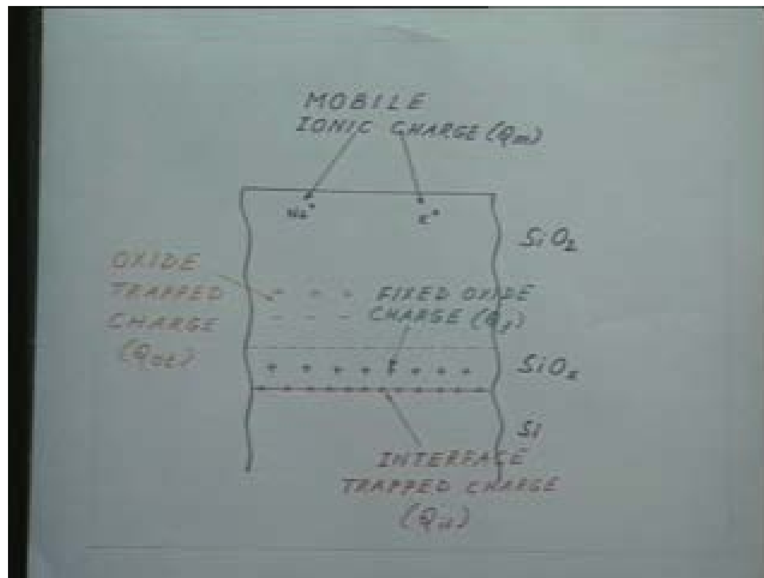
VLSI Technology
Dr. Nandita Dasgupta
Department of Electrical Engineering
Indian Institute of Technology, Madras

Lecture - 14
Oxidation IV – Oxide Charges and Oxidation Systems

So, in the last class we were discussing about the different charges which are present in the oxide and these charges should not be there in the oxide. So, they are in a way decide the quality of the oxide.

(Refer Side Time: 1:21)

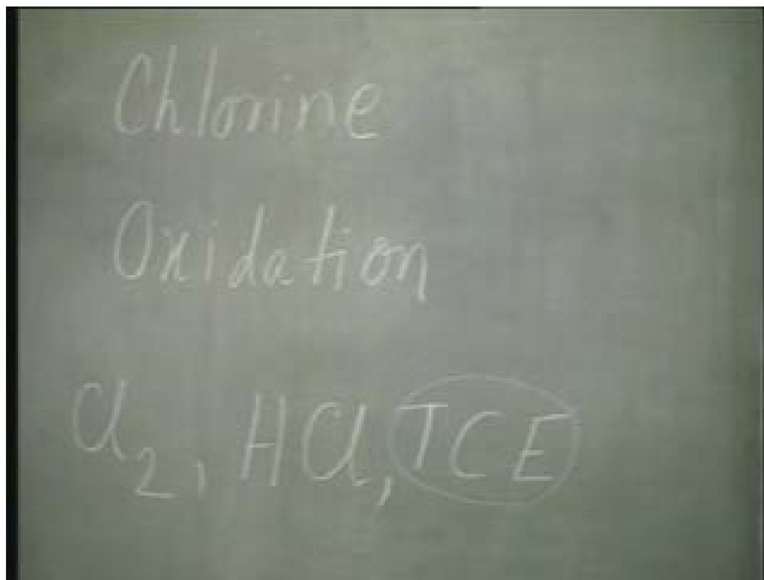


So, looking back again at this, presence of the charges, we have seen that first of all there are the interface trapped charges right at the silicon-silicon dioxide interface and then within a 100 Angstrom thickness, we have positive fixed oxide charges or Q f. Then we can have, all over the oxide we can have positive or negative oxide trapped charges which are present because of, may be exposure to the radiation environment or because of avalanche injection and finally we can have mobile ionic charges which are free to move all around inside the oxide and which are usually as a result of human body

contamination or contaminated water. Since these charges decide the quality of the oxide, it is imperative to minimize these charges.

So, how does one minimize these charges? First of all let us talk about the mobile ionic charges. Since these charges are free to move, it is easy to understand that whenever a gate voltage is applied, depending on the polarity and magnitude of the gate voltage, these ions will move towards the metal gate or move inside towards the silicon-silicon dioxide interface and in the process, it will affect the device performance. So, it is very important to reduce the mobile ionic contamination and one way, a very interesting way to remove or fix this mobile ionic impurity is by doing chlorine oxidation.

(Refer Side Time: 3:15)



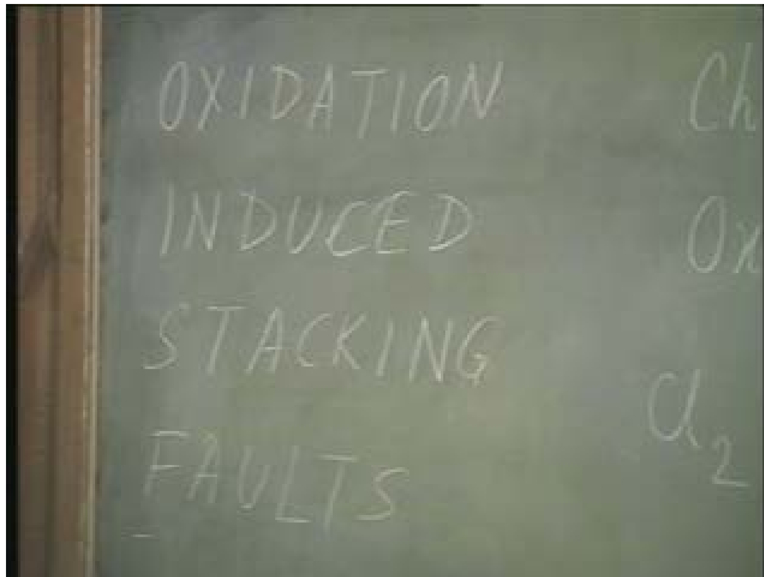
In chlorine oxidation one uses a small quantity of chlorine in the oxidizing ambient. This chlorine may be in the form of chlorine gas or it can be in the form of anhydrous hydrogen chloride. But, since both these are potentially hazardous material to deal with, a safer option is going to be trichloro ethylene, which is commonly referred to as T C E, trichloro ethylene. Trichloro ethylene is a liquid and the common practice is oxygen is bubbled through a container containing this trichloro ethylene, so that the gas that is

coming out of this container will be saturated with T C E vapour and from there a small amount of chlorine will be incorporated during oxidation.

How does that help in fixing the mobile ions? Well you know, chlorine can form sodium chloride and potassium chloride and thereby fix the mobile ions. The ions are no longer free to move inside the oxide and therefore their detrimental effect on the electrical performance of the device is therefore lost. Of course, about mobile ionic impurity another point must be stressed is that they, the presence of mobile ions directly refer to the cleanliness of the system. So, if mobile ions are present in the system, that means the cleanliness of the lab environment is suspect and something should be done to improve that cleanliness also. In addition to that chlorine oxidation is going to help.

Now chlorine oxidation is used for combating quite a few other major problems in oxidation.

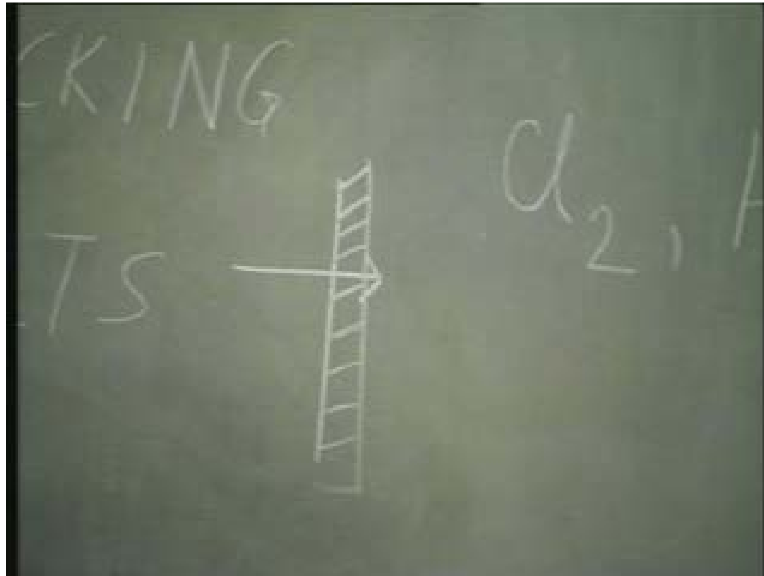
(Refer Side Time: 5:36)



One such problem is called oxide, oxidation induced stacking faults, oxidation induced stacking faults or the acronym that is commonly used is O I S F. What are oxidation induced stacking faults? At the time of thermal oxidation of silicon, you know the

oxidation is taking place at the interface at the last, because the oxidizing species are moving inside.

(Refer Side Time: 6:22)



As the oxidation process goes on, the oxidizing species move inside and this is where the oxide is formed. Therefore the oxidation reaction is taking place at the interface, finally. Now, what is there at the interface? To begin with, it had a lot of silicon-silicon bonds, because it was part of the silicon crystal lattice with all the silicon-silicon tetrahedral bonds. Now, these silicon-silicon bonds are breaking and new bonds are forming between silicon and oxygen. But since this interface is the last region to be oxidized, there is a possibility that even though the silicon-silicon bonds are broken, they still have not found time to react with the oxygen, to form new bonds with oxygen.

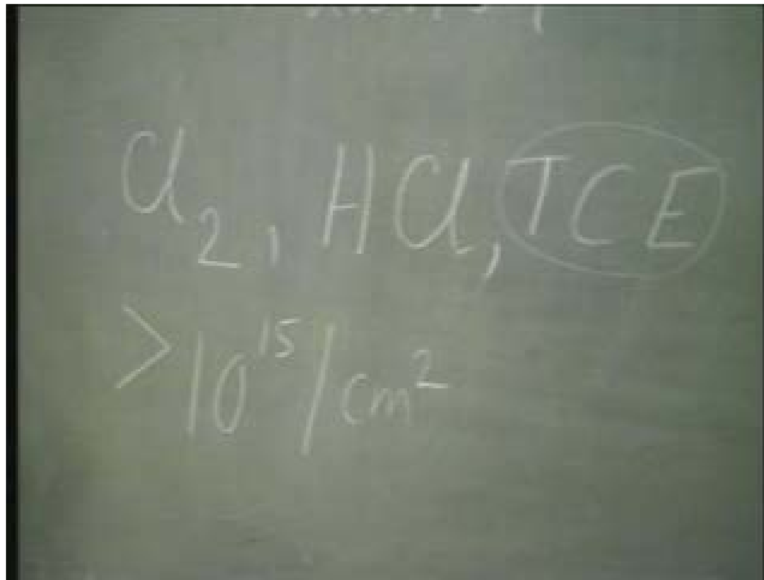
So, as a result of that we have a lot of broken silicon-silicon bonds and therefore these silicon atoms can move into the interstitial spaces, because their position in the crystal lattice where they formed the bonds, those bonds are broken. So, now these atoms are free to move into an interstitial site. Therefore in a practical case, at the end of the oxidation it is found that at the silicon-silicon dioxide interface there will be a large number of silicon interstials. Now, these silicon interstials or at least a small fraction of

these silicon interstials will now move into the bulk of the silicon crystal. In the bulk of the silicon crystal there may already be some defect sites present. These defects might have arisen at the time of crystal growth itself or it can be because of some mechanical damage on the surface or it can be say, prior to oxidation if an epitaxy was being performed it can be because of some growth hillocks, some defects in the bulk may be present. Now, these interstitial silicon atoms they move on to the defect site and coalesce there and that acts as the nucleation site for the fault formation. From there, these oxidation induced stacking faults are going to origin.

Chlorine oxidation can help in combating this oxidation induced stacking faults in a very interesting way. The presence of chlorine in the oxidizing ambient is going to create a lot of silicon vacancies at the surface. So, if a lot of silicon vacancies are present at the surface, then these interstitial silicons which are formed at the silicon-silicon dioxide interface they can move on to these vacant sites. They do not have to move into the defect site and cause the faults, they can be simply absorbed in these vacant sites and preserve the regular crystalline structure. So, it is found that a small quantity of chlorine during oxidation helps not only in fixing the mobile ions, but also reduces the oxidation induced stacking faults.

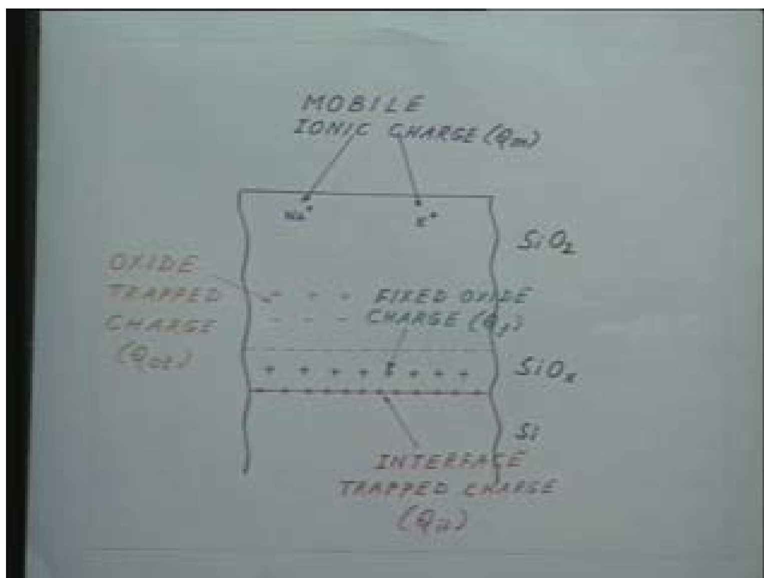
In addition to these things, chlorine oxidation also helps in improving the lifetime of the carriers. You know, lifetime in the device is dependent very much on the presence of deep impurities. Two very notable deep impurities are gold and platinum, which may be present there at the time of crystal growth itself. Now, these deep impurities, they kill the life time. Chlorine can getter these deep impurities. It can form compounds with these deep impurities and thereby improve the lifetime of the devices. But remember, chlorine should be used in very small quantity.

(Refer Side Time: 10:36)



If we use say, larger than 10 power 15 per centimeter square, then we are going to have what is known as halogen pitting. So, chlorine wherever it is used in VLSI processing, is used only in very small quantity. In small quantity its effects are beneficial, but if you use it in large quantity, then chlorine is going to cause pitting of the surface, severe damage to the surface and therefore use of chlorine should be restricted to a minimum.

(Refer Side Time: 11:16)

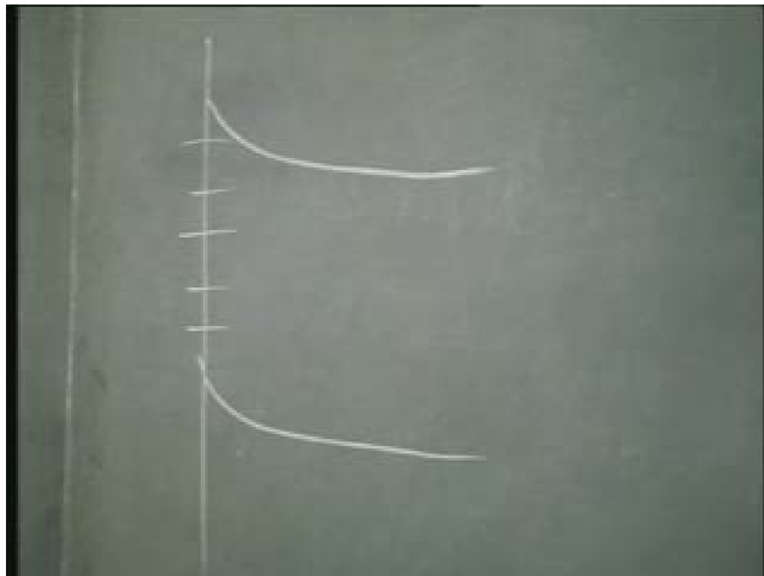


Going back to the charges in the oxide, we talked about chlorine oxidation because chlorine can help fixing the mobile ionic charges. Next comes oxide trapped charges. Oxide trapped charges are not so much processing dependent. They come about after processing that is after the device is fabricated. If this device is subjected to radiation environment or if it is subjected to very high field condition, so that avalanche injection of electrons can take place, then mostly the oxide trapped charges are formed. So, during processing technology itself, the formation possibility of oxide trapped charges are relatively low. That is not to say it cannot form during processing.

For particular processing, like for example say ion beam etching, when the semiconductor is being subjected to a focused ion beam, then in that case it is possible that some amount of oxide trapped charges are found during processing itself. But, by and large it is after the processing is over and the device is subjected to ambient conditions, adverse ambient conditions, radiation environment and such like, then the oxide trapped charges come into picture.

Next oxide charge to be discussed are the interface trapped charges.

(Refer Side Time: 12:59)



These interface trapped charges are a very problematic element particularly when a MOS device is concerned, because these interface trapped charges are spread all over the forbidden energy gap and depending on the position of the fermi level, they can be occupied or empty. So, as the bands bend, because of the application of a gate voltage as the bands bend up or down, these interface states, they change their occupancy; therefore, their electronic activity change and correspondingly they affect the device performance.

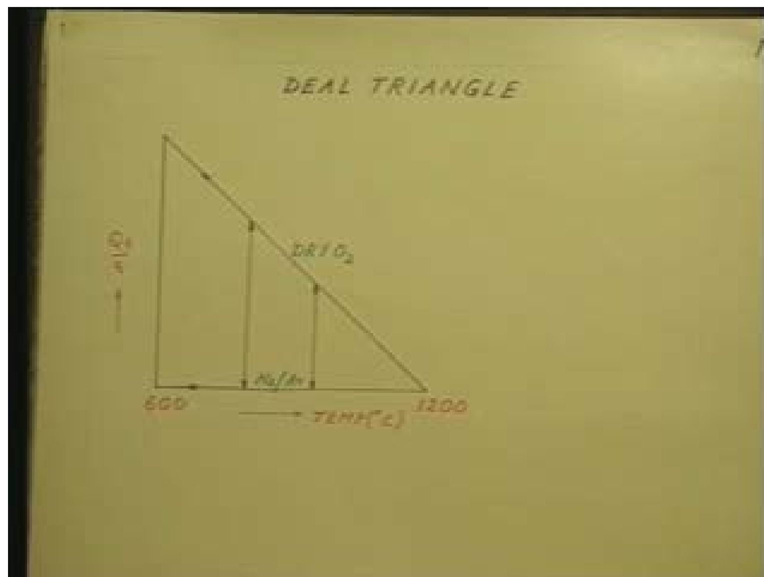
Now, the origin of these interface trapped charges are primarily silicon dangling bonds. You see, before the oxidation when we have the fresh silicon surface, the surface means loss of crystallinity, so on the surface of the silicon there are a large number of dangling bonds. As the oxidation goes on, oxygen reacts with silicon, satisfies most of these dangling bonds and the number of these dangling bonds, they come down. But still at the end of the oxidation, there will always be some dangling bonds present at the silicon-silicon dioxide interface and they are the major source of interface trapped charge. One way to reduce this interface trapped charge is to have a post oxidation annealing in hydrogen at comparatively lower temperature that is at around 450 degree centigrade. So, what happens in that case? The dangling bonds which are present at the interface, hydrogen reacts with silicon, satisfies these dangling bonds and therefore the interface charge density is reduced.

Now finally, we come to the fixed oxide charges and you remember that these fixed oxide charges are present within 100 Angstrom from the silicon-silicon dioxide interface, where the oxidation is incomplete. The composition of the oxide there is SiO x, not SiO 2, x being less than 2. The presence of this positive fixed oxide charge changes the threshold voltage of a MOS device pushing it to the negative side. Because the fixed oxide charge is positive, it pushes the threshold voltage to the negative and in fact that posed a major challenge to the early MOS technologists. You will remember that when one wanted to fabricate an n MOS, enhancement mode n MOS that means it should have a positive threshold voltage, but because of the presence of these fixed oxide charges, we ended up with a MOS device with a negative threshold voltage that is a depletion type

MOSFET. So, one first has to understand why the fixed oxide charges come about and then how to reduce the amount of fixed oxide charges.

Again, focusing our attention on the actual oxidation process, you remember that oxidation at the interface is taking place at the end. So, what really happens when the oxidation process is over, finished? The wafers were kept inside the furnace at a very high temperature, 1000 degree centigrade or 1200 degree centigrade and from there the wafer temperature is lowered either by pulling out the wafer or by lowering the furnace temperature itself. Whatever be the case, the sample temperature is being lowered from the oxidation temperature to say, 400 degree centigrade, before we take out the actual wafer. If this process of reducing the temperature is carried out in an oxygen ambient, then some amount of oxidation will still go on at the reduced temperature and since the oxidation rates depend on the temperature, the oxidation rates will be lower at lower temperature, the possibility of the oxidation being incomplete is greater at reduced temperature. Therefore, if one reduces the temperature of the sample in the oxidizing ambient itself, there one expects to have a very large density of fixed oxide charge.

(Refer Side Time: 18:00)



This is actually depicted here in this figure. This is called the Deal triangle, named after B.E. Deal who was one of the pioneers in the silicon-silicon dioxide oxidation process. What you have here is on this axis the oxide, fixed oxide charge Q f by q. This q is electronic charge, so it is normalized value and on this axis, we have the temperature. So you see, from 600 degree centigrade to 1200 degree centigrade, that is the scale. Now, let us say that the oxidation is taking place at a very high temperature, 1200 degree centigrade and now from this temperature, the sample temperature is being lowered and while reducing the temperature still an oxidizing ambient is being maintained. Therefore, there is the possibility of having incomplete oxidation. So, this is the path it traces. As you reduce the temperature, if it is done in dry oxygen ambient, this is the amount of fixed oxide charge you get which is very high, greater than say, 10 power 12 per centimeter square.

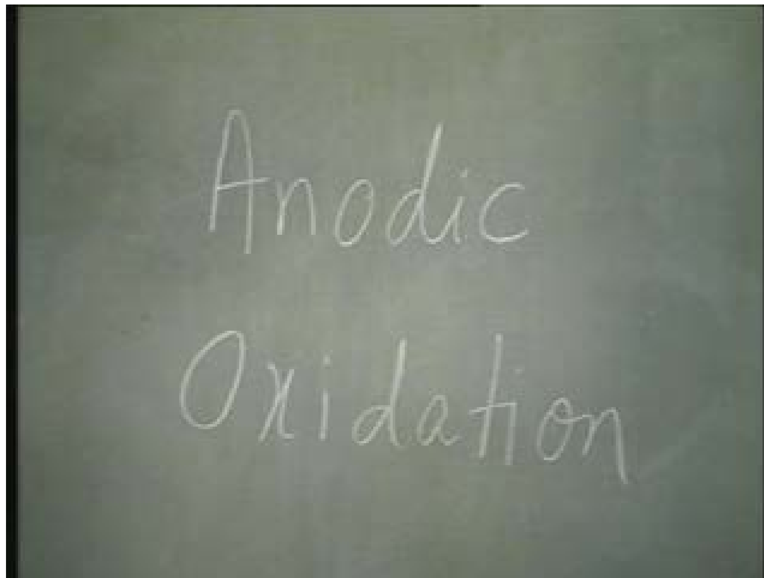
What can be done in order to stop this incomplete oxidation process? Well, what is done is this. That at the end of the oxidation process that is for the stipulated period of time you carry out the oxidation. At the end of this period, close the oxygen gas flow and create an inert ambient that is nitrogen or argon and now reduce the temperature in this inert ambient. So, in this inert ambient oxidation is entirely stopped, oxidation is not taking place. Therefore, there is no question of incomplete oxidation. Whatever oxidation took place, took place at very high temperature and at that high temperature since oxidation rates were much faster, there was much less chance of having incomplete oxidation. So, then one follows this base line and you end up with a much lower fixed oxide charge. So, this is the standard practice today.

Today at the end of the oxidation process, always before taking out the samples, nitrogen flushing for 15 or 20 minutes is carried out. This is done in order to remove the traces of oxygen before the temperature is lowered, so that when finally the temperature is lowered there is no possibility of having incomplete oxidation and by doing that, nowadays the fixed oxide charge density can be brought down to less than 10 power 10 per centimeter square, whereas if you do not follow this practice, it used to be greater than 10 power 12

per centimeter square. So, making an enhancement mode in MOS transistor today is not a problem, because the fixed oxide charge density can be controlled.

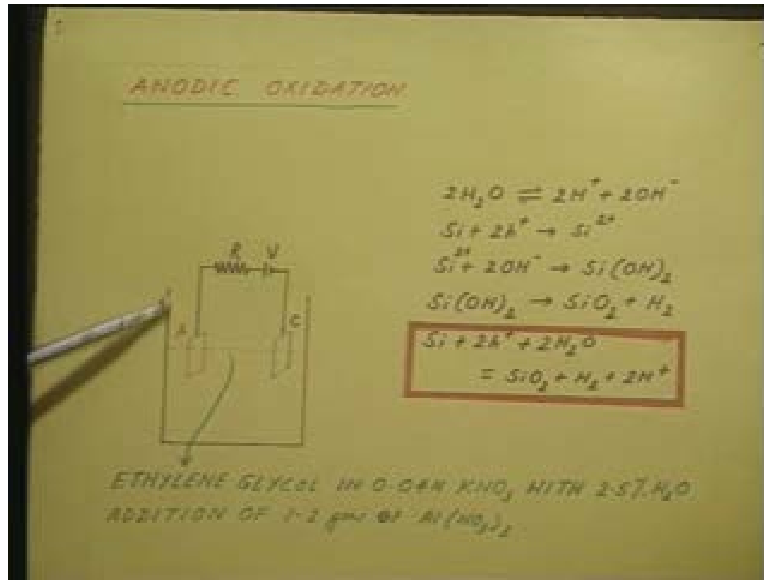
So far, we have talked only about the thermal oxidation of silicon. Thermal oxidation of silicon is a very important processing step, both from the point of view of MOS transistor gate oxidation as well as for isolation and masking. But, there are other techniques by which silicon can be oxidized, even though the quality of that oxide is not as good as the thermal oxide is. But still, those oxidation techniques also have their merits and demerits. One such oxidation process that we would like to discuss now is called the anodic oxidation.

(Refer Side Time: 22:16)



As the name itself suggests this is actually a process of electrolysis. In the anodic oxidation process, what we have is basically an electrolytic bath. You use an electrolyte and the silicon sample is used as the anode and a noble metal such as platinum is used as cathode. You apply a voltage between these two electrodes and an oxide film is formed on the sample. So, this is basically an anodic oxidation and it is depicted here.

(Refer Side Time: 23:07)



This is the electrolytic bath. We use an electrolyte which is ethylene glycol in 0.04 normal potassium nitrate with 2.5% water and usually there is an additional 1 or 2 grams of aluminum nitrate in order to ensure better uniformity of the oxide film. This is the sample which is used as the anode, platinum is generally used as cathode. A voltage source is connected between the two and a current limiting resistor and this is the basic reaction that takes place. This is the overall reaction, there are a number of sub reactions. First of all, water molecule is electrolyzed into hydrogen and hydroxyl ions. Then, silicon with two holes forms Si 2+ positive ions. Then, this Si 2+ reacts with the hydroxyl ions to form SiOH whole 2 and then this SiOH whole 2 dissociates to form silicon dioxide and hydrogen.

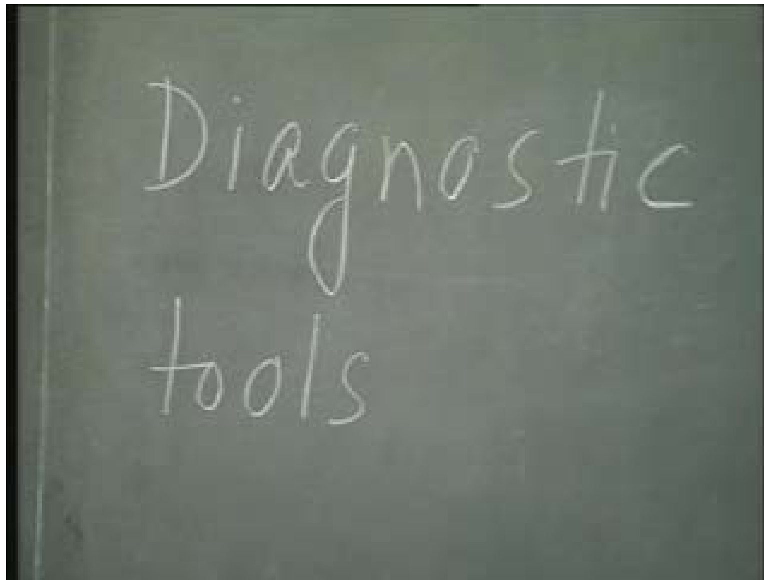
So, the overall reaction is silicon plus water plus holes forming silicon dioxide plus hydrogen. This hydrogen is evolved at the cathode. You can see bubbles of hydrogen gas forming at the cathode. So, from these reactions one thing is quite clear that is holes are an important part of the anodic oxidation reaction. In other words, it is much easier to oxidize p-type silicon using this anodic oxidation process, because in p-type silicon there are an abundance of holes. Holes are the majority carrier. So, it is very easy to anodically oxidize p-type silicon.

What happens in case of n-type of silicon? In case of n-type silicon, you know that electrons are the majority carriers, holes are the minority carriers. So, in order to have sufficient holes we must create them. That can be done either by shining light when there will be enough hole electron pair generation or it can be done by applying a large enough voltage, so that this junction between the silicon and the electrolyte, this is analogous to a metal semiconductor junction, it is a semiconductor electrolyte junction, this junction is broken and there are sufficient generation of holes. So by this way, you can anodically oxidize an n-type silicon wafer also.

Now, the greatest advantage of anodic oxidation process is that it is a room temperature process, unlike thermal oxidation where one has to raise the temperature to considerably high temperature, 1000 to 1200 degree centigrade. Anodic oxidation can take place at very low, I mean room temperature itself, but the quality of this anodic oxide will not be very good. One major reason for that is in the oxidation process itself. Unlike thermal oxidation where the oxidizing species move inside the silicon and therefore we have a fresh interface, here silicon moves out in order to react with the oxidizing species and therefore the original contamination which was present on the silicon surface still remains at the interface and that is one major reason why the anodic oxide quality is much poorer.

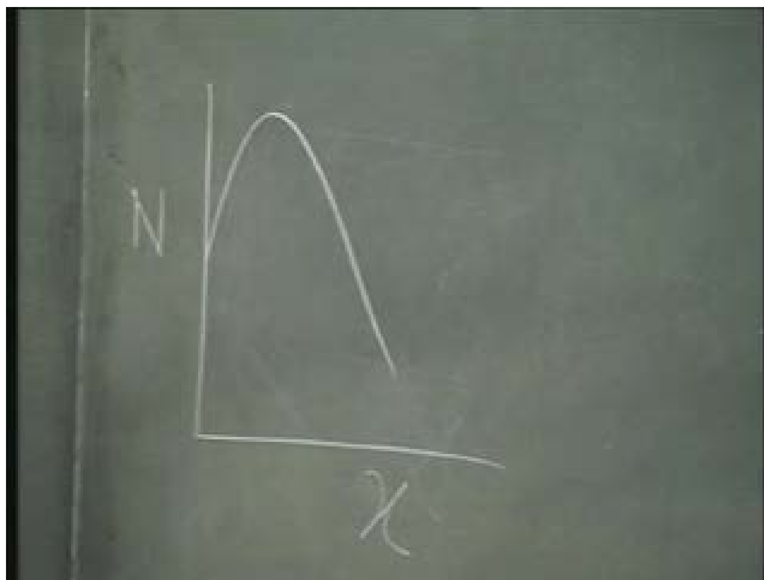
In addition to that you see, it is a liquid electrolyte process and therefore more prone to contamination. Because we are using a liquid electrolyte, it is more easier to, more prone to contamination.

(Refer Side Time: 27:24)



Still, anodic oxidation has this advantage of being a room temperature process, so it can be used as a diagnostic tool that is to say, by using anodic oxidation one can accurately determine various facts about the semiconductor. For example, suppose one wants to find out the doping profile in a silicon sample.

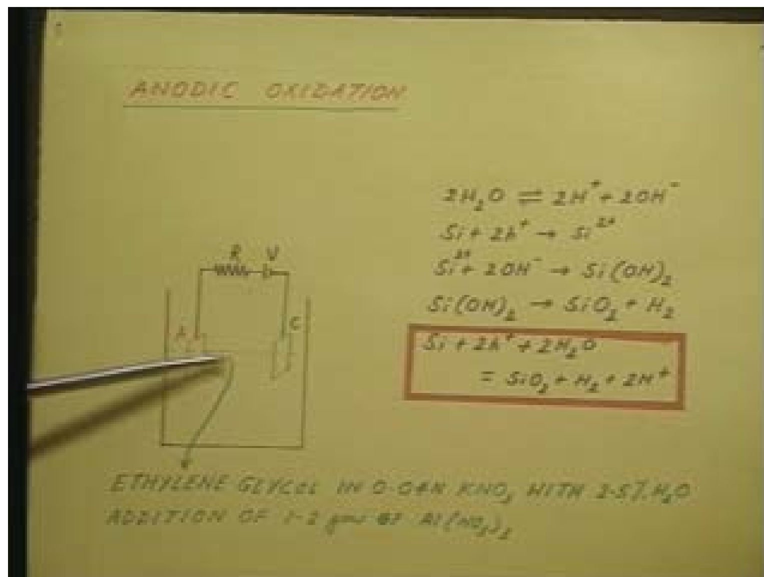
(Refer Side Time: 27:43)



Let us say the sample is doped and it has a doping profile like this, N versus x. This is a doping concentration with the distance. Now, if I want to find out about this doping profile, one way will be to measure the surface properties of the sample, measure the surface sheet resistivity for example, then etch the surface and again measure the sheet resistivity and from this difference in sheet resistivity, one can calculate the doping concentration. But, it is very difficult to etch silicon in a precise, controlled manner. So, what is done is you grow a thin layer of oxide and you know 100 Angstroms of oxide consumes exactly 45 Angstroms of silicon, so you can etch 100 Angstroms of oxide. So, that means you have gone inside by 45 Angstroms and again measure the properties. The only problem is you cannot use thermal oxidation to do so.

Why? Because, thermal oxidation itself will change the doping profile. As you raise the temperature, the dopants will also move. So, the result you will get will not be corresponding to the accurate doping profile. In cases like that anodic oxidation is found to be very useful, because it is a room temperature process. So, it does not affect the doping, it does not affect the doping profile. Added to that there is another advantage.

(Refer Side Time: 29:31)

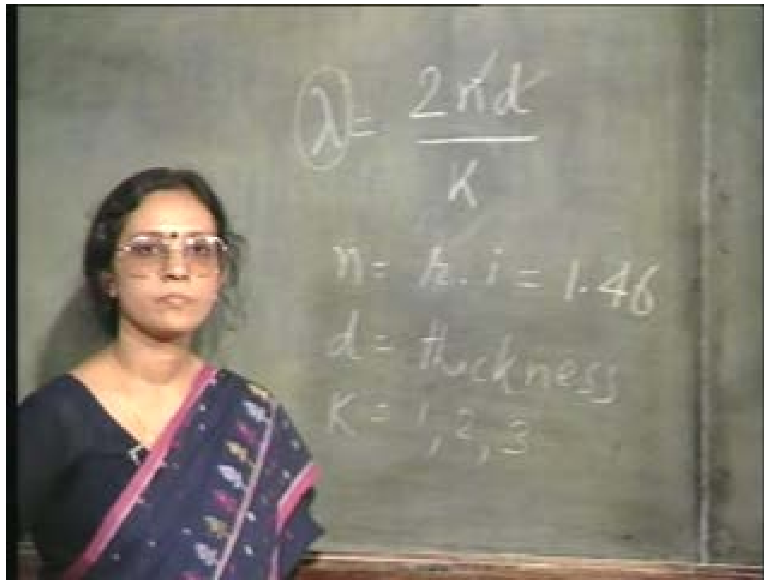


The thickness of the anodic oxide is directly proportional to the voltage that is developed here. See, as the oxide thickness grows, the voltage across the two electrodes also will build up and this can be accurately linked to the oxide thickness. In fact, it follows 3 Angstroms per volt. So, if you see that 10 volts have developed that means you have grown 30 Angstroms of oxide. So, that can be very accurately determined, the thickness of the oxide and etch that oxide and measure the surface properties of the fresh surface thus exposed and keep on, continue doing that in order to obtain the full doping profile. So, even though anodic oxidation is not really compatible with VLSI technology, for VLSI technology it is not a standard processing step, thermal oxidation is the standard processing step, not the anodic oxidation; so, even though it is not a standard processing step, it still has its usefulness. It can be used as a diagnostic tool.

A variation of anodic oxidation is called plasma oxidation, where instead of using the liquid electrolyte, an oxygen plasma is used. What is a plasma? Plasma is partially ionized gas. So, one uses a chamber at low pressure in which oxygen gas is introduced and then it is partially ionized to form a plasma and silicon is used as the anode. Silicon reacts with this oxygen plasma to form the silicon dioxide. So even there, the advantage of low temperature processing is present. In addition to that, because there is no liquid electrolyte, it is less prone to contamination. So, the anodic oxide properties are likely to be better, sorry the plasma oxidation properties, the oxide properties are likely to be better than the anodic oxide properties. But, in either case the properties of the thermal oxide are the best. Therefore in VLSI processing, except as diagnostic tools, these anodic or plasma oxidation do not play a major role.

Now we come to the question of evaluating the oxide properties. Now that we have grown the oxide on silicon, how to evaluate the properties of the oxide? The first property that we want to evaluate is of course, the oxide thickness. How can the oxide thickness be estimated, be measured? The easiest way to do that is by looking at it. You see oxide is actually a transparent film grown on an opaque substrate.

(Refer Side Time: 33:00)



So, when I have such a combination, a transparent film on an opaque surface, then if it is subjected to a monochromatic perpendicular light, then there will be an intensity enhancement at lambda equal to 2 n d by K, where n is the refractive index of the film, d is the thickness of the film and K is an integer, 1 , 2, 3, etc. Now what happens if this transparent film is subjected to white light? White light as you know is a composite. So, in that case that film will appear coloured corresponding to this particular wavelength. That is why when there is oxide on silicon film, depending on its thickness, it appears differently coloured - blue or pink, something.

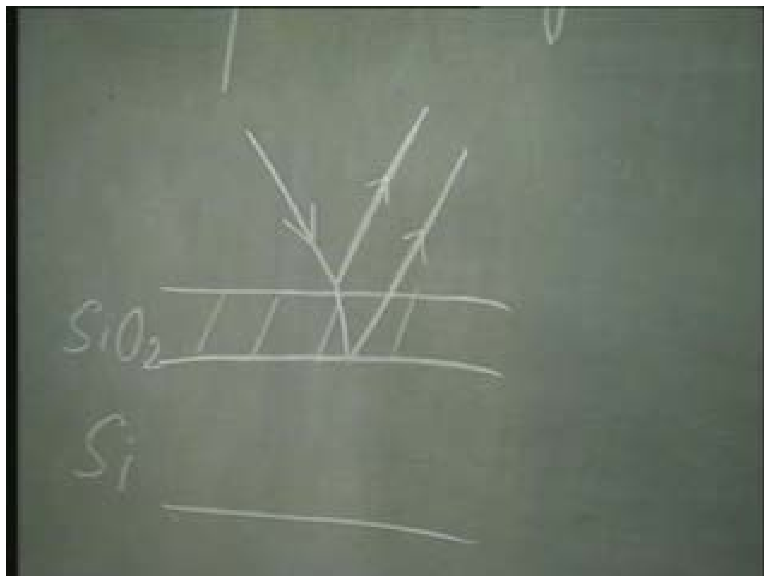
Now the refractive index of silicon dioxide is known and that is 1.46. So, by using that the refractive index value is 1.46 and looking at the colour corresponding to the lambda, one can evaluate the thickness. There is only one problem. That is depending on the value of K, we will have a repetition in the film colour. So, you may have blue coloured oxide film for more than one thicknesses. Only one thing you need to know. That is since you know the oxidation process time and temperature, you can estimate what is the expected oxide thickness and then, from there comparing with the colour chart, one can find out what the actual thickness is going to be.

Of course, there are other much more accurate methods of determining the oxide thickness, but this is nevertheless a very handy way of doing that. Simply by looking at it, looking at the colour and comparing it with the colour chart, one can say whether the oxide thickness is 600 Angstrom or 800 Angstrom, things like that.

(Refer Side Time: 35:14)

A more accurate method of doing so is by ellipsometry.

(Refer Side Time: 35:28)

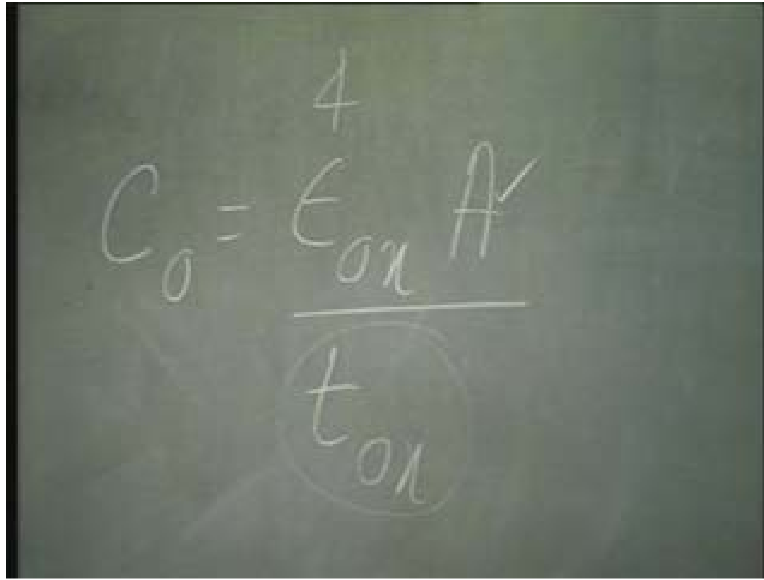


In this technique, a monochromatic plane polarized light is made to be incident on the sample surface. This is silicon dioxide, this is silicon and a monochromatic plane polarized light is incident on the surface. So, there will be a reflection from this surface and also there will be a reflection from this surface. There will be a phase shift as well as a change in the polarization and the difference in the path traversed by these two reflected beams will be dependent both on the thickness of the film as well as its refractive index. So, by measuring this path difference and phase difference, one will have information both about the thickness as well as the refractive index of the film. So, knowing the refractive index, one can find out the thickness of the oxide film and this is accurate within a few Angstroms.

That is a very large degree of accuracy. Within 5 Angstroms or 10 Angstroms, one can determine what the oxide thickness is, but even here the path difference will repeat every 2500 Angstroms. So, that means that one has to know the order of magnitude of the oxide thickness, whether it is 2000 Angstrom thick or 4500 Angstrom thick. Within this 2500 Angstrom limit, the oxide thickness can be determined with a very large degree of accuracy.

There are of course, other techniques like interference fringe method. In that case also, the film is subjected to, sample ..... oxide film is subjected to a monochromatic light and by observing the interference fringe pattern and measuring the distance between the fringes, one can calculate the oxide thickness. In the laboratory however, we use another very simple and electrical method of determining the oxide thickness and that is done simply by taking the C V characteristics of the oxide. You know, if one has a MOS capacitor, then in the accumulation region, the capacitance of this MOS capacitor is going to be the oxide capacitance and what is the oxide capacitance?

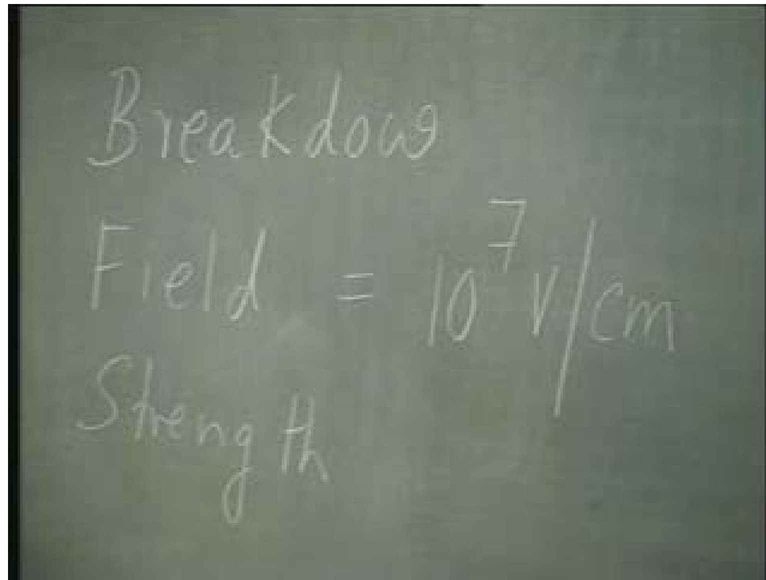
(Refer Side Time: 38:05)



That is nothing but epsilon oxide into area divided by t oxide. So, knowing the area of the capacitor, epsilon oxide for silicon is 4, so one can find out what the oxide thickness is and that is actually a very simple, ready to use method of determining the oxide thickness. In addition to measuring the oxide thickness, the capacitance technique can be used to evaluate the quality of the oxide film. That is it can be used to measure the different charges that are present in the oxide, namely the fixed oxide charge, the oxide trapped charge, the mobile ionic charge and the interface state density. By doing different C V measurement techniques, it is also possible to measure them.

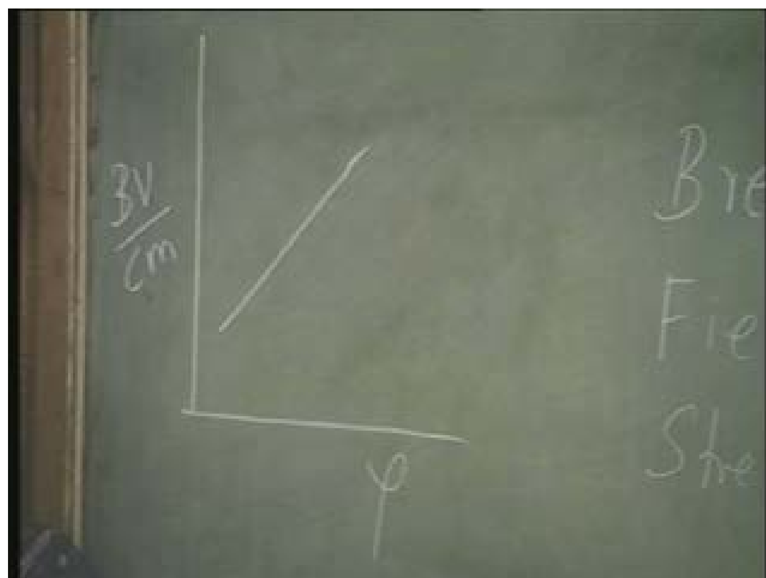
The other important factor regarding an oxide is its breakdown voltage strength. You know, oxide is actually an insulator. It is a dielectric material. So, when one applies a voltage on this oxide film, this oxide must be able to withstand so much voltage, it must not break down.

(Refer Side Time: 39:31)



The breakdown field strength of a good quality thermal oxide is about 10 power 7 volts per centimeter. That means if one has a 1000 Angstrom thick oxide, it should be able to withstand 100 volts. That is what it signifies that the breakdown field strength of the oxide is 10 to the power 7 volts per centimeter and there is also another interesting point about the breakdown field strength.

(Refer Side Time: 40:19)



That is it is proportional to the density of the oxide, so that if I plot the breakdown field strength versus the density, it follows a linear relationship. That is the more dense the oxide is, its breakdown field strength should be that much higher. So, therefore since dry thermal oxidation has the largest, the dry thermal oxidation gives the most dense oxide, one gets the highest breakdown field strength using a dry oxide. So, this is what I mean when we say that dry oxide quality is much better; this is one of the things we mean.

There is another problem associated with oxidation and these are called pin holes. Pin holes means there are microscopic holes in the oxide. For some reason, the oxidation has not taken place and there is a pin hole present. So, when there is such pin holes present, you know and when a voltage is applied, the oxide film just cannot block this. So, one has to dictate the presence of pin holes. But, this problem is going to be particularly severe for thin oxides, because the nucleation of oxide may not be uniform all over. So, there is more likelihood of pin holes being present there.

So, one way of finding out if pin holes are there is by subjecting the sample to a 10% copper sulphate solution. When the sample is subjected to a 10% copper sulphate solution, if pin holes are present, then copper will get plated around that pin hole, so that afterwards by seeing it under the microscope one can find out if pin holes are present there or not. So, pin holes are quite an important problem, particularly in thin oxides and their presence must be detected and avoided. So, this is about the oxidation of silicon.

Let me emphasize once again that thermal oxidation in silicon is the best, yields the best oxide, the oxide properties are the best for thermal oxidation. That is why thermal oxidation is the standard processing step in VLSI technology, not the anodic or plasma oxidation. Even though they can be used, they are used for other purposes like diagnostic tools, but not in, definitely not in forming the gate oxide or even for isolation and masking, one does not readily resort to anodic or plasma oxidation. They are thought to be not VLSI compatible, even though they may have other uses.