MITOCW | MIT6_774F04_lec04.mp4

JUDY HOYT: Point and start with our actual lecture. So we are in the middle of talking about chapter 4. Again, I hope you're keeping up with the reading in the text. That's part of the reason I don't assign homework continuously. Between now and Thursday, you don't have any homework. You can have time to be reading the text. Chapter 4, we're talking-- discusses wafer cleaning and gettering.

Last time, we talked about some important aspects. We talked about point defects in silicon, neutral and charged vacancies and interstitials. Other important properties of silicon wafers we discussed were the carbon content and the oxygen content. We're going to talk more today about oxygen and Czochralski silicon.

And we got a little bit of a taste last time of some qualitative information on clean rooms and how-- and that's-- we call that the level 1 introduction to controlling contamination on wafers. And these notes, we're going to talk about the two other major techniques for controlling contamination, which are to clean the wafers throughout the process. Every few steps, we do a wafer cleaning. And we'll talk about how that works. And then finally, the principles and modeling of a process called impurity gettering. So that's what this handout is about.

Let's go on to slide number 2 of this handout. And it shows schematically this level 2 approach of reducing contamination. And that's to do-- wafer cleaning. And there are two alternate paths here indicated on the left and the right that we might take a silicon wafer through. The path on the left is primarily what we'll be talking about in this course because it pertains to the front end of the process.

In the front end, by definition, you have no-- you have not done any metallization or have not put any metal contacts on the wafer. That's what we mean by front end. And that's what this course is about. So typically, throughout a process, we talked about there might be 16 or 20 mask levels. Each time you do a masking step and you do photolithography, you have this organic gloop, this photoresist, on your wafer. It needs to be removed.

So before you can go on to the next process step, you need to remove that organic material, that resist. And we do that in what's called a resist strip. That's if necessary. That is if there's photoresist on the wafer. And then after you remove this bulk gross macroscopic organic contamination, then trace organics and trace metals are removed in a step called the RCA clean. And that's named for where it was invented, the old RCA Sarnoff labs in New Jersey. That's where that comes from, the RCA clean.

And then you execute the next process step. So typically, the RCA clean would precede any high-temperature processing step that we'll talk about in this course, such as a diffusion step in a furnace or an oxidation step. You would have to do this RCA cleaning. By contrast, if you're in the back end, something similar happens. You do have to strip the photoresist. And you do have to do some kind of post-metal cleaning, usually before another step.

But the solutions that you would use are much less caustic than the ones we're going to talk about here in the acidic and basic solutions because metals can-- you don't want to etch all the metal wiring off your wafer. So the difference with backend cleaning, it's generally more gentle. And it's not designed to remove metals because that's what you have in the back end. You have metal wiring.

So let's go on to slide number 3. And this, I've taken out of your text. It talks about these-- this front end cleaning process and the steps involved. So if we start at the top, this section up here, these top few steps illustrate photoresist stripping. Again, every time you do a masking level, you typically have this organic film that's somewhere in the order of several microns of this material that needs to be stripped. And it's fairly gross. It's quite thick. It's macroscopic material.

So it can be stripped in a number of ways. This is a very common way of doing it is to take hot, very hot, mixture of sulfuric and hydrogen peroxide. That's a very, very strong oxidizer. It really attacks organics, especially photoresist. Sometimes in fabs today, prior to doing the actual wet chemistry to remove the photoresist, sometimes people do a dry process. And we'll talk later in this course on dry etching.

But there are ways to create in the gas phase reactive species that will also attack the resist and remove it. And an oxygen plasma is often a way that people use to remove photoresist, sometimes prior to doing the wet chemical strip. So after you do this sulfuric peroxide step, there's typically a water rinse, which is shown in here as implicit. And then you dip the wafers in a mixture of water and HF. HF is an etchant that etches oxide, which is all forms of silicon dioxide.

So in this step, there's a strong oxidizer. You end up forming a thin chemical oxide on the silicon surface. This strips it off. And then you do a final DI rinse. So at that point, you have all the gross macroscopic photoresist film removed from your wafer. You still have trace impurities, trace organics, and trace metals that have gotten onto the wafer during the process step that you just did prior to this.

So in order to remove those trace impurities, we do what's called the RCA clean. There are a lot of different variations on the RCA clean. But this is-- the one I'm outlining here is the most classic RCA clean. It has two steps, typically, major steps, one called silicon clean 1, SC-1, and one called silicon clean 2.

The major purpose of SC-1 is to strip organics. It does attack a certain amount of metals, and it also-- to remove particles. So it's good at etching off particles. And SC-1 itself consists of a solution of 511. So this 511 represents the concentration ratio of water to hydrogen peroxide and ammonium hydroxide. And it's usually heated to this temperature range, about 80 degrees C for about 10 minutes.

After that SC-1, after that treatment, then it goes through a deionized water rinse. Again, we talked last time. You have to use very pure chemicals, a water rinse at room temperature and then very often, not always, but very often, right in this step here, in between SC-1 and SC-2, there's an HF dip. Again, the purpose of that is to remove the chemical oxide that you grew in this step to create a fresh silicon surface and then a DI rinse. It wasn't indicated in your text. So I don't-- I put a box for it. But it is very common to have HF dip in the middle there.

And then you go to SC-2. This is-- SC-1 was an alkaline solution, like a basic solution. SC-2 is low pH. It's acidic. And it consists of 611 or something in that ratio, water peroxide to hydrogen chloride. So it's very strong acid. And the key for SC-2, its main role is to remove any other remaining trace metals that were not effectively removed by SC-1. And there are some metals that don't come off very well in SC-1. And also alkali ions. We'll talk about those. But the mobile ions, sodium, potassium, calcium, are very effectively removed in the SC-2 and not so effectively removed in SC-1.

And then there's a DI water rinse. Sometimes there's a final HI dip right at the end, where you remove the Chemox. So you create a pure silicon surface. Sometimes there's not, especially if the next step-- if the wafer is going directly into an oxidation furnace, people leave the surface passivated with a little bit of chemical oxide on it anyway. So again, this is very much a rough estimate of what people do. There are a lot of variations on this theme.

Let's move on to slide number 4 and talk a little bit more about the chemistry. There's more details about this in your text. So if you want to read about that, that will be helpful. An important thing we need to do in these solutions is to remove metal atoms from the surface of the wafer. Again, trace metals are very bad. We'll talk more about this. We saw last time they can reduce the lifetime. They can cause all kinds of problems for logic and memory devices.

So the way that we remove them from the surface is to convert them into ions so that they're no longer neutral. And these ions end up being soluble in the cleaning solution. So they dissolve into the wet solution. And they stay in the solution instead of on your wafer. So to do this, we need to oxidize the metal atoms. That is, we're going to use the term oxidation as referring to removing electrons from the metal atoms.

So this top equation here, shown on this line, shows that the oxidation process of silicon, taking silicon atoms at the surface, combining them with an oxidizer, such as water, and forming SiO2, plus evolving some hydrogen ions and some electrons. So this is an oxidation reaction. It ends up removing electrons.

Similarly, this M here, the second equation, M is meant to represent a generic metal. It could be iron. It could be copper, whatever. Just say iron. By an oxidation reaction can have its electrons removed or stripped and form a positive ion, along with some electrons in the solution. And if you read through your text, this is a table I took directly from your text. This is a table of oxidation reduction reactions for various elements, say some metals. It has silicon dioxide itself. Here's iron and chromium.

And what's shown here is the standard oxidation potential in volts and the oxidation reduction reaction, the actual chemical reaction, in the third column. So example, the second row here shows the oxidation of silicon to form SiO2. This row here with the iron shows the oxidation of iron to form Fe3 plus, a positive to remove electrons.

In general, a stronger oxidant, something that's better at stripping off electrons from an atom, has a more negative oxidation potential. So as you go down in this table, you're getting to more and more negative oxidation potentials. These are very strong oxidants, oxidizers. So H2O2, hydrogen peroxide, is a very strong oxidizer. Ozone, O3, is a very strong oxidizer.

So the way to interpret this table is we're going from fairly weak oxidizers to very strong oxidizers as we go to the bottom. In general, if you have a lot of these reactions taking place on the wafer or in the solution at once, the lowest reaction in the table is going to dominate. And it's going to tend to go towards the left.

So for example, say I have a solution that contains hydrogen peroxide. It's quite down in the-- far on the table. It's going to tend to be pushed-- it's a strong oxidizer. It's going to tend to be pushed to the left. And that tends to drive all the reactions above it to the right, basically, just as a rule of thumb.

So for example, hydrogen peroxide, which is well below iron, tend to take electrons from iron, forming this Fe3 plus ions, and create ions that are soluble in the solution. So it's an effective way of removing-- of ionizing iron and removing it from the wafer. At the same time as this happens, this is happening, the silicon itself will be oxidized by the H2O2, by the peroxide.

So we're forming these ions on the wafer, these metallic ions that can then dissolve. And at the same time, we're growing a little bit of SiO2, a very thin layer of this silicon dioxide. And that gets stripped, that silicon dioxide, in the HF strip-- the HF dip, of a next step. So that's some idea of the chemistry. Again, there are-- your book has a reference to chemistry texts if you want to read that. That's involved in some of this cleaning, cleaning technology.

Let's go on to page number 5. Now this is an actual-- some actual data I pulled out of-- some laboratory data on slide 5. And what it's showing, it's actually taken from a wafer mapping tool made by a company called KLA-Tencor. It's a picture. It's a face-on picture of a wafer. Up here, it's a six-inch wafer. Here's a flat on top. It didn't reproduce very well. I'm sorry. So this is the flat of the wafer. This is the six-inch diameter edge of the wafer.

And this is a machine that takes a laser and scans it across the surface of the wafer. And it looks for spurious reflections. Every time you hit a particle, you no longer have specular reflection. And it notes that. And it says, aha, there must be a particle or what they call a light point defect, LPD. And this is to prevent having to have people sit in the clean room and literally count the particles on a wafer. That would be too tedious and inaccurate. So it's all automated these days.

And it even plots out, the computer plots for you, the point on the wafer. These little black points are places where it found light point defects, where it found nonspecular reflection, something that looks like as if there's dirt or some kind of macroscopic little particle on the wafer. And in fact, it not only does that, it actually tells you roughly the size of the particle.

So if you look at these bins, there's bin 1 through 8 here. So it bins them for you. And these are in microns. So I apologize. It wasn't written down, if you want to put that on your notes. So bin 1 goes-- is any particle that is between something like 0.2 microns to 0.3 microns. So in that size range, it found 16 particles. In fact, the system only found 21 on the entire wafer. So there aren't very many particles on this wafer. It's pretty clean.

16 of them are in this range, reasonably small, 0.2 to 0.3 microns. I don't know which-- where they are exactly on here. But these 21 points are the particles shown on the right. And it found five of them in the range of 0.3, 0.4. And it found zero in all the other bins. Bin 3 is 0.4 to 0.5 all the way up to, say, bin 8, which is very large, a micron or 1 to 7 microns. So it didn't find any particles that large.

So this is a pretty darn clean wafer. This actually happens to be a wafer map from a wafer that's fresh from a box. So this is a prime grade wafer taken right out of the wafer manufacturer's box. It didn't spend more than a few seconds in the clean room before it got inserted into this machine. And it counted what the wafers looked like. So that just gives you an idea of how sophisticated these systems can be these days. And they're getting ones that can go to shorter and shorter wavelengths.

So that's a way of mapping particles. Again, particles are only one thing we're trying to remove, though. Remember we talked about, they're just randomly distributed. A particle is a macroscopic object. It has some diameter or size, 0.2 to 0.3 microns. There may be at the atomic scale atoms of iron, atoms of copper distributed across the surface. They're not in particulate form. But they're just atomic level contaminants. Those are the things we need to remove as well. In fact, those are more of a problem than these particles. The particles are fairly localized. And they don't hurt you in-- for the purpose of doing a lot of experiments.

Let's look at the next slide, though, slide number 6. Here's another example of a wafer map. This is actually the same wafer. Actually, ironically, I'm just trying to give you a little bit of practical education here. This is after an RCA clean. But it was a non-optimal process. Now, obviously, it was non-optimal because it looks like there's a lot more dots on the wafer now than when it was fresh out of the box.

And this can happen if you're in a clean room that's not that clean or if you have water-- particles in your DI water. The last step is a DI water rinse. Or if you put it through a dryer, which heats up the wafer and spins it around. And if that dryer, the spin dryer, is very dirty, it can throw particles at the very end throughout your whole wafer.

So just to make sure we understand the RCA clean, there's chemistry behind it. But there also has to be a lot of practical things that have to be done properly. And so as it turns out, now instead of 21 particles, this wafer now has 1,020-- 1,012 or so, quite a few, most of them added in a very small range. You'll notice they're quite small particles. Most of these are below 0.2-- 0.4 microns or so in diameter.

They are perhaps the most difficult to remove of all things. At the atomic level, removing copper, iron, sodium, we can do that with the-- the particles have a certain mass to them. And they also tend to get stuck. They're electrostatically attracted because they can be charged. So they're attracted to the wafer. So it's a little bit trickier to get particles off. And people are always working on new particle reduction technologies.

This is some other experimental evidence on slide 7, just to show you that the RCA clean does work and that it can work quite effectively. This is not tracking particles. But what it's looking at is how well it can remove certain mobile ions and metal contaminants. And this is a picture of a wafer and a process or a measurement technique called TXRF. And next lecture, we'll talk about total X-ray fluorescence.

It's a metrology technique that can scan across the wafer. And in certain areas and one-centimeter spots, it can measure by X-ray fluorescence the atomic density of certain contaminants on the wafer. And these five points, center, top, right, left, and bottom, were taken. And they correspond to these xy coordinates here.

So TXRF was performed at five points on the wafer. And these elements on the left column here are the ones that were measured, sulfur through bromine. And there are two different types of wafers that were measured. Here on the right in red, these are wafers that were subject to a fairly dirty process called chemical mechanical polishing. That's actually tends to be part of back end. But it can actually be used in front end in shallow trench isolation, STI, process.

CMP uses a slurry of particles and all kinds of various solutions, one of which contains a lot of potassium and perhaps even some calcium. So it's a fairly dirty process. So after CMP, you really have to get the wafer clean. So this was a test on the right in the red, CMP plus a dilute sulfuric clean only. So it didn't do the RCA clean.

And you can see, look at the amount of calcium here, reasonably high quantities. These are all measured in units of 10 to the 10th atoms per square centimeter. So this is looking-- if you looked in a square centimeter, so imagine going into a square centimeter and counting the atoms, you'd find 10 to the tenth of them, roughly. Well, actually, in this position, you'd find 17 times 10 to the 10th, or roughly 1.7 times 10 to the 11th per square centimeter, sitting on that surface. So it's a way of quantifying that.

So here it's in the tens, twenties to hundreds on the right-hand side in the solution that used sulfuric only. On the other hand, when the sulfuric was followed by an RCA clean, all the calcium went down below the detection limit of the measurement. The measurement technique can't measure below about 2 times 10 to the 10th atoms per square centimeter.

If you compare these numbers to what's required in the ITRS roadmap, in fact, I think some of your homework was about what's required in the ITRS roadmap, you know that on the right-hand side, 17, 50, whatever, a hundred and some times 10 to the 10th, that's well above what's allowed to be processed on a wafer for the starting material. So we really need to get this down into this range. And the RCA clean does that here for calcium.

Another example is iron. It's not quite as obvious. But on the right-hand side, there's a little bit more iron. It's measured here. Again, it varies from point to point because there may be some particles on this wafer. And it's picking up particles, perhaps spuriously, which contain iron. This particular spot on the wafer looks very dirty. It's got a high count, about 12 times 10 to the 10th.

And zinc comes down uniformly. Look at the amount of zinc that's introduced by the CMP process that is not removed by sulfuric acid. It's not until it goes to the RCA clean that you get that down below the background. So the RCA clean can be very effective in removing contaminants, even ones that come from a very dirty process, which is basically like a sandpaper type of process almost, the CMP slurry. So it's certainly required.

So let's go on to page number 8. There are actually-- the RCA clean is very old-fashioned. It's been around for, I don't know, 30 years, a long time. And since its invention, a lot of people have tried to improve upon it. And there are some improvements. There are a lot of advanced cleaning processes that are being developed over the years.

And in fact, this is one that's-- I took from your text by a professor named Professor Omhi in Japan, who has proposed this cleaning process. And his goals were to try to get a process that would operate at lower temperature, at room temperature, and that would not remove, perhaps, so much silicon. And it also wouldn't use so many chemicals. This is an amazing thing.

But there's a real concern about environmental factors. There's a lot of research today trying to improve the environmental friendliness of silicon IC processing because the chemicals I just mentioned, all of that concentrated sulfuric acid, hydrochloric acid, it's very concentrated. All that has to go down, be diluted somehow and treated before it goes out into our wastewater stream. So people are trying to minimize the amount of these chemicals that are dumped into the environment by inventing new cleaning methods that are more environmentally friendly.

So he feels this is more friendly. But it has a lot of the similar principles. Notice the first step is to put it in ozonized water, so water that has ozone in it. So if we go back a couple of slides, let's see, if I go back to slide number 4 here, look at ozone, O3. It's the strongest oxidizer on the list, so even stronger than peroxide. So he creates ozone and puts it in the water. And a very strong oxidizer to strip organics that photoresist and even some metals.

The next step-- again, and these are all at room temperature, which is nice. So you don't have to heat things up because, when you boil acids, they tend to vaporize and they go up into the atmosphere. And that also creates a lot of atmospheric pollution. So not having to boil is nice. Then he has a little bit of HF, peroxide, and water in a surfactant. Surfactant helps remove particles from the surface. And he puts it in a megasonic. This is a very high-frequency vibrational technique that actually tries to vibrate the particles off.

So he puts a soap-like substance, which is a surfactant, a very high purity, to remove particles and more metals. He then goes back to this strong-- oxidizes the ozonized water to strip any chemicals that adhere to the wafer and then does an HF dip to get rid of the chemical oxide, produce a passive surface.

So there's a lot of variations on the RCA theme. I've included here one reference from 1999 that I-- from Mark Hines, I think is a nice reference, it's a little old now, but on alternate cleaning processes and the efficiency of cleaning processes with respect to how well they clean the surface prior to gate oxidation. Gate oxidation is one of the most sensitive high-temperature processes you have to do. We don't want to incorporate any mobile ions, or iron, or anything in our gate oxide. So he's got some-- it's an interesting article if you want to know more about details about surface cleaning.

So that's basically cleaning. Let's go now-- now I'm going to talk a little bit, the last topic, on contamination control, which is gettering. But before we go to gettering, I've just been indicating certain elements are bad or whatever, have properties. Let's say a few words about the elements, which we call deep levels in silicon.

And this is a chart that I took from Simon Sze's famous book on the physics of semiconductor devices. I took it right out of his chart-- his book. And what it shows is the silicon bandgap here is being-- at the top, the conduction band, the bottom of this-- straight line at the bottom, the valence band. And the dashed line is the mid-gap point or the gap center.

And what he has is a series of elements that, when you put them into silicon, what kind of states do they create in the bandgap? There are some shallow levels here. For instance, if we put boron in silicon, we all know that's a p-type doping. It's very shallow. It lies very close to the valence band. And it can accept electrons, and create holes, and dope your p-type. So boron is not too much of a problem.

But the things that I have put a box around here, for instance, zinc and gold, have deep levels very close to the middle of the band gap here right near the dashed line. Copper as well has a mid-gap level. Iron has a level right next to mid-gap. It turns out when you have this deep level, this state, this electronic state that's close to mid-gap, it creates a generation and-- it creates a center where a lot of generation and recombination can happen. So it can trap carriers and reduce carrier lifetime.

So the worst place, the worst types of elements then, the bad actors, are the ones that have deep levels right near mid-gap. And you see there are some of the things that people commonly talk about in clean rooms as you need to avoid. For example, copper, right here, it's got a number of deep levels, one at mid-gap, iron, zinc, and gold. So they're very-- deep levels are very undesirable. You should take this, though, summary chart with a grain of salt. It gives you a rough idea. Don't take it too literally. It's representative, however.

If you want to look at it a different way, you move to the next slide. And I took this chart. This is actually a colored version of the periodic chart that I took from your textbook. And it characterizes the elements according to their position in the periodic table. And here in column 4 is where silicon is, of course. That's our semiconductor.

The shallow acceptors, which make the semiconductor p-type which we need to add in, are typically boron. They're shown here in column 3. The shallow donors are in column 5. Of course, they have one extra valence electron compared to silicon. So those are fine elements. The problems are here. The problem elements, as you notice, tend to be in the transition metals. These form these deep levels, these Shockley-Read-Hall recombination centers. And a lot of them are in the transition elements.

The other problems are shown over here on the left, the alkali ions, lithium, sodium, potassium, as well as calcium is also included in that category. These are a problem not so much because they create deep, deep levels in the silicon, but because they impact the quality, the electrical quality, of the gate oxide, its ability to act as a good insulator. They can cause shifts in the threshold voltage, their presence. They tend to be mobile in oxide under bias. So that's a problem. So we certainly want to get rid of these.

So the idea here in gettering, which is a process we're going to talk now, is we're going to take all these unwanted elements, the transition metals, some of the heavy metals, the alkali ions. And we want to collect them in a region of the chip. Either we want to avoid introducing them, which we'll try to do, or if we can't totally, we're going to collect them in the region of the chip where they won't be harmful.

In order to do gettering, and gettering is most effective for the transition metals, we need to take the elements wherever they are during the process and make them mobile. So they have to get away from their position near the device and be attracted to another part of the wafer. So they have to then diffuse. So they have to be made mobile. They have to then mobilize and diffuse to the trapping site. And then they have to stick there. So there's three processes that you have to do in order to getter impurities.

So let's go on to slide number 11. And this is a schematic illustration and discussion of the methods of gettering for alkali ions. And this is a cross-section, meant to be a cross-section of your wafer starting here at the bottom in red, the back side of the wafer. And up at top is the front side of the wafer. There are some devices here shown. These little regions are supposed to be p-n junctions, perhaps.

And in pink at the very top is what's called-- someone has put on the wafer a layer of PSG, also called phosphosilicate glass. It's actually an SiO2 layer, which it has a lot of phosphorous in it, up to 5%. So it's an alloy. This is an old-fashioned technique. People used to use it because alkali ions tend to be trapped in that glass. It's not so much in use anymore because it tends to absorb water later on in the process. And if you are using aluminum metallization, it can cause corrosion.

So the approach to alkali ions these days as opposed to trying to gettering them is a little different. People just try not to get them in during the process. And then once the chip is fabbed, they don't want them to get in after fabrication. They put a protective layer. So the alternative of putting PSG in here is to put silicon nitride. It's a relatively tough layer. It doesn't scratch very well.

And it protects the chip from alkali ion contamination after processing. So once the chip is finished, there's a scratch mask put down of silicon nitride. And then alkali ions that would come on the chip surface later after fabrication can't get through it. Now, so this isn't really so much gettering anymore. It's more like preventing it from happening. And this assumes, of course, that the processing steps are free of alkali ion contamination.

That's why it's so important in processing, if we go back to this periodic chart, that we try to keep people out of the lab because all of us are full of sodium and potassium. Without sodium and potassium, we'd all die immediately. Our hearts would stop beating. So we're just exuding sodium, potassium, calcium, our bones, calcium, everything. So people are one of the worst, worst things in the lab. And that's why you would never touch a wafer or do anything where anything from your body can get onto the wafer because we have a lot of alkali ions in us.

So trace metals are a little different. There are some pretty effective ways of gettering trace metals. In fact, there are two methods. One, they're both shown here schematically on slide 12. The first one is called extrinsic gettering. What does extrinsic mean? Well, extra. You externally put something-- do something to the wafer. Usually, on the back side, you often put a layer of something, which is created to trap ions or trap metal ions down here. So it's usually extrinsic to the wafer. That is, it's extra. It's been added on.

The second way is to do what they call intrinsic gettering. And that's to, in the center of the wafer, in the middle region here that's shaded, is to intentionally form a region full of oxygen precipitates. And these precipitates create little traps where the metal atoms can diffuse to and get stuck. And they're within the bulk of the wafer.

Generally, for most ICs, the bulk of the wafer is not part of the active region. It's just a handle just holding the thing. So it's OK to have the metals there. Now, of course, if you're making a power device, where the currents flow through from the top of the wafer to the bottom of the wafer, this would be a problem. But for most microprocessor memory, this would be effective.

Intrinsic gettering is more popular these days than extrinsic. For these reasons, it's a little bit-- has a little better control. The gettering region here is actually closer to the devices. It can be within 10 or 20 microns. So it's not so hard to get the bad elements out of the devices and into the gettering region. And the other thing about it is that these SiO2 precipitates tend to be thermally stable once they're formed. So it's usually works throughout the entire process. Not always the case with the backside gettering.

Let's go on to slide number 13 and think about, what are we talking about? We're taking elements away from-- impurities out of the device regions during the processing. And we somehow have to get them to diffuse to these trap sites. So the first thing we need to know is, how rapidly at a given temperature are these elements going to diffuse? If they don't diffuse fast enough, there's no way you're ever going to get them to the traps.

So there's some good news and some bad news. The good news is that most metals diffuse as interstitials, so in the interstitial spaces in silicon very rapidly, which means you have a chance of grabbing them, of getting them to go from the device region to the back of the wafer or into the bulk of the wafer. And in fact, here's some numbers for you quantitatively on this plot.

The left axis shows the diffusivity. And we measure that in units of a length squared per second. We're going to talk about the process or per unit time. We'll talk about the process of diffusion later in the course. But the bigger this number, let's put it that way, the faster the elements will move at a given temperature throughout the wafer.

And this is what's called an Arrhenius plot. So it's plotted versus 1,000 or 1 over temperature on the bottom axis. If you want to read temperatures in actual units of centigrade, you can read it up at the top axis from here, from 800, all the way up to 1,200. And the different curves are for different elements. And all the metals here are indicated in red. So here's some common metals, things that we mentioned as being very bad for lifetime killers. Here's copper. Look at copper, almost irrespective of the temperature in this range. Copper is zipping through the wafer faster than any of these others.

Here's gold, Au, as an interstitial form. It diffuses quite rapidly. Gold here, when it's substitutional on the lattice, so the gold atom ends up being substitutional, doesn't diffuse quite so fast. It's orders of magnitudes down. Looking here at these, these are the diffusion coefficients in blue here for the shallow impurities, like arsenic, boron, phosphorus. They diffuse much slower.

And this dark, I guess, stippled region that's colored in tan is roughly the region where people believe the silicon interstitial diffuses. Again, no one's been able to actually measure exactly because you can't see a silicon interstitial. So you can't profile it, how rapidly it diffuses. But there's indirect ways of inferring it. So somewhere in this band, this brown band, is where the silicon interstitial diffusivity lies. It's not bad. It's not as fast as the fastest metals. It's not as fast as copper or sub-- or interstitial gold or iron. But it's in the range of some of the metals.

Just to give you an example here, this bottom bullet shows, let's say take copper at 900 degrees. You can read off this chart that diffusivity, D, is about 10 to the minus 4 centimeters squared per second. We'll learn that a rough estimate of how far a profile diffuses can be the square root of D times the time. So if I put this wafer that accidentally has some copper by accident on the surface, in the furnace at 900 degrees in one minute, the square root of dt is 780 microns. And the wafer is only typically 600 microns thick.

So basically, that copper can go right through the wafer and be anywhere in the wafer. That's bad because, if you accidentally get copper on the back of the wafer, let's say where you say, oh, it won't hurt anything, you put it in the furnace, well, it sure will. It'll be in the front real quickly. It's good if you're trying to getter because, if the copper happens to be in your device region, you can try to get it to the back of the wafer. And the key is to hold it there. So it's both good and bad. But metals are very fast diffusers. Well, and that's-- copper is an extreme example.

So let's go on to slide 14, which talks about metal gettering a little bit more. And this plot on the vertical axis is temperature. And the bottom, the horizontal axis, is atoms per cubic centimeter. It represents the solubilities of the fast-diffusing metal as a function of temperature. So if you look at any point on this curve, if you have a concentration higher than that at that temperature, then the metal will tend to precipitate out, will not be in solution.

So metals preferentially like to reside in sites on the silicon lattice where there's imperfections. So this is key. We can use this, this idea. Why might this be? Well, the metals don't fit because of their atomic size different from silicon. They don't fit into the silicon lattice. The other thing is that if you have a fault or a disordered region in the crystal, it might be able to accommodate this size difference and trap the metal atom there.

And it turns out that dislocations and stacking faults, these are crystal imperfections that can exist in the wafer, sometimes, they are well known as trapping sites for-- or sites that where copper, gold, and iron tend to accumulate and precipitate. So if you have dislocations, depending on where they are, that could be a good thing. You could try to create dislocations.

And these, they tend to be decorated, so to speak, with these copper, gold, and iron. So the trick, the name of the game, is to form these kind of defects intentionally, not in your device region, because then you'd have copper, and gold, and iron there, but away from the active region. And so the idea of metal gettering is how to form these types of things.

So on slide 16, we talk a little bit about how extrinsic gathering can take place. Usually, in the extrinsic case, they try to form, people try to form, these sites on the back side of the wafer. And there are a number of ways of doing it, some very crude, grinding and sandpaper. Sometimes you'll look at the back of a wafer, especially an older wafer. And you look at it and there's this ground pattern, almost look like it was put in a lathe. And it goes around, this pattern that goes on the back side.

That was intentionally ground into the back of the wafer to make a lot of these gettering sites. They used to use sandpaper abrasion. That's not so popular anymore. People these days use the cleaner processes or the cleaner methods. Ion implantation can be used to damage the back of the wafer, a deposition of a polycrystalline film. It's not that unusual to find polysilicon on the back of the wafer or a region very highly doped with phosphorus on the back of the wafer.

These are all regions where metals, if they diffuse back there, will get trapped and stick. So the whole idea is, let's make extended defects that are stable throughout the high-temperature processing. The problem is if they tend-- these extended defects tend to heal, the metals will then be released and they'll go back to the front of the wafer.

So that's extrinsic gettering. So if we go on to page or slide number 16, we have a schematic of this intrinsic gathering, which is a little more common. Remember the last time we talked, we said that oxygen is present in all Czochralski silicon. It comes from the crucible. There's nothing we can do about it. And it can be introduced in this supersaturated form. That is, oxygen can be dissolved in there because the wafer was cooled at a relatively rapid weight-- rate.

But there's more oxygen there than the crystal wants to hold at that temperature in equilibrium. So if you were to heat that crystal up and give it a long enough time, the oxygen could then tend to form-- precipitate out and form SiO_2. So oxygen forms these SiO_2 precipitates in the crystal. And the interesting thing is that these precipitates are usually accompanied by some kind of defect, mechanical defect, like a stacking fault or an extended defect, like dislocations.

And that's because of the volume mismatch. SiO_2 wants to expand. These little SiO_2 precipitates want to expand. They compress the silicon lattice around. If they compress it enough, it can pop out these extended defects. So the key point of gettering is to get this precipitation to occur in the bulk, in the bulk of the crystal, throughout the bulk of the wafer, but not in the near surface region.

So how are we going to do that? Well, have to have a near surface region that doesn't have high oxygen so it can't precipitate. So what we do is we create something called a denuded zone, which is low in oxygen. And there's a couple of methods of doing it. This denuded zone, which is pictured schematically here, typically might be 10 to 20 microns thick. It sits on top of the wafer, the wafer, again, being full of oxygen.

So we can do that-- there's a process we can use called epitaxial growth, which we're going to talk about later in this term, where we can grow material on top that's very low in oxygen. Or if you don't want to do epitaxial growth, because it's fairly expensive, you can just take the wafer. And the very first step you can do is go to a very high temperature and cause the oxygen to outdiffuse, out the surface, out the front side of the wafer. So whatever oxygen was in the top 10 microns can go out.

Oxygen buried deeper below that at this high temperature can't make it because it can't diffuse-- it doesn't have enough time to diffuse that amount of distance. So one of the first steps people use for doing intrinsic gettering is sometimes a high-temperature step to create this denuded zone or to grow an epitaxial silicon layer.

So let's go to the next slide, slide 17. And the next couple of slides are going to outline the thermal processing. So just by doing these thermal processing steps, you can do intrinsic gettering. And this plot shows schematically on the vertical axis the wafer temperature and time. Again, time isn't really indicated here quantitatively, just to give you a rough idea.

But the first step is to take the wafer at a very high temperature, say 1,100 degrees, and do the outdiffusion step. So we're going to take that wafer, put it in the furnace, and let the oxygen go out, not all of it, but whatever is in the near surface region. And this expression in the middle of the slide on slide 17 shows the diffusivity as a function of temperature. It gives you a rough idea.

And the nice thing is you don't have to remove all the oxygen. You just have to get the level below a certain level. So for instance, you need to get it from about 20 part per million down to less than 10. If it gets low enough, the precipitation won't take place because, again, it'll be below the solubility limit. So you don't have to worry about precipitation. So you're just trying to reduce it down. And usually, 1,100 to 1,200 is sufficient to create a denuded zone. And that's in the order of 10 to 20 microns deep. It has-- the lower oxygen concentration then will precipitate. So that's the first step is a very high-temperature step.

Next step in the process, shown on slide 18, is called nucleation. So here we go from this high temperature, where we got rid of the oxygen in the surface region. And now we want to nucleate the precipitates in the bulk. So we go down to a lower temperature, say 700 degrees. And we want to cause very, very small precipitates to take place.

And the optimum temperature for this is somewhere around 700. And this is discussed in chapter 3 of your text. It's optimum because you need these nuclei to grow to a minimum critical size to prevent them from shrinking later when you start to ramp up the wafer temperature. So the critical size is about two nanometers, say, one to three nanometers in diameter, this precipitate of SiO_2. And you need to-- you typically shoot for a density about 10 to the 11th per cubic centimeter. So there aren't too many of these. But you definitely have a density of them.

So we create these nuclei. And then we actually do-- we grow them to be a little bigger so they become stable. So once you've created this nucleation step, you then take it to higher temperature, say 900 degrees. And you grow these nuclei to make them a little bit larger. So you're not creating more. You're just making them larger.

And you typically want a size, a minimum size, say, in the range of 50 to 100 nanometers. You have to be careful how you ramp this up so you don't cause them to all shrink. The basic idea is then you've enabled yourself to create these SiO_2 precipitates in the bulk of the wafer. So we go on to the next slide, slide 20. This, I took from your text. These are some actual cross-section scanning electron micrographs.

And you're actually looking at the cross-section of a full wafer here in each picture. Each picture is a full wafer. So this top here at the very top is the top of the wafer. This bottom surface here at the bottom of one little rectangular region is the bottom of the wafer. And you'll notice, what you see are all these little black dots, are these SiO_2 precipitates.

And there's a bunch of snapshots here shown, starting from the top. And there are two different directions. If you go down vertically, we're reading the nucleation time. So remember, the nucleation time is the time that we nucleate these. And so if you-- starting at the top here in the upper left and moving down the first column, we're increasing the pre-anneal time. And what you're seeing is-- what you're doing is you're getting a higher and higher density of these precipitates as I move from the upper left down to the lower left.

And then as I move right on this chart, I'm increasing the growth time. That was called the precipitation time or the amount of time we're growing these. Now these are grown at about 1,000 degrees. And you can see them growing here. And so by the time I get to this upper right, we have a certain density of these.

But in general, if I move to longer nucleation time, so down here, and longer growth times, so I'm moving to the right, I get the highest density and a reasonable size of these precipitates in the wafer bulk. And you can see all of them here in the lower-- the bottom right corner. So these are precipitates. And look at the very surface of the wafer, the top 10 microns. You don't see any. That's the denuded zone. And that denuded zone was formed, of course, at the top and bottom.

Denuded zone was created by doing an initial 1,000 degree C 10-hour anneal before any of this pre-annealing. So during that initial annealing, the oxygen diffused out at the top and bottom of the wafer for about 10 microns distance. And then it didn't diffuse out in here. And that's why when you go through this processing step, you can create these precipitates. And these are the precipitates that we want there so that metal ions can then diffuse from the devices and get stuck. And then keep them away from your devices. So that's the pictorial schematic of the process.

So then let's go on to slide number 21 and now talk about way people model this process. And there are typically three steps people need to model. The first step here, shown number one, happens to be-- I'm showing a specific case for the gettering of gold because gold is a very, very common element. It needs to be gettering. And it has some unique properties.

Step number 1, we need to mobilize the atom. So remember we talked about gold when it was substitutional. It diffused very slowly. If it's interstitial, it diffuses fast. So to make it mobilized, I need to somehow get it off substitutional sites and make it interstitial. The second step is shown schematically, step number 2 here, which is the gold has to actually diffuse away from the device region down to the trapping site, wherever it might be. And the third is, of course, the trapping mechanism, number 3.

So people have tried to model these different aspects, the making mobile, the diffusion, the trapping, with mixed success. The models are actually not universally accepted. There is no really good program you can sit there and run that describes gettering, and very quantitatively, in the way that you can for oxidation that we'll see in this course. But they give you a physical mechanistic picture of how the getting operates. But they're not very quantitative.

So let's look at those three steps. The first step was to make the atoms mobile. And remember, gold and a number of atoms can actually exist either on the lattice or in interstitial form. The diffusion is-- the diffusivity is much higher when it's interstitial. So to make the mobile really means to get them off the lattice into the interstitial spaces where they can diffuse.

So then these-- just to show you some examples here of common metals and the different types of solubilities they have in interstitial or substitutional form. So copper and nickel have higher solubilities in interstitial form. So they tend to exist already interstitial. So mobilizing them is easy. They're already in the interstitial space. Gold and platinum actually have higher solubilities in substitutional form. But they can diffuse very rapidly once they become interstitial. So these have to be mobilized.

Titanium and moly, molybdenum, are actually issue-- problems. They're primarily substitutional, but they have relatively slow diffusion rates, whether they're substitutional or interstitial. So Ti and moly are going to be hard to getter because their diffusion rates are not very large. So to get them all the way to the back of the wafer is going to take a long time. So it's just not very effective.

So to getter Ti and moly in these intermediate or slow-diffusing metals, we prefer to use intrinsic gettering. So instead of going 500 microns, all it has to getter through is the denuded zone. All it has to diffuse through is about 10 microns. So that's another reason why intrinsic gettering is more popular. You don't have to diffuse as far.

So let's go on to slide 23. And we'll just do an example of how people try to model the gettering of gold and silicon. Or this also applies to platinum. So the first thing, we said that gold can react with a silicon interstitial. And this Si sub i is meant to represent a silicon interstitial and form an interstitial gold.

So an interstitial silicon atom can come along, take the gold which is on the silicon lattice, and knock it off, and take its place. And that's called the kickout mechanism. So that's one way of getting the gold off of lattice site because there are other mechanisms. An interstitial can react with a vacancy and form substitutional gold.

So basically, the idea, if you look at this first reaction, this first equation, any process that creates excess interstitials should be helpful in gettering. If I have a lot of excess silicon interstitials around whenever there's a gold, it can just knock it off the lattice and get it moving. So process that-- any process that creates excess interstitials should be helpful in this gettering, this mobilization process.

Processes that create excess vacancies will tend to hinder gettering because they're-- if you have a lot of excess vacancies, it's going to drive this back to the left, the second reaction, and cause gold to go back to substitutional. So interestingly, just from an empirical point of view, a lot of the things people do for gettering happen to also be known to create excess interstitials. So for example, backside phosphorus diffusion, when we talk about phosphorus diffusion, we'll see how it tends to inject a lot of excess interstitials.

Ion implantation, in which we shoot ions into the crystal, tend to knock silicon off lattice sites and create a lot of interstitials. And internal gathering or intrinsic gettering involves oxidizing silicon, which we will find out. Creating SiO2 precipitates tends to punch out a lot of excess silicon interstitials. So qualitatively can we say that the things that we know work for gettering also tend to create a lot of interstitials. And that would be consistent with this type of model, the fact that some of the metals need to be able to be mobilized.

So then go on to slide 24. And the second step is, once its mobile, is to get the metal to diffuse to the gettered site in this interstitial form. And so it has to diffuse to the back side or into the intrinsic region. Let's say we're doing diffusion to the back side. This plot on slide number 24 is a plot of the concentration of that gold. It has a function of depth in the wafer, where the [? 00 ?] point, or this point right here, is supposed to represent the wafer backside. So that's the backside of the wafer.

As you move to the right, you're going through the thickness of the wafer. So you're going into the center of the wafer. And I move all the way to the window, I'd be at the front side of the wafer. So imagine the wafer is standing up sideways. And this is a plot of different time contours of what the gold concentration would look like. Initially, at time t equals 0, it's a flat line. It's just at 10 to the 15.

So let's say, somehow, the wafer got 10 to the 15th per cubic atoms of gold in it uniformly distributed. Now you heat it up and you increase the amount of time. And as I heat it, basically, the gold is diffusing out of the wafer. It's lowering here in the center. And at the edge, it's diffusing down. So this is what we expect the gold profiles to look like over time, just in a rough sense.

If we go on to slide 25, actually, there are some plots which-- there's a plot which shows, as a function of time, at 1,000 degrees, what happens to the gold concentration profiles, actually, experimentally. And this represents the observed profiles. They're actually not quite like what was shown in slide 24. If I go back here, you see this is a much more-- in slide 24, it's a much more gradual drop off on this plot. In slide 25, in fact, it's very abrupt. It's very sudden than predicted by this simple model.

Well, it turns out that people explain the sudden drop off by the fact that the silicon interstitial diffusion is fast. But it's actually slower than the metal atom diffusion. So the rate-limiting process in this step is for the silicon interstitials to diffuse in the back or inside into the wafer and mobilize the gold. So as soon as the silicon interstitials arrive at a particular location, the gold is converted to interstitial form. And then it's got-- it's diffusing so fast, it can rapidly diffuse out the wafer. So what's rate limiting here is the indiffusion of the silicon interstitials.

So this sharp drop off point here, you see at 12 seconds-- 12 minutes, it's dropping off here, 20, in 25, 30 minutes. That sharp drop point corresponds to the depth people believe which the silicon interstitials must have diffused in a given time. In fact, in the old days, people used this as a marker for how to measure or estimate silicon interstitial diffusion.

They couldn't see the interstitials, but they hypothesized they were having this effect on the gold. So they would model the gold diffusion and the silicon interstitial in diffusion that's associated with that. And so that was one method of estimating silicon in diffusion So anyway, that's just a practical issue with the way gold diffuses through wafers.

And in fact, on slide 26, this is just to remind you of how the relationship or the relative magnitude of the gold diffusion, look at the gold interstitial diffusivity up here, compared to the silicon-- the diffusion of silicon self-interstitials. If you have an element like titanium, which diffuses slower than silicon interstitials, it's going to have a more classical type of profile, as shown here on the lower right, for its diffusion profile of titanium outdiffusing or diffusing to the back of the wafer.

So we've made the mobile. We got them diffused to the site. The final thing is trap them. It doesn't do any good to make them diffuse there and then have them come right back out again. So they have to be trapped. And this tends to be very, very empirical. But people have experimentally observed that certain types of backside damage trap metal atoms. If you measure the concentration in that backside region after you've damaged it, it has a higher concentration of metals in it.

It's hard to model this. But one approach is to model it using the mathematics of segregation and to talk about the gold as it exists in the bulk and the gold as it exists in the trap site and compare those two concentrations. Remember, in chapter 3, we talked about segregation, described the doping behavior as it segregated between a liquid phase and the solid phase during Czochralski growth. Well, the same concept here really applies here. We have the segregation of these gold or this metal atom between the silicon bulk and the trapped region, or the backside trapped region. So we can use similar mathematics.

And that's what's shown on slide 28. This is a relatively simple way of looking at a segregation model for metal atom trapping during gettering. And what we do is we write down the solubility of gold in two states. On the upper equation is the solubility of gold in bulk silicon. So this is not in the gettered region, just existing in the silicon. And it has some exponential dependence on an activation energy, Ea1 over kT.

And the second equation down is the solubility of gold in the gettered site, G. So it's got this G next to it in the gettered region. And here Ng, N sub g, is the density of gathering sites. N sub silicon here is the density of silicon atomic sites. And what we define as the segregation coefficient, just as we did before in liquid solid segregation, K0, as to be the ratio of the concentration of the total gold in the wafer divided by that in-- that's in the silicon. So that's our segregation coefficient. And we can just simply ratio-- plug-in these exponentials and ratio them, as shown in the equation.

So if Ea1, the activation energy a1, minus Ea2 is a positive quantity, then K0 is going to decrease as temperature increases. So this tells us something about where the metal wants to reside. In fact, empirically for phosphorus backside gettering, it's been found that the segregation coefficient can be written like this, a simple equation, where it has a positive activation energy.

So what this says is we need to increase the number of gettering sites, Ng, which, of course, just means keep the amount of phosphorus in the backside region high. And we keep the temperature low, trap it there with a relatively low temperature. So we don't want to go too high to get the highest amount of segregation. So again, it's fairly qualitative. But it gives some estimate of what-- of how gettering works as a function of temperature trapping.

That's one example of a model. The second example is shown on page 29. And this actually goes back to work that Shockley and Mole did very long time ago back in the 1960s. And this work was to understand the enhanced solubility of metals in heavily-doped silicon. It was observed that certain metals are more soluble. They tend to want to be in heavily-doped silicon. They have a higher concentration in heavily-doped silicon than they would in lightly-doped silicon. And this is a model that helps to explain that.

So in substitutional form, we know that gold introduces deep levels in the bandgap and near the mid-gap position. In fact, there's one here, an acceptor level marked here. And there's one here, a donor level, marked there. So this Au minus level is going to be created whenever the gold can capture a free electron. So if there's a lot of free electrons, we can create this Au minus state, so this acceptor level.

So I can just write down this simple chemical equation, gold, one gold atom, plus an electron goes to an Au minus ion. Now, if you're familiar with chemistry, you know you can write any reaction. You can write an equilibrium constant associated with that reaction rate that depends only on temperature. So we write this number, K sub equilibrium. It depends only on temperature.

It says that the reactants, the concentration of reactants, Au minus, divided by the concentration of the product of the concentration of the-- or this is the products divided by-- I multiply the concentration of the reactants. That has to be a constant. So the right-hand side of the equation, the concentration of Au minus divided by the gold concentration times the electron concentration, has to be a constant. It depends only on temperature. So that's an important equilibrium relationship.

If we go on to slide number 30, here just at the upper left-hand side, just repeated that equation. This equation has to hold both when we have intrinsic silicon-- an intrinsic silicon, the electron concentration is just ni-- or when it's extrinsic. In extrinsic, the electron concentration is just n, whatever you've doped it to, the donor density level.

So if that's the case, I can write this equation for those two cases. And it has to hold. So basically, I can write down, as it turns out, that the ratio of the gold, Au minus, in N-type material to its concentration in intrinsic material just has to go like n over ni. So then what that means is the solubility of gold as an acceptor is going to be higher in N-type material.

So as I raise n over ni, let's say I dope-- I'm at 1,000 degrees. ni is 7 times 10 to the 18th. And I create a heavily doped region, 10 to the 21. Then just looking at n over ni is a factor of 100, roughly. So that means the gold solubility is going to be 100 times higher in this heavily doped 10 to the 21 region compared to the intrinsic region of the wafer.

So again, this is looking at it from the electrochemical point of view. And it's a little bit hand-wavey. But it's consistent with the fact that when people do heavy n plus diffusion regions on the back of the wafer, like phosphorus, they can very effectively hold metals because it has a higher solubility, just because of the electrochemistry involved.

So on slide 21, it's qualitatively listed as a series of explanations of why n plus silicon is a good getter. It's a good trap, trapping area. We just saw the statistics of the acceptor and donor levels versus doping from Shockley's paper that showed-- that is consistent with that. There's also people that have an ion pairing model, that a large atom like gold, well, gold is very large, it might want to pair with a small atom like phosphorus and form some kind of AuP complex because this minimizes strain.

The other issue is that point defect concentrations are much higher in doped silicon than intrinsic. Remember, last time we talked about, as I move the Fermi level around, I can create more interstitials or vacancies. Gold diffuses primarily by an interstitial mechanism. Once it arrives in the getter region, it needs to find a lattice site to become substitutional. So you can imagine an interstitial gold might react with a vacancy in the lattice. And it gets stuck there to make it substitutional because, remember, substitutional gold doesn't diffuse very fast.

So in n plus silicon, we saw last time, it has a much higher population of vacancies than in intrinsic silicon. And so that's going to tend to drive this to the right. So once the gold gets-- it diffuses by interstitial mechanism, it gets to the back side to the n plus, there's plenty of vacancies around to give it a lattice site to get stuck onto. So that's another, again, somewhat qualitative explanation.

So that's for backside gettering with an n plus region, such as phosphorus, which was the classic method. How about intrinsic gettering? What kind of models do people have for gettering atoms or-- metal atoms near SiO2 precipitates? Well, actually, this tends to be a lot more qualitative. But if you go back to chapter 3, chapter 3 talks about how SiO2 precipitation takes place. And in fact, it involves a net-- volume expansion. The SiO2, when you add-- when you form it, it tends to increase the lattice locally, wherever that is. And it actually compresses the lattice around it.

So people have written this-- it looks fairly complex. But each term has a meaning. This equation, which describes the formation internally of silicon reacting with oxygen interstitials. So here's a silicon lattice site here. A certain number of lattice sites are involved, reacting with oxygen interstitial, O sub i, and forming SiO2. It could be at the surface. It could be internally. But you're forming a small SiO2 region, along with some stress that gets involved.

Look at this fairly complex equation. But some of the elements of it are identified. Gamma here is the number of interstitials that contribute to the precipitation process per oxygen atom that join into the SiO2 precipitate. So a certain number of silicon-- silicon is a lattice site. A certain number of silicon lattice sites participate in this reaction.

Interestingly, look at the process. It consumes vacancies. So vacancies on the left have to be consumed. And it injects on the right interstitials. So the formation of SiO2, and we're going to see this in a lot more detail when we talk about oxidation, planar oxidation, involves the injection of excess interstitials, which is interesting because excess interstitials, we know, are important for mobilizing certain metal atoms. So that could be one reason why SiO2 precipitation is important.

We know that the stress term is generated because it compresses the crystal around it. This can cause macroscopic defects. And the metal atoms tend to be located around these stacking faults or dislocations. If you want to do it in a quantitative or semi-quantitative manner, you can use the same model. Same mathematical segregation model was used as was used for previously. And again, this reaction creates interstitials, which will help free up substitutional gold to make it ready, interstitial for diffusion.

So on page 20 or slide 23, I just want to summarize gettering. It's still fairly qualitative. I think you can tell. These gettering methods really haven't been changed much in the last 10 to 15 years. There is a better understanding of how it works. But we really still need better models. In the future, intrinsic gettering, this process of creating the denuded zone either by epi or whatever, is going to dominate because the thermal budgets for IC processing will decrease.

And trying to diffuse those atoms all the way to the back of the wafer, especially titanium and moly, which have lower diffusion coefficients, is not very practical. So having the gettering site close is better. Clearly, the amount of oxygen, the oxygen concentration in the wafer, is going to be critical for this whole gettering process. So we're going to need tighter controls so that we can really control the gettering process, and also to control other things like wafer warpage. We talked earlier about how oxygen actually increases the mechanical strength of silicon.

So newer techniques are needed to do gettering at lower temperatures. So that's an area for work. And there is really no accurate simulation tool. In this class, for homework, you're going to learn how to use simulation tools called SUPREM4 to model processes like diffusion, oxidation, whatever. There is no such simple tool for gettering. It's still very hand-waving qualitative arguments, like we've given today.

And I'll finish up by summarizing chapter 4. The last couple of lectures, we talked about particle control, wafer cleaning, and gettering processes. These three things, all of these are crucial for successful manufacturing and even for doing the research that you do here in MTL. There's a three-term tiered approach. It was to keep the air and the water clean, continuously clean the wafers in these chemicals, and use the RCA clean, and finally, to do gettering.

Gettering involves releasing the impurities, diffusing them to the trapping site, and getting them trapped. The good thing is that most metals diffuse relatively fast. That's both good and bad. There are a lot of means. And next time, we're going to talk about the physical, electrical means for characterizing the wafer contamination. And how effective is your gettering? And how contaminated are your wafers?

And I showed you some data today without explaining TXRF and things like that. We'll talk next time a little bit about how some of those contamination measurement techniques work. So that's all I have for this lecture. Please come up and hand in your homework in the orange folder here, homework number 1. Thanks.

[SIDE CONVERSATIONS]