Electronic Materials Devices and Fabrication Layering: Thermal Oxidation Dr. S. Parasuraman Department of Metallurgical and Materials Engineering Indian Institute of Technology, Madras

Module - 01 Lecture - 23 The Diffusion Equation

In last class, we looked at an over view of the patterning process. We saw that in the case of IC fabrication we can divide the processes into four main kinds: we had layering, lithography or patterning, doping and heat treatment. If you think of an IC fabrication as an assembly line process, we start with a blank silicon wafer silicon is typically the material that is used Gallium Arsenide is also used.

So, we start with the blank wafer which goes through these various processes and you get finished IC circuit out. We also looked at an example of fabrication of a MOSFET metal oxide, semiconductor, field effect transistor, where we started with the blank wafer which went through this different steps could be layering, could be patterning, doping, heat treatment, is included in all of this to give you the finished MOSFET.

So, starting from today for the next few lectures we look at each of these processes in detail. Today we are going to look at layering and in layering we are going to look particularly at grown film; that is films that consume the underline silicon.

(Refer Slide Time: 01:42)



So, today we are going to focus on the oxidation of the underline silicon film. So, the ability of silicon to form silicon dioxide or SiO2 is very important. If you remember when we talked about the introduction to integrated circuits, the first circuits were made of germanium, but later when ICs came into being silicon was used as the material of choice. 1 of the reasons of course, is that silicon is abandoned.



(Refer Slide Time: 21:21)

So, it is easy to manufacture, but more importantly silicon can also form SiO2 and SiO2 is a very good insulator. SiO2 is used as the original dielectric material for the

MOSFETs. So, it is access a dielectric between the gate and the semiconductor of course, later which scaling Silicon dioxide was originally replaced with oxynitrides. And finally, we have high cay dielectrics, when we look at Silicon naturally has a native oxide on the surface.

(Refer Slide Time: 03:00)

The Native oxide has a thickness typically around 3 nanometers, Silicon dioxide performs a couple of functions: 1 it helps in passivating the surface so, this passivation can be both physical and chemical. So, this is very important because surface passivation will also affect the electronic properties. So, whenever we think of a surface it could be silicon 1 1 1 or silicon 1 0 0 there are always some dangling bonds.

Dangling bonds are usually defect states, that lie in the middle of the band gap and they can affect the electrical properties. So, Silicon dioxide helps in passivating these dangling bonds so, that that improves the properties. So, this passivation can be physical and chemical. Silicon dioxide is also important because, it is used for patterning the substitute.

(Refer Slide Time: 04:36)



So, it is an example of hard mask so, Silicon dioxide is a hard mask because it can basically survive at high temperatures. So, in the example of MOSFET that we saw in last class the first layer we group was the field oxide which is nothing, but silicon dioxide. So, if you want to do doping at high temperatures and you want to do doping only in specific regions then, usually some sort of an oxide layer is used for pattern.

But the native oxide layer we saws 3 nanometers thick is usually too thin for doing any of the patterning processes. So, a thicker oxide layer has to be grown on silicon and this is essentially the oxidation process. We also saw briefly that there are 2 kinds of oxides that can be formed.

(Refer Slide Time: 05:36)

The first 1 is dry ox in this particular says silicon reacts with oxygen gas usually this is at high temperature to form SiO2. The other kind of oxide which called your wet ox in which case silicon reacts with steam again at high temperature to give you Silicon dioxide plus Hydrogen so, both of these processes required high temperature and they both consume the silicon.

Now because, silicon dioxide has a different density compare to that of silicon there is always some volume expansion, when we have an oxide layer forming on the surface. To look at, that considers a silicon substrate with an oxide layer on top.

(Refer Slide Time: 06:57)

So, this oxide layer is a grown oxide so, that it consume some of the silicon. So, this is SiO2 so, the dotted line represents the original interface of the silicon. So, this much amount of silicon has been consumed in order to grow the oxide. So, this is the original silicon interface d is the amount of silicon that is consumed typically a thickness and d prime is the thickness of the oxide layer.

Then, d is the silicon that is consumed informing this oxide layer. So, one can find a relation between d prime and d while simply looking at the number of moles of silicon consumed in the number of moles of SiO2 formed. To do this we just need some physical properties.

(Refer Slide Time: 08:30)

So, first is density in grams per centimeter cube, silicon in SiO2 silicones density is 2 33, SiO2 is 2.65. We also need the molecular weight for silicon this will be the atomic weight so, silicones atomic weight 60.08. So, if d is the thickness of the silicon consumed and A is the cross sectional area; A times d is the volume of the material. The mass is nothing, but density of silicon times the volume or the number of moles.

So, Z of Si is nothing, but the molecular weight let me just make a notation here so, z. So, Z of Si is the atomic weight of silicon and this is the mass of silicon that is consumed. Similarly, we can write another expression for the number of moles of SiO2 formed, the thickness of the SiO2 layer is d prime. So, d prime is the thickness of SiO2 and d is the thickness of the silicon that is consumed.

These numbers of moles are essentially equal because, 1 mole of silicon gives the 1 mole of SiO2. So, if you equate these two area is common we can rearrange the other terms and substitute the values from, which we get d prime by d is nothing, but 1.88. So, if you take an example let us say we want to grow 100 nanometers SiO2 that is the target thickness. So, this will be d prime corresponding value of d is just d prime divided by 1.88.

So, it 53.2 nanometers of silicon is consumed. So, there is always some volume expansion when you consume silicon and you grow a layer of Silicon dioxide on the

surface. Silicon dioxide growth on silicon is called a thermal oxidation process so; we can have a simple model of this thermal oxidation process.

(Refer Slide Time: 12:57)

So, we saw that there were two ways of drawing Silicon dioxide you can either have a wet oxidation or a dry oxidation. In both cases, the typical operation temperatures are somewhere between 900 to 1200 this again depends upon the thickness of the silicon dioxide layer that we want to form and also the process time and so on.

So, later we look at some numbers and try to compare both wet oxidation and dry oxidation; in both cases a Silicon dioxide layer is formed on the surface. So, for any further oxidation the oxidizing species, whether it is oxygen in the case of dry ox or steam in the case of wet ox has to diffuse through the Silicon dioxide layer reach the silicon interface in order to form further oxide.

So, we can look at a basic model for oxide growth so, we first consider the gas phase; to the gas phase has the species that forms the oxide. Then, there is an oxide layer that forms on a silicon and then ultimately you have the silicon. Let me, just write this here so, we have 2 interfaces: 1 is the interface between the gas and the oxide layer. So, this oxide layer can either be a pre existing layer or you could be somewhere at the middle of the process. So, that you have an oxide layer of certain thickness.

Let us, call this thickness d0 we have another interface between the oxide layer and the silicon. So, if you look at the various steps of oxide growth, the first step is that the oxidizing species has to be transported from the bulk to the oxide gas interface. So, I will use iF in future to denote an interface. So, the first step is the oxidizing species has to be transported from the bulk.

So, let CG be the bulk composition and CS is the surface composition so, they must be some transport of the species to the interface so, that it can diffuse. So, let F1 be the flux corresponding to it, the next step is the oxide has to be the diffusing species has to be transported through the oxide layer.

(Refer Slide Time: 16:45)

So, you need to have diffusion through the oxide layer so, this could have some Flux F2. So, C0 could be the concentration at the gas oxide interface this is the concentration in the oxide layer and from here it goes to some concentrations Ci, which is the concentration at the interface. And then finally, we can have reaction with the silicon to form a new oxide layer this has the Flux of F3. So, if you look at this system in steady state.

(Refer Slide Time: 18:00)

(1) Reaction ALL Si Flow F
809	

F1 is equal to F2 equal to F3 so, that whatever species that diffuses from the gas phase to the gas oxide interface gets dissolved in the oxide. And diffuses uses to the interface, were it reacts with the silicon in order to form a SiO2. 1 of the assumption that the model is that, there is no dissociation of the species in the oxide. So, that whatever is in the gas phase get directly transported or directly diffuses across the oxide to the interface.

So, we can actually model these various processes and using the relation between F1, F2 and F3. Derive some numbers for how the oxide layer goes a grows as a function of the diffusion coefficient and also the thickness. So, let us take a look at that so, let me just redraw the diagram.

(Refer Slide Time: 19:18)

So, I have a gas phase, I have an oxide phase and then I have silicon; the oxide has some thickness d0, the gas has a bulk concentration CG and there is a surface concentration CS within the oxide. There is the concentration C0 at the surface and Ci at the interface, we also saw there were 3 fluxes: F1 F2 and F3 and F1 equal to F2 equal to F3 in steady state.

So, consider the first process which is the movement of the diffusing species from the bulk of the gas phase to the surface. So, this can be simply written a some constant hG times CG minus Cs hG is nothing, but the mass transfer coefficient in the gas phase. So, we can rewrite this in terms of the composition of the diffusing species within the oxide layer.

If you do that, F1 is h times C star minus C0; C0 we have already seen is the concentration at the interface of the gas oxide layer. It is the concentration of the diffusing species in the oxide layer; C star is the equilibrium bulk concentration in the oxide. We can usually think of this a some solubility limit so, on the case of dry ox, where dry oxidation the gas species that is diffusing is O2 in the case of wet oxidation the species that is diffusing a steam.

So, C star is your bulk concentration in the oxide h is related to hG by Henry's law. So, is hG divided by hK above when time temperature where h is the Henry's law constitute.

So, Henry's law typically relates the amount of gas that is dissolve in any solid to the partial pressure of the gas.

(Refer Slide Time: 22:48)

So, this by these expressions we have written F1 which is the flux in the gas phase in terms of the concentration of the oxide layer, concentration of the diffusing species in the oxide layer. Next we can look at diffusion within the oxide layer.

 $\frac{C_0 - C_1}{d_0} \quad D = diffusion \ coef}$ $\frac{D}{d_0} \quad D = f(\tau) = D_0 \exp\left(-\frac{E}{d_0}\right)$

(Refer Slide Time: 23:15)

So, this we can usually write in terms of some diffusion coefficient F2 which is D time C0 minus Ci by d0. So, d0 here is the thickness of the upside layer and I say mentioned earlier you can either, start with an upside layer of the surface or you look at the process

at a specific time, where you have a certain oxide layer that is grown. C0 and Ci are the concentrations at the 2 interfaces D off course, is the diffusion coefficient.

D is the function of temperature and is usually written as D0 exponential minus Ea over KBT which is the standard expression. So, D0 is a constant is diffusion constant and Ea is the activation energy. So, the values of D0 and Ea will change depending upon the diffusing species. The last flux term is F3; F3 is related to the rate of oxide formation.

So, it is a rate at fix silicon reacts with a diffusion species to form the oxide an F3 is usually written as some rate constant ks time the concentration at the interface. So, ks is the rate constant for the Silicon; Silicon dioxide reaction.

(Refer Slide Time: 25:18)

So, the equating F1 equal to F2 equal to F3 it is possible to get the expressions for both C0, which is the concentration of the surface and Ci and also, how the thickness of the oxide layer changes as a function of temperature. There are usually 2 extreme cases when we at oxide growth on the surface.

So, if look at oxide growth on the silicon surface there are usually 2 limiting cases now, diffusion in the gas phase which usually fast. So, that that is usually not rate limiting step, but 1 kind of process is where your rate limiting step is the diffusion of the species within the oxide layer.

(Refer Slide Time: 26:00)

This is called the diffusion controlled case or simply diffusion controlled role, in here the rate limiting step is the diffusion of the species through oxide layer to the silicon interface. So, the supply of the oxidizing species to the silicon SiO2 interface controls the over al rate of the process. So, this is usually the case when we have a thick oxide layer and at high temperatures where the rate of reaction is passed.

The other extreme case is called Reaction Controlled so; in this case the conversion of silicon to Silicon dioxide; that is the rate of the process is the limiting step. So, Si to SiO2 is limiting usually; this is the case when we have a thin oxide layer and may be not a sufficient temperature for the reaction to happen. So, in this case both are essentially extreme conditions and they can actually be obtained with solving the general equation by equation the 3 fluxes.

So, we can actually write a general equation that relates the thickness of the oxide layer to the time. So, this is obtained by solving for F1 equal to F2 equal to F3 the assumption is there is starting oxide layer of a certain thickness on the silicon before we start oxidation. So, the final solution I will just write the general equation will not go through the steps relates the thickness d0 to the time, A,B and tau are some constant.

So, d0 here is the oxide thickness at time t A is equal to 2d so, d is a diffusion coefficient ks is the rate of reaction and h is the mass transfer coefficient, which is related through the mass transfer coefficient in the gas phase by Henry's law. B is also a constant B is

2D C star over N1. So, d again is the diffusion coefficient, C star is the bulk concentration of the diffusing species in the oxide and N1 is a number of oxide molecules incorporate is number of molecules incorporated in the oxide layer.

So, this is again the diffusing species or let me just write diffusing molecules. So, this will be different whether you have Oxygen as a diffusing species or water as the diffusing species. Tau is nothing, but di square plus A di by B and di is the oxide thickness at time t equal to 0 so, this is the initial oxide thickness.

So, this is a general expression that relates the thickness of the oxide layers to the various parameters in these right equations. So, we saw there are two extreme cases: 1 is diffusion controlled case; the other is reaction controlled case if we can take this expression and simply it to obtain various relations or different relations for these 2 cases.

(Refer Slide Time: 31:25)

In the case of diffusion control so, here the diffusion of the oxide species through the oxide layer is what that matters. So, this is usually for a thick oxide in this particular case t is much larger than tau and t is much larger than A square over 4B it will uses assumptions in the general equation. So, let me write the general equation down.

If we use assumptions in the general equation, this essentially simplifies to d0 square equal to Bt this is called a Parabolic rate law. Because, if we look at this equation

resembles the equation of a parabolic; to see how we get this from the general equation we can rearrange the general equation.

(Refer Slide Time: 32:54)

You write this is d0 A over 2 is equal to 1 t plus tau a square over 4B whole power half minus 1. So, this is just rearrangement of this equation. So, if t is much larger than A square over 4B and t is much larger than tau. This equation will essentially simplify, that ignore this 1 and the other 1 and we are left with this d square equal to Bt or a parabolic rate law.

(Refer Slide Time: 33:45)

On the other hand, when you have a reaction controlled case. So, here the diffusion is passed, but, the reaction with the silicon is slow. So, in this particular case t plus tau is much smaller than A square over 4B. So, the final expression is d0 equal to B over A t plus tau which is a linear rate law. So, based upon these equations conclusion is as you have a thicker oxide layer it takes more time to grow.

(Refer Slide Time: 34:51)

So, let us put some numbers in so, consider the case of dry ox so; in dry ox the oxygen is a specie that is defusing for a typical temperature of 1100 degree centigrade, the constant B symbol 0.0117 micro meter square per hour. So, these values are usually tabulator for different concentrations, in different pressures and so on. So, have to glow and oxide layer d0 of 100 nanometers thickness the time that is required just by using the parabolic rate law equation is d0 square over B which is proximately 51 minutes.

Now, instead of 100 nano meters if I want to do 200 nanometers so, d0 is 200 the time is not just double. So, it is not just 2 times 51, but because of this parabolic rate law. If you calculate is the actually 3 hours and 25 minutes or it is more than 3 times.

(Refer Slide Time: 36:24)

On the other hand instead of dry ox if you height wet oxidation B0 is higher. So, if you have wet ox B is higher B is usually 0.287. So, that the time to grow 100 nanometers is just 2 minutes.

(Refer Slide Time: 37:00)

$d_0^2 + Ad_0 = B(t+2)$ (1) DATUSION CONTROL thick ONLINE $t \gg t + \frac{1}{48}$	
$\frac{D_{W} \circ h}{100} = \frac{1000}{B} = \frac{100}{B} = \frac{100}{$	

So, comparing these two numbers 51 minutes for dry ox and just 2 minutes for wet ox shows you that wet oxidation is usually much faster than dry oxidation. Because, a diffusing species and the reactions there are involved are different. The trade of course, is that if you are looking for very uniform process control 2 minutes is usually too lesser time to get good process control.

So, here you may either go for lower temperature in order to increase the time or you may go for a dry oxidation in case of wet oxidation. So, there always trade of involved in choosing the appropriate temperature, the appropriate diffusing species depending upon the oxide thickness that is required. So, the oxidation also depends upon the orientation of the silicon wafer. So, this is because for example, if you look at silicon 1 1 1 usually the planer density is higher. Because, there is more packing so, this is actually a faster growth rate.

(Refer Slide Time: 38:03)

Just to compare at 900 degree c for wet oxidation, the value of B which is related to a diffusion coefficient is 0.151 micro meters per hour for the 1 1 1 plane, for the 1 0 0 plane it is lower 0.143. So, again with increase in temperature this difference become smaller, but at lower temperature is always a significant difference into growth depending upon the plane that is exposed to the service.

The presents of the oxide layer will also have in effect on the concentration of dopants. If you have N type dopants as the oxide layer is being formed, these N type dopants have a lower solubility in the oxide and they get rejected. So, if you have an N type silicon the oxide layer will lead to a pile up of a dopants in the silicon which is usually near the silicon; silicon dioxide interface.

Remember the oxide layer is consuming the silicon as it grows. So, that any dopants there are if there are N type they basically get rejected. On the other hand, if you have P

type silicon the oxide as layer as it grows will tent to consume the dopants. So, that there is a depletion this is significant because, we know that the dopants effect the electrical properties of silicon.

So, by growing an oxide layer you can actually change the concentration of the dopants and effect the local electrical properties of the silicon. Usually, oxide layers there are grown especially for patterning are thick so, there are few 100 nanometers in thickness. Thin oxide layers are grown these days especially, when we have device scaling so, we have smaller and smaller dimensions.

So, these thin oxides can also be grown at low temperatures. So, thin oxides usually they have a thickness of we just say thickness less than 30 nanometers or 300 angstroms. So, 1 example is the case of and oxide layer that is used for the gate oxide usually high k dielectrics is used for the gate. But during the patterning process initial oxide layer is grown, which is then replaced by the high k dielectrics.

This is a thin oxide layer is usually has a thickness of a few tense of nanometers; thin oxides can be grown by chemical reaction at temperatures close to room temperature. For example, an ultra thin oxide layer so, the thickness is less than 20 nano meters and we grown by silicon reacting with Nitric acid at temperature is around of 100 degrees. So, this is much lower than the conventional thermal oxide layers, were temperature is usually around the 1000 degrees.

So, 10 orders of magnitude higher the advantage of growing thin oxides in these low temperatures is that these can be used with the regular patterning process. So, that we can grow oxide layers at really small areas so; thermal oxides are usually gone grown in a furnace it is a batch process.

(Refer Slide Time: 42:57)

So, that multiple wafers are grown at the same time so, you can usually have either Horizontal Furnace which is used for small wafers typically 3 or 4 inch wafers can have an Horizontal Furnace or a Vertical Furnace. So, Horizontal Furnace is are usually for 3 or 4 inch wafers they are usually present in save research labs were work is done on the small wafers. Vertical furnace is usually used for large wafers so, almost always in all production kind of pads you will find vertical diffuse inferences.

(Refer Slide Time: 44:09)

If you look at a typical profile of a furnace you would have some you can divided into 3 zones there is a source zone. So, the source zone is where you have your gas source there

is a center zone, where usually the wafers are loaded. Then, there is a load zone the load zone is the 1 there is facing outer most.

	Load zone	
() NOTE		

(Refer Slide Time: 44:54)

If you look at the temperature profile across the furnace is usually a Flat profile in the center zone. So, this is temperature so the wafers that need to be processed are usually loaded in the center zone.

(Refer Slide Time: 45:24)

In the case of furnace growth for thermal oxide there is usually a Boat; the Boat contains the wafers for example, in the case of vertical diffusion furnace your boat could be having something like say 100 to 125 wafers. The wafers that need to be process so, these are call the product wafers are loaded in the center.

(Refer Slide Time: 46:03)

Product mafers = center blank mafers = edges OL baffles)

So, the product wafers are in the center and then a bunch of blank wafers are usually loaded at the ends. So, these are loaded in order to ensure temperature uniformity and also usually gas uniformity. So, these are called Blank wafers were also called Baffles, sometimes baffles are made of other materials are well and these are usually loaded at the edges. So, the boat is loaded and then fed into the furnace and then, there is a whole programming cycle that goes about.

(Refer Slide Time: 46:47)

So, the initial temperature is at room temperature so, these are heated to the process temperature expose to the gases and after the oxidation process is done they are cool and the boat is taken out and the wafers are unloaded. We only look at the oxidation process so far. Another example of a grown layer is a Nitradation layer or the Nitride layer that forms on silicon.

So, the processes performing the Nitride layer is similar so, we can also think of it as a grown layer where you have a Nitride layer at the surface. So, we can think of similar models for diffusion and growth. We also have oxy nitride layers where we have a layer with both oxygen and nitrogen in them.

(Refer Slide Time: 47:41)

