JUDY HOYT: And then on Thursday, we'll have these. I also want to mention, towards the end of lecture today, we'll have the course evaluations for you to fill out. OK, before we start the formal lecture for today, for those of you who were here last time, I wanted to go over something that we had talked about.

We didn't have time to go through how you would do these calculations from the last handout, which was handout 36. On slide 27, there was a couple of questions that we were asking. And I thought we'd just spent about five minutes before we start today's formal lecture going through this, at least on the board, so people have an idea of how to make these back-of-the-envelope calculations.

So what you're being asked to do here is you're looking at a silicon MOSFET with-- a these are the source drain extensions, these shallow extensions that have a certain depth XJ. These are the deep source and drain that have been silicided. Remember, the last lecture was all about siliciding. So this blue region is supposed to represent a silicided region.

And then there's a metal contact and this is highly schematized but imagine this little thing coming down here as your aluminum metal or whatever coming down and touching the silicide and is spaced a certain distance the metal contact is by an insulator from this point here where the silicide starts. And so what we're asked to do is to calculate separately three resistors that are in this problem.

There's this resistor here, which represents the resistance of the source drain extension as the current flows through the channel into the extension, which is underneath the spacer. And then there's a resistance associated with this sheet resistance of the silicide arrangement because the current has to get from this point here to the point where the metal contact is up here. And then there's the interfacial resistance or the contact resistance associated with this metal contact. So those are the three.

And in a few minutes here, we can just go through that. And I thought it might be useful for us to calculate those. So the first one, we're asked to calculate the-- let's figure out the contact resistance.

So you're given the specific contact resistivity, which we talked about in the last lecture as 2 and 1/2 times 10 to the minus 7. And that has units of ohm centimeter squared. It's a resistance times an area.

And if I take that-- if I were to get the contact resistance R of that contact, I just have to take that rho sub S value and divide it by the area of the contact. And are told that the area of the contact in number one there on the slide is 0.2 by 1 square micron.

So that's just going to be 2.5 times 10 to the minus 7 over 0.25-- so the numbers are easy-- times 10 to the minus 8 square centimeters because 1 micron is 10 to the minus 4 centimeters. And you have to square that. So you just calculate this out. And you end up with 100 ohms for the contact resistance.

So that's the resistance between the metal and this contact right here flowing through this face, given a contact area of that size. So that'd be kind of a state of the art contact. So you add about 100 ohms associated with that.

You also want to calculate the resistor shown here by this little resistor here, estimate what this resistor is for the current flowing from this point to the point where it just gets into the contact. Well, that's going to be dominated by the sheet resistance of the silicide of this blue region. So if we know the sheet resistance, when we divide by the thickness, we can calculate. For any sheet, you can always calculate the resistance.

So to get the sheet resistance of the silicide, well, you're given its resistivity. It's 15 times 10 to the minus 6 ohm centimeter. So it's just going to be its resistive rho divided by the thickness of the sheet. So that's 15 times 10 to the minus 6 ohm centimeter. Remember, that's a resistivity now. That's not a specific contact resistance.

And you are told the thickness of that sheet, roughly. And I'm going to ignore the doped region down here because this is a metal. It's a silicide. It has a much lower resistance than the silicon. So I'm going to ignore that. We're just going to calculate the resistivity of that metal sheet.

And that thickness is 4 times 10 to the minus 6 centimeters. So the centimeters go out. And you end up with a certain resistance, about 3.75 ohms per square. Remember, that's a sheet resistance. So that's an ohms per square sheet

As for current flowing through a sheet, we talked about this, that is a square, so going in this space, coming out that face, where this length and this width are equal. That's the definition of the sheet resistance. So it's about 4 ohms per square, something like that.

Then we need to know-- once you know the sheet resistance, the nice thing about once the sheet resistance, that tells you the resistance going through a square. You just use geometry considerations, how many squares does the current go through in flowing from this point here to the contact? Well, it's a half micron wall. And it's a micron into the page. So that's a half a square, basically, because it looks something like this.

It's a half micron long. But it goes a micron into the board, into the page. So this is 0.5. And this is 1. So by definition, that's half a square because if it were a full square, there'd be two of these right next to each other.

So then you can just get the resistance of a silicide to be a 3.8 ohms per squared times 1/2 a square. And you end up with about 1.9 ohms, very small, comparatively speaking, couple of orders of magnitude, maybe a factor of 50 less than the contact resistance.

OK, so that seems reasonable. And then because it's a metal and so it has a very low sheet resistance. So the resistance is small. So on the final one, number three, what you're asked, now we're going to be calculating the source drain extension resistance. So what's the resistance associated with this very thin, shallow junction depth, XJ, so this little resistor right here?

So the current comes out of the channel, has to flow in that source drain extension before it hits the silicide. Well, again, that's a very similar situation where we just need to calculate a sheet resistance. And from the sheet resistance, we figure out the number of squares from the geometry. And we get the resistance.

So again, so for the source drain extension, for that reason, we're told the resistive rho is 5 times 10 to the minus 4 ohm centimeter. And we know the number of squares. And we have a thickness of that region.

So I can get the sheet resistance. Again, it's just a resistivity divided by the thickness of the region. So that's 5 times 10 to the minus 4 ohm centimeter divided by-- I think it was an easy number-- 500 angstrom. So that's 5 times 10 to the minus six centimeters. And that goes out. And we end up with about 100 ohms per square.

That's a typical shallow junction for a shallow junction, shallow being 500 angstroms or less. You dope it very heavily. We said this is doped 210 to the 20 with arsenic or whatever. And it's about 100 ohms per square as the sheet resistance. Now, again,

Looking at this, how many squares is this? Well, the distance from this point in the channel to the point where it contacts the silicide that we're told, that spacer width is 0.1 microns. So that's 0.1. And into the board is 1 again. So that's a tenth of a square. So then we can just get the resistance of the source drain extension.

Then it's just the sheet resistance times the number of squares. So that's 100 ohms per squared times 0.1 squares. And then we end up with 10 ohms. So that resistors, roughly-- I mean, this is a very crude estimate, that resistor here. So this resistor here is about 10 ohms. This little sheet here represents about a couple of ohms.

The big resistance is what? The big resistance is the 100 ohms. It's the contact resistance because we have a finite specific contact resistivity, 2 and 1/2 times 10 to the minus 7 ohm centimeter squared. And we're trying to make contact in a very small area. So what's the solution? Well, I can make the area bigger, make the devices bigger on the chip. I get fewer devices. That's not good. Or we have to scale this number.

According to the ITRS, remember, this number is going to have to go down. In 2003, ITRS wants this number to be about 2 times 10 to minus 7. So it's already supposed to be lower than that. And over the next 10 years, it's supposed to be lowered by a factor of 3 or 4. So this number of 100 ohms gets lower or it certainly does not increase as we scale the devices. But just kind of gives you a rough idea of what the big contributions are if you do it right.

Again, if you didn't get enough doping in this shallow source drain extension, this 10 ohm number could go up, depending on your junction depth, how you scale that, and your doping. But these are simple calculations you can do to back of the envelope. And we didn't get to go through that last time we presented it. But this gives you an idea of how those numbers work out.

Any questions on that or on the contact? We only spent one lecture on contacts. So there's a little bit of this in chapter 11, although I don't think there's a specific example of a MOSFET like this.

OK. So let's go on then to today's actual material, which is going to be a little bit of a divergence, to a certain extent, from the other lectures. This material is not in the text. This is all more up-to-date research material that's come out and has not yet been put into a textbook. But you have the handouts.

So what I want to talk about-- what we've talked about so far for semiconductors is just using pure silicon, pure silicon wafers and pure silicon materials to make CMOS, to make MOSFETs. It turns out, if we use epitaxial growth, we did spend a little time talking about epitaxial growth of silicon by CVD or some similar technique.

If we use epitaxial growth, you can mix in small amounts of germanium into the silicon material and make an alloy called silicon germanium. Or you can mix an even smaller amounts of carbon, maybe 1% or so, and make an alloy of the silicon and carbon. And people are doing this. And some of these alloys are used today in modern production devices in manufacturing.

So I just wanted to go through some of the important material electronic properties of these newer semiconductors and how the layers are processed to a certain extent and what some of the issues in research and development that people are considering. So here on slide two, what I'm showing are some applications of these silicon germanium alloys and related materials like silicon germanium carbon.

The first bullet is for a number of years now, maybe the last 10 years, silicon germanium has been used in production as the base layer in a heterojunction bipolar transistor, in a bipolar transistor, the base layer being quite thin. We haven't talked much about bipolars in this class. This class is mostly about making CMOS devices. But it's a thin layer, a couple hundred angstroms, maybe 200 to 500 angstroms thick.

This is used in high frequency analog applications, generally telecommunications applications, base stations, and things like that. And there are a number of companies, maybe half a dozen or more, that manufacture these and sell these devices. So for the purpose of growing thin films, this is something that has been done now in manufacturing for some time.

The second bullet refers to a little more advanced type of material. And this just coming in to production now. And these are layers of relaxed silicon germanium or strained silicon germanium-- and I'll talk about what that means in a little while-- that are used to induce strain in silicon.

So these layers are used to take silicon channel and to either stretch it or compress it mechanically. And by stretching and compressing the lattice in the channel, you can enhance the mobility as it turns out. Enhancing the mobility means you get a higher current drive for a given voltage so the devices will switch faster.

So for enhancing CMOS, some of these materials are now in production. Intel has a new Pentium that's in production using this type of technique, these materials. Sometimes, silicon germanium can be used as a boron diffusion barrier. We spent a lot of time in this class talking about the fact that boron is a fast diffuser in silicon. That's one of the problems with making shallow junctions.

Well, it turns out, in silicon germanium, for reasons that are not completely understood, the more germanium you add, the slower the boron diffusion coefficient. So if you put in 20% germanium the lattice, the boron diffusivity drops by about a factor of 10 in silicon germanium. So that's a huge factor. So people may want to use silicon germanium, just as a means of slowing down boron diffusion.

It's used in micromachining. In MEMS technology, it happens to be an excellent etch stop. So there are etches, wet chemical etches, primarily wet, maybe a few dry, that will stop on silicon germanium and etch through silicon and stop on silicon germanium and vice versa. So it can be used as a sort of an etch stop layer.

Some people are advocating using-- I don't think it's in production but I know of-- but polycrystalline silicon germanium as the gate material. We've talked in this class the last few lectures about metal gates. Before they get to metal gates, people were advocating using a polycrystalline silicon germanium instead of polycrystalline silicon. I don't know, again, if that really is making it into production. Silicon germanium is also used as an infrared photo detector material for integrating photodetectors with silicon CMOS.

Silicon germanium carbon is in production now, maybe not silicon carbon, but silicon germanium carbon is in production with very small amounts of carbon, maybe a half a percent, as a way of suppressing transient enhanced diffusion in heterojunction bipolar transistors. So think somebody's doing a report on that.

One of the people signed up for that. So that's kind of a hot topic, the use of carbon to completely suck up all the interstitials. And then that completely changes the TD effect. So these materials have quite a bit of application.

So on slide 3, this looks like a long outline. But I'm just going to touch on-- after I introduce these the materials. I'm going to touch a little bit on the MOSFETs and the applications of strained silicon and strained silicon germanium to CMOS. I'll talk a little bit about how the materials are grown and about how the dopants diffused in these materials and a little bit about some newer material where we're putting them on insulator. Strained silicon on insulator is a very kind of a hot topic of research these days.

Well, if we go to slide 4, this is more of a motivational slide. The question we're trying to ask is, if you have a CMOS circuit, maybe inverter or something, how quickly can it switch from one state of the zero state to on state-- the zero state to the one state, for example? And this the delay in doing that is something called the gate delay. It's usually measured in picoseconds. And that gate delay is something we want to minimize.

We want to minimize the switching of each gate in the logic sequence so that the overall circuit can do computation faster. So the gate delay generally, if you've taken classes like 6012 or you've had some other undergraduate classes on CMOS, as you scale the gate length-- that's why we're making the gate length shorter and shorter-- the gate delay goes down. This is actually some data showing how it does indeed scale at 0.1 micron. The gate delay is down in the 5 or so picosecond or some 5 picosecond range.

Well, it turns out that the gate delay is inversely proportional to the current. So if you can pump more current through each device, then as one device charges, the device after it has to charge up a device. The more currents going through, the more quickly it will charge its neighboring device. So the name of the game in getting higher speed in logic devices, at least for digital, is at a given voltage, to get more drain current out of the device.

And so you say, OK, at a given voltage, how can I get more drain current? Well, this is what the drain current looks like. There are several physical variables it depends on. It depends on the mobility of the carriers, directly proportional. It's inversely proportional to the oxide thickness. So if I make the oxide thinner, I get a higher capacitance, and the current goes up. That's why we're talking about gate dielectric scaling all the way down to below 10 angstroms. So that's where the gate dielectric scaling comes from.

And if I make the channel shorter, it does also go up. So how do people do conventional scaling to increase ID or speed for logic? They make the gate length shorter. They reduce the gate oxide thickness. The real question has been in the last five or 10 years, what about the mobility?

Is that a variable that we can tweak? We're pretty much done with scaling the gate oxide because we can't get much thinner, only a few atoms thick. Making the transistor shorter is getting harder and harder lithographically and for electrostatic reasons, don't have too many more variables in this equation. So mobility is the last thing to touch. And that's where the silicon germanium alloys come in.

Here on page five, what I'm showing is some classic data on the literature of what the electron mobility that number. Mu N is not a single number in a MOSFET. So this is a plot of the effective electron mobility in a channel.

And again, as this number goes up, the current goes up at a given voltage. As a function of the vertical field in the device, the vertical field is induced by the gate. So as I put a gate bias on the device, as I bias the device, the gate bias, I get higher inversion charge. And so in fact, the mobility actually goes down.

But so you notice it's a function of the vertical effective field in the silicon and that modern MOSFETs are operating somewhere in this range right now. I can just tell you, based on the gate oxide thickness and the typical voltages, 1 to 3 volts that MOSFETs are operating in, if you calculated out the effect the field in the channel, it's about a mega volt per centimeter. So modern MOSFETs these days, the electron mobility is right about on this curve, somewhere between 200 and 300 but, closer to 200.

A number of years ago when we used lighter doping, so let's say you're using much lighter doping like 317, and the oxides were thicker, the vertical field was much lower, maybe half of what it is today. And the mobility was twice. It was what 400. So this is the problem with mobility. As we march down the ITRS, we scale that gate oxide thickness.

So we get higher capacitance. That's good. But we also increasing the doping in the channel, and we're increasing the vertical effective field. So we're marching down this curve. And the mobility is going lower and lower and lower with each generation. That's bad. That takes away from our current drive, to a certain extent. So that's a problem.

So what people have decided to do is OK, it turns out, if I can change this mobility curve, in fact, I can push the whole curve up quite significantly if I put stress on the silicon. And I'll say a little bit more in the next few slides about how that works. So people have tried a number of different ways to stress out the devices or to put stress on the silicone. You have to stress it in a certain way. And this schematic kind of gives you an idea of a good way of inducing stress.

For electrons, this looks like just like a regular nMOSFET. But you notice I've highlighted this channel region in yellow. And it's called strained silicon. A classically strained silicon was generated by taking a wafer of silicon germanium or a very thick layer of silicon germanium, say with 20% or 30% germanium, silicon germanium has a larger lattice parameter than silicon.

Pure germanium is about 4% bigger. So this 30% layer is going to be a little bit larger in its lattice parameter. So if it's relaxed, it's bigger. And if the silicon layer is thin, when you grow it epitaxially, the silicon lattice will actually stretch in the X and the Y directions to match that of the substrate, as long as it's thin enough.

It's energetically favorable to stretch the bonds instead of breaking them. So the silicon is stretched. This lattice parameter in the X and Y plane is about 1.5% bigger than it would be in normal silicon. That's a huge amount for it to take a crystal lattice and stretch it by 1.5%. That's a lot.

As a result, when you stretch it in the X and Y, it actually compresses a little bit in Z. So in fact, we call this particular configuration uniform biaxial tensile stress. It's biaxial and it's tensile, is being pulled in X and Y. So this is sort of the classic way that people have introduced strain and made MOSFETs. And they've looked at the mobility as a function of the amount of the strain.

There are other methods though for straining. Most of what I'll talk about in this lecture is this epitaxial growth method. It's the easiest one to get a handle on the physics. But there are other methods. People have used, for example, when you make a MOSFET, you don't just make an individual MOSFET on a device. All around the MOSFET, we do something called shallow trench isolation that we talked about in this course.

We dig out a trench on all the way around the MOSFET. And we fill it up with oxide or something, some insulator. Well, IBM show back in '97, depending on how you fill that up, you can potentially induce stress, either compressive or tensile, depending on what material you put in because you're digging a moat around the device, and you're putting some material in it. So you can actually use the STI, so-called, to induce stress.

There are other processes during fabrication that Intel is using now where they actually dig out the source and drains. They cut them out with etching. And they fill them back up with silicon germanium, which compresses the channel, which helps the hole mobility. So people are doing things like that. You can actually stress the device. On top of the device, they put a layer of high stress nitride and try to induce stress by applying films that are high stress.

You can also bend the chip. There's a small startup company in the southeast of the United States where they're taking chips, thinning them, and actually trying to pin them mechanically during stress. It seems a little crazy. But it may be possible. But epitaxial growth is in some ways the most well controlled. And the physics is a little bit easier to get a handle on.

So let me just mention here on slide 7. This is the IBM work in 1997 where they talked about in STI-- we learned in this course, you fill up the shallow trench with-- well, first, you oxidize the liner. Then you fill it up with LTO or some low temperature oxide. What they say was they grew an oxide liner for isolation.

And then they fill the trench with polysilicon, which they thought would induce a lot of stress, compressive stress, pushing in all around the edges of the device. And then they looked at the hole mobility. And indeed, what they found, particularly on SOI layers, which are very thin, when they made the device width very small, say 1 micron compared to 10 microns, so you make the device width small, that brings the shallow trench in closer and closer to the center of the channel so it can have a bigger effect as it's pushing in.

And they saw that the hole mobility went up quite a bit. On a 10-micron device, maybe the mobility was here, about 100 or 110. It went up by about 34% to about 140 just by doing this. Now, maybe it wasn't such a great idea to put that much stress and use polysilicon. It wasn't necessarily a practical thing. But it just shows that depending on how you do your processes, how you do your STI, it can have a pretty big effect, particularly on the hole mobility.

So this was an indicator early on that people need to pay dramatic attention to exactly what thermal budget you use, what materials you use in STI every layer you put on that device. Now that the devices are small-- the width of devices today is less than a micron. Now that they're so small, edge effects can have a huge impact on the stress in the center of the channel. And that can end up impacting the mobility.

So I refer to the fact that we have this strain. And I guess I should have introduced slide 8 a little bit earlier. But this just gives you a picture, a ball and stick picture of what we're talking about. So this ball and sticking up in the upper left is meant to represent cubic silicone because it has a lattice parameter of about 5.4 angstroms. It's the diamond cubic structure. But nevertheless, it has a cubic symmetry to it. X, Y and Z in those three dimensions, the lattice constant parameter is the same number.

Cuprate germanium is larger, same cubic structure, but it's larger by about 4%. If you make an alloy anywhere in between, and you have it thick enough that it relaxes, pretty much can use a linear interpolation. It's not quite perfect. But you can use a linear interpolation. So at 50% silicon germanium or 50% germanium, the lattice mismatch is about 2%.

So there are two different types of silicon germanium that people tend to talk about. One is when you start with a silicon substrate and you grow a very thin layer, say a few hundred angstroms of silicon germanium on silicon, what happens is you're taking this lattice here, this cubic silicon germanium wants to be cubic. It's larger. It wants to be larger in equilibrium than silicon.

And you're squeezing it though. It turns out, again, if the layer is thin enough, it's energetically favorable for the in-plane lattice parameter of this material that you're growing to match that of the substrate. So the silicon germanium is kind of squeezed in X and y. And it pushes up in Z. So this layer then is what we call a strained silicon germanium on relaxed silicon. And it's under biaxial compressions. Being compressed in the XY plane.

Now, if we make it thick enough, eventually, there's too much stored energy in those stretched bonds, and the bonds will start breaking and get dislocations, and it'll relax back to its cubic structure. But for thin layers, you can achieve this sort of pseudomorphic structure. You can also do kind of the opposite, or the upside-down experiment. You take a wafer of silicon germanium that's very thick, that's already relaxed, that has a larger lattice parameter and try to stick silicon on it in epitaxial growth.

Again, if the silicon is thin enough, it will now stretch in X and Y to match the lattice parameter of the substrate. And it'll compress in Z. So we call this silicon layer, epi layer as being injured biaxial tension. And again, if you grow it too large. The bonds will no longer stretch. They'll break. You'll get dislocations. So it can go back. Instead of being stretched, it'll go back to its normal lattice parameter. At that point, it'd be relaxed.

So this is the type of material you'll see a lot in FETs. And you see this material in bipolar transistors and some p-channel MOSFETs. So what are some of the important properties of this strain? I already mentioned to you, for MOSFETs, when you start stretching the silicon lattice parameter or compressing the parameter, you can improve the mobility. But there are other parameters that you change when you introduce strain or when you put silicon germanium in.

Probably one of the most important parameters you need to know about is the band gap. So there is an energy band gap difference between silicon and silicon germanium. If this represents the band gap of relaxed silicon, so EC is the conduction band, EV is the valence band, this is the band gap of strained silicon germanium. You see it's much smaller. And most of the difference occurs in the valence band. The valence band energy is larger.

So the silicon germanium has a smaller band gap. How much smaller? Well, you can look on this curve. This is the difference between the band gap of the two as a function of germanium content. So here, if I have strained silicon germanium, you should use its upper curve, not the one that's as unstrained. So a 20%, the band gap shrinks by about 150 millivolts roughly. So and you can read that right off the curve. Here at 50%, it's another number, maybe 350.

So you have an idea of how it moves. All these data points are our data from the literature. There are people extracted the band gap by various devices, diodes, bipolar transistors, photodiodes. And the solid line is some theory that was published a long time ago. So now, the other thing, that's if it's strained. If it's unstrained, if you grow the silicon germanium so thick that it relaxes back, so it comes back to this cubic structure, then what happens to the band gap?

Well, in fact, the band gap difference is a lot smaller. When it's unstrained, the band gap difference say a 20% is only about 100 millivolts. So a big fraction of this band gap offset, this band gap difference, is due to the strain. It's due to the lattice parameter change. A certain fraction is due to the fact that we're just adding germanium. So you need to know whether the material is unstrained or strained or somewhere in between in order to figure out what the band structure looks like.

And in fact, on slide 10, I don't have to go through this in great detail. If you haven't had energy band theory, it's a little tricky. But slide 10 gives you an idea of how the strain does affect the energy bands and the lineup. And in fact, if we look at the case for strained silicon germanium on silicon, so it'll be on-- remember, we said when we grow strain, so if the germanium is under biaxial compression in the X and Y planes being compressed, strain the silicon germanium on silicon, the band lineup looks like this, as I mentioned.

Most of the band offset is in the valence band. So the holes will tend to be confined in the silicon germanium. There's very little band offset. The lowest energy conduction band is about aligned in strained silicon germanium versus cubic silicon. So there's not much opportunity for confining electrons. So this material, it turns out, is ideal for a heterojunction bipolar transistor for an NPN. This material has been in production for about 10 years. People use it in high speed RF devices.

So that's kind of a more well-characterized and well-known structure where the germanium fraction in a typical HBT you would buy off the shelf right now is probably about between 10% and 20% of the atoms would be germanium in an HBT. On the other hand, if you grow the opposite structure, so you have relaxed silicon germanium substrate and you grow a thin layer of silicon-- remember, we said this was biaxial tension-- the band line up is quite different.

In fact, you do get a sizable difference in the conduction band position. And in fact, the conduction band is lower. The energy is lower in strain. So the electrons will tend to go into the strained silicon layer. So if you make a FET, and you make this as your channel, the electrons want to go into the strained silicon. The holes want to be in the silicon germanium side. So this is what's called a type II band alignment. So you can make lots of interesting devices just by taking silicon and straining it on unrelaxed silicon germanium.

The other thing that happens here that's important, and we'll talk about when we talk about the mobility enhancement, you notice what's happened is the conduction band is actually split into two levels. Originally, there was only a single energy level. And that splitting is what's really responsible, to a large extent, for the mobility improvement.

If you really wanted to get nitty gritty, I won't go through it. But here on page 11, I've made a reference to a famous paper by Chris Van de Walle and Richard Martin published in the mid 1980s. Van de Walle and Martin actually theoretically did a pretty good job of trying to calculate what the energy band structure looks like for strained silicon germanium on cubic silicon, which we're showing over here and also for strained silicon germanium on cubic germanium. They have both of these.

And I won't go through it in any great detail. But if you're interested in the solid state physics and you end up working in this area, this is a good paper for you to look at. Again, the band gap would be the distance between this point here in the conduction band and a point here in the valence band. So they actually calculated how the valence and conduction bands move as you add strain and germanium to the system.

On slide 12, I'm also giving you a reference for not only has the band gap changes, but people have actually made some reference some measurements of the energy band offsets when you have strained silicon on relaxed silicon germanium say for a MOSFET. This is taken out of Jeff Welzer's thesis from 1994. And what you can see here, this is the basic structure. You have an oxide. You have a thin layer of strained silicon. And you have these energy offsets here.

The energy conduction band energy is a little lower. And the valence band energy is slightly different in strained silicon. And these numbers, delta, you see how different they are, these energy band discontinuities were actually measured and delta EV here as a function of the germanium fraction in the substrate.

The bullets are the measurements. And the lines are the calculations from Van de Wall and Martin. And indeed, you get pretty good agreement. So we have a rough idea now of what these energy band structures look like when I add strain or when I add silicon germanium to the substrate. So that's important for figuring out how all these devices work.

I mentioned that when you squeeze silicone, either you stretch it in a biaxial sense, you can improve the mobility. And the question is, well, why is that? I don't want to go into detail in the solid state physics. But it turns out, if you put silicon in biaxial tension, you break the cubic symmetry of the lattice. Ordinarily, in the lattice, the XYZ directions are all completely symmetric when silicon is not strained. When it strained, they actually become asymmetric.

X and Y lattice parameter is larger than in the Z. And what happens by that breaking of the physical symmetry is that the sixth energy of ellipsoids in constant energy space that are all at the same energy actually split up when you introduce a strain. And two of them, these two ellipsoids here where the electrons can be, these two red ones, the so-called perpendicular valleys labeled Del 2, they actually end up being at a lower energy than these other four.

So we get this strain-induced splitting in bulk silicon. All six of these valleys that represent states that electrons can occupy, all six of them are at the same energy level. But by group theory, when we break the symmetry, we break that up into a two-fold degenerate level, which is lower energy, and a four-fold degenerate level.

So by breaking this up, actually, what happens is now, most of the electrons want to be in these red valleys. And the scattering between the red and the green is suppressed because it costs you energy. So instead of scattering between all six, you can only scatter very effectively between these two red ones. And so the scattering time goes up. And the mobility is directly proportional to the scattering time. The in-plane effective mass also goes down. So the mobility also goes up for that reason.

So just by breaking the symmetry, we can change the occupation of the energy bands in silicon. And we can make the mobility higher for electrons. In the valence band, I'm not going to go through in great detail. But it's quite a bit more complicated. But it turns out, with strain, we also split the degeneracy in the valence band. We also suppress interband scattering. It is a lot more tricky. But it is possible to also induce improved hole mobility or PMOS mobility by introducing strain.

And in fact, here on slide 14, I'll just walk through some of the older data. This was from 1992 by Welzer. These were the first strain of MOSFETs that were fabricated where he made a thin layer of strained silicon on relaxed silicon germanium and then just implanted it and made a silicon MOSFET with a source and drain. It was an nMOSFET. And this, he called the surface channel MOSFET because the bands aligned so that the electrons wanted to be at the surface. So the channel is at the surface.

This he called the buried channel MOSFET because if you look at the energy band diagram here, there's a little bit of silicon germanium on top. And that confines the electrons to be a little bit buried below the surface. It's like a MOSFET but where the channel is slightly buried.

And on slide 15, he actually measured some of those mobilities. These were the first mobilities measured. And if you look at this solid line, this is the mobility measured in a MOSFET that had no strain. And for a surface channel MOSFET, the mobility that he measured was about, oh, I don't know, 80% higher, about a factor of 1.8 improvement in the surface channel. So right off the bat, as a function of vertical field, you see we have parameter. By changing the strain, you can up the mobility by about 80%.

If you do the buried channel, you can get even higher mobilities at low gate bias. But the mobility seems to go down at high gate bias for various reasons. So kind of the first evidence that strain makes a big difference.

These are the actual structures that one can grow shown on slide 16. Again to make an nMOSFET, you can grow relaxed silicate germanium with a thin layer of strained silicon. And this becomes a high electron mobility channel. This slightly different structure is needed for the PMOS, very similar but with slight variations, so you can make CMOS devices using this type of technique.

And I'll skip through the energy band diagrams for now. Slide 17 shows some metrology. The question you may be thinking, well, how do you really know? All right, your mobility went up. But how do you really know that the layers are strained? Especially when they're so thin, how do you measure strain in a layer that's only 100 angstroms thick? You really can't use X-ray diffraction. X-ray diffraction on 100 angstrom layer could be really tough because the layer is very thin. So how do you measure that change in the lattice parameter?

Well, it turns out there's a technique called Raman scattering, which is an optical interaction. There are certain vibrational modes of a lattice that will scatter a scatter of photons. And the Raman effect can be used to measure the strain or the change in the lattice parameter, even a very, very thin layer.

So here's an example of a MOSFET that was made, a Raman spectrum. So you're measuring the intensity of the scattered light as a function of the frequency in inverse centimeters. And you see a large peak here associated with relaxed silicon germanium. And there's a small peak here, which is the thin strained silicon layer.

Now, if that layer were unstrained, bulk silicon, if you just take a wafer of bulk silicon and subject to Raman scattering, you get a line at 521 inverse centimeters. So it turns out the shift of the strained silicon peak, the shift between that peak and 521 is a measure of the amount of strain. So this is a very nice way you can oxidize a wafer, strip the oxide of strained silicon. You can see if the strain is still there. In fact, you can figure out the relaxed silicon germanium content as well using this.

So slide 18 just shows you some measurements that people have made, the Raman peak shift and the frequency shift in inverse centimeter as a function of the amount of germanium fraction of the substrate. For strained silicon layers, you can see there's a theoretical line here. And the bullets represent measured points.

So using Raman, if you have the right kind of Raman setup, you can actually get a peak shift. So for 20% germanium here, we expect a Raman peak shift of about 6 to 7 inverse centimeters. So that gives you an idea of how you can measure the strain in the material.

Slide 19 shows you a material that's been oxidized. You might be thinking, OK, fine. You can grow these epitaxial layers. They'll be strained. What happens if you oxidize the silicon? Is it going to somehow relax the strain because of that oxidation process and the breaking the bonds and the oxidation? Well, it turns out it doesn't. In fact, the as-grown Raman spectrum here, this is for the epitaxial layer as it was first grown, you do get a strained silicon peak right here near to 510.

If you oxidize it at 850 for 10 minutes, you still get that strained silicon peak in. Although, the peak is smaller height because, of course, the layer is thinner. When you oxidize, you consume silicon. But the peak position, it remains at 510. It remains at the same point. And this has been verified very many times.

And, of course, in the MOSFET, you do see the enhanced mobility. So there doesn't seem to be any significant strain relaxation, at least when you oxidize in these type of temperatures. If you were to go up to 1,000 or more, then you could start running into trouble when the germanium starts diffusing around.

So on slide 20, I'm just reviewing for you, giving you an idea of some of the early measurements people made of how the mobility was enhanced. And the left side is for nMOSFETs. And you see this black line, this dashed line is for control devices. And these other lines on top are for 10% germanium in the substrate, 20 and 30. You see, as you add more germanium in the substrate, the whole of the electron mobility curve all moves up. And you get a nice enhancement.

On holes, for PMOS vets, there was some enhancement measured, although the enhancement is diminished. You see the enhancement factor is actually going down as you go to higher vertical fields here. So at higher vertical fields is where a lot of modern MOSFETs operate. So that's a bit of an issue for pMOSFETs for using this technique.

Those are research devices. This on slide 21 is some more recent publications in a manufacturing style process. Although it's not in manufacturing, the process was a full manufacturing style process that IBM did. It had all the attributes of a real manufactured device. It had shallow trench isolation. Had it had weld implants, halo implants, raised source drains by EPI, and a thin gate oxide. So this is a real device as opposed to the university style.

And again, you can see the mobility for electrons goes up with strain, about 100% improvement. That is a factor of 2. Holes, there is an improvement at low gate bias. But here at high effective field or high gate bias, the improvement is pretty small, only a few percent. That just shows the state of the art as it exists a couple of years ago.

And if you were working at relatively low gate bias, and you just want to get an idea of how the mobility enhancement varies with the amount of strain, you can look at this curve on page 22. This is the electron mobility enhancement factor. So it's the peak mobility in a strained silicon FET divided the peak mobility in an unstrained FET. And you see here with no strain, that it's 1. And then it goes up to about a factor of 1.8, maybe 2, something like that. There's a lot of data have been published in the last 10 or 12 years. And they all seem to agree about this saturating behavior.

On the right is for loss for hole mobility. The data does show an increase, especially as you get above about 30% germanium. You can get larger enhancement factors. And this is the theoretical curve. The data seems to be a little bit below the theoretical curve but shows roughly the kind of strain dependence you would expect.

And so what this is telling you, if you're saturating out, you don't really need to put more for electrons, more than about 20% germanium in the substrate. That's enough strain. So the bands are split enough, if you go beyond that, you're inducing more strain and more splitting. But the mobility itself is not actually improving because something else is limiting the mobility. But it gives you a rough idea of the amount of strain need to induce in some of these devices.

Page 23 is a summary of some of the mobility enhancements now for pMOSFETs. It's the mobility enhancement ratio. Again, this is the number you would like to be large as a function of vertical field. If you operate the device at very low vertical fields, you can get large ratios. But most modern devices operate over here. So the enhancement factor is only about 1.1 or 1.2. It's approaching 1. That's a problem, unless you go to very high germanium contents.

There's some data here with a 40% substrate. So that means the strained silicon layer is very thin, and it's strained to a 40% substrate. Here, in mobility enhancement factor, there's not a full curve available.

The highest field is people went to was about 0.6. But they did get pretty large enhancement factors. So it looks like if you're going to use biaxial tensile strain in the silicon, you need to go to very high levels of strain if you want this to work. So that's one big disadvantage for the pMOSFET in this technology.

Oh, just another graph on page 24 to give you an idea. We talked about mobility. How about things that you care about electron velocity or GM that a lot of circuit designers are worried about? Well, indeed, that does improve. And even as you scale devices, this is data back from 1998 now. It's fairly old. But it still gives you an idea of if you look at the GM here for an unstrained device shown in red, as you scale, it goes up like this.

For a strain nMOSFET shown in blue, it goes up. And this enhancement factor or the ratio between these two lines stays about constant. In this plot all the way down to about 90 nanometers or so was the shortest devices that were made, just around 100 nanometers. Since that time, shorter devices have been made. So the nice thing is that the enhancement doesn't seem to go away when you go to short channels. So that's very important for the technology.

So let me talk a little bit now, I want to spend a few minutes on how the material is grown and some of the issues related to how dopants diffuse in this material. Slide 26 just shows some typical gas sources that are used. A very common way to grow this is by low pressure CVD or sometimes ultra high vacuum CVD, depending on the particular lab that's doing the work.

For epitaxial silicon germanium, a very common silicon source is silane. Sometimes people use dichlorosilane. We've talked about dichloro earlier in the class. For germanium, people use germane. What is germane? It's the same as silane. But you take out the silicon, and you put in a germanium atom. So it's GeH4.

Hydrogen is often used as a carrier gas, although not in UH BCVD but in low pressure CVD. And these are common dopants. Diborane and arsene are used for P and N type doping, sometimes phosphine.

For silicon germanium in a selective process, you need to add-- so if you want to deposit silicon germanium on a wafer that has oxide on it and little openings in the oxide where there's windows that are patterned so that you can reach down into the silicon, this is called selective deposition. You need to add hydrogen chloride, HCl. And this is commonly done for both the bipolar and CMOS applications.

For silicon germanium carbon, you do the same type of thing. But you add one more gas, which is called methyl silane. And that the methyl group, the CH3, brings in the carbon, which is used to dope those devices.

We talked a little bit about strain. And here on slide 27, I'm showing a little bit of idea what that means. This ball and stick diagram in the lower right was our pictorial picture in our heads.

When the EPI layer is thin, we said the silicon germanium, if it's thin enough, it will compress in X and Y. It will match the lattice parameter of the thick substrate. That's only when it's thin enough and it's energetically favorable to bend the bonds. In that case, we say it's fully strained.

If you build up more and more layers, you build up more and more stress, at some point, it becomes energetically favorable to break bonds. So at this interface, bonds are actually broken. And it's represented here by these misfit dislocations that are going into the board or into the board here.

And so these misfit dislocations end up relieving the strain on the thick layer. They have a certain spacing. And if you have enough of them and their spacing is tight enough, in fact, you will end up with a fully relaxed layer. The silicon germanium will have a last parameter that it has an equilibrium.

And this is a cross-section view in the upper right. And this transmission electron micrograph is a planar view. So if you look down through an epitaxial layer that's been grown to be too thick, and here is a bird's eye view. This is a misfit dislocation segment along the interface. And there's a dislocation arm that threads to the surface.

And these dark lines end up having extra electron diffraction contrast. And these dark lines correspond to misfit dislocations at this interface. And they run along the easy slip planes in silicon and silicon germanium, which they run along the 110 direction, so just typically parallel and perpendicular to the flat. So this film was grown to be a little bit too thick. And it started to form dislocations. It's no longer fully strained.

In fact, you may be wondering, oh, well, how thick is too thick? Well, there is no one number, unfortunately. There for a long time, people thought there was a quote unquote critical thickness so that at a particular germanium fraction, when you go above that thickness, bingo. You get dislocations. Doesn't quite happen that way. In fact, there is a whole metastable regime. And the critical thickness depends on the temperature at which you're doing the growth because it's a kinetically limited process.

And this is a good paper, Derek Houghton's paper from 1991, who describes the kinetic model and explains the fact that there's no one critical thickness at a given germanium fraction. What he plotted here was a summary of the critical thickness above which dislocations form as a function of germanium fraction of the silicon germanium layer for different growth temperatures. So if you're growing at 900 degrees, you're very hot. And it's very easy to nucleate dislocations. And they travel very rapidly.

You tend to then get the equilibrium critical thickness that that's this line here. So if you're growing at high temperature, or you're annealing at high temperature, if you're at 20%, the equilibrium critical thickness for silicon germanium is about 120 angstroms, 110 angstroms, something like that. So it says that if you're growing in high temperature, and you grow above about 120 angstroms of 20% silicon germanium, it'll start to form dislocations at that point.

However, if you are growing that same germanium fraction at 550, so you're in a very low temperature epitaxial growth, you're are kinetically suppressing the formation of dislocations. In fact, the critical thickness is like this curve here. It's about 700 angstroms. So there's a metastable regime, if you grow the layers thin enough, you can keep them strained. And they won't form dislocations until a much thicker layer.

The problem with that is, you can grow them very thin at high temperature, very thick, and they won't have dislocations. But if you go to process them it to make a device, at a higher temperature, dislocations can pop in if you're above the equilibrium critical thickness. So this region is called metastable because you can grow them at low temperatures without dislocations. But if you were to anneal them at a higher temperature dislocation, it would fall, and it would start to relax.

Generally, this region up here, up in the upper right, so it's very thick films, are almost always relaxed. They'll always have dislocations. I refer you to this 1991 paper for a good reference on understanding the formation of misfits in strained silicon germanium layers.

Slide 29, it turns out, in the early 1990s, there was some breakthroughs on, you say, well, growing strains silicon germanium. That's very interesting. People want to do that. There were actually some breakthroughs on how to grow relaxed silicon germanium. This seems like it should be very easy to make the silicon germanium relax. You just grow it thick, right? We just saw that paper from Derek Houghton.

Well, if I'm at 30%, I just grow it to be 3 or 4 microns up here. It'll definitely be above the critical thickness. It will be relaxed. It will have its own lattice parameter. The problem with that, if you just grow it right off the bat like that, a box layer, you end up with a huge number of threading dislocations. So if you don't grade the profile, you end up with more than 10 to the 9th, maybe 10 to the 10th of these threading arms.

For every misfit, you have two arms that go up to the surface. And you have so many of them that the layers aren't very useful for very many devices because there's too many defects. But in the early '90s, there were breakthroughs on how to grow these layers, how to grow them in high enough temperature so the threading arms could run to the edge of the wafer and how to grade the germanium content. So you start growing at silicon lattice. And then you gradually add germanium as you go up and you grade it over many microns.

It turns out, when you do that, you can relax the silicon germanium. But you can reduce the number of arms that thread to the surface. All the arms run out to the edge of the wafer where you don't care about them you're not going make any devices at the edge of the wafer. So the threading dislocation density can go down from 10 to the 9th to less than 10 to the 5th if you do the grading right.

So a nice thing about this is all of a sudden, on silicon substrates in the early '90s, people had a way to take a silicon substrate, grow a relaxed silicon germanium layer, and have it be pretty high quality and reasonable enough high quality that you could make devices on it that work. There's still a lot of threading dislocations, but not enough to screw up the devices too badly. So that was kind of the breakthrough that enabled people to make relaxed silicon germanium and then strained silicon MOSFETs on top of it.

What are some of the materials issues if you're growing this relaxed silicon germanium? Well, I already mentioned the threading dislocation density. It actually decreases if you go to higher growth temperatures because the dislocations can move faster. The problem, though, is you also tend to get a very strange field that creates a crosshatch. In fact, if you look a plan view optical micrograph of a relaxed silicon germanium layer, it looks like a plaid shirt.

This plaid is that buried strain field causes the surface to undulate ever so slightly, maybe 100 angstrom, high hills and valleys, very smooth. And the period of these undulations are the spacing is quite large. It can be microns. So these are smooth undulations. They don't really mess with the electron mobility. The electrons and holes don't care about them.

It makes the wafers look ugly, however. And it makes lithography difficult because optical machines don't like looking at plaid. They like things that look smooth. So it is a practical problem. People have found a ways around it. You can CNP the layer, and you can remove all the undulations. And it's not necessarily a big problem. But it is an issue.

If you grow the silicon germanium too hot, it turns out it doesn't even grow as a planar film. It grows as big little clumps, little mountains or islands. It doesn't wet the surface properly. And this type of growth mode is generally considered not very useful because it's not really a smooth film. So you cannot grow silicon germanium at too high a temperature for a given germanium fraction.

What do these relaxed silicon germanium graded buffer layers look like? I said, well, they have very few defects that thread to the surface. That's indeed true. These are some cross section TEM micrographs of relaxed silicon germanium. See at 10%, this is the silicon substrate. This is the graded layer. There are lots of dislocations in the graded layer. That's what you want. You want dislocations down in the graded layer because you want to relax the lattice parameter. You want it to expand to be its equilibrium parameter.

But what you see is as you get up to the surface, if the graded layer is thick enough into the box layer, there are too many dislocations of visible in cross-section PEM. Again, this is up here is where you're going to build your device. So that's what you care about. But if you compare a 10% to a 20% to a 30% germanium, you see a higher and higher density of these buried dislocations, of course, as you increase the germanium fraction because you have to relieve more lattice mismatch.

So the layers get pretty ugly. Even though the device is up on top, there's still a lot of integration issues. You can imagine trying to make CMOS circuits on a substrate that looks like this. All these buried dislocations are going to worry and they're going to bother the device engineers because after all, we do have CMOS wells.

We do have deep junctions in some of these structures. So this silicon germanium can affect our my shallow trench isolation fabrication, formation of silicides, and the way dopants diffuse. So it certainly is an issue.

In fact, there's a whole field of research going on right now, which is the study of the diffusion of dopants in silicon germanium. And I'm showing here on slide 32, this is some data for silicon germanium alloys with 20% germanium in them. And what we're looking at is the diffusivity of different dopants. So this is the effective equilibrium diffusivity.

So its diffusion coefficient in centimeter squared per second as a function of inverse temperature, well, the usual kind of Arrhenius plot. And if you look at boron, so the dashed line here is for diffusion in silicon. And the solid is for diffusion in silicon germanium. If you look at boron here, in fact, in 20% silicon germanium, the diffusivities reduced. I'd said before, a factor of 10, looks more like a factor of 6 or so, depending on the temperature.

But so the boron diffusivity goes down in silicon germanium, which is nice because boron is a fast diffuser in silicon. But if you look at arsenic, this is arsenic in silicon, the dashed line. Arsenic in silicon germanium is the solid line here with the green arrow. So the arsenic in silicon germanium is diffusing a lot faster. In fact, arsenic in silicon germanium is diffusing at about the same rate as born in silicon germanium at 20%.

So we've taken the boron diffusion coefficient, it goes down. Arsenic, which is usually very slow in silicon, now goes up. Now they diffuse about equally, which is kind of interesting. You might make some interesting device structures when you have symmetric diffusion coefficients between the N and the P type doping. Unfortunately, what it means is that in silicon germanium MOSFETs, we actually are much more worried about the n-type dopant diffusion because that's what's going fast. And we're less worried about the p-type doping diffusion because the boron is going slower.

Phosphorus doesn't help the situation. Phosphorus is also enhanced, maybe about a factor of 2 or 3, depending upon the temperature. So and this line here gives you an idea roughly of the diffusivity of germanium. Well, this is at a very low germanium fraction as a function of temperature. It's diffusing slower than the dopants, at least at this germanium fraction.

So I mentioned to you that arsenic moves faster in silicon germanium. This is actually a big problem if you're trying to make a strained silicon mMOSFET with silicon germanium in the substrate. Let's say you're trying to make a shallow source drains. So you implant 50 keV at 3 times 10 to 15, may be a typical source drain implant. And you anneal it at 1,000 degrees for various times.

And the dashed lines here are for a silicon substrate and the solid are for silicon germanium. So if you just look at, say, this dashed line here, this red dashed line, for the silicon substrate with this 22nd anneal at 1,000, you have a junction depth of about 500 angstrom. That's perfect. For the silicon germanium substrate, which is the solid line, 22nd anneal, your junction depth is more than twice that. It's like a 1,200 angstroms.

So the arsenic and the silicon germanium is going much more rapidly. You see this big, big shift. And that's a problem. If you use very short anneal times, you can minimize that difference. But 1,000 degrees, one second is not really enough to really repair all the damage and to activate all the dopant.

So there's a very interesting issue here is, how do we optimize the annealing time for anti dopant? So we get good activation in the silicon germanium. But we don't introduce too much diffusion, otherwise, we can't control the device fabrication.

Here's another issue on slide 34. I'm not trying to make silicon germanium look difficult. But it is a new material. And a lot of things change. Slide 34 shows you that indeed, the germanium itself can diffuse. Well, we're talking about taking a silicon germanium thick layer and putting a thin layer, say 100 angstroms, of strained silicon on it. That's my channel. OK, that sounds fine. We can grow that epitaxially.

But now, I need to implant the source trays and anneal it at 1,000. What's going to happen to that silicon germanium, the germanium and silicon germanium? Well, it's going to diffuse up into the silicon. Once we get germanium into the channel, that's a no no. The electrons don't like germanium in the channel. The electrons can be scattered by the germanium. Mobility goes way down. And your device is ruined. You lose a lot of your strain too.

So you have to keep the germanium out of the channel. Well, it doesn't sound too bad. But if you just look at some of the diffusion coefficients, in fact, these different lines, these are measured. This is the germanium self-diffusion in silicon germanium alloys for different germanium content. So this is for silicon. So this will be silicon self diffusion. This is 10% germanium, 20%, 30, 40 and 50.

So if you just look at a given temperature, say at 950 degrees. Look it's what's happening to the diffusion coefficient. In pure silicon, germanium diffusing very slowly for a low germanium content, here, at around 10 to the minus 17th centimeter squared per second. But if you go up to a 40% alloy, the diffusivity goes up by two orders of magnitude.

So what this is saying is basically, you get an exponential increase, it looks like, in the germanium diffusion coefficient as you increase the amount of germanium in the alloy. So it's a big issue. If you want to do strained silicon on 20%, you're going to have a silicon germanium. You have a certain diffusion coefficient. If you do it on 30%, all of a sudden your diffusion coefficient, factor of 5 or so higher.

So that's a big issue because you don't want the germanium diffusing into the channel. So this data is not exactly interdiffusion data. That whole problem needs to be worked out. But interdiffusion is definitely observed for like 1,000 degrees 10 seconds at 20%. So that that's a no no. So that's going to be a big issue in how we're going to activate source strains in those kind of structures.

That's an interesting research area. You might be saying, well, what is the mechanism of diffusion of germanium in silicon anyway? Actually, it's not really known. On page 35, there's an example of people are just trying to get a handle on this. This was published about a year and a half ago. And this particular experiment, they did a process where they had strained silicon on relaxed silicon germanium.

And they did an anneal in a rapid thermal annealer where they put in ammonia. So they're doing rapid thermal nitratation. So what do we know about nitratation? Well, nitratation, we believe injects vacancies, right? That's what we use it. So in the samples that were nitrided, the germanium diffused right through the strained silicon all the way to the surface. And the sims profile for the blue here shows a germanium went right through that set 100 angstrom layer. It was gone.

The red line is for the germanium that was annealed in an inert ambient. It diffused a little bit during this rapid thermal anneal. But ignore the surface peak because that's sort of a sims artifact. But at least it didn't go all the way. So this is a very preliminary result. And it has to be repeated. But it's sort of indicates that processes that inject vacancies probably should be avoided if you're trying to control the diffusion of germanium.

And so there are a lot of interest in doing processes to inject interstitials, inject vacancies. All the same kind of tests that have done in silicon can now be done in strained silicon on silicon germanium to see how the germanium diffuses or even how the dopants diffuse.

I want to say a little bit now about these MOSFETs on insulator because I've talked a little bit about how the silicon germanium is becoming a problem. How can we get around some of these problems by doing strained silicon on insulator?

Well, on page 37, I'm showing an example of a potential thing you might be thinking of. Well, let's say we could get rid of-- if we go back here, let me go back a few slides, and just look at this on slide 31, this ugly looking thing with all these dislocations and that thick silicon germanium substrate. I need them to relax the strain so I can grow a thin layer of strained silica on top. But I don't really need them once the strain is relaxed.

So that idea was, what if you could take this layer? The top portion looks pretty good. If you could take this whole layer and bond it to another wafer and then get rid of, etch away, all the dislocated material, then you'd have relaxed silicon germanium on insulator, for example. And all this dislocation material would have been etched off or maybe removed by a smart cut like process. I think some of you are studying smart cut. We talked about a little bit we talked about SOI formation.

So, in fact, that's what this cartoon on slide 37 shows where you might have a thin layer of relaxed silicon germanium that has been generated by growing a thick layer and bonding it on. And then you could grow strained silicon on top and make your nMOS and your pMOS that way. So this is one idea that people have been pursuing to get around some of those dislocation issues.

In fact, on slide 38, if you're interested, I've listed a number of different references here on how people are trying to form relaxed silicon germanium on insulator. It turns out, there are a lot of different processes, not just the bonding process. There is a SIMOX-like process where people grow silicon germanium on silicon epitaxily, and then they ion implant the oxygen. That's exactly how SOI is made. So SOI, they ion implant into pure silicon. Here, they ion implant the oxygen into a silicon germanium layer.

So the SIMOX-like process, Toshiba has talked about that. There are some problems though. For high germanium contents of above about 15%, it doesn't work very well. There's another very interesting technique, also came out of Toshiba, called germanium condensation. They actually start with an SOI wafer. So they start with a wafer you buy from a commercial vendor. It would have silicon on insulator. They epitaxially grow a thin layer of silicon germanium with a relatively low germanium content. And then the oxidized the whole thing.

It turns out, in the oxidation, the germanium is snow plowed forward into the layer. And the layer relaxes. So they end up with a higher germanium content after oxidation. And they have written in the layer-- originally strained silicon germanium layer ends up relaxing by the oxidation process at very high temperatures. So that seems to be interesting.

There's a very similar technique called melt solidification by another group. And the bond and etch back is what I mentioned before. You grow the relaxed silicon germanium, bond it to an oxide wafer, and etch it back or smart cut it off. This is just that same process on slide 39. I won't go in great detail.

But you can imagine, by bonding and selective etching, you grow your graded silicon germanium relax layer, you grow your relaxed cap, and then you flip it over and bond it to an oxidized handle wafer, just like you would in a bond an etch SOI. And then you etch everything back.

And you end up with just relaxed silicon germanium after you've done CMP, relaxed silicon germanium on insulator. So that's a technique that's being pursued to form relaxed silicon germanium uninsulated. Then you can grow strained silicon afterwards.

You can also do bonding and hydrogen-induced delamination or smart cut. If we study smart cut for silicon, here, we start with a relaxed silicon germanium epi layer. We ion implant with hydrogen. And at the peak of the hydrogen, bubbles form.

So after we've done the bonding, we heat it up. And those bubbles end up cleaving it off. You get the delamination process. This is exactly analogous to what happens in SOI fab. And now you have silicon germanium on insulator. Of course, it has to be polished and smoothed. But it is a way to form that material.

I'll skip through this at this point because we're basically, it shows, when you make that, and you make a strained silicon MOSFET, you get the same mobility that people have seen in the past. Again, I'll skip that, interest of time. And what I want to end up with is there's another approach people thought, well, why don't you just transfer the strained silicon layer itself? Why do we even have to transfer?

Once you grow the relaxed silicon germanium, you can grow strained silicon on top of it. Bond that to oxide, and then remove everything except the strained silicon. And if the strain remains, then you have strained silicon directly on the insulator. You remove the silicon germanium completely. So you don't have to worry about silicon germanium diffusion and diffusion of dope into silicon germanium and all of that. So this is an ultra thin strained silicon directly on insulator sort of concept.

This is just a cartoon. It's actually been demonstrated fairly recently here at MIT on page 44. In fact, this is the process that was used. A very complicated stack of epi layers was grown. But the important thing is the top of the stack was strained silicon. And that whole stack was bonded to an oxidized silicon wafer. And then everything was removed by selective etching.

So when you're done, all you have left is strained silicon directly on the oxide. And you use etch top layers, so you can etch back the entire wafer and just leave strained silicon behind. And in fact, on page 45, what you see is that the strained silicon is bonded to the oxide. It's low defect density material. And if you do Raman scattering, the strain remains.

Even after you remove the silicon germanium, you annealing at high temperatures, even up to 1,000 degrees, the strain doesn't go away. So the bond is very strong. So now, you have strained silicon directly on insulator. Looks just like SOI, you can process it just like SOI. But the strain is already built in. You don't have any germanium that you need to worry about.

So I think that's going to be a pretty interesting area for the future. Just yesterday, I read in the paper that it had been announced that a company called Soitec is now manufacturing 300-millimeter strained silicon directly on insulator SSDOI. And once it becomes available 300 millimeter, a lot of companies will be able to use it for their production.

So that's about all I have to say on that. I know I need to leave you time. I'm glad I see the surveys here. I hope you haven't been waiting outside. But please take 10 minutes or so and go ahead and fill out the course evaluation survey.