Biomedical Ultrasound: Fundamentals of Imaging and Micromachined Transducers

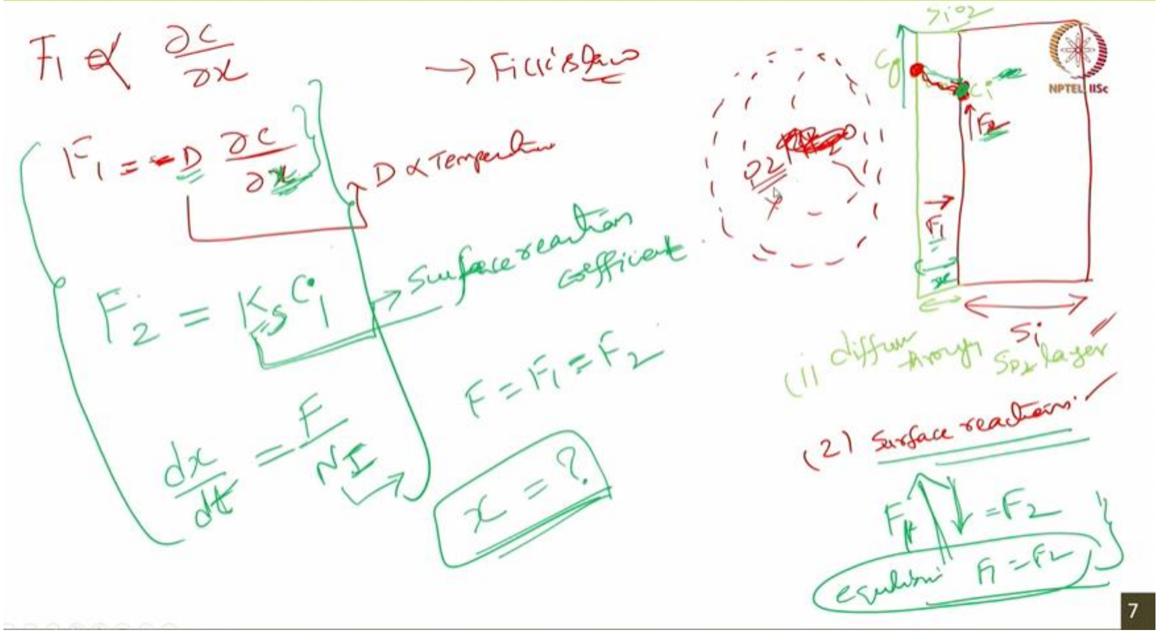
Prof. Karla P. Mercado-Shekhar, Prof. Himanshu Shekhar, Prof. Hardik Jeetendra Pandya

IIT Gandhinagar, IISc Bangalore

Lecture: 18

Thermal Oxidation II

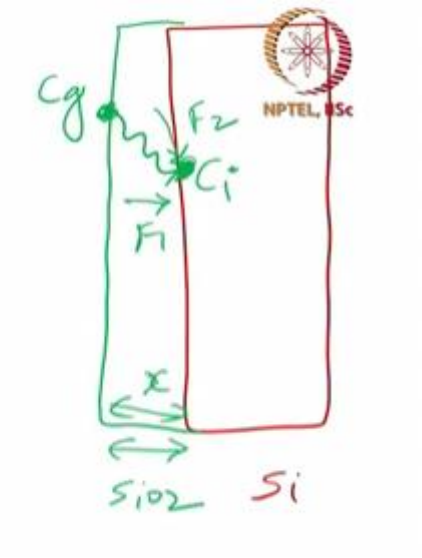
Hello, welcome back to the course again on Biomedical Ultrasound: Fundamentals of Imaging and Micromachined Transducer. This is Srinivas Bhaskara, TA of this course. We were discussing about thermal oxidation and in the previous lecture, we have covered multiple topics. If you look at the slide, we discussed about dry oxidation, wet oxidation, and its different aspects. if you get X thickness of silicon dioxide how much silicon is consumed, all those things we have seen and then we have also established the basics that are required for understanding Deal Grove model. This is where we stopped previously. So, if you look at it, there are definitely two flux and there is a one more flux in the ambience. We are neglecting this because we are thinking that there is infinite number of molecules are available for that.



So now let us try to model one by one. So I have flux 1 which is a diffusion flux

$$F_1 = -D\frac{\partial c}{\partial x}$$

if I recreate the picture you have silicon, with the oxide thickness of x (SiO_2), is grown on it. So the molecules has to diffuse to the interface and then react. So, there is flux 1 in the oxide layer and at the interface there is flux 2 which is getting oxidized. Now let us say the oxide concentration is Cg and the silicon concentration is Ci.



$$F_1 = -D\frac{\partial c}{\partial x} = -D\left(\frac{C_i - C_g}{x}\right)$$

$$F_1 = D\left(\frac{C_g - C_i}{x}\right) \qquad (1)$$

Also,

$$F_2 = K_s * C_i \qquad (2)$$

This F_2 is the flux at the surface that is getting oxidized So, what is F_3? The ambience flux we are neglecting because oxygen in oxygen, water in water, not so much will diffuse. That will not be a bottleneck. At equilibrium

$$F1 = F2 = F$$

Both has to balance out each other right. So then the rate of growth is given by

$$\frac{dx}{dt} = \frac{F}{N_I} = \frac{K_s * C_i}{N_I} \qquad (3)$$

N_I is just a number.

Now, generally what is on the surface is known. The C_g is decided by the solid solubility limit. What is this solid solubility limit? Now, if you have a glucose for example, if you take 100 ml of water, typically if I remember correctly, you can only dissolve 9 grams of glucose powder. If you put more, what will happen? It will precipitate. It will not get dissolved. Now, suppose if you increase the temperature of the solution, then you can dissolve more. So, this is a typical concept of understanding solid solubility. Now, in the similar way, how much silicon dioxide can handle? This depends on the temperature at which you are operating it. This is also one more advantage of operating at higher temperatures right.

Now why wet oxidation gives very thick oxide films is because, the solid solubility of water is very high compared to solid solubility of oxygen in SiO_2. This could be your assignment question. So, I am just raising this part. I am just giving a hint why you get a thick oxide films using wet oxidation than the oxidation. So, in solid solubility limit, there is a limiting factor; there is a limit on the number of atoms or molecules that can be diffused through the silicon dioxide. C_g is the limit and it is typically known.. This is known for every material inside silicon dioxide. But for C_i is not clearly that much known. It is very difficult since it is random. So better if you can get C_i in terms of C_g using this equation.

So if I equate equation 1 and 2, it becomes,

$$D\left(\frac{C_g - C_i}{x}\right) = K_s * C_i$$

$$C_i = \frac{C_g}{1 + \frac{K_s x}{D}}$$

If I substitute this in equation 3, we get

$$\frac{dx}{dt} = \frac{K_s * C_g}{N_I\ (1 + \frac{K_s x}{D})}$$

Now if I can rearrange these terms, we get,

$$\frac{N_I\ (1 + \frac{K_s x}{D})dx}{K_s C_g} = dt$$

Now, if you integrate this part assuming no initial conditions we get

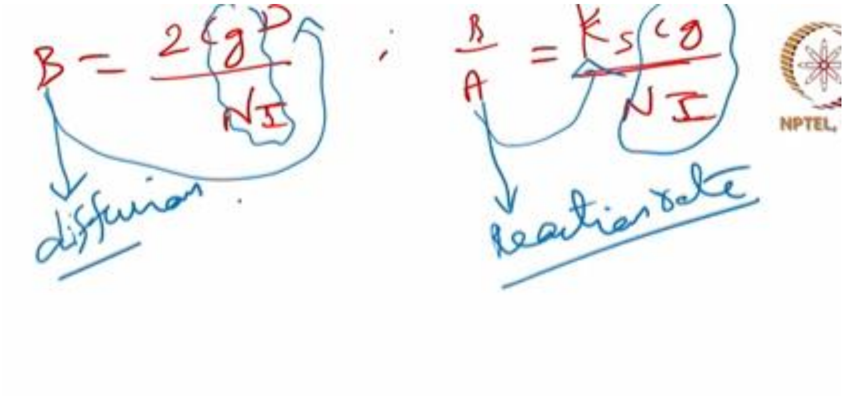
$$\frac{N_I x}{K_s C_g} + \frac{N_I x^2}{2 C_g D} = t$$

We can write this equation as ,

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

Where, $B = \frac{2C_g D}{N_I}$, and $\frac{B}{A} = \frac{K_s C_g}{N_I}$

Now if you look at these B and B/A then, C_g/N_I is there in both equations, and if I ignore that, B depends on the diffusion, and B/A depends on the K_s. That means somehow B is connected to the diffusion and B/A is connected to the reaction rate.



Somehow this B/A signifies about the reaction and that is the reason why we did all this kind of rearrangement and shuffling to get the above equations.

Now, the equation we have is a quadratic expression. So the growth is not going to be linear that is very clear. Now what is x. So let us say you have a expression

$$ax^2 + \mathrm{b}x + c = 0$$

And we know that the solution is,

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Here, $a = 1/B$, $b = \frac{1}{B/A}$ and $c = -t$

Now if I substitute, we get,

$$x = -\frac{1}{B/A} \pm \frac{\sqrt{\left(\frac{1}{B/A}\right)^2 - 4\left(\frac{1}{B}\right)(-t)}}{2\left(\frac{1}{B}\right)}$$

Here, x is the thickness and hence it is always positive.
So the only possible solution is,

$$\mathrm{x} = -\frac{A}{B} + \frac{\sqrt{\left(\frac{A}{B}\right)^2 + \frac{4t}{B}}}{\frac{2}{B}}$$

Simplifying this, we get

$$\mathrm{x} = \frac{A}{2}\left[\sqrt{1 + \frac{4tB}{A^2}} - 1\right]$$

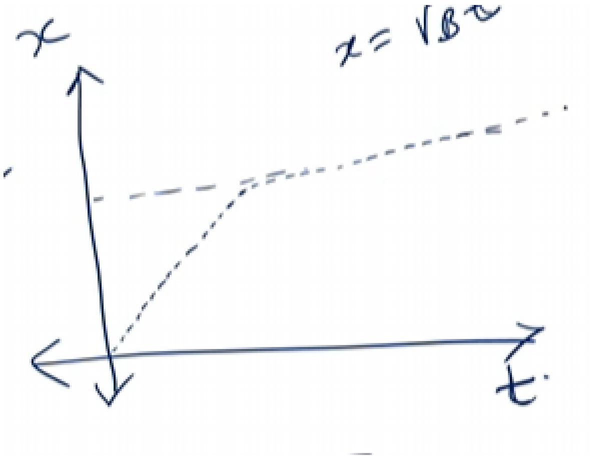
Now let us see how this x varies, with respect to time. So if I plot the x with respect to time and if I say t is very very small then the term with t becomes very small, and you can use the $(1+x)^{1/2}$ expansion, and we will get,

$$x = \frac{B}{A}t$$

Since B/A is a constant, x will grow linearly with time. That means, imagine if there is no growth, if the x are initially at t = 0, you do not have anything that is formed. So whatever that is formed is because of the reaction, not because of the diffusion. So at, t=0, what exactly is formed is because of surface reactions and not because of the diffusion that is what it is saying. Now what is this B/A suggesting? B/A depends on K_s and it is related to surface reaction. That is why this model has become extremely popular across the community. Also, suppose if the time is very high, then you can neglect the one in the equation, and if you can rearrange, x will become,

$$x = \sqrt{Bt}$$

And if you look at it, x is not linear, instead it is parabolic and it shows that somehow it is going to saturate. Of course, there is a growth, but not linear growth. It is a non-linear growth or parabolic growth.

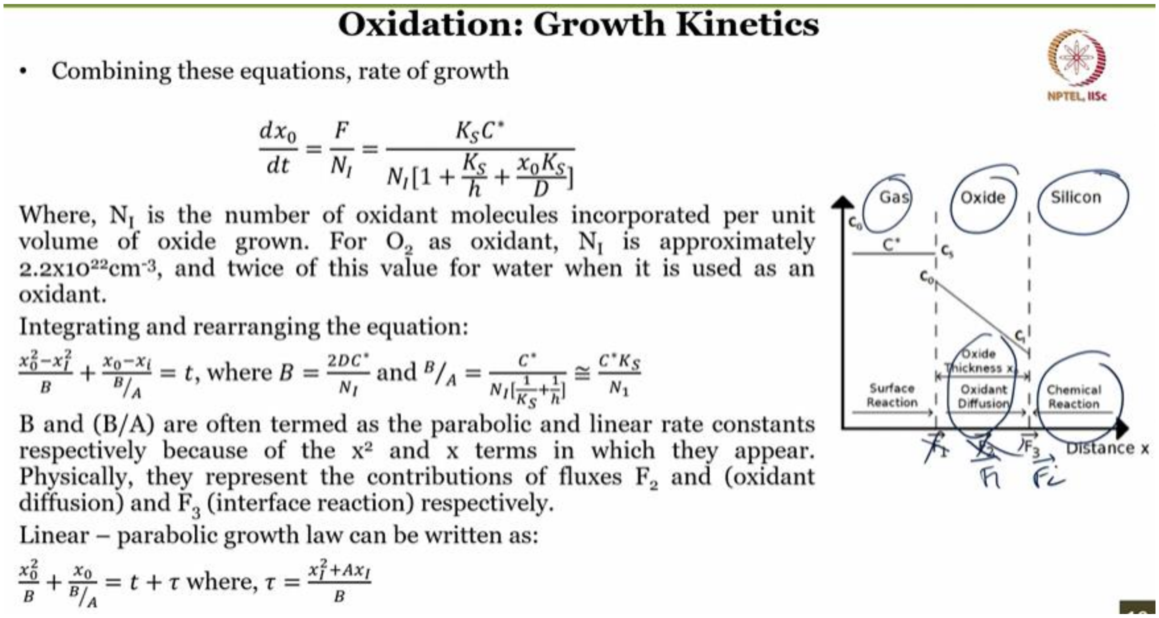


Now, we know that $B=2C_gD/N_I$. Here, D is the diffusion.

Suppose, if there is some considerable thickness or you have kept it for oxidation for long time and what happens there is a thickness that is already formed. Now whatever the new interface that has to be formed like silicon dioxide thickness that has to be formed is through the diffusion. Means diffusion will decide how the increase in the thickness happens. Initially, reaction rate regime will decide how the thickness growth happens. At this time, the growth is decided by B/A and in the nonlinear phase, the growth is decided by$\sqrt{Bt}$. That means that initially when t=0, the initial phase comes under the reaction rate limited regime and after that, during the nonlinear phase, it comes under the diffusion limited regime.

Let us understand a few more aspects and then we will get into the problems. This is how the modeling is, and this holds good for considerably larger thicknesses and the range of thicknesses and range of temperatures as well.

Now, here if you look into that what I have done, I have done approximations. Now what they are saying, you have a gas, you have an oxide and you have a silicon. Now we have only considered the flux in the oxide and silicon.



Oxidation: Growth Kinetics

- Combining these equations, rate of growth

$$\frac{dx_0}{dt} = \frac{F}{N_I} = \frac{K_S C^*}{N_I[1 + \frac{K_S}{h} + \frac{x_0 K_S}{D}]}$$

Where, N_I is the number of oxidant molecules incorporated per unit volume of oxide grown. For O_2 as oxidant, N_I is approximately $2.2 \times 10^{22} cm^{-3}$, and twice of this value for water when it is used as an oxidant.

Integrating and rearranging the equation:

$\frac{x_0^2 - x_I^2}{B} + \frac{x_0 - x_i}{B/A} = t$, where $B = \frac{2DC^*}{N_I}$ and $B/A = \frac{C^*}{N_I[\frac{1}{K_S} + \frac{1}{h}]} \cong \frac{C^* K_S}{N_1}$

B and (B/A) are often termed as the parabolic and linear rate constants respectively because of the x^2 and x terms in which they appear. Physically, they represent the contributions of fluxes F_2 and (oxidant diffusion) and F_3 (interface reaction) respectively.

Linear – parabolic growth law can be written as:

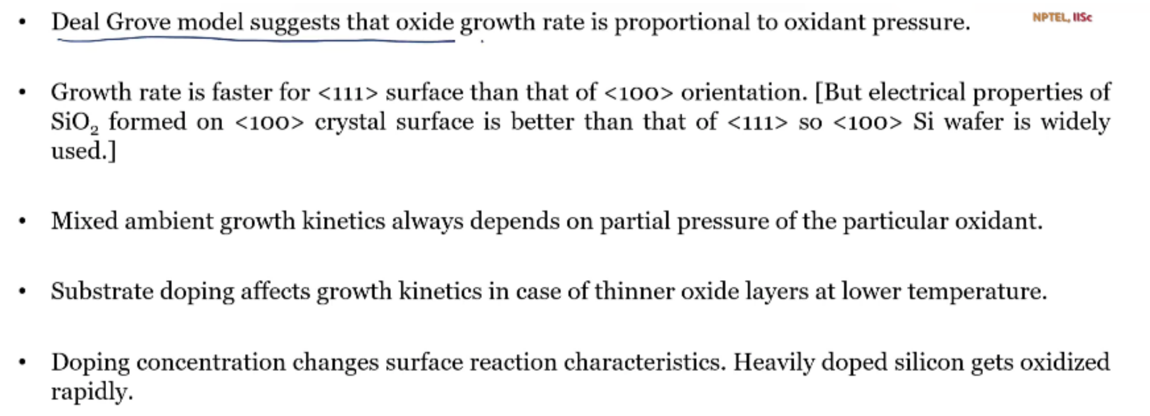
$\frac{x_0^2}{B} + \frac{x_0}{B/A} = t + \tau$ where, $\tau = \frac{x_I^2 + A x_I}{B}$

We have considered the flux in gas to be very high. It means that mass transport coefficient (h) is very high. Thus the following equation becomes approximated as what we have initially derived.

$$\frac{dx_0}{dt} = \frac{F}{N_I} = \frac{K_S C^*}{N_I[1 + \frac{K_S}{h} + \frac{x_0 K_S}{D}]}$$

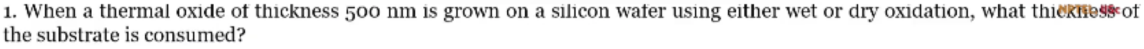
This happens because we have considered the flux in the ambient is very, very high compared to diffusion flux and the reaction flux. That is the only difference. So, if you wanted to ignore this assumption, the equation will become more complicated.

So based on what we derived, initially the growth depends on B/A, later the growth depends on B. B/A means it depends on K_S, and B depends on D. So, this empirical or mathematical model will match the experimental outcomes, and that is why it is very popular. Now, what is the Deal grove model is suggesting? Deal Grove model suggests that the oxidation growth is proportional to oxidant pressure. Growth rate is faster for 111 surface than 100 what does this mean? So, initial growth rate depends on the K_s and this K_S depends on surface reaction. Now we all know that the number of bonds available on the 111 surface is more compared to 100 orientation. That means the energy required to break that will be less. So the reaction kinetics will be different from 111 surface compared to 100 surface, and for 111 surface it is more compared to 100 surface. This is what this Deal Grove model is also suggesting.



- Deal Grove model suggests that oxide growth rate is proportional to oxidant pressure.
- Growth rate is faster for <111> surface than that of <100> orientation. [But electrical properties of SiO_2 formed on <100> crystal surface is better than that of <111> so <100> Si wafer is widely used.]
- Mixed ambient growth kinetics always depends on partial pressure of the particular oxidant.
- Substrate doping affects growth kinetics in case of thinner oxide layers at lower temperature.
- Doping concentration changes surface reaction characteristics. Heavily doped silicon gets oxidized rapidly.

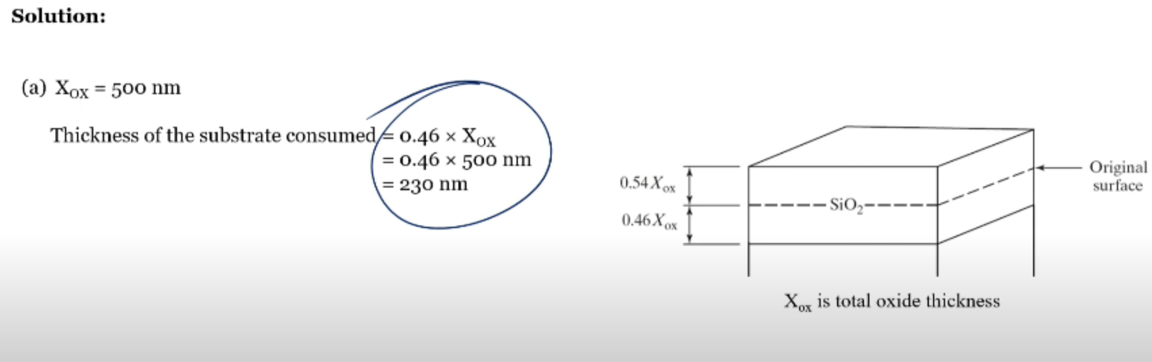
So, let us look into some problems now.



1. When a thermal oxide of thickness 500 nm is grown on a silicon wafer using either wet or dry oxidation, what thickness of the substrate is consumed?

This is a very typical problem like when a thermal oxide of thickness 500 nm is grown on a silicon wafer using either wet oxidation or dry oxidation, what is the thickness of the substrate that is consumed?

If you want X thickness of silicon dioxide then 0.46X will be consumed and 0.54X is formed on top of this.



Solution:

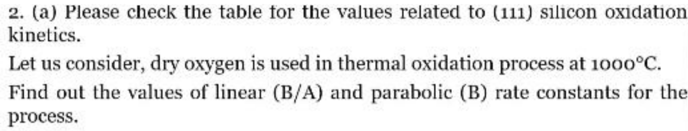
(a) X_{OX} = 500 nm

Thickness of the substrate consumed $= 0.46 \times X_{OX}$
$= 0.46 \times 500$ nm
$= 230$ nm

X_{ox} is total oxide thickness

So, already they have grown 500 nm , ie. X= 500 nm. So, the thickness that is consumed is 0.46 X.
So thickness consumed=0.45*500 =230 nm

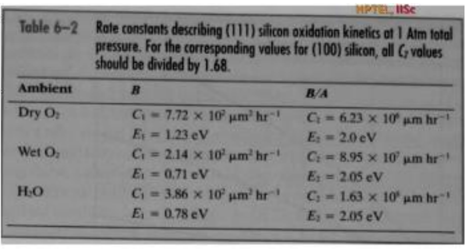
Next,



2. (a) Please check the table for the values related to (111) silicon oxidation kinetics.

Let us consider, dry oxygen is used in thermal oxidation process at 1000°C.

Find out the values of linear (B/A) and parabolic (B) rate constants for the process.



NPTEL, IISc

Table 6–2 Rate constants describing (111) silicon oxidation kinetics at 1 Atm total pressure. For the corresponding values for (100) silicon, all C_2 values should be divided by 1.68.

Ambient	B	B/A
Dry O_2	$C_1 = 7.72 \times 10^2$ μm² hr⁻¹ $E_1 = 1.23$ eV	$C_2 = 6.23 \times 10^6$ μm hr⁻¹ $E_2 = 2.0$ eV
Wet O_2	$C_1 = 2.14 \times 10^2$ μm² hr⁻¹ $E_1 = 0.71$ eV	$C_2 = 8.95 \times 10^7$ μm hr⁻¹ $E_2 = 2.05$ eV
H_2O	$C_1 = 3.86 \times 10^2$ μm² hr⁻¹ $E_1 = 0.78$ eV	$C_2 = 1.63 \times 10^8$ μm hr⁻¹ $E_2 = 2.05$ eV

I told you one more thing that K_s and D are dependent on temperature. K_s can be modeled as

$$K_s = K_o e^{\frac{-E_G}{KT}}$$

And,

$$D = D_o e^{\frac{-E_{G2}}{KT}}$$

So instead of K_0 , I can write C_1 since for B, they are saying C_1 right.

Now, if you look at B depends on D and B/A depends on K_s.

$$B = \frac{2C_g D}{N_I} \text{ , and } \frac{B}{A} = \frac{K_s C_g}{N_I}$$

Means, B/A depends on temperature, and B also depends on temperature, but they are different. Now, B/A depends on surface reaction kinetics, B depends on diffusion kinetics. Now, accordingly you have the values. So, B can be modeled as

$$B = C_1 e^{\frac{-E_1}{KT}}$$

$$\frac{B}{A} = C_2 e^{\frac{-E_G}{KT}}$$

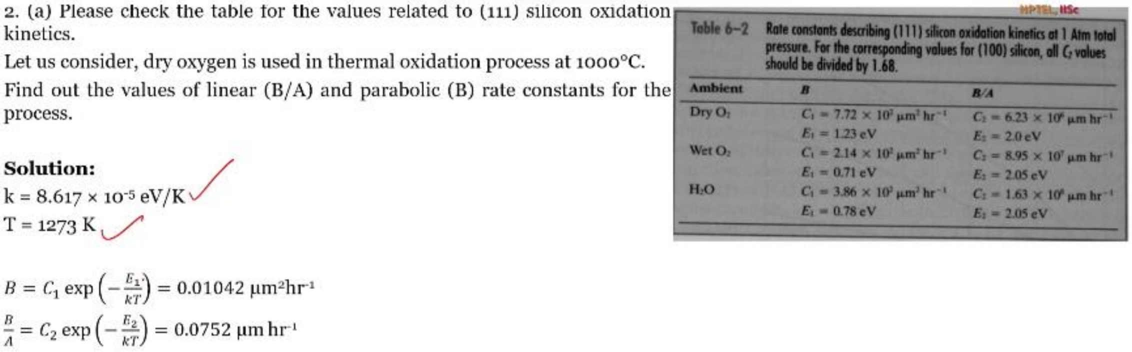
That is where they are connected, okay. So do not think why we are suddenly jumping into this B/A and how the dependence of temperature has come into picture. So let us ignore that.

Now what is the question? Now please check the table for the values related to 111 silicon oxidation kinetics. This is for silicon oxidation. Let us consider dry oxidation is used in thermal oxidation process at 1000 degrees. So the temperature is 1000 degrees. So here you have to take everything in absolute values. So 1000 degrees is 1273 Kelvin and find out what is the value of B/A and B.

So B/A according to this table for 111 growth,

$$\frac{B}{A} = C_2 e^{\frac{-E_G}{KT}}$$

So K is a Boltzmann constant and it is around $8.821 * 10^{-5}$ eV/K. You can find out that value and you know the E_2 this is 2 eV. So, if you know these values, substitute in that and then you will get this answer.



2. (a) Please check the table for the values related to (111) silicon oxidation kinetics.

Let us consider, dry oxygen is used in thermal oxidation process at 1000°C.

Find out the values of linear (B/A) and parabolic (B) rate constants for the process.

Solution:

$k = 8.617 \times 10^{-5}$ eV/K ✓

T = 1273 K ✓

$$B = C_1 \exp\left(-\frac{E_1}{kT}\right) = 0.01042\ \mu m^2 hr^{-1}$$

$$\frac{B}{A} = C_2 \exp\left(-\frac{E_2}{kT}\right) = 0.0752\ \mu m\ hr^{-1}$$

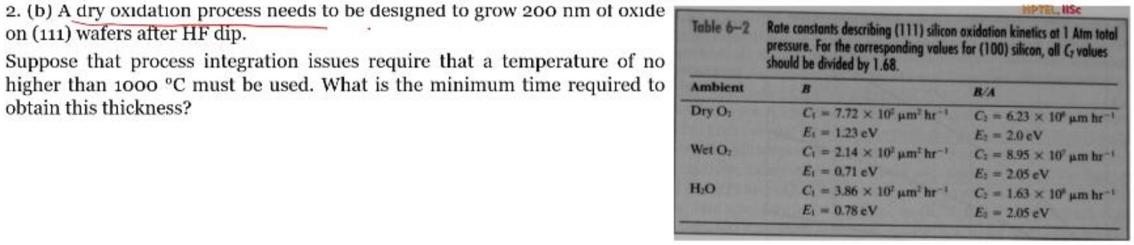
NPTEL, IISc

Table 6–2 Rate constants describing (111) silicon oxidation kinetics at 1 Atm total pressure. For the corresponding values for (100) silicon, all C_2 values should be divided by 1.68.

Ambient	B	B/A
Dry O_2	$C_1 = 7.72 \times 10^2\ \mu m^2\ hr^{-1}$	$C_2 = 6.23 \times 10^6\ \mu m\ hr^{-1}$
	$E_1 = 1.23$ eV	$E_2 = 2.0$ eV
Wet O_2	$C_1 = 2.14 \times 10^2\ \mu m^2\ hr^{-1}$	$C_2 = 8.95 \times 10^7\ \mu m\ hr^{-1}$
	$E_1 = 0.71$ eV	$E_2 = 2.05$ eV
H_2O	$C_1 = 3.86 \times 10^2\ \mu m^2\ hr^{-1}$	$C_2 = 1.63 \times 10^8\ \mu m\ hr^{-1}$
	$E_1 = 0.78$ eV	$E_2 = 2.05$ eV

Now, one more thing is, interestingly if you see this point on the table, all C_2 values you see should be divided by 1.68 for 100. For the corresponding values, say for 100 silicon, all C_2 values should be divided by 1.68. Why only C_2 values should be divided by 1.68, why not C_1? This is because, B/A depend on K_s and B depends on D. Diffusion really does not matter what the orientation is, but K_s matters. If 111 is there, you will have different K_s, for 100, there is a different surface kinetics. So, for 100, we know that it is less number of bonds available for breaking. Obviously, K_s will be less. That is the reason why they are telling you to divide by 1.68, only for B/A. Really does not matter for B.

Next problem,



2. (b) A dry oxidation process needs to be designed to grow 200 nm of oxide on (111) wafers after HF dip.

Suppose that process integration issues require that a temperature of no higher than 1000 °C must be used. What is the minimum time required to obtain this thickness?

NPTEL, IISc

Table 6–2 Rate constants describing (111) silicon oxidation kinetics at 1 Atm total pressure. For the corresponding values for (100) silicon, all C_2 values should be divided by 1.68.

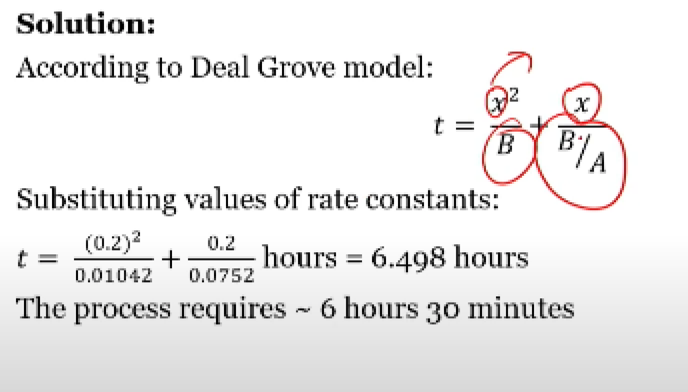
Ambient	B	B/A
Dry O_2	$C_1 = 7.72 \times 10^2\ \mu m^2\ hr^{-1}$	$C_2 = 6.23 \times 10^6\ \mu m\ hr^{-1}$
	$E_1 = 1.23$ eV	$E_2 = 2.0$ eV
Wet O_2	$C_1 = 2.14 \times 10^2\ \mu m^2\ hr^{-1}$	$C_2 = 8.95 \times 10^7\ \mu m\ hr^{-1}$
	$E_1 = 0.71$ eV	$E_2 = 2.05$ eV
H_2O	$C_1 = 3.86 \times 10^2\ \mu m^2\ hr^{-1}$	$C_2 = 1.63 \times 10^8\ \mu m\ hr^{-1}$
	$E_1 = 0.78$ eV	$E_2 = 2.05$ eV

So a dry oxidation process needs to be designed to grow 200 nm thickness of oxide on 111 wafer. Thickness that they are looking is 200 nm. Try to convert everything into micrometer, then that will be easy because these are in $\mu m^2/hour$ units. Suppose that process integration issues requires that a temperature no higher than 1000 degree. So, your temperature is 1273 Kelvin. What is the minimum time required to obtain the thickness?

You have

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

Now you need to find out t. You substitute these values. Find B like in the previous question, and substitute x=0.2 µm. Finally, you will get 6 hours 30 minutes.



Solution:

According to Deal Grove model:

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

Substituting values of rate constants:

$$t = \frac{(0.2)^2}{0.01042} + \frac{0.2}{0.0752}\ \text{hours} = 6.498\ \text{hours}$$

The process requires ~ 6 hours 30 minutes

Now one can do same thing for wet oxidation as well and figure out for yourself whether the wet oxidation time is less or dry oxidation time is less. So, if you do the same calculation the time required for wet oxidation is less than time required for dry oxidation. That is why this Deal Grove model is extremely popular.

I think we are coming to the end of this. So, the thing is we have covered multiple aspects, we have seen how the flux F_1 and F_2 will balance with each other. Then we tried to form a mathematical modeling called Deal Grove model and we have looked at some problems. So, thank you see you again bye.