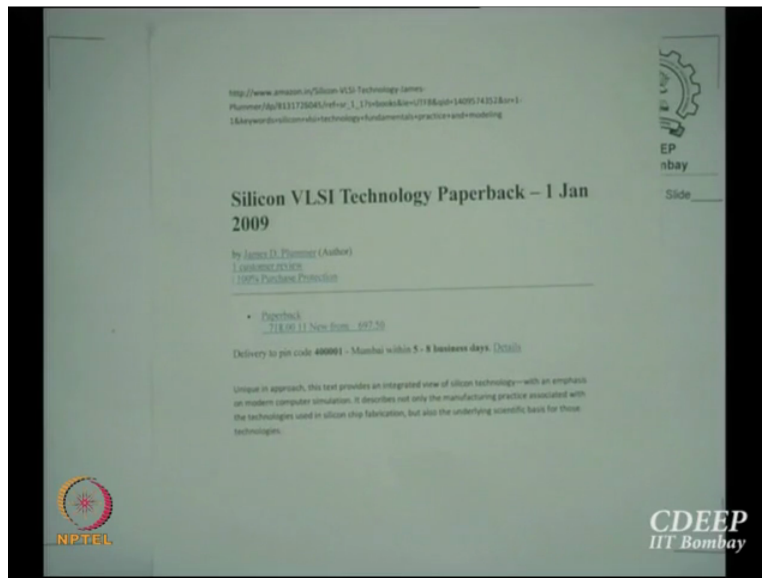
Fabrication of Silicon VLSI Circuits using the MOS technology
Professor. A N Chandorkar
Department of Electrical Engineering
Indian Institute of Technology Bombay
Lecture 12
Thermal Oxidation of Silicons

(Refer Slide Time: 0:17)



I have actually checked on the website the Amazon India has this book, this is the Ad for that. I am, if you have not purchased it so far, or could not, this is what recent I saw yesterday. It is 718 rupees or whatever possible and this is second edition, 1st January 2009. This is available at website www.Amazon.in Silicon VLSI Technology. This is written here, you can take, they say five days or five to eight business days.

So in case you have not purchased or do not want to purchase, is yours. But in case you want to, this you can note down or maybe you can go on Google and say Amazon India. So whichever way it is, so do not tell me that book is not available or something, something. Is that okay? Coming back, and those who want to write that website, maybe you can. In case you feel that is what I am marking.

(Refer Slide Time: 1:38)

Bruce Deal and Andy Grove were first to model the process of Thermal Oxidation.
(Ref: J. Appl. Phys., Vol. 36, p 3770, 1965)
Hence the Model is called 'Deal-Grove' Model.
This model provides kinetics of growth of SiO_2 (X_{ox}) as a function of temperature and time.

Model Assumptions

(a) Initially there is a finite (Thin) oxide layer, before oxidation ($t = 0^-$) Starts

(b) The Oxidant Gas specie impiges on this SiO_2 layer and diffuses through the Oxide-layer and reaches $Si-SiO_2$ interface.

CDEEP
IIT Bombay

We were other day looking for Bruce Deal model, Bruce Deal and Andrew Grove model for oxidation. The first paper published in this area way back in 65 and this model is very popular in the literature called Deal-Grove model. And I just told you other day, Grove was the one of the CEO or Chief of Intel from say 1992 to 1997. He is the (())(01:54) scientist or (())(01:55) CEO of Intel. Andy Grove also has a book on physics and technology of semiconductors or rather technology of semiconductor which is old book.

But many things remain same irrespective whether today or yesterday. So one can also look for Grove's book in the library. There is another book on VLSI technology by Wolf, which is also very good which is rather recent compared to Plummer of course is older. But around 2006 Wolf has also published. So any of these books are good enough, please do not say that the books are not there, because of that whatever happened.

At least do not blame books. So the assumption what he did was that initially there is a finite thin oxide and at t is equal to 0 minus, this is the oxide available and t is equal to 0 plus, oxidant attacks this silicon layer, silicon dioxide layer and starts the oxidation. We did last time, I am just trying to recapitulate.

(Refer Slide Time: 3:05)

© The Oxidant specie then reacts with Silicon atoms at the Interface and creates newer oxide layer.

This process continues and Sidioxide layer keeps growing to a thickness by end of time of Oxidation. This process is depicted in the Fig. below:

Gas SiO_2 Silicon

There are Three Fluxes in Gas, SiO_2 and Silicon regions.

Let F_1 be oxidant flux in Gas phase

F_2 be oxidant flux in SiO_2

F_3 be oxidant flux for reaction at Silicon.

Then the oxidants we see then, reacts with, diffuses through the thin oxide, reaches to the silicon-silicon dioxide interphase, reacts with silicon and forms SiO2. And we have already last time said there are three fluxes, F1 in the gas phase, F2 in the oxide, and F3 near the interphase of silicon-silicon dioxide. And we also discussed other day that in steady state if the oxide, oxidant is coming, diffusing, reacting, the system has to be in steady state. And in steady state all fluxes must be equal. This we have done last time, I am just trying to push it again.

(Refer Slide Time: 3:49)

Ⓐ Transport of Oxidant from the Furnace ambient to the Oxide surface takes place with Flux F_1.

Ⓑ Oxidant specie Diffuses through SiO_2 layer with a flux F_2.

© Oxidant reacts with Silicon with Flux F_3.

However in Steady State we can say that

$$F_1 = F_2 = F_3 = F$$

800°to1200°c

And we say F1 is equal to F2 equal to F3 and that we call it F. And we will show you this process again but this is just to show you how oxidation proceeds. Inside a furnace where temperature is kept 800 to 1200, wafers are vertically start on quartz rack. Oxygen is entering here and oxidizes the silicon. So let us start with the model, this is what we did last time. So now we actually look for the model.

(Refer Slide Time: 4:21)

Here is the model, same thing what I wrote but there are now few more definitions. Let us say CG is the oxidant gas or oxidant concentration in the gas stream. That is not near the surface but inside the whole tube actually. This is called CG, then corresponding to this CG we have a oxidant concentration at oxide interface is CS or rather C0. Or sorry C0, forget it, CS we have, CS is the concentration here. So there is a gradient because there is no enough oxygen here or oxidant here, so there is a gradient from CG to CS.

Now these are called gas phase concentrations. So equilibrium concentration in, at the solid also can be found and we define C star as the equilibrium oxidant concentration in solid related to this CG. CG is the gas phase concentration and C star is the equivalent of that in the solid phase. Whereas C0 is similarly is equilibrium constant, oxidant concentration in solid related to CS. So this relates to here and this relates to here.

This is, then we say initial oxide thickness is CI, initial concentration at the interface is CI. And we say already xo oxide thickness exist or xi rather but xo is the oxide thickness which we want

to find at a given time and temperature. So let us look at the fluxes now. According to Deal-Grove model flux F1 which is in the gas phase from the ambient to the surface, we know the flux will be proportional to the gradient.

If CG is the gas concentration in the bulk, CS is the gas concentration near the surface, then F1 must be proportional to CG minus CS. Whatever at the surface and what is available in the stream, the gradient is set and that difference between the two is the flux F1 and the proportionality constant is called mass transfer coefficient hG. So F1 is hG times CG minus CS. This is the first flux which we are looking into, that is the gas phase flux.

However we are more interested to know there are terms which probably you should learn from thermodynamics but there are words called pressure, total pressure and there are words called partial pressures. Partial pressure is defined as the pressure of a gas inside a stream in a given temperature in a given area. So if I know F1 is hG, this I want to replace this CG and CS which we are not known to me. CG and CS cannot be monitored, so I will try to know is there equivalence of these in terms of partial pressures.

And if I know the partial pressures, then I also know correspondingly to that partial pressure, using Hertz equation we will find out what is the solid phase concentration for this partial pressures. So at the end I am interested to correlate CG, CS with C star and C0. This is what I am looking for. So the first thing I do is after I write hG into CG minus CS as the first flux and whereas I repeat hG is called mass transfer coefficients.

(Refer Slide Time: 8:06)

We define CG by ideal gas law, P is equal to RT, there is nothing great about. This is concentration, so P is equal to RT. So CG is equal to PG by kT. N is number one, molecular this, so we, RkT is, NkT can be written as N is equal to 1. Therefore RT is kT. So C, sorry CG is PG by kT. CS is PS by kT where PG, PS are the partial pressures of the oxidant at gas ambient and oxide surface. Again same, PG in the gas stream and PS is the at near the interface of oxide and gas stream.

So if I substitute CG and CS in the last equation by PG and PS term, then I get F1 is equal to hG PG by kT minus PS by kT or is equal to hG by kT into PG minus PS. Please remember if I fix temperature and if I fix the gas flows, total gas flows, hG is the constant which is the mass transfer coefficient proportional to the total pressures. We will see in case of CVD. We invoke Henry’s law for gases and fluids.

There is Henry law which actually relates the partial pressure to solid state concentration. This is called Henry’s law. According to Henry's law the C star which is the equilibrium oxide concentration in solid state is proportional to the partial pressure in gas stream. Similarly C0 which is at the surface of SiO2, C0 is proportional to partial pressure at the surface which is PS. And there is a constant associated with this equilibrium, equivalent to this which is called Henry's constant.

So C star is H times PG, C0 is H time PS where H is called Henry's constant. So now we can see I have relationship with C star and C0 in terms of PG and PS through Henry's constant, so I can use these two equations to go back into 3 and find the flux F1. What is the method? First I converted CG and CS into equivalent partial pressures by mass transfer coefficients. The partial pressures are related to their equivalent solid state concentrations by using Henry's law or Henry, H is the Henry's constant.

So now I have C star and C0 which are at the surface. Please remember C star and C0 are at the surface of SiO2 which are in solid phase, that is the oxidant we are going to actually diffuse through. So F1 therefore, is that okay? I repeat first we relate CG to PS, PG. CS, CG, PS, PG and then we wrote this equation. Then we say okay, using, invoking the Henry's law for fluids C star is proportional to PG, C0 is proportional to PS. Then C star is H time, H is H PG and C0 with H time PS. And now I substitute PG, PS from here in this equation 3.

(Refer Slide Time: 11:41)

From equation (iii)

$$F_1 = \frac{h_G}{kT}\left(\frac{C^*}{H} - \frac{C_0}{H}\right)$$

or $F_1 = \frac{h_G}{HkT}(C^* - C_0)$ — (V)a

where we can further define $h = \frac{h_G}{HkT}$ & then

$F_1 = h(C^* - C_0)$ — (V)

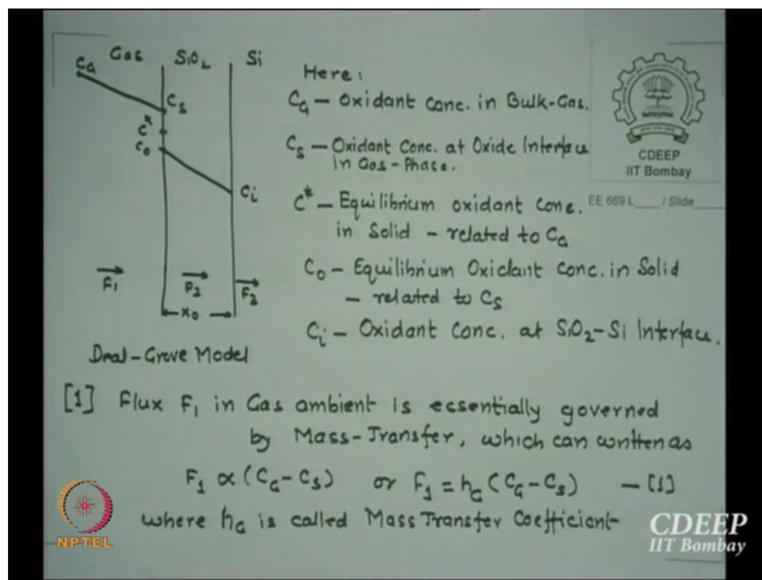
[2] Once the Oxidant reaches Oxide Surface with concentration of C_0, It diffuses into Oxide (x_0). By Ficks law

$F_2 \propto \frac{C_0 - C_i}{x_0}$ where C_i is the Oxidant Conc at SiO_2-Si interface.

If I do that, F1 is hG by kT, C star by H minus C0 by H. Or F1 is hG by HkT C star minus C0. And I can redefine this hG upon HkT which is all constant at a temperature and marks total flow and pressure. Then h is hG upon HkT is redefined as term h, which of course is proportional to mass flow. So F1 is h times C star minus C0. So first flux in the gas phase is related to solid phase concentration difference, h times C star minus C0, so this is the first flux we obtained.

Now as I say what is our game is to find F2, find F3 and then what do we do? We write F1 is equal to F2 equal to F3 and then solve. Ultimately what I am really trying to do is I want to find dx0 by dt. What is dx0 by dt? Rate of oxide growth, that is what I am interested in, rate of oxide growth. So if I know the time, I know how much oxide, I will be growing in case of oxidation. So that is my purpose. So first flux, F1 I got. Now secondly if we look at our figure again, just a minute.

(Refer Slide Time: 13:15)



If we look at this figure, F1 we have found in terms of C star, C0. F2 is the flux which is diffusing inside thin oxide. F2 is the flux of oxidant passing through the oxide thickness.

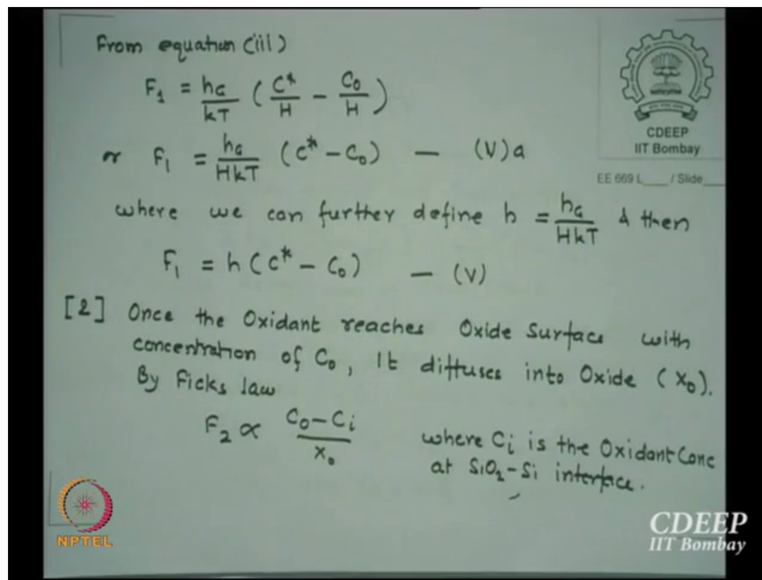
(Refer Slide Time: 13:29)

And that is equal to C0 minus CI. CI as I say is concentration of oxidant at silicon-silicon dioxide interface.

(Refer Slide Time: 13:44)

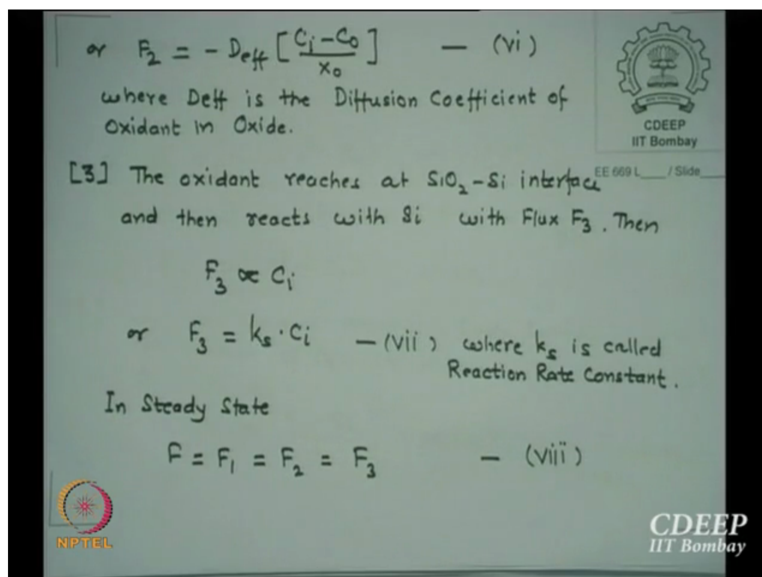
CI, so there is a gradient diffusion.

(Refer Slide Time: 13:48)



So F2 is then proportional to C0 minus CI by x0, this is gradient. C0 minus CI by x0 is essentially a gradient. And what is the gradient constant should be? Since it is diffusing what should be the constant of proportionality here? Diffusion coefficients of that oxidant in the oxide. Please remember diffusion coefficient is for the material oxide or whatever specie going in that material. Diffusivity is different in different material for different gas flows going in. So specifically you have to, earlier we have talked about impurities in silicon. Now we are talking of equivalent solids concentration of gases in the oxide and how do they diffuse through.

(Refer Slide Time: 14:39)



So if I write that, I write F2 is minus D effective. Minus is because I am subtracting CI minus C0 by x0, where D, why I did this? Because final concentration minus the initial concentration divided by x is actually slope. It is essentially how we define, that is going down. So that is the final concentration minus the initial concentration divided by x0 is the gradient. So minus D effective CI minus C0 by this where D effective, diffusion coefficient of oxidant in oxide.

This is flux, F2. Now this is now made available to react with silicon, oxidation, oxidant is now diffusing. Please remember what is the process we are saying. The gas, from the gas phase oxygen is coming, equivalently it gets to the silicon oxide surface, then diffuses through and then reacts with silicon to form fresh oxide. Is that clear? That is what Deal-Grove model is suggesting. Gas phase through oxide and to react with silicon, that is the process we have defined.

So we want to know what is the flux 3 which is (propo) and we know flux 3 is essentially as many silicon atoms are available or concentration of silicon available, CI at that place, that is the only one which it can react with. Oxygen cannot react more than CI because whatever available only can react. So we say the flux F3 is proportional to CI and the proportionality constant is called reaction rate constant k, ks.

Reaction, ks is called reaction rate constant. So F3 is ks CI. So now I have three fluxes. F1 is from gas phase to the silicon dioxide surface. From silicon dioxide surface to silicon surface diffuse and then reacts with silicon to form fresh oxide. And as I say if I assume it is in steady state, which will happen, I put the wafer in, I have constant temperature, constant flow, system will go into steady state. In that case the flux F is F1 equal to F2 equal to F3.

(Refer Slide Time: 17:00)

Here:
C_G – Oxidant Conc. in Bulk-Gas.
C_S – Oxidant Conc. at Oxide Interface in Gas-Phase.
C^* – Equilibrium oxidant Conc. in Solid – related to C_G
C_0 – Equilibrium Oxidant Conc. in Solid – related to C_S
C_i – Oxidant Conc. at SiO_2-Si Interface.
Deal-Grove Model
[1] Flux F_1 in Gas ambient is essentially governed by Mass-Transfer, which can written as
$F_1 \propto (C_G - C_S)$ or $F_1 = h_G(C_G - C_S)$ — (1)
where h_G is called Mass Transfer Coefficient

"Professor-student conversation starts."

Student: C star remove kar diya?

Professor: No, what we are available concentration in the solid phase from the bulk is C star, of which only C0 is really going to diffuse because that is at the surface.

Student: Cs is available.

Professor: Cs is available which correspond to C0 in solid.

Student: Means?

Professor: Correspond I just now said, gas concentration equivalent in solid concentration is related. I just now said through the mass transfer coefficient. And through that to partial pressure, through that to Henry's constant to the actual concentrations. So we are saying the gas is coming, equivalently how much gas is available overall, of which how much will diffuse from the surface is C0. It will go to CI because there is a gradient, this is very few oxidant here and very large oxidant here, so it will diffuse and when it reaches here, it will start reacting with the silicon surface and form oxides.

"Professor-student conversation ends."

(Refer Slide Time: 18:09)

First we take $F_1 = F_2$, then

$$h(C^* - C_0) = -D_{eff}\left(\frac{C_i - C_0}{x_0}\right) \quad \text{— (ix)}$$

Then $F_2 = F_3$ can also be used to give

$$-D_{eff}\left(\frac{C_i - C_0}{x_0}\right) = k_s C_i \quad \text{— (x)}$$

Solving equs. (ix) and (x), we get

$$C_i = \frac{C^*}{\left[1 + \frac{k_s}{h} + \frac{k_s x_0}{D_{eff}}\right]} \quad \text{— (xi)}$$

$$C_0 = \frac{\left[1 + \frac{k_s x_0}{D_{eff}}\right] C^*}{\left[1 + \frac{k_s x_0}{D_{eff}} + \frac{k_s}{h}\right]} \quad \text{— (xii)}$$

So we do little maths, there is nothing very serious maths. Since we have already done, F1, F2, F3 are equal, we say, first we equate F1 equal to F2 and next time we will equate F2 equal to F3. So if I write F1 equal to F2, I get h C star minus C0, minus D effective CI minus C0 by x0. And if I write F2 equal to F3, I get minus D effective CI minus C0 by x0 is ks CI. So I have equation 9 and 10 which are, I got it from equating F1 equal to F2 and F2 equal to F3.

Two equations to announce, I am interested to know C0. I am interested to know CI in terms of C star available gas phase concentration. And correspondingly I derive, solve these two equations 9 and 10. How do I eliminate CI out of this? Substitute here and find C0 or get C0 from here, substitute here to get CI. This maths little longer, so I just wrote down the final equation. I repeat, solve 9 and 10 to get CI and C0. Substitute one of them to the other, you will get the other value or substitute how much C0 into the second equation, you will get CI first, whichever way you do.

So I get CI is equal to C star upon 1 plus ks by h plus ks x0 by D effective. Then C0 is 1 plus ks x0 by D effective times C star and the denominator is same. Please remember denominator is same, 1 plus ks x0 by this, ks by h. So I have now CI and C0. I may not be very, I do not need to know C0 very much but I am interested to know CI, why? Because that is the concentration is going to react with silicon. So there I figured out both can be calculated but I am more interested to know CI in terms of C star which I got.

So what Deal-Grove model says now, what is the oxidation rate. Can anyone suggest what is the oxidation rate? If F is the flux and N is the number which is available for reaction, so F by N is essentially equal to dx0 by dt. If, what is dx by dt? The available flux divided by available concentration where it can react, the ratio of that is essentially dx0 by dt.

(Refer Slide Time: 21:18)

This Deal-Grove Model can then be used to get x_o as function of Temperature of oxidation and time of Oxidation.

If N_1 is conc. of Oxidant molecules, then

$$\frac{dx_o}{dt} = \frac{F}{N_1} = \frac{F_3}{N_1} = \frac{k_s C_i}{N_1} \quad \text{— (xiii)}$$

Substituting C_i from (xi) into (xiii), we get

$$\frac{dx_o}{dt} = \frac{k_s C^*}{N_1} \frac{1}{\left[1 + \frac{k_s}{h} + \frac{k_s x_o}{D_{eff}}\right]} \quad \text{— (xiv)}$$

We know $N_{1O_2} = 2.22 \times 10^{22}/cc$ and $N_{1H_2O} = 4.44 \times 10^{22}/cc$

So that is Grove-Deal's model which says if N1 is the concentration of oxidant molecules, then dx0 by dt is F by N1 and F is equal to F1 equal to F2 equal to F3. So F3 is the smallest term, Ks CI, so I used F3 instead of, I can write anyone of them. But then there will be two variables coming, so I just want to remove, I have got CI value anyway. So I got ks CI by N1. Please remember this N1 for oxygen is 2.22, 10 to power 22, N1 for H2O HOH molecule is 4 point double of that.

OH, OH double the concentration whereas pure oxygen has N1 as 2.22, 10 to power 22. This has been chemically monitored, measured by many experiments, by atomic mass spectroscopy to FTIR everything may, one is monitoring that. So if I have dx0 by Ks CI but this I substitute CI from the last equation. I just derived expression for CI, so substitute CI here, so it gets k upon C star upon N1, 1 upon, 1 plus ks by h, k sorry, ks x0 by D effective. So how does it look like? We will do some simple looking expression for this, first you note down this. Yes.

So I get dx0 by dt is ks C star by N1, 1 upon 1 plus ks by h plus ks x0 by D effective. So this equation now I know, I can modify this equation to suit some kind of good-looking expressions.

What is good-looking? Some A, B constant, so it looks very simple equation, algebraic. Of course not algebraic, first order differential equation. Is that noted? This one, dx0? So I have now the oxidation rate and I will modify exactly in form, that is what I am going to do. Is it okay?

(Refer Slide Time: 23:32)

From eq. (xiv)

$$\frac{dx_0}{dt} = \frac{2C^* D_{eff}}{N_1} \cdot \frac{1}{\frac{2D_{eff}}{k_s}\left(1 + \frac{k_s}{h} + \frac{k_s}{D_{eff}} x_0\right)}$$

$$= \frac{2C^* D_{eff}}{N_1} \cdot \frac{1}{2D_{eff}\left(\frac{1}{k_s} + \frac{1}{h}\right) + 2x_0} \quad \text{---(xv)}$$

We define

$$A = 2D_{eff}\left(\frac{1}{k_s} + \frac{1}{h}\right)$$

and

$$B = \frac{2D_{eff} C^*}{N_1}$$

Then

$$\frac{dx_0}{dt} = \frac{B}{A + 2x_0} \quad \text{--- (xvi) a}$$

Okay, so dx0 by dt is 2 C star D effective by N1. I actually multiply it, 2 D effective by ks both sides, this and here and I get this expression. Readjust the terms, 2 D effective, 1 plus ks by, 1 by ks plus 1 upon h plus 2x0, into this constant. And now I define some terms. I define term A as 2 D effective 1 upon ks plus 1 upon h as A. And 2 D effective C star by N1, I define as B. Okay? Same expression, this I have defined as B and this I defined as A. So I get a very nice looking simple differential equation first order, dx0 by dt is B upon A plus 2x0.

Please remember ks is a function of temperature, h is a function of mass transfer coefficients or mass flows. In normal case h is much higher than ks but let us see how much. dx0 by dt, B upon A plus 2x0 is the simplest equation we get, with good-looking. Assuming that B and A are constants at a given temperature for a given mass flow. If you change that, these terms will, B, and A values will also correspondingly change. Is it okay?

So that is a very trivial looking this but it is important. Why are we doing this again? Because at the end when I monitor the oxide, I have the furnace, I actually oxidize the wafers and then I may do some characterization to know the oxide thickness. But if probably I am doing something process on a system where there is no furnace and there is nothing to this. I must be able to get

relative oxide thicknesses every now and then for change in oxide thickness or whatever I am doing in process which is running on a CAD tool. That is called T-CAD for technology tools are available, earlier ones we used to have program from called Stanford called Supreme process stimulator.

There are now many, one is Intra, other is (())(26:10). There are many such program, this is which is device plus process simulator. Since we have many process simulators now, we like to know what models they use because they will also find out what is the oxide capacitance every now and then. So they need to know oxide, how much grown. So you may specify temperature, you may specify gas flows, you may specify but then at the end software has to do something to evaluate that. We need models to do that, so all this effort is to see a model which is going into a T-CAD tool.

So whenever you are using central or others, you probably do not even look at the models, you just, it adds data and you just substitute and you just ahh, good! But actually what I have gone through is this and maybe tomorrow for thin oxide, better models will be required. So you must know how do we actually derive models, what is the physics behind, chemistry behind, materials science behind and you again behind.

(Refer Slide Time: 27:18)

or $A\frac{dx_0}{dt} + 2x_0\frac{dx_0}{dt} = B$ — (xvi)

Initial Condition

At $t=0$ $x_0 = x_i$ (By assumption)

or $(A+2x_0)dx_0 = B\,dt$

Integrating

$$Ax_0 + \frac{2x_0^2}{2} = Bt + B\tau \quad \text{— (xvii)}$$

where τ is time taken to grow oxide thickness x_i

or we can say $\tau = \frac{x_i^2 + Ax_i}{B}$ — (xviii)

CDEEP IIT Bombay

NPTEL

So I rewrite the same term, A dx0 by dt. So if I solve the, if I see this equation, this is the quadratic equation. So we put an initial condition to solve this differential equation. Sorry, not it

is a simple differential equation. I say at t is equal to 0, according to Deal-Grove's model there is initial oxide x0 is xi. That is how we started with. There is an initial oxide. So x0 is xi and corresponding to this time, xi, if I use, okay, so we substitute here.

Let us say tau is the time taken to create this xi. Let us say tau is the time taken, equivalently actually it is existing. We do not know what time we did, we did not do anything. So we say okay, equivalently if we have to go this much oxide, how much time? So that time I declared as tau. So I say now Ax0, if I put at this, A 2x0 dx0 is Bt. And I integrate this, then I get Ax0, 2x0, so I have Bt, plus initial condition I will put it at x is equal to xi, tau is the time taken.

So I rewrite this term, Ax0 by 2 x0 square, Bt plus B tau. Tau is the time taken to grow oxide thickness of xi and there is, this is only a fictitious number. Why fictitious? Because initial oxide is already there, we are just trying to equate it into a time frame. Okay, or to say tau is xi square plus A xi upon B. Axi, xi square Axi by B is essentially this. xi also I do not know in fact. So we assume normally tau should be very small because thin oxide is there.

But if I somehow figure out how much was xi, then I should be able to know how much time equivalently it would have been. Okay, so this is my equation and this looks what? A simple quadratic equation and therefore I can find x0 terms. I repeat the term which I am getting is Ax0 plus 2 x0 square. 2, 2 of course will cancel. x0 square is equal to Bt plus tau.

(Refer Slide Time: 29:37)

So this is the expression you will get, x0 square plus Ax0 is Bt plus tau. For thicker oxide growths tau can be neglected, why? Because t will be much larger than tau. Tau is very small but it is existing. But for thinner oxide that may be comparable, so we must figure out how much is actually thin oxide during initial time must be actually evaluated. I will show you how. So if this quadratic equation can we have solution of minus A plus minus A square plus 4Bt plus tau by 2, minus A by 2 plus minus A by 2, 1 plus t plus tau upon A square by 4B.

And of course negative solution is neglected. Why? Because there is nothing called negative oxide growths. So we say it is only minus A by 2 plus A by 2 terms and assumption is, and time should be such that this term should, because this is 1 plus, so obviously this term will be larger than A by 2 and therefore positive growths are expected. Is that clear? I repeat if this term is, even if it is point something, 0.1 plus something is there, which means A by 2 times this will be always larger than A by 2. And therefore positive growths are expected.

So if I write only plus sign, I get x0 is and I put minus inside, so A by 2 into 1 plus t by tau, A square by B to the power half under root minus 1. Now we define some whatever we aim, we define, we actually give some nomenclature to them. And why, we will see soon. We call B as a parabolic rate constant, B is called parabolic rate constant. And B by A is called linear rate constant. This is definition namewise and why we name linear and parabolic, will be soon seen when we will take the two cases.

I repeat B is defined as parabolic rate constant and B by A. In many models, this is given kp and this is given kl, parabolic kp and kl, so most of the earlier CAD tools may be using or even central uses Kp means parabolic rate constant. Kl subscript all is essentially is linear rate constant which is same as what I have been using. This is Deal-Grove model, I cannot change, k is there because that is Deal's model. If it is my model, I can do any other names but this is Deal's model and that is how they are defined and way back in 65.

(Refer Slide Time: 32:43)

Two limiting cases, okay, just look at these terms.

(Refer Slide Time: 32:48)

This term is smaller or larger, depends on this, t plus tau is larger than this, t plus tau is smaller than this, two extreme cases.

(Refer Slide Time: 32:59)

Two Limiting Cases :

(i) Case - 1 : Time of Oxidation 't' is Small, Such that $(t+\tau) << A^2/4B$

Then $x_0 = \frac{A}{2}\left[1 + \frac{1}{2}\frac{(t+\tau)}{(A^2/4B)} - 1\right]$

or $x_0 = \frac{B}{A}(t+\tau)$ — (xxii)

Since for a Process Condition (Gas Flow, Temp. etc) $(\frac{B}{A})$ is constant

$\therefore x_0 \propto (t+\tau)$

or $x_0 \propto t$ x_0 has Linear Growth with Time

CDEEP IIT Bombay

EE 669 L____ / Slide____

NPTEL

So first we say t plus tau is much smaller than A square by 4B. So we can then expand it by binomial term, so 1 plus x to the power half. If x is less than 1, 1 upon half x. So this 1 plus half t plus tau by A square by B minus 1. 1,1, cancels. So x0 is B by A t plus tau. For a given temperature, given gas flow, B and A both are constant. How is x0 related to time? Linear. B is, x0 is proportional to tau, tau is very small. So x0 is proportional to time, what is this growth is? Linear. It is linearly increasing. So now we understood why I named B by A as linear rate constant, because x0 is B by A times t and therefore B by A is named as linear rate constant.

So initially what will happen that what does that mean? If t is smaller, what does that mean? Initially oxide will grow linearly with time and as time increases we believe it will become parabolic. And let us see how that can happen. Is that clear? So initial growth of silicon dioxide is linear with time and then starts, it reduces the rate. Can you think why it will reduce the rate? As the time increases oxide thickness will increase, so the gradient will decrease. Is that clear to you?

So if the flux available is less, whatever is available may react but available is less. So the oxidation rate will go down. So that is the essential as you go. Thicker oxide, it will, much smaller impurities I mean oxidant will reach there, so smaller thickness relative to the time will wave off. Is that point clear why this is happening? Before we come to parabolic, these are few terms to be explained.

(Refer Slide Time: 35:21)

B by A is 2, now we substitute B and A, 2 Defective C star by N1, 2 Defective ks plus this can be rewritten as C star upon N1, ks h by ks plus. In general for a given mass flow in oxidation furnace and for the temperatures which we normally use from 800 to 1200, h is much larger than ks. ks is e to the power minus e by kT kind, minus 4, minus 5. Whereas B will be or h, sorry, h will be order of few centimeter per seconds which is much higher.

So what happens? In most normal cases h is, so which is it limited by? h or ks? Smaller the one limits it. Larger one does not limit it. So h is, I do not say every time, in CVD we will say no, it is mass limited but in this particular case we say h is much larger than ks. So if I do this, I can neglect h, 1 upon h it can goes. So we get B by A is C star by N1 into ks. Is that okay? If h is larger than ks, this can be neglected. 1 upon h is smaller compared to 1 upon ks and therefore we neglect h and we get C star upon N1 into ks.

And as I say ks, how do we define ks? Reaction rate constant. What is reaction rate? Available oxidant reacting with silicon. The rate with which that reaction takes place, it is a chemical process which essentially means which is essentially e to the power minus temperature dependent term. B by A is also linear rate constant which also follows ks dependence of temperature. Is that clear? Bs, B by A is a function of ks, ks follows temperature dependence e to the power. So B by A will also follow temperature dependence same as ks.

So we write then, this is the x0 C star by N1, ks t plus tau. This is called, so initial oxide growth is related to with time, linear growth and it is also limited by available reaction rate at CI. How much is available only can, please remember it is availability of CI I mean is the silicon atoms which can it react. I may have enough oxidant but not all is possible to react. Initially enough oxidant has come but there are not enough bonds where oxygen can, SiO-Si bond can be formed, so the growth rate is limited by available bonds.

So it is essentially decided by ks which is temperature dependent. Typically we will show you the term but I just now want to clarify. Please remember oxidant is reaching enough amount, thickness is very small, much of it diffuses through. Whatever is available, most of it will reach CI. But available oxidation will be limited by the reaction there which is temperature dependent which is essentially how many atoms can react. This is the case I, what is that case we talked? Time is very low, smaller times, t plus tau.

(Refer Slide Time: 39:00)

(ii) Case-II:

When oxidation time is large enough, we have

$t >> \tau$ and $t >> A^2/4B$

Then $x_0 = \frac{2\sqrt{t}\sqrt{B}}{A} \cdot \frac{A}{2} - \frac{A}{2}$

or $x_0 = \sqrt{Bt}$ or $x_0^2 = B \cdot t$

Clearly x_0 is increasing parabolically with time.

$\therefore$ B is called Parabolic Rate Constant

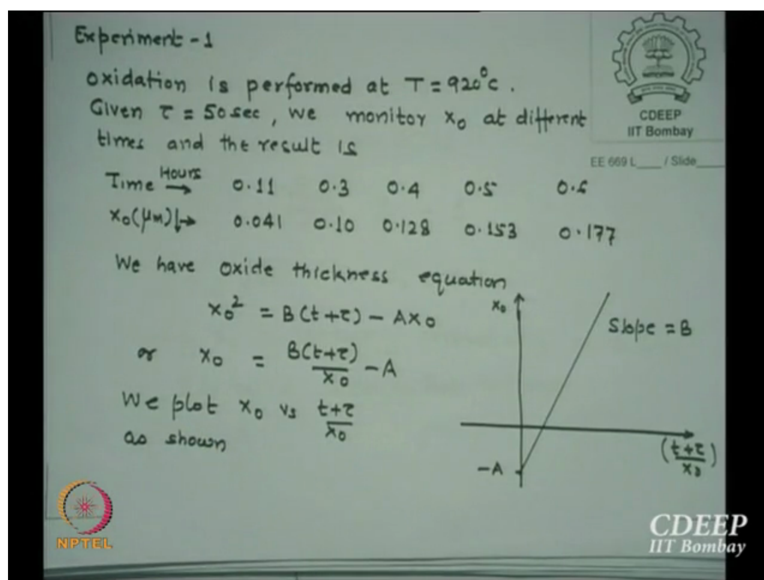
However if t is much larger than tau, and t is much larger than A square by B, longer time oxidation is performed. t will be always larger than tau, tau is very small. So if t is much larger time, t plus tau is t. And then we say t is larger than A square by 4B. Then we can neglect 1 there and rewrite the term x0 as 2 root t B by A into A by 2 minus A by 2. If I readjust these terms, this is just substitute into the equation which we wrote earlier in quadratic solution.

Substitute, t is much greater than this, 1 can be neglected there. Therefore under root of that is root t B and into A by 2 minus A by 2. And if I do this, it essentially comes to be equal to root Bt. We are neglecting small terms, so compared to this everything is small. This term is smaller than this, we already said through this. So we always say x0 is root of Bt which means x0 square is Bt. What is this law? Parabola.

This x0 square is Bt is parabola. So x0 growth is now under root of Bt means it reduces the oxidation rate as, therefore oxide thickness as time is larger and larger initially. So initially linear, and then parabolic growth starts in the case of oxidation. So this is the, and since B is parabolic constant, means is giving a constant of parabolicity, we call B as parabolic rate constant. B is parabolic rate constant. I told you this A square by 4B is much smaller compared to t, so A is much smaller. Obviously this term is much larger than this, so that is neglected.

In actual model you do not neglect anything but it is 100 minus 0.5, think of it whether you want to retain 0.5. If you wish, fine. 99.95 is the answer or 100, your choice. What decimal accuracy you want, 61 precision, 128 bits of precision, computer can do any basic.

(Refer Slide Time: 41:52)



Here is some experiments are performed just to compare those values which I am talking about. I have done an oxidation, experiment means actually experiment was performed. I have grown an oxide or silicon at 920 degree centigrade and assumed that tau given to me is 50 seconds. Now we monitor x0 at different times. I keep growing 5 minute, say time in hours, so this is some 6

minutes and this is 18 minutes, this is 24 minutes, this is 30 minutes, this is 60, 40 minutes. Roughly 36 minutes.

So I have different oxide thickness or oxide times, oxidation times. This is in hours, please remember this is in hours. And I have monitored by some way the oxide thickness. We will see how to monitor oxide thickness. And I monitored, measured them. For 0.11 hour, it is 0.041 micron. 0.3, 0.10; 0.4, 0.128; 0.5, 0.153; 0.6, 0.177. Since you did not want to leave, I used it to show I can take care. How do you get that value, also is as per you used to.

Since x0 square is Bt plus tau minus Ax0, so I get x0 is Bt plus tau by x0 minus A. What is this equation looks like? This is y is equal to Nx plus C kind of equation, linear. So that if I plot t plus tau by x0, this term versus x0, so some way at 0, that is t plus tau, at tau by x0 at this point, whatever is the constant is minus A. So you can monitor that minus A, how much is minus A. So in actual I just want to make parabolic, I said neglected but in calculation I have, I really take care because I need to know A.

So by extrapolating this curve, I will get minus A. And the slope is B, so I could get B and BA and therefore B and B by A, that is kl and kp are monitored if I know actual oxide thicknesses at different growth temperatures at different growth times. Assumption everywhere is temperature is constant and also the mass flows are constant. So if, we can see there that if I do experiment, I will be able to measure B and B by A. And then what do I have to do? This I have to repeat for many temperatures to get B and B by A dependence with temperature.

And then I figured out later, maybe I do not know whether I have graph, I have, okay I will show you. It shows that it is actually following the physics that is ks and diffusivity, is essentially whatever their temperature dependence is, same happens to B and B by A as well. We will see this little later. So is that okay? So I can monitor B and B by A by actually monitoring the oxide thicknesses at different times of oxidations. Is that point clear?

(Refer Slide Time: 45:24)

From experimental data $x_0 - t$ as above,
our plot of x_0 vs. $\frac{t+\tau}{x_0}$ gives

$B = 0.2\ \mu^2/\text{Hour}$ — (i)

and $A = 0.5\ \mu$

or $\frac{B}{A} = 0.4\ \mu m/\text{Hour}$ — (ii)

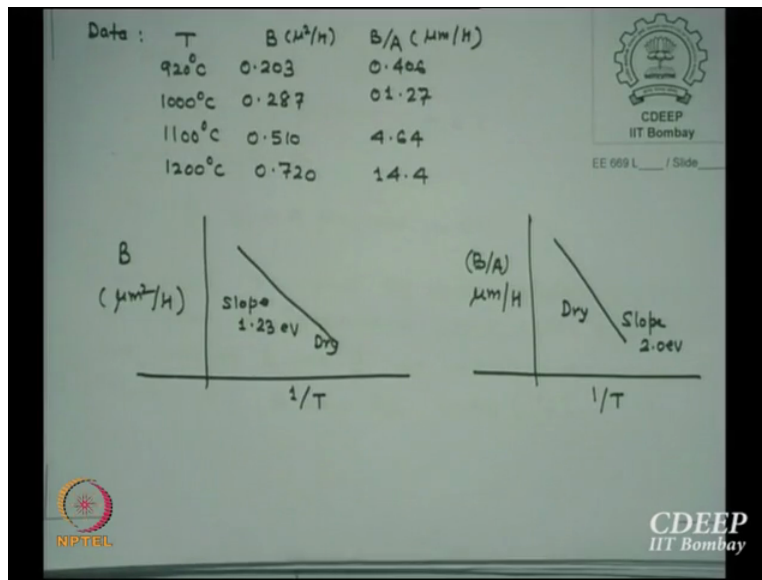
Experiment-II We repeat the above oxidations at same times for Temperatures $1000^\circ c$, $1100^\circ c$ & $1200^\circ c$ and evaluate B and $\frac{B}{A}$ at various temperatures
We then Plot B and B/A versus $(1/T)$

So let us say by typical experiment which I did or rather someone else has done, of course I have calculated but this data was taken from our lab many years ago. So B is 0.2 micron square per hour, A is 0.5 micron and B by A therefore is 0.4 microns per hour, 0.4 micron per hour. So I just now said if I know the data, I will be able to plot time versus x0 by this. And if I know this, I will be able to evaluate B as well as B by A.

Now as I said you I will repeat this experiment, what do I do? At different temperatures. Again oxide thicknesses for, only thing is now I may do it for a given time, I mean same time so that, but even if you do different time, graph will show the slopes. So it does not really matter as long as you, but preferably you do same so that you know where the slopes are moving, just to see them. Why we want to do this? Because I want to know whether B and B by A really follows D and ks. That is what we are doing from the experiment, from the theory we are looking, that B by A is following ks and B is following D effective.

So can, does that temperature dependence appears, so we like to see that. So we actually do repeated experiment at 1000, 1100 and 1200, monitor oxide thicknesses at same times and replot B and, replot x0 versus time and get B and A for all of them, for different temperatures. This is for 920, repeated for 1000, repeated for 1100, repeated for 1200.

(Refer Slide Time: 47:33)

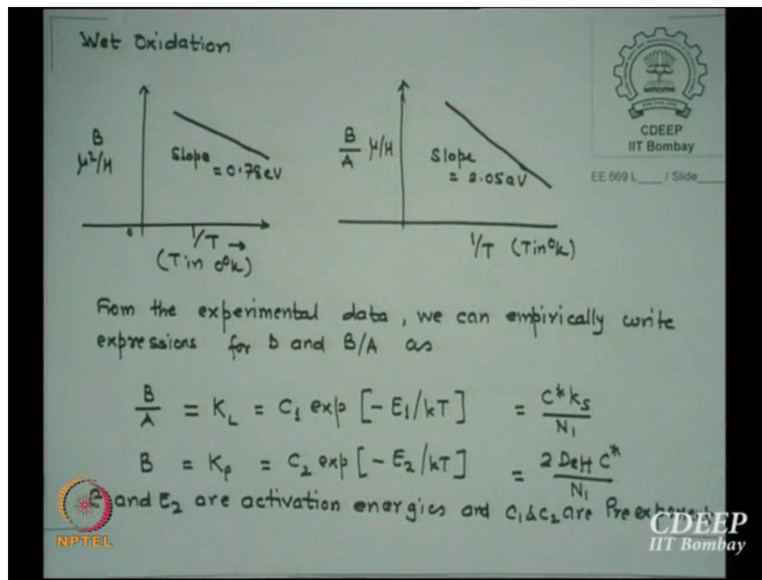


If I do this, the data I get is 920 is this and 1200 is this. This is B by A, please remember units. B is expressed as micron square per hour and B by A is expressed as micron per hour. Linear and parabolic words square. So if I plot now B versus 1 upon T for dry oxidation, of course I have not talked about this but we will come, dry means only pure oxygen is passed. So for the dry oxide and I plot B versus T, 1 upon T, I see a straight line.

And its slope is 1.23 electron volt. If I plot D versus temperature, 1 by T in fact, and I actually see for dry oxide case, the slope is 1.23 electron volt. If I repeat the same graphical this for B by A versus temperature, I plot B by A versus 1 upon T and I get slope of 2 electron volt. This is experimental because what I did, I actually went in and I did oxidation, monitored the thicknesses at different temperature for different times and plotted them to get B and B by A at different temperatures.

So this is no model, this is essentially what I can measure. And now I want to prove that what I said in a model probably fits to what I monitored. Therefore Grove-Deal model is reasonably good. Is that point clear to you?

(Refer Slide Time: 49:38)



So B and B by A for dry, we do wet oxidations in 95 degree water wafers. So I did it, we did same experiment for wet oxides and I find for B the slope is 0.78ev and B by A has slope of 2.05ev. Let us look at this term again, B and B by A. Okay, if you have noted down, as I say B by A is also called KL and B is also called Kp in many simulators. And that can be written as C1 exponential minus E1 by kT, C2 exponential minus E2 by, this is the model they have used.

And if you see here, these are the yields, E1 and E2. This is the temperatures, slope you have got it. C star into ks by N1 is, that is what we have just derived. B by A is C star ks by N1. B is D effective by C star. ks, C is constant. N1 is constant. So if this temperature dependence has to come, ks must have similar relations. If this relation has to be followed in this, D effective must be for, I mean E2 must be following the activation energy associated with D effective. Is that point clear?

If these are equals from the graphs if I say, then the since here they are constants, only D effective is temperature dependent, ks is temperature dependent. So obviously this E1 must confirm to ks and E2 must conform to D effective. And yes, this experiment was further extended and once we did this, we found that they did which essentially means what? The linear rate constant B by A essentially is governed by reaction rate constant which has temperature dependence of E to the power E1 by kT where E1 is 1.23 for dry oxidation, 0.78 for wet oxidation.

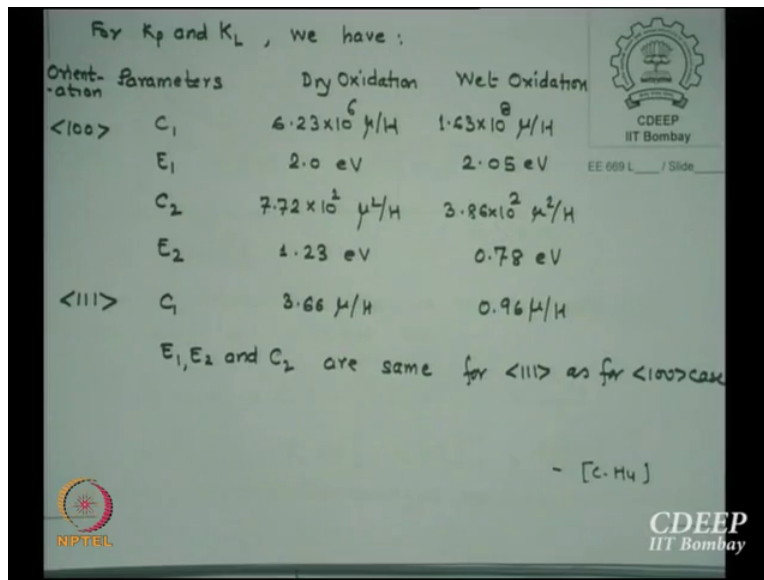
So ks actually follows what the growth is, this. Diffusivity of oxidant for dry or wet, you can see from here, sorry, this was ks and this one. For B we find it is 1.23 and 0.78 and we figure out D effective have the same energy of oxidant in oxide as such we know and same is reaction rates we know about in real life. We actually monitor by different methods. So we figured out that this process is reaction rate limited and this process is diffusion limited and which is obvious.

Initially when the oxide thickness is small, the available oxidant is enough, it is the reaction possible at a given temperature to convert silicon into silicon dioxide. Si plus 2O2, SiO-SiO bond has to formed. Now this at a given temperature is, this is the reaction, so it has to, this reaction will be temperature dependent. That is what exactly we did experiment and we found yes, it does depends on this.

However when I increase the time, that means I have larger time, by then already oxide thickness has grown. So the available oxidant now at the interface of silicon, silicon oxide is smaller and it is this as much as you can diffuse through is going to available for oxidation. Any amount coming I will oxidize but available is only what you will supply. So it is the diffusivity which decided how much oxidant I can provide, initially everything available, so reaction is following.

In next time reaction can be done any amount but available kitna, so that means the reaction rate constant is dominant in linear initial times and parabolic rate constants are dominant in thicker oxide times.

(Refer Slide Time: 54:05)



This is called Hu's paper way back in 71. Very famous person, now we also we figured out that the oxide thickness for 111 plane is different from 100. We have shown the last time that oh, maybe we have to show, sorry we have not shown. So why do you think that C1 constant is larger for compared to this? The reason why 111 shows oxide thickness thicker, can anyone suggest why it cannot be thicker? Because more atoms on that plane is, the question asked was exactly this Miller planes.

Miller plane stain how many atoms are available on that to react. So 111 shows the maximum silicon atoms, 4 of them in fact are four corners and 3 inside. Since they are the largest number, the growth rate is highest along the plane. This is exactly what is data has been shown, taken from (())(55:13).

(Refer Slide Time: 55:21)

Based on this I can again show you this, same time once again and again. The slope activation in ks and activation energy related with the diffusion is found identical to what is measured for B by A and B which verifies B by A is ks limited and B is D effective limited, identical. The energy associated with ks and 1 upon T and D effective 1 upon T, their slope essentially matches with the actual data measured by B and B by A which means the Grove-Deal model to a great extent is valid, except for the assumptions which we may have to modify as things go.

But for a thicker oxide less than say, thicker oxide around 100 Armstrongs or above, Grove-Deal model fits very well. Anything below 100 is not true. For D effective and ks, for whom, ks will be had different activation energy which is essentially what B by A we got. B by A is proportional to ks. So if I monitor ks by, not by B by A method, by actual reaction by thermodynamics, so if I evaluate thermodynamically this equation, I get whatever activation energy associated. I figured out that is same as what B by A I got by experiment, so which means B by A is case limited.

I did same thing for diffusivity experiments from the chemical point of view and whatever energies I found for both wet and this, I matched it with B values and I finally I got the same within errors, within experimental errors.

"Professor-student conversation starts."

Student: Which energy should they have in experiment?

Professor: Activation energy is the energy required to react, is essentially, it is enthalpy. See the, yeah, any reaction whether binding or dissociation or enthalpy formation is essentially related to energy. If A plus B has to go to C plus D, then the reaction is favored forward if the Gibbs energy is plus, that is enthalpy minus entropy, T delta S is positive. If enthalpy minus T delta is negative, dissociation B, C plus D will go back to A plus B.

“Professor-student conversation ends.”

We will see in CVD. This is where we adjust the growth. If I want growth, what should I do? A plus B should be stronger and delta G must be positive. So A plus B will react. When I want to etch, what should I do? I do not want reaction, I want etching. So I see to it, C plus D goes back to A plus B. That is exactly what we do to depositions and etching are identical, the reaction are favored or not favored. Is that point, if something is formed, I reverse it, do not form. Remove it. So that is exactly what I am trying to say. So what essentially I am trying to say you, that Grove-Deal model to a great extent is a good model except for as I say very thin oxides.

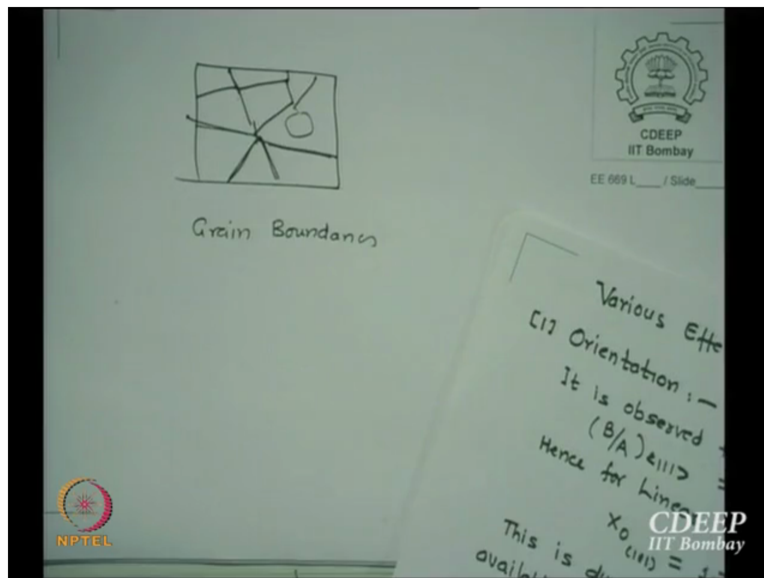
(Refer Slide Time: 58:45)

If I see orientation, as I already said B by, orientation is only, afterwards it is the diffusivity. In thick oxide, how much is available is going to decide. Only in thinner oxide, available bonds will decide available reaction. So we know B by A is 100, 111 is at least 1.7 times B by A of 100

constants, I already given C1, C2, you can see value. And this is also essentially the ratio of bonds of 111 to 100. So initial oxide will be thicker for 111 compared to 100 by 1.7 times.

Afterwards why they will, it will not this because availability of oxidant is going to decide and not the rate. So then 111 and 100 will have same thing as diffusion limitations. But initially available bonds to react at a given temperature will be decided by which kind of planes you have. We also have in real life polysilicon as gate and we will see that later, the growth of oxide is different from poly compared to crystallines. We are not done poly so far, so I do not want to preamp but just for the heck of it, I may show you something which is of…..

(Refer Slide Time: 60:12)



See poly crystals, say let us say this is crystalline, but its orientation is different at different x, y, z. There are crystallites but there are many of them, poly, large numbers. So in this part they may be crystalline but their orientation will be different compared to orientations here and here. Some may be larger crystallites, some will be smaller. The line between two such crystallites is called grain boundary.

Each crystallite is called grain, so these are grain boundaries. That is two crystallites are meeting at that point, that is why it is called grain boundaries. Now grain boundaries do not have crystals there. So it is like a void in the system. So if you push something, it may not go through crystallite but may actually go through the grain boundaries. So more and more atoms may be possible to be oxidized because much of the oxidant now will not necessarily go through silicon

or silicon dioxide layer on the top but through grain boundaries and react. So do you expect thicker oxide for poly? Because more deeper, at least thicker oxide because it will, oxygen can go deeper in the polycrystallines. Polyoxide is mini-crystallite and hence many grain boundaries.

(Refer Slide Time: 61:51)

Various Effects on SiO_2 Growth Rate

[1] Orientation :-

It is observed that

$(B/A)_{<111>} = 1.7\ (B/A)_{<100>}$

Hence for Linear growth regime for same time

$x_{o(111)} = 1.7\ x_{o(100)}$

This is due to the fact that no of atomic bonds available on Surface of Si at SiO_2-Si interface are 1.7 times larger for (111) case than (100) case.

[4] Poly Oxide : Poly Silicon has many Crystallites and hence many Grain Boundaries. Oxidant diffuses faster through Grain Boundaries leading to enhanced oxidation rate. $x_o = at^n$ $(n > 0.5)$

Oxidant diffuses faster through the grain boundaries leading to enhanced oxidation rate and the model which is suggested is a t to the power n where a is fit constant, n is also fit constant. Kyaa karenge? Iska oxide lenge, time se plot karenge aur iss function ko fit kar denge. So we say okay, this is the law polyoxides follow. It does not follow Deal-Grove model. It follows another law. Is that okay? Why it is so? Because we have not yet studied very strongly the diffusivity through grain boundaries.

There are many micro crystalline theories, some other day, if you are really working PhD for that, then I will show you how even this model is not correct. But as of now we will not discuss. Is it okay? So we say polycrystallite, crystal, poly when oxidize, it oxidizes thicker compared, faster therefore compared to normal or silicon. And the formula it fits into is x0 a t to the power n, n is typically more than half which essentially is closer to parabolic. n, greater than half is parabola. So it is slightly higher than parabola, faster growth.

This is like Bt, okay, so as root Bt. So essentially it is saying if you do that, it is essentially parabolic growth because thicker oxide is growing. But n is not half, but it is larger than half. Is

that clear? Thicker. And there is no thinner oxide, because by then grain, enough oxidant will be made available and it will be decided only by this available this.

(Refer Slide Time: 63:53)

The next effect for us, under pressure the crystals are stressed leading to enhanced effect. Whenever you put a wafer inside pressures and there is an experiment done for water wafer or water solution and we actually put high pressure on that and their oxidation rate, Jayraman's experiment very popular one. This pressure effects are coming back in some other way now in FETs as well as MOSFETs. For the pressure we are talking, strains.

Pressure is essentially stress proportional to strain. So when the lattice is not matched, then it is a strain which is essentially due to the pressure of, that is stress. Stress is pressure. So force per unit area is pressure. So strain is essentially now coming back once again, to improve your motilities. But in oxide, this is not crystalline, it is oxide. In oxide more bonds are available because under pressure lattice is heavily pressed and it breaks actually.

Lattice porosity is not 100 percent, so more oxidation is possible. So formula was figured out in linear growth regime, B by A at any pressure is B by A at atmospheric pressure times p to the power n. Typically the pressure which we used by Jayraman earlier was 1 to 3 atmospheric pressure and n was found to be 0.7 to 0.8, again fit function. So all modeling people very happy, they fit. Experimental data fit.

Then you can say that I could have fitted something, p should be some numbers, n should be some, I can fit any in such combination. But I know roughly that how much pressure I can put actually before it actually breaks. So I figured out up to 3 atmospheric it sustains. So I fixed those value in between and for those values how much n it can fit to the data? So some physics was brought afterwards. After I really see what is happening, I say okay, isko funda bhi maar dete hein, which is not true because there are n combinations I can make out of it. But that is all modelers do including person like me.

Then there is an oxidation which many times we will have to do which is called oxidation due to impurities present in the oxide. Doped have impurities present in like, heavily doped silicon has different oxidation rate compared to lightly doped silicon. So there is a difference in growth rates and typically the fitting function. For doping greater than 10 to power 9 per, 19 per cc. Where do you think these numbers will appear? Which area of MOSFET? Source-drain.

So in source-drain the field oxide growths will be thicker naturally because of B by A is 20 times B by A undoped, 2 times B by A for p-type. Beta for B is doped, is twice B type undoped. And B doped is 0.04 times undoped for n and p-type. This is the fit data again. These numbers do not actually matter but close to these numbers can be fitted. In real life it may become 0.30, 0.0385 but it is okay to say 0.04.

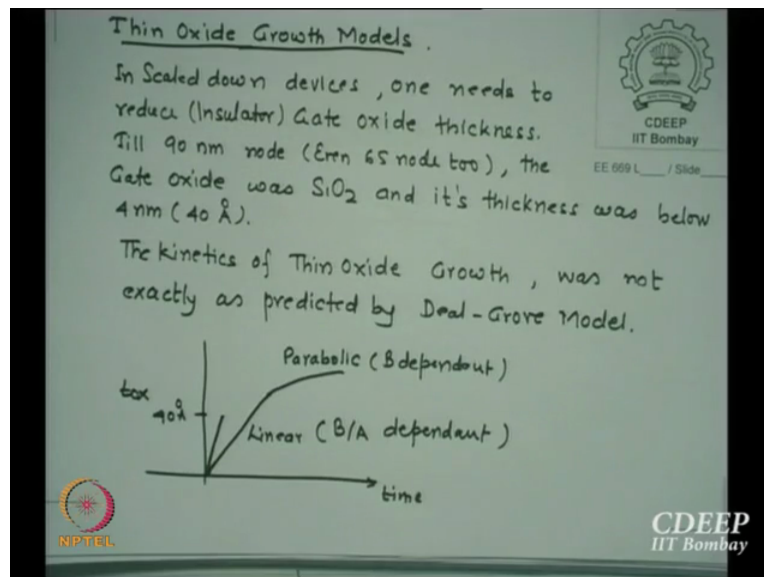
The real data may not be exactly this numbers but this is good enough because most of the times we are not very keen, we not actually just do by, we actually monitor. So I know how much oxide thickness I have there. So I may not make mistake in future this but to model I cannot do very very accurate model. If I have to, then whatever, every time I get experiment data for the lab, I really try to fit in this and actually get the fit models.

For this lab or these furnaces, these everything I know this data which may not be universal, only for this people. This is how all people do. We get spice parameters in the circuit simulator, all kinds of effects are taken care, all physics is introduced by fitting. What do you mean by fitting? I mean the physics cannot be fit, but that is what it happens. As long as IV, IDS, VDS, characteristics appears, circuit performance appears, who cares whether it is k to the power or not to the power.

This function fits, that is what we want. But if I have to understand what is happening, then I, is the mobility dependency is how much is correct, is the oxide thickness is varying, how much. Then I will start Bt is constant or not constant, how much available there. So I will now put physics and since it does not fit, I will put some constants, exponential constants so that it fits to the lab data for a given technology.

Then I have a circuit simulator with modified model card, no one starts to be putting my model card, so I will put my model card. That will give me the accurate result, this is how all designers do. Whatever actuality I may, they will fit to a value which fits to the data and then of course many of us will like to physics. So force physics on it but that is more interesting to see.

(Refer Slide Time: 69:40)



Before we quit last this, all these models since it starts with thin oxide assumption, if you are going thin oxide, then what is the model? Since we are scaling down the technologies at least below 90 and even 65, the gate oxide thickness is now reducing below 40 Armstrongs and in the 28 nanometer or 22 nanometer, it is less than 5 Armstrongs, it is even close to 1 Armstrong now. 1 Armstrong of course you cannot grow and that was the question asked in your quiz.

If you scale down technologies, the oxide thickness are becoming less than 5 Armstrongs and then there is one monolayer which you cannot grow. And therefore you need high k, so that capacitance remains constant, epsilon A by t. So increase 1 epsilon, increase t, the ratio remains

same. So use high k, here we are not talking of high k. We are just saying it is reducing. Till very late we were working on thin, 65 even there is SiO_2 thickness, gate insulators have been used.

Below 14, 4 nanometers, I started using nano because these days everyone must talk nano, not Armstrong. So I also wrote 4 nanometers. The kinetics of thin oxide growth was not exactly as predicted by Deal and Grove. So for example, thicker it may still work because then it is not thin oxide. So we are not interested to know something about parabolic because by then it is already thick. Where is the worry? The linear portion because there the thickness is smaller.

So we figured out that if I see the Grove-Deal model and if I see the actual thickness, it is somewhere different slope which essentially means thinner oxides do not follow exactly in the very thin oxide regime the Deal-Grove model. To reduce the oxide thickness what should I reduce? One is of course time, but time cannot be 1 second. I mean you cannot push and take out, you can do, so you will put some minutes. So what can I reduce? Temperature.

So I actually reduce the temperature, I do not go oxide at 900, 1000, 1200 but I go at 800 possible. You can say why not 600, there is no reaction. The reaction minimum temperature for SiO bond to form is 800. So the thinnest of oxide can be grown around 800 degree centigrade which is around 1073 degrees kelvin. Please convert everywhere into kelvins. So there are models which are available in market. Quickly we will show you and…..

So is that clear to you? For thinner oxide regime, this is what we see, this is what Grove-Deal predicts. One of my PhD student way back in 80s did this work which model I will show you later.

(Refer Slide Time: 73:05)

There is first model which appeared was Reisman's model which has some modification they did. x0 is a t plus t in to the power this. t in is a fit parameter which is essentially xi by a. a and b are constants. As a fit parameter xi is also fit parameter. I get the data and fit it into this model and we say okay, Reisman says if you grow this and fit into this, it will fit. So fit get a, b and xi from the actual data and use this model for your lab, for your CAD tool.

You can see everywhere what we are doing is fit. So someone should say is it b by a or is it what? No, I do not know what is there. It is a and b. Then there was Han and Helms model which, he say there are two parallel reactions are going on in this. So he says B1 upon 2x0 by A1 and B2 upon 2x0 by A2. Now this you have this B1, B2, A1, A2 as fit parameters and try to adjust x0 versus time using these two parallel terms.

So if this does not fit the data, you add from here some term and see it may be minus as well, sign B2 may be minus, depends on the fit you want. And you see that it fits to available data. This is called Han and Helms model. The most important model which was accepted for quite some time was from Plummer and his student Massoud and Massoud has the best of data as well in way back in 90s. Massoud published much of the experimental data for thin oxides as well to thick oxides but and his data has been taken as most standard data by almost all industries. This was work done at Stanford. Yeh doh term hain, yeh aapne pehle likha naa, ekk aur term add kardi meine.

(Refer Slide Time: 75:34)

Then Massoud and Plummer has a model, they say okay, the second term instead of B2 by something, they I added a term called e to the power exponential x0 by L, again to, and they expected that the L value which they will use should be less than 70 Armstrongs beyond which this is not valid, thicker oxides. So they added another parameter here C and L to fit the data and that is, that was accepted for many years or rather even now the first attempt is to use Massoud's data and Plummer's model. We also did some more kinetics on that for thin oxide. Our work was published in JAP 1989, myself, Vasi, and More, my PhD student. So we suggested some equations, that S is called the available site.

There is a term which we created called site, so we say oxidant plus site may form oxygen site combination. Oxygen's O2S plus another site may form 2OS at, with a constant, I mean proportionality constant K2, reaction rate constant. Then we say Si plus, Si-Si bond will react with OS to form Si-O-Si bond and create another site. That is how oxide will keep growing, thin oxide will keep.

We also introduced many constant in this expression, P is some constant, omega into silicon-silicon bond times n K3 into K1, K2. These are called constants of reactions. Q is K1 and R is K1, K2. This of course those who are very keen can see our paper of 1989. We fitted this data for 800 degree and 900 degree centigrade with Massoud experimented data, which he published for different pressures but 0.1 atmospheric pressure is the best result they claim.

So we fitted our model to this and by making proper choices of, K1, K2 also we derived what values we should have. And based on our analysis of this we could get P, Q, R values which fits into thin oxide. So we could fit data up to 20 Armstrongs of oxide thickness. Below 20 of course our model also did not fit. 20 to 40 Armstrongs or 60 Armstrongs our model fitted very well with the experimental data then known. So if tomorrow someone wants another model, you can always try something.

So you should be able to find some reactions, what should be the real materials going on, what bonds can it create, what is the binding energies available, what is the space charge around. So there are many things which you can think and add on to a model. Is that okay? So this finishes the modeling part, next time we will do oxidation techniques and characterization.