$\sqrt{2}a^2$	4	8	4	19.18	9.59	1.000
<111>	$1/2\sqrt{3}a^2$	2	4	3	15.68	11.76	1.227
<100>	a^2	2	4	2	13.55	6.77	0.707

CityU
香港城市大學
City University of Hong Kong

Table 4.4: Rate constants for silicon oxidation in H_2O (640 Torr).

Oxidation temperature (°C)	Orien- tation	*A* (μm)	Parabolic rate constant *B* (μm²/h)	Linear rate constant *B*/*A* (μm/h)	*B*/*A* ratio <111>/<100>
900	<100>	0.95	0.143	0.150	
	<111>	0.60	0.151	0.252	1.68
950	<100>	0.74	0.231	0.311	
	<111>	0.44	0.231	0.524	1.68
1000	<100>	0.48	0.314	0.664	
	<111>	0.27	0.314	1.163	1.75
1050	<100>	0.295	0.413	1.400	
	<111>	0.18	0.415	2.307	1.65
1100	<100>	0.175	0.521	2.977	
	<111>	0.105	0.517	4.926	1.65
					Average 1.68

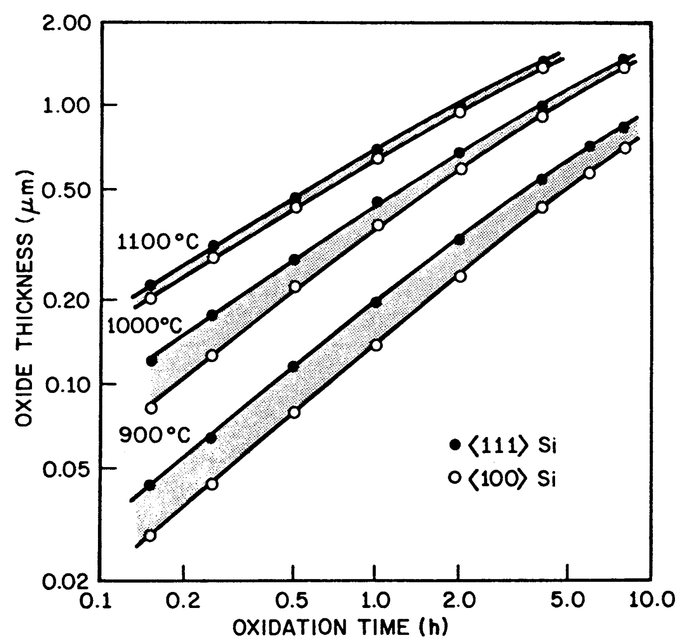


Figure 4.3: Oxide thickness versus oxidation time for silicon in H_2O at 640 Torr.

4.4 Effects of Impurities

As aforementioned, the wet oxidation rate is substantially higher than that of dry oxidation. Hence, any unintentional moisture in the furnace accelerates the oxidation process. For instance, at 800ºC, a 30 nm oxide film is grown in a period of 700 minutes if the ambient water content is less than 1 ppm, compared to an oxide thickness of 37 nm if the moisture content is 25 ppm.

Doping impurities are redistributed at the growing Si-SiO_2 interface. If the dopant segregates into the oxide (e. g. boron), the bond structure in the silica weakens, thereby permitting an enhanced incorporation and diffusivity of the oxidizing species through the oxide, resulting in a larger oxidation rate, as indicated in ***Figure 4.4***. The oxidation process in this case is diffusion control predominant. However, impurities that segregate into the oxide but then diffuse rapidly through it, such as gallium, indium, and aluminum, have no effect on the oxidation kinetics. For oxidation of phosphorus-doped silicon, a concentration dependence is observed only at lower temperature, where the surface reaction becomes important. This dependence may be the result of phosphorus being segregated into the silicon (***Figure 4.5***).

Oxidation rates are also influenced by adventitious impurities. High concentrations of sodium influence the oxidation rate by changing the bond structure in the oxide, thereby enhancing the diffusion and concentration of the oxidizing species in the oxide. Certain halogen species are intentionally introduced into the oxidation ambient to improve both the oxide and the underlying silicon properties. For instance, chlorine is instrumental in converting certain impurities in the silicon to volatile chlorides, resulting in a reduction in oxidation-induced stacking faults.

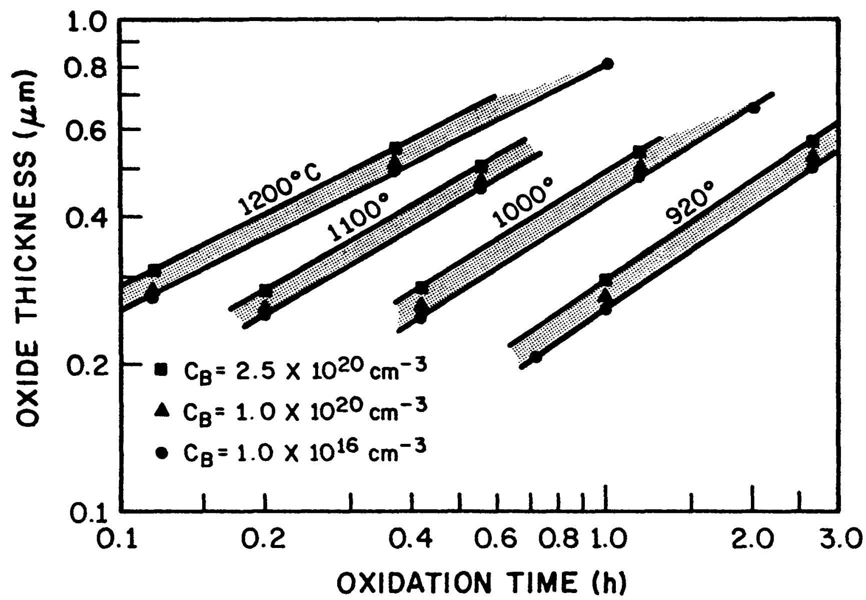


Figure 4.4: Oxidation of boron-doped silicon in wet oxygen as a function of temperature and boron concentration.

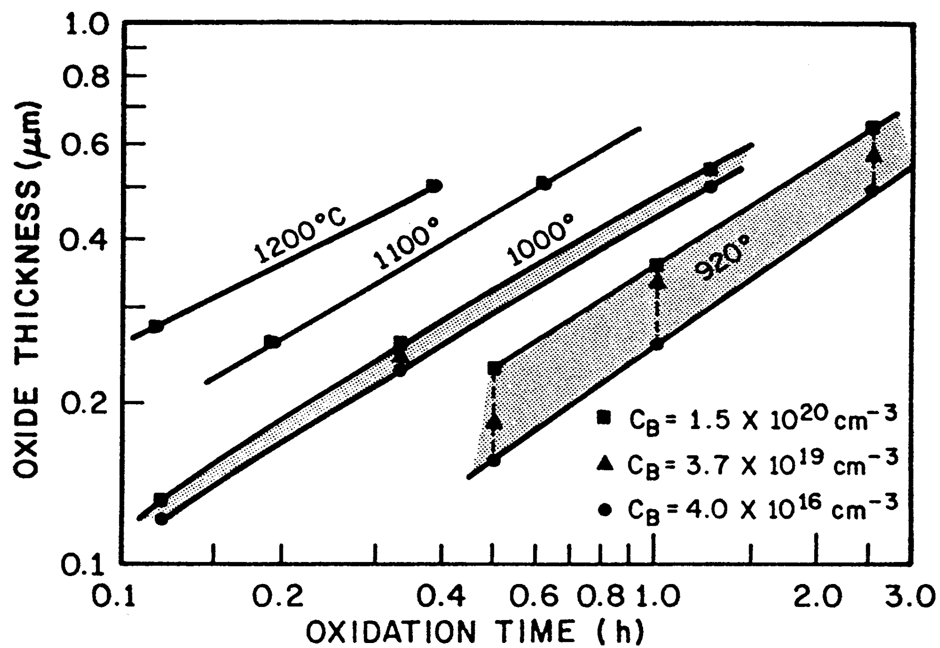


Figure 4.5: Oxidation of phosphorus-doped silicon in wet oxygen as a function of temperature and phosphorus concentration.

4.5 High Pressure Oxidation, Plasma Oxidation, and Rapid Thermal Oxidation

As stated in *Equation 4.14*, the parabolic rate constant, B, is proportional to C^* which in turn is proportional to the partial pressure of the oxidizing species in the gas phase. Oxidation in high pressure thus produces a substantial acceleration in the growth rate, as illustrated in ***Figure 4.6*** and ***Figure 4.7***. Thermal oxide layers can therefore be grown at low temperature in run times comparable to typical high temperature in order to reduce dopant diffusion and suppress oxidation induced defects.

Anodic plasma oxidation has all the advantages associated with the high-pressure technique and also offers the possibility of growing high-quality oxides at even lower temperatures. Plasma oxidation is a low-pressure process usually carried out in a pure oxygen discharge. The plasma is sustained either by a high-frequency or DC discharge. Placing the wafer in the uniform density region of the plasma and biasing it slightly negatively against the plasma potential allows it to collect active charged oxygen species. The oxidation rate typically increases with higher substrate temperature, plasma density, and substrate dopant concentration.

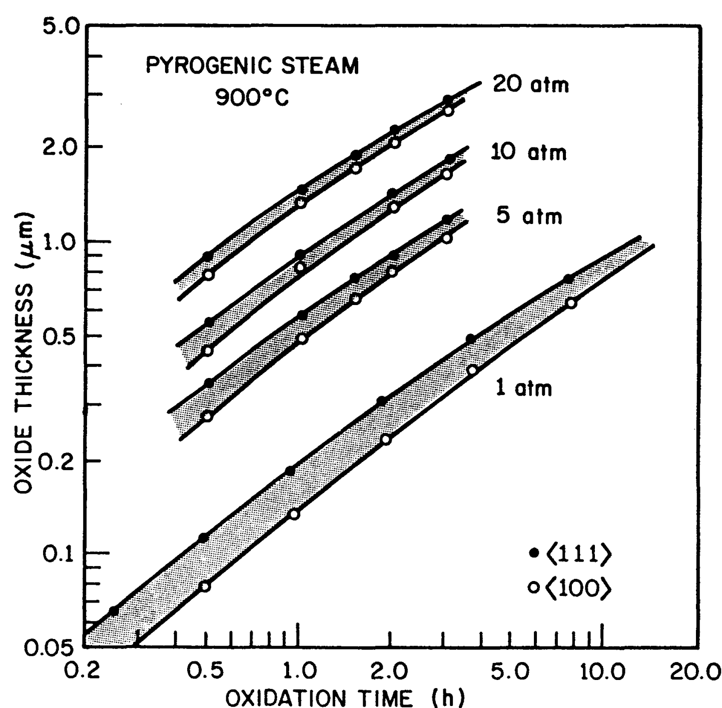


Figure 4.6: Oxide thickness versus oxidation time for pyrogenic steam at 900ºC for <100> and <111> silicon and pressures up to 20 atmospheres.

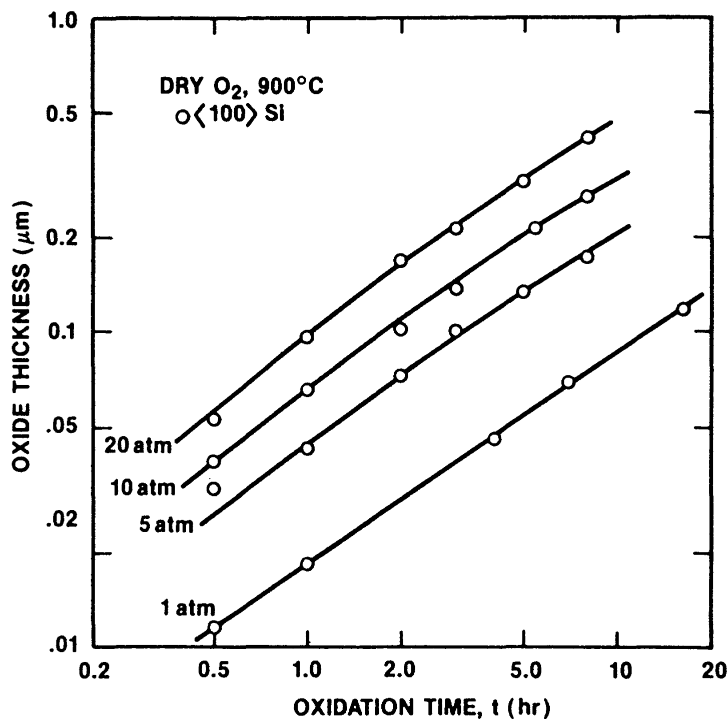


Figure 4.7: Oxide thickness versus oxidation time for <100> silicon oxidized in dry O_2 at 900°C and 1, 5, 10, and 20 atmospheres.

Rapid thermal oxidation (RTO) is increasingly used in the growth of thin, high-quality dielectric layers. The primary issues that differentiate RTO from conventional thermal oxidation are the more complex chamber design, radiation source, as well as temperature monitoring. From the point of view of oxide-growth kinetics, RTO may be influenced by both thermally activated processes and a non-thermal, photon-induced process involving monatomic O atoms generated by UV and creating a parallel oxidation reaction that dominates at lower temperature.

RTO growth kinetics exhibit activation energies differing from those measured in conventionally grown oxides. In the initial stage (on the order of 20 seconds), the RTO growth rate is linear followed by nonlinear growth. The duration of the linear region is hardware dependent, particular the heating source.

4.6 Oxide Properties

A silicon dioxide layer can provide a selective mask against the diffusion of dopant atoms at elevated temperature, a very useful property in IC processing. For it to work, the dopant diffusion rate in the oxide must be slow with respect to that in silicon, so that the dopant does not diffuse through the oxide in the masked region into the silicon. The masking oxide thickness must also be large enough to prevent it from reaching the silicon substrate. ***Table 4.5*** lists the diffusion constants of the common dopants in silicon. The often used n-type impurities as well as boron have very small diffusion coefficients in oxide and are compatible with oxide masking. However, this is not true for gallium, indium, and aluminum.

Table 4.5: Diffusion constants in SiO_2.

Dopants	Diffusion constants at 1100^oC (cm^2/s)
B	3.4×10^{-17} to 2.0×10^{-14}
Ga	5.3×10^{-11}
P	2.9×10^{-16} to 2.0×10^{-13}
As	1.2×10^{-16} to 3.5×10^{-15}
Sb	9.9×10^{-17}

Various charges and traps exist in thermally grown oxide films. If a charge is present close to the Si/SiO_2 interface, it can induce a charge of the opposite polarity in the underlying silicon, thereby affecting the ideal characteristics of the device, such as the threshold voltage of a MOS capacitor. ***Figure 4.8*** shows the general types of charges that can exist in a thermal oxide film. Interface-trapped charges (Q_{it}) can interact with the underlying silicon. These charges originate from structural defects related to the oxidation process, metallic impurities, and bond-breaking processes. A low temperature hydrogen anneal at 450^oC effectively neutralizes most interface-trapped charges. Values of $10^{10}/cm^2$-eV and lower have been observed. Fixed oxide charges (Q_f) are located in the oxide within approximately 3 nm of the SiO_2 / Si interface. Q_f cannot be charged or

discharged easily. Its density ranges from 10^{10} cm^{-2} to 10^{12} cm^{-2}, depending on the oxidation and annealing conditions and orientation (Q_f <111> is larger than Q_f <100>). Mobile ion charges (Q_m) are attributed to alkali ions such as Na, K, and Li, as well as negative ions and heavy metals. Densities range from 10^{10} cm^{-2} to 10^{12} cm^{-2} or higher and are related to the processing materials, chemicals, ambient, or handling. Common techniques employed to minimize Q_m include cleaning the furnace tube in a chlorine ambient, gettering with phosphosilicate glass (PSG), and using masking layers such as silicon nitride. Oxide-trapped charges (Q_{ot}) may be positive or negative, due to holes or electrons being trapped in the bulk of the oxide. Densities range from $< 10^9$ cm^{-2} to 10^{13} cm^{-2}. They can be annealed out by low-temperature treatment.

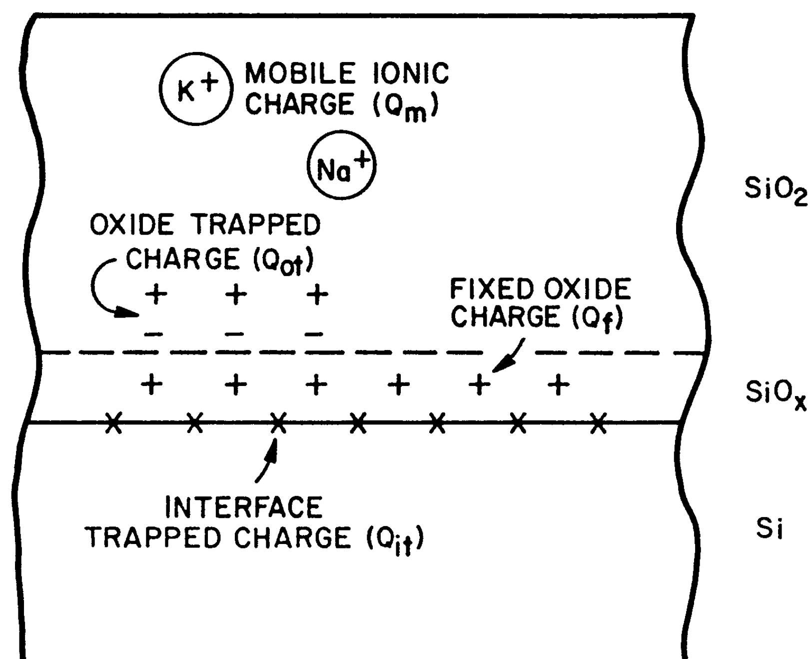


Figure 4.8: Charges in thermally oxidized silicon.

4.7 Dopant Redistribution at the Interface

During thermal oxidation, the interface advances into the silicon substrate, and doping impurities will redistribute at the interface until its chemical potential is the same on each side of the interface. The ratio of the equilibrium concentration of the impurity in silicon to that in SiO_2 at the interface is called the equilibrium segregation coefficient. Two additional factors that influence the redistribution process are the diffusivity of the impurity in the oxide (if large, the dopant can diffuse through the oxide rapidly, thereby affecting the profile near the Si - SiO_2 interface) and the rate at which the interface moves with respect to the diffusion rate.