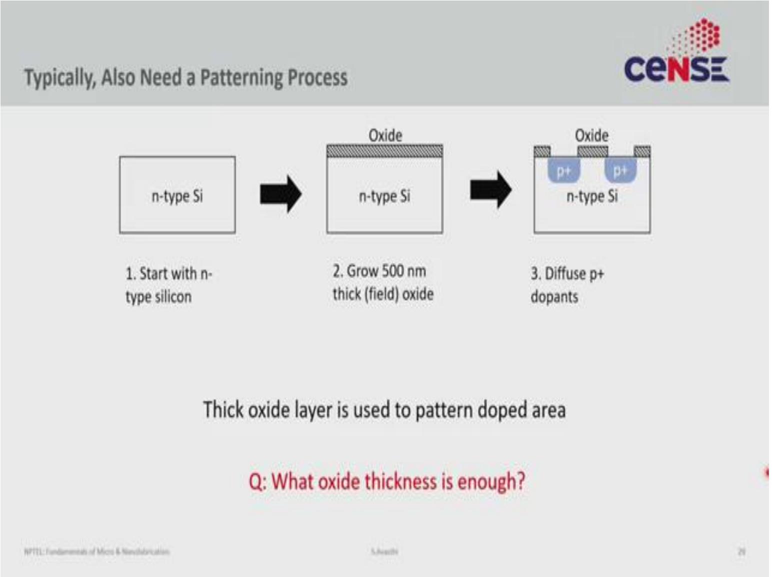
Fundamentals of Micro and Nanofabrication
Prof. Sushobhan Avasthi
Centre for Nano Science and Engineering (CeNSE)
Indian Institute of Science, Bengaluru

Lecture - 09
Diffusion -- Advanced Concepts

Welcome back. This is Fundamentals of Micro and Nanofabrication. My name is Sushobhan Avasthi, I am from IISc, Bangalore and today we shall discuss some of the advanced concepts of diffusion which are under the general module of additive processing.

(Refer Slide Time: 00:41)



Typically, Also Need a Patterning Process

1. Start with n-type silicon

2. Grow 500 nm thick (field) oxide

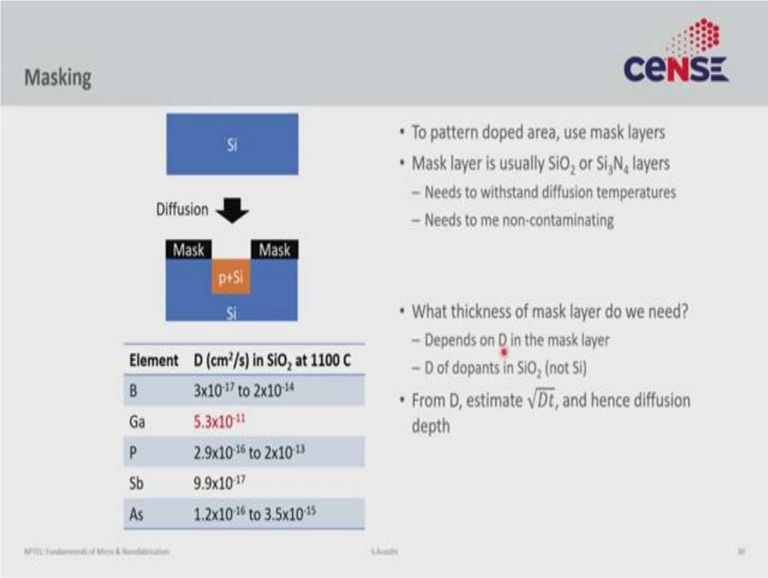
3. Diffuse p+ dopants

Thick oxide layer is used to pattern doped area

Q: What oxide thickness is enough?

Till now, we have only discussed diffusion to create a doped layer at the surface of a semiconductor wafer. However, blanket diffusion is typically not useful if you want to make a device; to have dopants under selected regions, you have to pattern it. You may want to start with N-type silicon, grow some oxide and dope the silicon on either side of this oxide, but not below. If you want to do that with diffusion, you need a method to prevent diffusion from happening in this oxide region. One advantage is that the oxide can act as a diffusion barrier. The oxide being an amorphous material, diffusion through it is very slow. But then what oxide thickness is enough to prevent the dopant diffusion through it? Is 5 nm enough? Is 100 μm thickness required?

(Refer Slide Time: 01:44)



Masking

- To pattern doped area, use mask layers
- Mask layer is usually SiO_2 or Si_3N_4 layers
  - Needs to withstand diffusion temperatures
  - Needs to me non-contaminating
- What thickness of mask layer do we need?
  - Depends on D in the mask layer
  - D of dopants in SiO_2 (not Si)
- From D, estimate $\sqrt{Dt}$, and hence diffusion depth

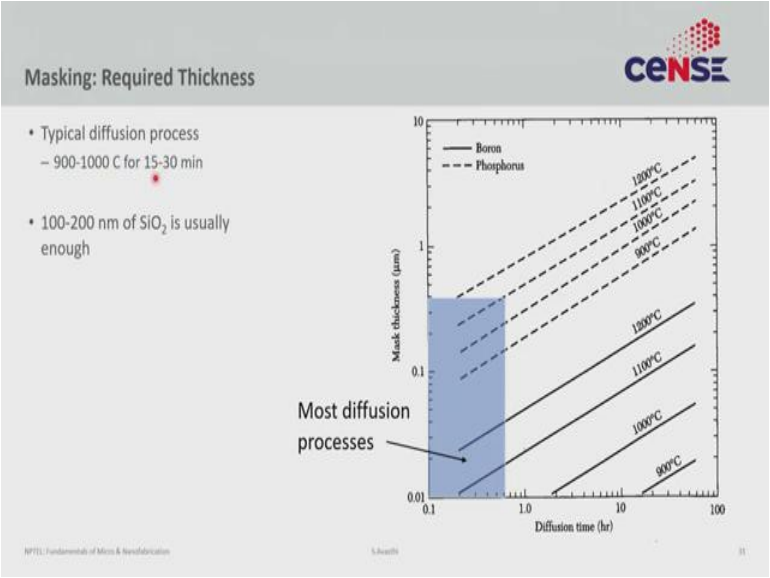
Element	D (cm^2/s) in SiO_2 at 1100 C
B	$3x10^{-17}$ to $2x10^{-14}$
Ga	$5.3x10^{-11}$
P	$2.9x10^{-16}$ to $2x10^{-13}$
Sb	$9.9x10^{-17}$
As	$1.2x10^{-16}$ to $3.5x10^{-15}$

To estimate that, we need to understand the concept of masking. Typically, we want a material that can prevent the diffusion of the dopants. If you put a mask of the barrier material and then diffuse, you only get the doping in the region where the mask was not there. Common masks are silicon nitride and silicon dioxide. Remember, the mask must be non-contaminating because the contamination will grow faster than the dopants inside your material. There must be a way to remove the mask later so that it does disturb your process flow. To calculate the mask thickness, look at the diffusion coefficient of the dopant in that mask. This is not the D of these elements inside silicon, but inside the mask (silicon dioxide).

Let’s look at the typical numbers. Boron diffusivity in the oxide ~ 10^{-17} to 10^{-14} cm^2/s. which is lower than in silicon. Similarly, for phosphorous, antimony and arsenic, the diffusion in silicon dioxide is slower than in silicon with the same thickness. One notable exception is gallium; the diffusion of gallium is very fast through the oxide for reasons that are beyond this course. This is one of the reasons why gallium is not very popular as a P-type dopant of silicon as it is very hard to mask it with oxide.

Now, what is the mask thickness we need? To first order, the thickness of the diffusion layer is $\sqrt{(Dt)}$. So, the thickness that will prevent diffusion must be $\geq$ a few multiples of $\sqrt{(Dt)}$.

(Refer Slide Time: 03:51)



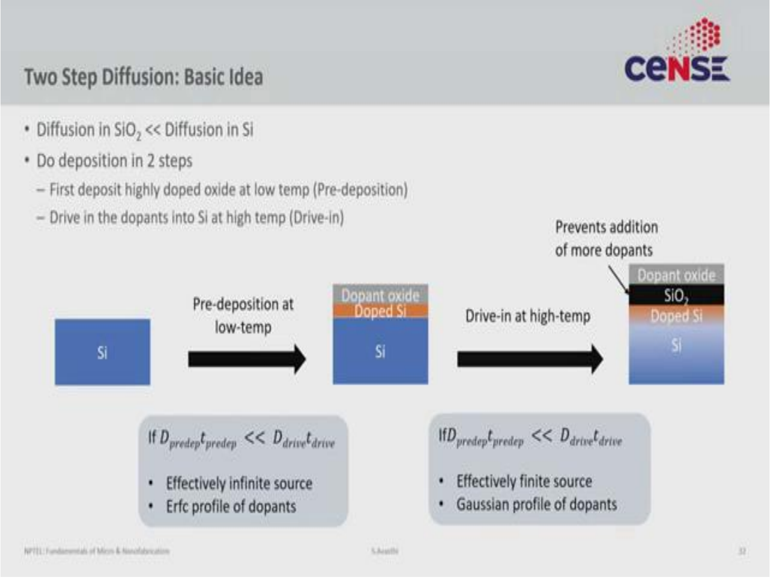
Masking: Required Thickness

- Typical diffusion process
  - 900-1000 C for 15-30 min
- 100-200 nm of SiO_2 is usually enough

Most diffusion processes

Most diffusion processes tend to last between 1 min to 30 min or 1 hour. Diffusivity values you already know from the last table. The thickness you need is given √(Dt), plotted here. In most cases, where the diffusion takes 15-30 minutes, 100-200 nm silicon dioxide is enough to stop any dopants except gallium, as gallium is not stopped by silicon dioxide.

(Refer Slide Time: 04:28)



Two Step Diffusion: Basic Idea

- Diffusion in SiO_2 << Diffusion in Si
- Do deposition in 2 steps
  - First deposit highly doped oxide at low temp (Pre-deposition)
  - Drive in the dopants into Si at high temp (Drive-in)

Pre-deposition at low-temp

Drive-in at high-temp

Prevents addition of more dopants

If $D_{predep}t_{predep}$ << $D_{drive}t_{drive}$

- Effectively infinite source
- Erfc profile of dopants

If $D_{predep}t_{predep}$ << $D_{drive}t_{drive}$

- Effectively finite source
- Gaussian profile of dopants

Another advanced concept is two-step diffusion. Typically, diffusion processes are very hard to control. We solved two cases of Fick’s equation, one with finite source and one with infinite. Any practical diffusion is probably somewhere in the middle, but the exact mathematical profile is hard to model and even harder to reproduce. While diffusion is a very easy process to do, getting it repeatedly the same in every way every time is very tricky.

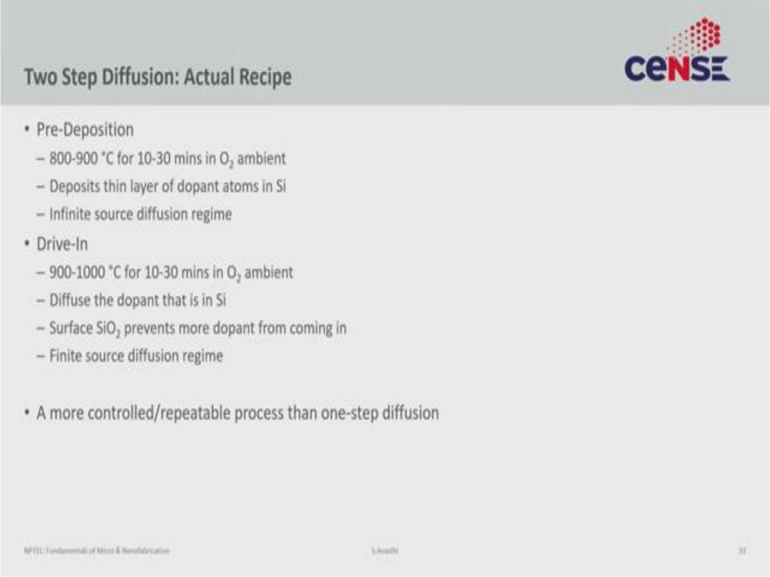
To solve this problem, you can use a two-step process, where the deposition of the dopants at the surface and inside silicon is split into two steps instead of doing it all in one shot. There is a pre-dep or a pre-deposition step and a drive-in step. During the pre-deposition, often at low temperatures, you deposit a dopant source on the surface. For example, during the pre-dep, $POCl_3$ will form phosphorus oxide or phospho-silicate glass at the surface. You form a dopant oxide on top of silicon, and because it is in contact with silicon, some amount of dopant would diffuse in even at this low temperature, forming a very thin layer of doped silicon at the surface.

Next, you take this wafer and do a drive-in step, typically under oxygen and at high temperature. During this step, you would not just diffuse the dopants further inside silicon, but you would also form a layer of silicon dioxide, which as we have discussed is a masking layer and does not allow diffusion of dopants from the source any further. So, this silicon dioxide will isolate the dopant oxide formed on the top from the silicon below. Anything that has diffused in the pre-dep will continue to diffuse further, but fresh dopants will not be able to diffuse from the dopant oxide into the silicon.

The pre-dep step is sort of an infinite source diffusion, where you have a lot of dopant on the top and a very little amount of diffusion is occurring. Drive-in is a finite source, where there is a fixed amount of dopant that has already diffused in and that dose of dopants will go in further, but extra dopants cannot diffuse into the silicon.

Tuning the temperature and the diffusivities at these temperatures, you can make a recipe that is a completely infinite source or completely finite source and anything in the middle.

(Refer Slide Time: 07:35)



Two Step Diffusion: Actual Recipe

- Pre-Deposition
  - 800-900 °C for 10-30 mins in O_2 ambient
  - Deposits thin layer of dopant atoms in Si
  - Infinite source diffusion regime
- Drive-In
  - 900-1000 °C for 10-30 mins in O_2 ambient
  - Diffuse the dopant that is in Si
  - Surface SiO_2 prevents more dopant from coming in
  - Finite source diffusion regime

- A more controlled/repeatable process than one-step diffusion

This is a common recipe that we use at IISc. We pre-dep phosphorous silicate glass at 800-900 ^{o}C for ~ 10 min in oxygen ambient (infinite source regime) and then a drive-in at 900-1000 ^{o}C for 10-30 min depending upon how deep we want the dopants to be (typically finite source regime). Overall, these two processes are much more repeatable and controllable than a single step diffusion process.

(Refer Slide Time: 08:16)

Interstitialcy Diffusion

- Variation of substitutional (vacancy) diffusion
  - Needs Si interstitials
  - E_a of substitutional dopant to form interstitials should be low

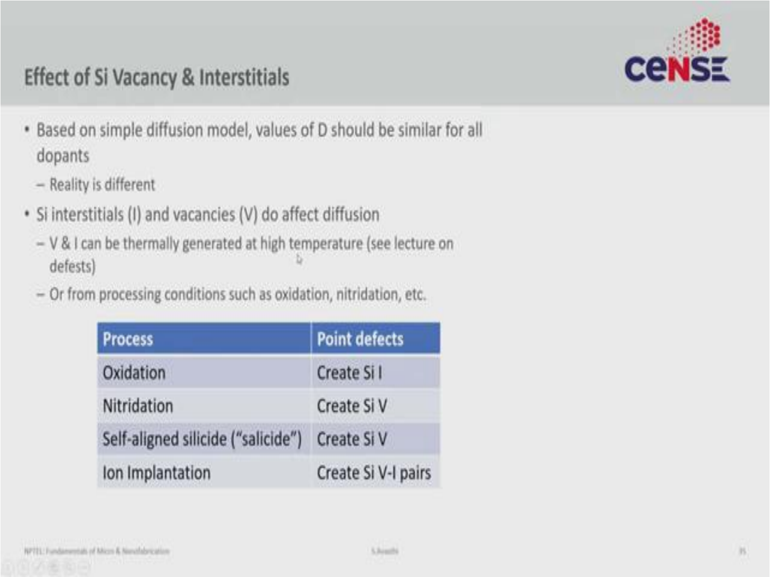
- Important diffusion mechanism for dopants
  - Especially P & B

http://www.iue.tuwien.ac.at/phd/puchner/node28.html

Another advanced concept is interstitialcy diffusion. We have talked about two diffusion mechanisms, substitutional and interstitial. This is a mixture of the two and needs a silicon interstitial. The silicon interstitial interacts with a substitutional impurity and exchanging its position. This substitution impurity is now an interstitial impurity, which interacts with another substitutional silicon, creating a new silicon interstitial and puts itself in a new substitutional position. You start with a substitutional impurity and a silicon interstitial and end with both of them at a different place. So, the concentration of interstitial does not change, but the position of the impurity atom does. Using this process, a lot of impurities that are substitutional can go through the material much faster than their substitutional diffusion mechanism would suggest because for this impurity to move by substitutional mechanism alone, it must need a silicon-vacancy which is harder to come by. By using a silicon interstitial, it can move via an interstitial mechanism, which tends to be faster.

This mixed or hybrid diffusion is very important for impurities such as boron and phosphorus; you cannot model their doping profiles accurately without accounting for this mechanism. This is just one of the possible complicated mechanisms. There can be way more complicated mechanisms, but for the purpose of this course, this is enough.

(Refer Slide Time: 10:04)



Effect of Si Vacancy & Interstitials

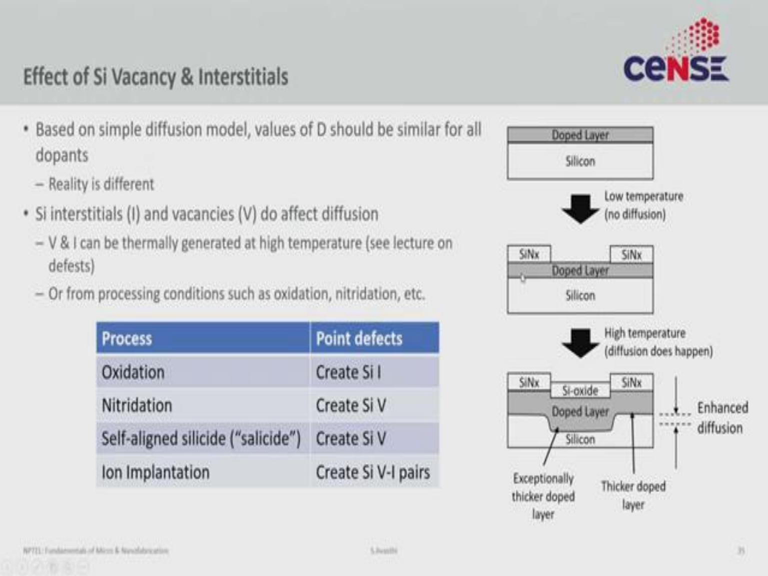
- Based on simple diffusion model, values of D should be similar for all dopants
  - Reality is different
- Si interstitials (I) and vacancies (V) do affect diffusion
  - V & I can be thermally generated at high temperature (see lecture on defects)
  - Or from processing conditions such as oxidation, nitridation, etc.

Process	Point defects
Oxidation	Create Si I
Nitridation	Create Si V
Self-aligned silicide ("salicide")	Create Si V
Ion Implantation	Create Si V-I pairs

As a consequence, diffusion is a little more complicated than the simple Fick's model for a simple substitutional or interstitial mechanism would suggest. The diffusion of impurities is affected by silicon interstitials and silicon vacancies, depending on the mechanism. Both interstitialcy and substitutional mechanisms require the presence of a vacancy or an interstitial. The vacancy and interstitial concentration changes with temperature, adding to the temperature dependence of the diffusion coefficients.

Any other process that creates vacancies or interstitials, for example, oxidation or nitration would also enhance the diffusion. This is a second-order effect. Even though you would not expect oxidation to make any change in the diffusion profile, it would because it would create interstitials and increase your diffusion.

(Refer Slide Time: 11:25)



Effect of Si Vacancy & Interstitials

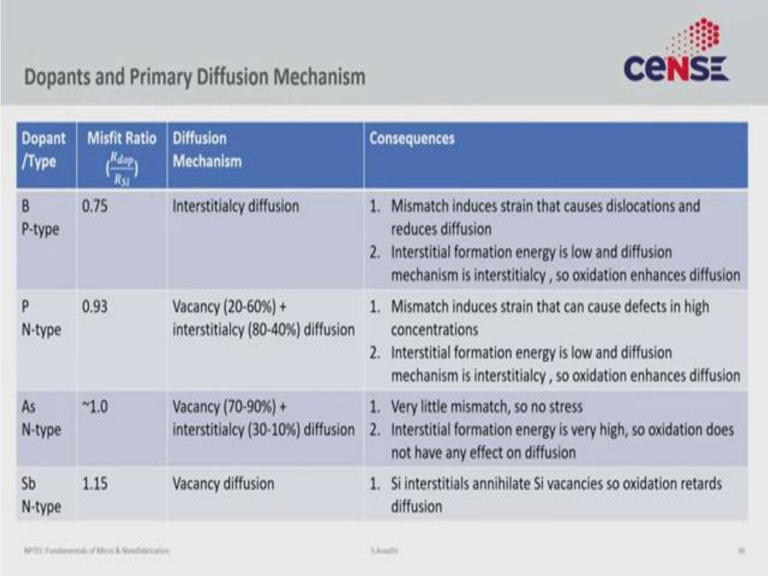
- Based on simple diffusion model, values of D should be similar for all dopants
  - Reality is different
- Si interstitials (I) and vacancies (V) do affect diffusion
  - V & I can be thermally generated at high temperature (see lecture on defests)
  - Or from processing conditions such as oxidation, nitridation, etc.

Process	Point defects
Oxidation	Create Si I
Nitridation	Create Si V
Self-aligned silicide ("salicide")	Create Si V
Ion Implantation	Create Si V-I pairs

A practical example to explain this process is on the right. Through diffusion, you have created a doped layer, deposited some silicon nitride masking layer and then you started oxidizing your wafer. This process does happen in the middle of fabricating a CMOS transistor in silicon. As you are oxidizing this doped layer, you create interstitials that will enhance the diffusion. As you are doing this at a higher temperature, you expect the dopant layer to become thicker everywhere. However, it will be especially thicker right underneath the area where the oxide is being formed.

This is something you would not expect unless you understand the mechanism by which the diffusion is occurring. Diffusion is enhanced because of vacancies interstitials. This is a table to show what silicon defects are created during certain processes. Implantations create both vacancies and interstitials, oxidation creates interstitials and nitridation creates vacancies.

(Refer Slide Time: 12:46)



Dopants and Primary Diffusion Mechanism

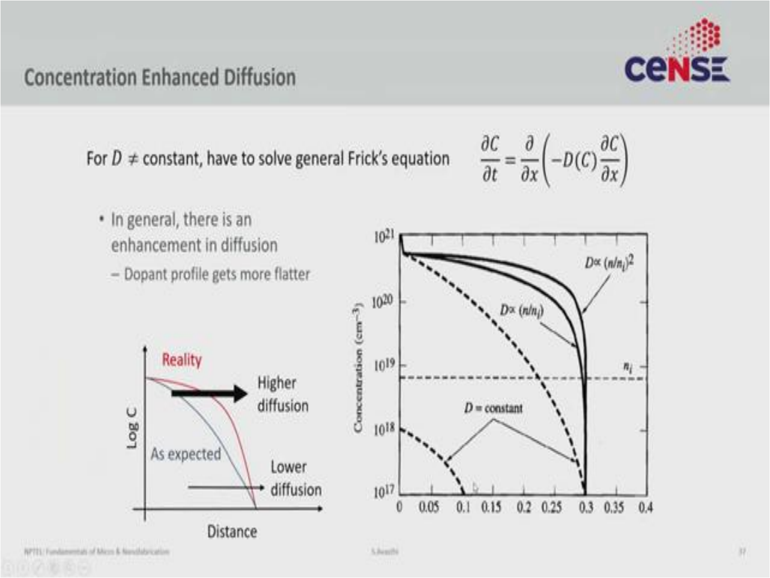
Dopant /Type	Misfit Ratio $(\frac{R_{dop}}{R_{Si}})$	Diffusion Mechanism	Consequences
B P-type	0.75	Interstitialcy diffusion	1. Mismatch induces strain that causes dislocations and reduces diffusion 2. Interstitial formation energy is low and diffusion mechanism is interstitialcy , so oxidation enhances diffusion
P N-type	0.93	Vacancy (20-60%) + interstitialcy (80-40%) diffusion	1. Mismatch induces strain that can cause defects in high concentrations 2. Interstitial formation energy is low and diffusion mechanism is interstitialcy , so oxidation enhances diffusion
As N-type	~1.0	Vacancy (70-90%) + interstitialcy (30-10%) diffusion	1. Very little mismatch, so no stress 2. Interstitial formation energy is very high, so oxidation does not have any effect on diffusion
Sb N-type	1.15	Vacancy diffusion	1. Si interstitials annihilate Si vacancies so oxidation retards diffusion

This is a table for what complicated mechanisms can exist for a given dopant. Boron and phosphorus, you expect substitutional mechanism, but they also do something else. For example, boron almost always does interstitialcy, even though you would expect it to use substitutional diffusion. Also, boron always diffuses faster and point of note when you are doing oxidation boron will diffuse even faster because oxidation creates interstitials.

Phosphorus tends to do both interstitial and substitutional, around 20-60 % is vacancy dependent and 40-80 % is interstitial dependent, is somewhere in the middle. Another extreme case is arsenic. The misfit ratio is 1, making it extremely ideal, and is almost always substitutional. Antimony does not do interstitialcy at all; it is way too big for it.

And each of these mechanisms has consequences, for example, mismatch causes strain, dislocations and reduces diffusion. At the same time, oxidation enhances the diffusion of phosphorus but does not enhance significantly the diffusion of arsenic or antimony.

(Refer Slide Time: 14:34)



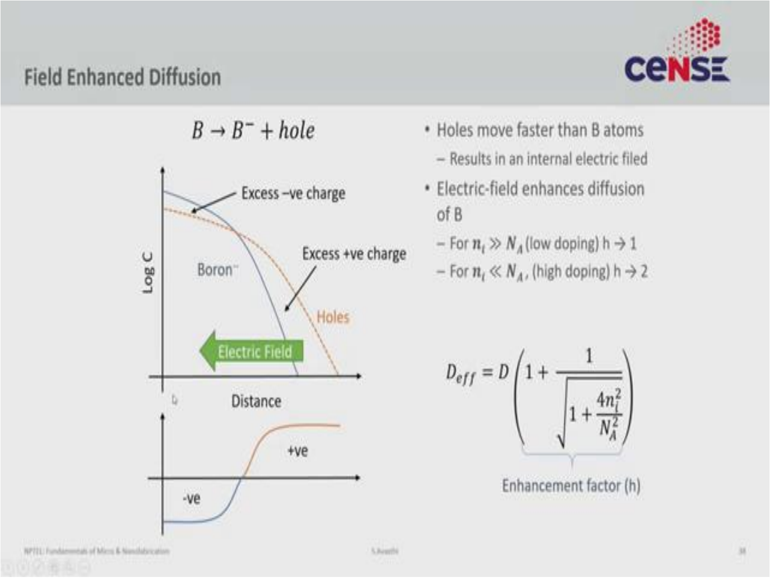
Concentration Enhanced Diffusion

For $D \neq$ constant, have to solve general Frick's equation $\frac{\partial C}{\partial t} = \frac{\partial}{\partial x}\left(-D(C)\frac{\partial C}{\partial x}\right)$

- In general, there is an enhancement in diffusion
  - Dopant profile gets more flatter

Another advanced concept is to understand that the diffusivity is dependent on the concentration. For easy mathematics, we assume that D is independent of C, making the second law easy to solve. However, D depends on concentration because of the complicated mechanisms that we have not discussed here. You have to account for it in order to model your profile accurately. The diffusion is faster in the areas of higher doping. Because of that, you get the red line profile instead of the blue line, which is the expected Gaussian or error function profile. Your profile tends to be squarer (sharper) than you expect.

(Refer Slide Time: 15:52)



Field Enhanced Diffusion

$B \rightarrow B^- + hole$

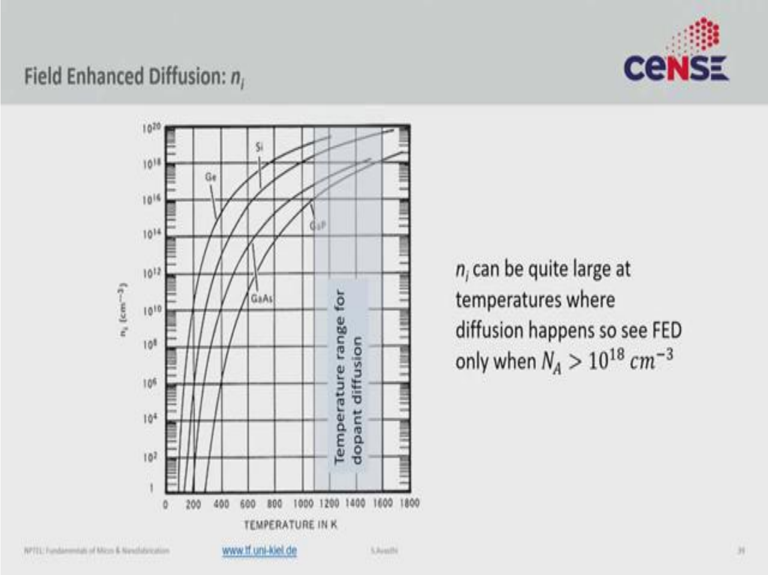
- Holes move faster than B atoms
  - Results in an internal electric filed
- Electric-field enhances diffusion of B
  - For $n_i \gg N_A$ (low doping) h → 1
  - For $n_i \ll N_A$, (high doping) h → 2

$$D_{eff} = D\underbrace{\left(1 + \frac{1}{\sqrt{1 + \frac{4n_i^2}{N_A^2}}}\right)}_{\text{Enhancement factor (h)}}$$

There is also field enhanced diffusion. Boron would form a B^- impurity and a hole with a positive charge. At higher temperatures or rather at any temperatures the diffusion of the B^- (the substitutional impurity) and the hole (just a charge) can be different. The hole tends to diffuse faster. So, the hole concentration will have a profile that looks like the orange line and B^- will diffuse slower and will have a profile that looks like the blue line.

This difference in diffusivity causes concentration difference and charging in the silicon, creating an electric field. This electric field will affect the diffusion of B^-; it will push the B^- towards the right (positive side), enhancing the diffusion. The attractive force from faster-diffusing holes will increase its diffusion. How much? It depends upon the concentration of holes and the temperature. The intrinsic carrier concentration n_i is typically very low at room temperature.

(Refer Slide Time: 00:25)



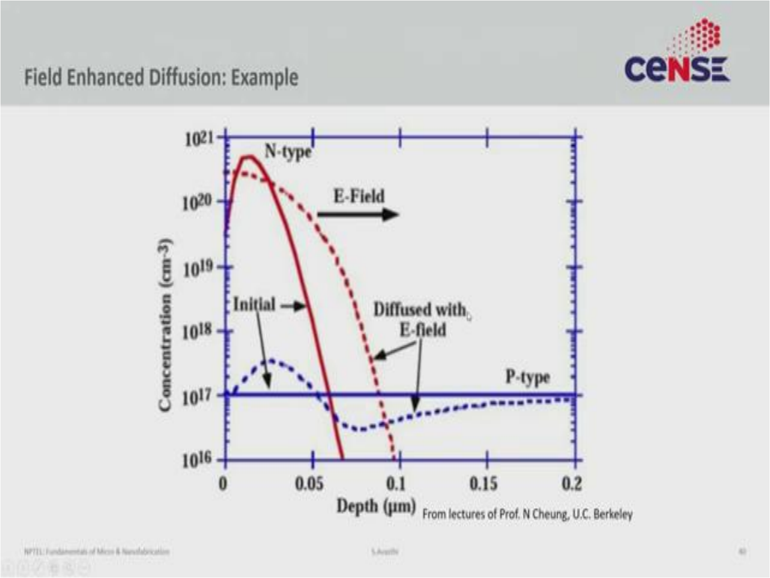
Field Enhanced Diffusion: n_i

n_i can be quite large at temperatures where diffusion happens so see FED only when $N_A > 10^{18}\ cm^{-3}$

www.tf.uni-kiel.de

However, at high temperature, the values of n_i can be significantly higher, for example, 10^{18} cm^{-3}. If the dopant concentration > 10^{18} cm^{-3}, you must expect field enhanced diffusion and not expect if the concentration is lower.

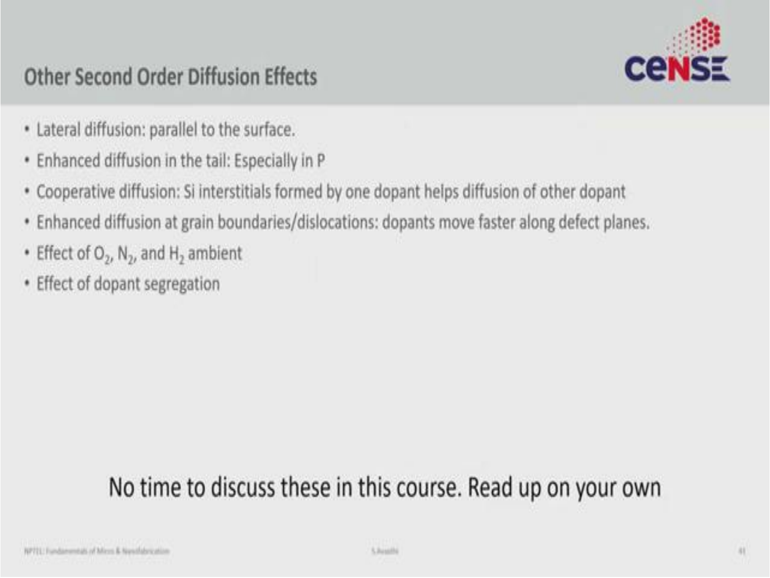
(Refer Slide Time: 17:53)


Field Enhanced Diffusion: Example

From lectures of Prof. N Cheung, U.C. Berkeley

Here is an example of a field enhanced diffusion. Solid lines show the expected profile of N-type dopants in a P-type wafer, but because of the electric field, the profile changes (represented by the dashed lines). The boron profile that was initially flat now has a certain shape because the B^- ions move left and P^+ (phosphorous ions) diffuse in at an enhanced rate under the E-field.

(Refer Slide Time: 18:41)



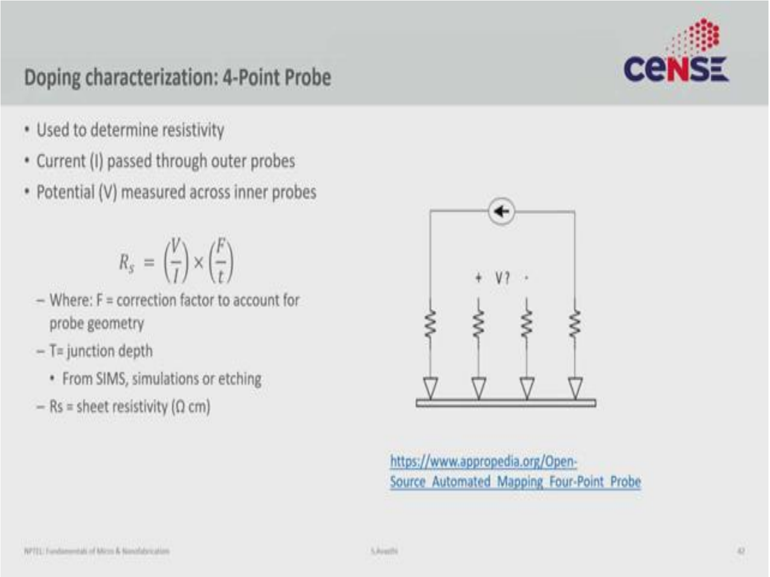
Other Second Order Diffusion Effects

- Lateral diffusion: parallel to the surface.
- Enhanced diffusion in the tail: Especially in P
- Cooperative diffusion: Si interstitials formed by one dopant helps diffusion of other dopant
- Enhanced diffusion at grain boundaries/dislocations: dopants move faster along defect planes.
- Effect of O_2, N_2, and H_2 ambient
- Effect of dopant segregation

No time to discuss these in this course. Read up on your own

This is just a small subset of various diffusion effects we know. As you would have noticed, silicon is a very well studied material and the diffusion in silicon has been studied to death. Here is the list of other types of diffusion that we are not covering in this course. You may read up these topics as an optional exercise if you are interested to look into diffusion in detail. In the modern days, we do not rely on the hand calculations to accurately predict the dopant profile but put this all in a simulator. The simulators have all these models built-in and would do all the calculations. As long as your simulations are correctly set up, you would get a very accurate profile.

(Refer Slide Time: 19:40)



Doping characterization: 4-Point Probe

- Used to determine resistivity
- Current (I) passed through outer probes
- Potential (V) measured across inner probes

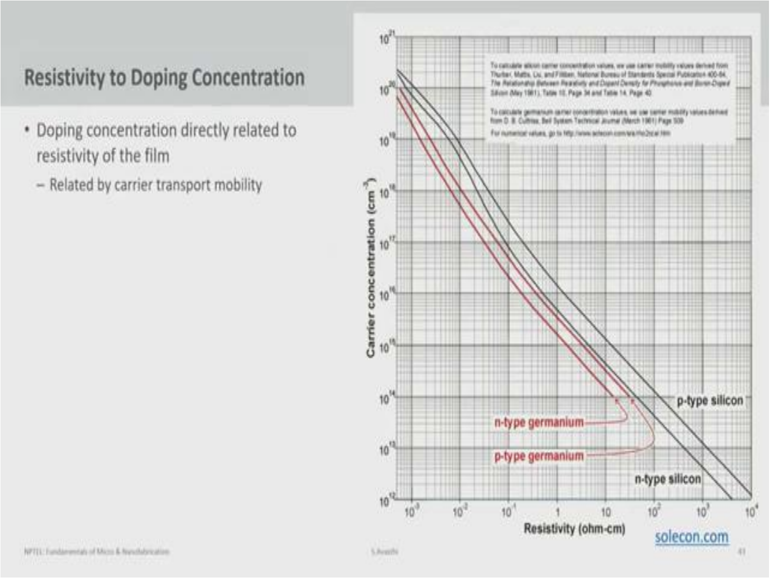
$$R_s = \left(\frac{V}{I}\right) \times \left(\frac{F}{t}\right)$$

- – Where: F = correction factor to account for probe geometry
- – T= junction depth
  - From SIMS, simulations or etching
- – Rs = sheet resistivity (Ω cm)

https://www.appropedia.org/Open-Source_Automated_Mapping_Four-Point_Probe

How do you characterize the dopant profile that you have made inside the material? The simplest thing is to measure resistivity. Adding dopants changes the resistivity of the surface. The resistivity of the surface can be easily measured using a four-point probe that simply pushes some current and measures voltage, and through a very simple formula, tells you what is the sheet (surface) resistance. We will not go into more detail about the four-point probe method. However, in order for this to work, you must know the thickness which is often unknown because you do not know how deep the dopants have diffused.

(Refer Slide Time: 20:25)



Resistivity to Doping Concentration

- Doping concentration directly related to resistivity of the film
  - Related by carrier transport mobility

To calculate silicon carrier concentration values, we use carrier mobility values derived from Thurber, Mattis, Liu, and Filliben, National Bureau of Standards Special Publication 400-64, *The Relationship Between Resistivity and Dopant Density for Phosphorus and Boron-Doped Silicon* (May 1981), Table 10, Page 34 and Table 14, Page 40.

To calculate germanium carrier concentration values, we use carrier mobility values derived from D. B. Cuttriss, Bell System Technical Journal (March 1961) Page 509

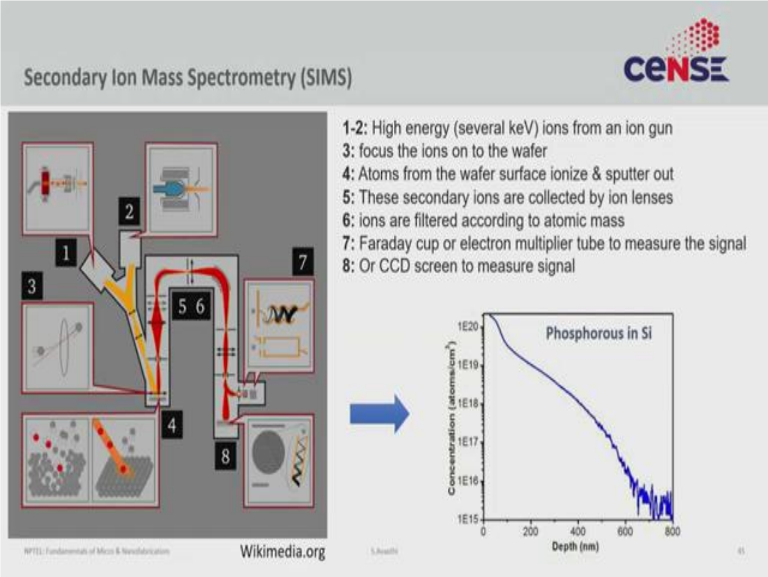
For numerical values, go to http://www.solecon.com/sra/rho2ccal.htm

solecon.com

NPTEL: Fundamentals of Micro & Nanofabrication S.Avasthi 41

For that, often people use SIMS. Once you know the resistivity, you can back-calculate the doping by the simple look-up graph, depending on the material and the doping type.

(Refer Slide Time: 20:45)



Secondary Ion Mass Spectrometry (SIMS)

CeNSE

1-2: High energy (several keV) ions from an ion gun
3: focus the ions on to the wafer
4: Atoms from the wafer surface ionize & sputter out
5: These secondary ions are collected by ion lenses
6: ions are filtered according to atomic mass
7: Faraday cup or electron multiplier tube to measure the signal
8: Or CCD screen to measure signal

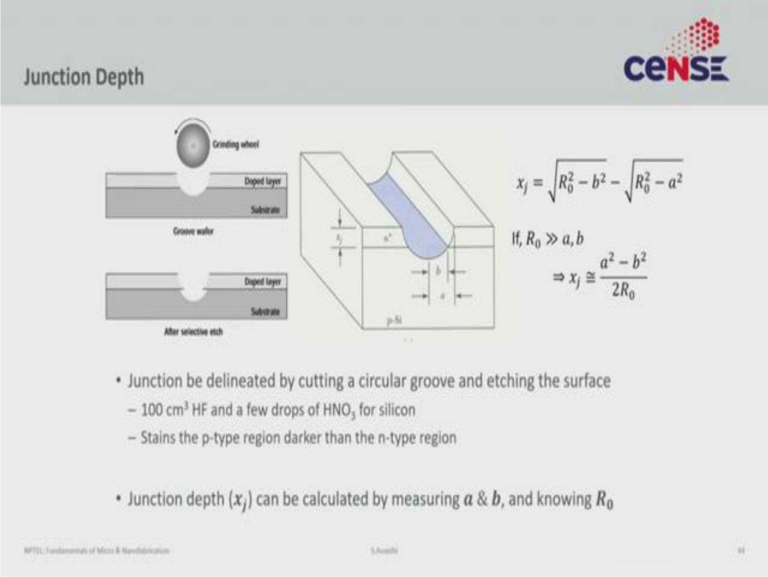
Wikimedia.org

NPTEL: Fundamentals of Micro & Nanofabrication S.Avasthi 41

SIMS is secondary ion mass spectroscopy, an advanced technique to actually profile the dopants inside your silicon. Very simply, it showers the silicon surface with a heavy ion which causes sputtering on the silicon surface. A stream of particles that hit the silicon surface and etch (sputter) the silicon surface. The silicon atoms that are sputtered, come off as vapor into the plasma. These vapors are analyzed by a mass spectrometer that tells you their mass and from their mass, you can back-calculate what element came out.

So, if your silicon surface has phosphorus, in the sputtering plasma you will measure silicon signal because you are sputtering silicon, and also phosphorus signal from the dopants. With respect to time, you know how deep you have etched. So, you can convert the time axis of etching into a depth axis. So, you would get these figures were you know the concentration of phosphorus as a function of depth. This is SIMS which is one of the workhorses to find the dopant profiles.

(Refer Slide Time: 21:58)



Junction Depth

$$x_j = \sqrt{R_0^2 - b^2} - \sqrt{R_0^2 - a^2}$$

If, $R_0 \gg a, b$

$$\Rightarrow x_j \cong \frac{a^2 - b^2}{2R_0}$$

- Junction be delineated by cutting a circular groove and etching the surface
  - 100 cm^3 HF and a few drops of HNO_3 for silicon
  - Stains the p-type region darker than the n-type region

- Junction depth (x_j) can be calculated by measuring a & b, and knowing R_0

The old school method of figuring out junction depth was selective etching to stain your doped areas and then use microscopy to measure that. Personally, I think it is outdated and do not know of any setup that does this process anymore. Other than academic interest, there is no need to understand it. But the concept is fairly cool. You start with the substrate with a doped layer on top and using a grinding wheel, you grind us a cylindrical trench into the surface. Then you do some selective etching such that the un-doped regions get etched faster and look at it under a microscope. This delineation helps in estimating the thickness of the doped layer.

You would see some strips, and the thickness of those strips can be geometrically related to the depth. Using simple geometry, you back-calculate the depth X_j of this layer the doping thickness. This idea is really cool, but practically of limited use if you can just do SIMS.

With that, we come to the end of this lecture. This takes care of the first set of lectures on additive processing where we discussed diffusion. Next, we will look at the more modern technique called implantation.

Thank you.