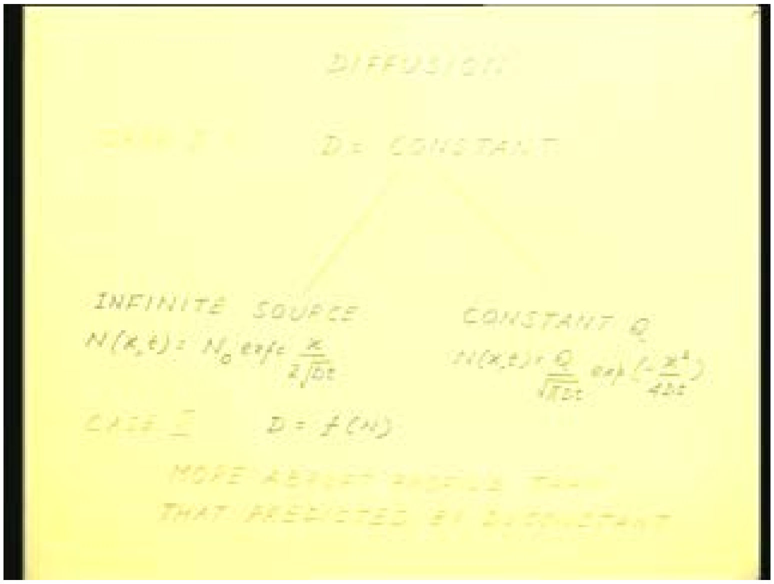
VLSI Technology
Dr. Nandita Dasgupta
Department of Electrical Engineering
Indian Institute of Technology, Madras

Lecture - 17
Diffusion III - Actual Doping Profile

In our discussion about the diffusion process, so far we have seen that it is possible either to have a constant diffusivity, in which case we will call it case one - D is constant and in that category it is possible to have two branches that is the infinite source diffusion which will be given by the complementary error function profile and the constant total impurity case, constant Q case, in which case the impurity profile will be given by a Gaussian function and we can also have a second case, case two where D is a function of the concentration.

(Refer Slide Time: 1:20)



DIFFUSION

CASE I : D= CONSTANT

INFINITE SOURCE

$N(x,t) = N_0 \,\mathrm{erfc}\, \frac{x}{2\sqrt{Dt}}$

CONSTANT Q

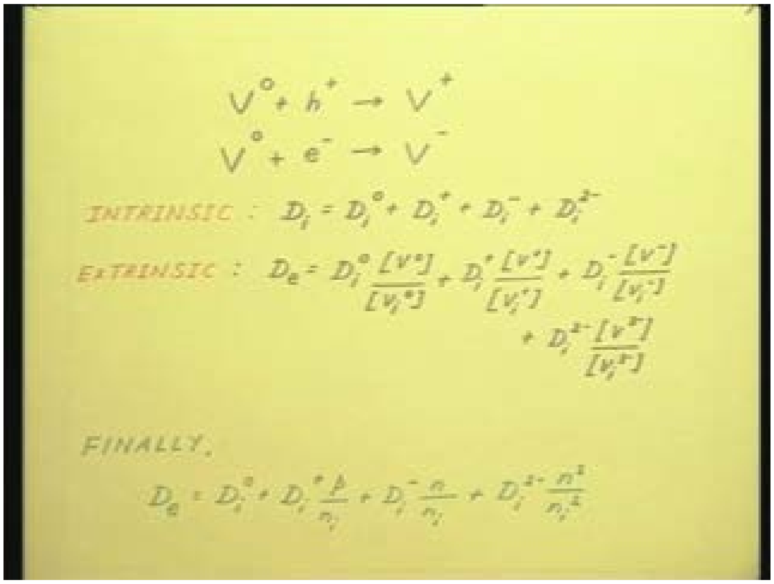
$N(x,t) = \frac{Q}{\sqrt{\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right)$

CASE II D = f(N)

MORE ABRUPT PROFILE THAN THAT PREDICTED BY D=CONSTANT

D is no longer a constant, but D is a function of the concentration and in that case we have seen that the profile will be more abrupt than that is predicted by the D equal to constant case, right. Now, at this point we would like to know what is the real life situation? In an actual diffusion system what really happens? Is the diffusion co-efficient D a constant or is it a function of the doping concentration? In order to look at that we have to look at what really is happening during diffusion. That is you know during diffusion, the impurity moves inside the semiconductor by interacting with the point defects. If there are, let us talk about substitutional impurities, because they are the most important electronically. So, if we are talking about substitutional impurities that is where the dopant atom is substituting for a host atom, there the movement is primarily governed by the interaction of the dopant atom with the available vacant sites, right.

(Refer Slide Time: 3:37)



$V^0 + h^+ \rightarrow V^+$

$V^0 + e^- \rightarrow V^-$

INTRINSIC : $D_i = D_i^0 + D_i^+ + D_i^- + D_i^{2-}$

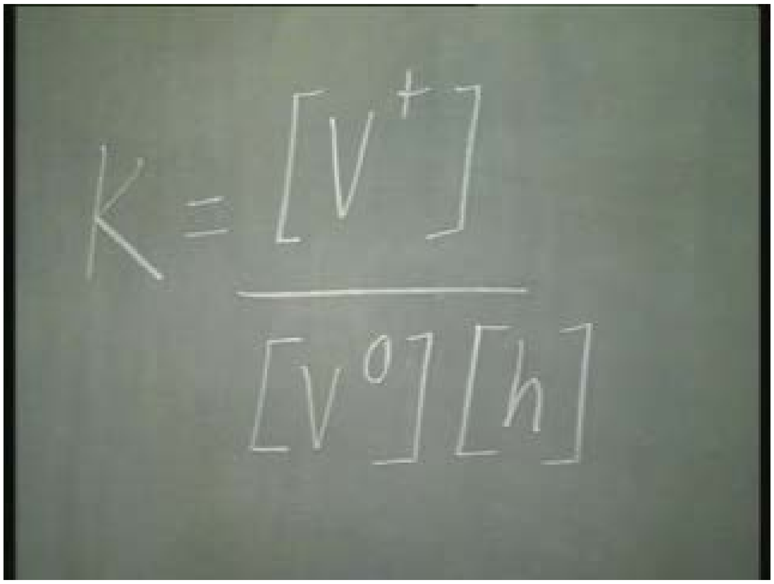
EXTRINSIC : $D_e = D_i^0 \frac{[V^0]}{[V_i^0]} + D_i^+ \frac{[V^+]}{[V_i^+]} + D_i^- \frac{[V^-]}{[V_i^-]} + D_i^{2-} \frac{[V^{2-}]}{[V_i^{2-}]}$

FINALLY,

$D_e = D_i^0 + D_i^+ \frac{p}{n_i} + D_i^- \frac{n}{n_i} + D_i^{2-} \frac{n^2}{n_i^2}$

If we look at curve two, you will see that these interactions can be expressed like this. Let us say, to begin with we have vacancies which are neutral. So, V with a zero on top, it means these vacancies are neutral. These neutral vacancies can act as a hole trap or as an electron trap. That is it can accept a hole or it can accept an electron and if it reacts with a hole, accepts a hole that is, then it becomes a positively charged vacancy. If it accepts an electron, then it becomes a negatively charged vacancy, right. So, it is possible to have either a neutral vacancy or a positively charged vacancy or a singly negatively charged vacancy or even doubly negatively charged vacancies, right.

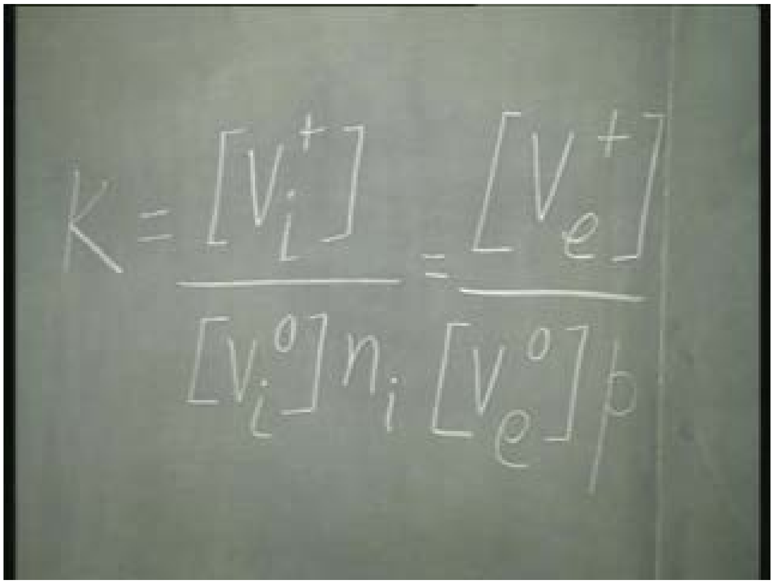
(Refer Slide Time: 5:05)



$$K = \frac{[V^+]}{[V^0][h]}$$

Looking at these two equations, if we apply the law of mass transfer to that, then we can say that for the first equation, then I can define a mass transfer coefficient K, given as .... and the concentration of hole, right. Let me for the time being write it like this. The terms under the third bracket sign are the holes. This is how you write the mass transfer coefficient from the equation, isn't it.

(Refer Slide Time: 6:24)

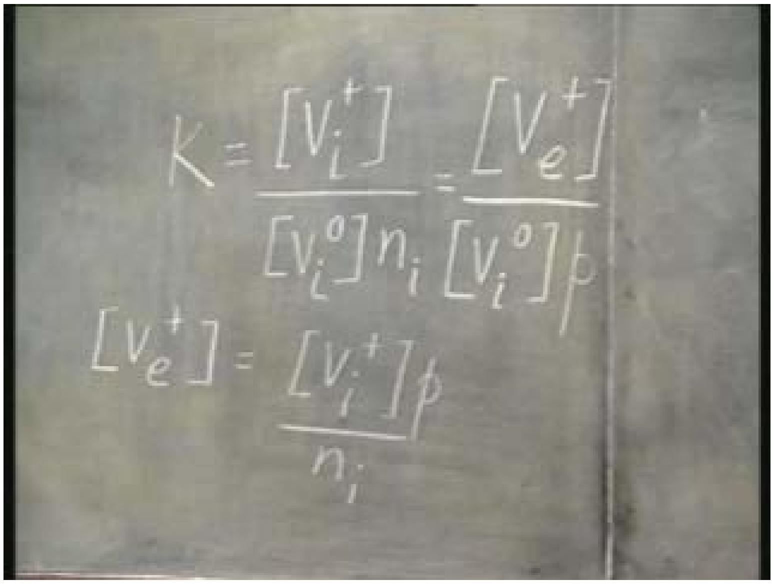


$$K = \frac{[V_i^+]}{[V_i^0]n_i} = \frac{[V_e^+]}{[V_e^0]p}$$

Now, if at the diffusion temperature the semiconductor is behaving like an intrinsic material that is the temperature is such that the intrinsic carrier concentration at that temperature is much higher than the dopant concentration, if the semiconductor remains intrinsic at that temperature, then I add the suffix i and what do I write for the concentration of holes? What is the concentration of holes in an intrinsic semiconductor? Simply given as n i, right. Now, however if this is an extrinsic case that is at this temperature the extrinsic impurity concentration is much higher than the intrinsic concentration, then .... and what is the concentration of holes in an extrinsic case? Agreed.

Look at these two situations. Remember one more thing. That is the concentration of the neutral vacancies V i 0 and V e 0, as I have written, they are going to be the same. Why? Because, it depends only on the temperature. The creation of neutral vacancies is going to be dependent only on the temperature of diffusion.

(Refer Slide Time: 8:19)

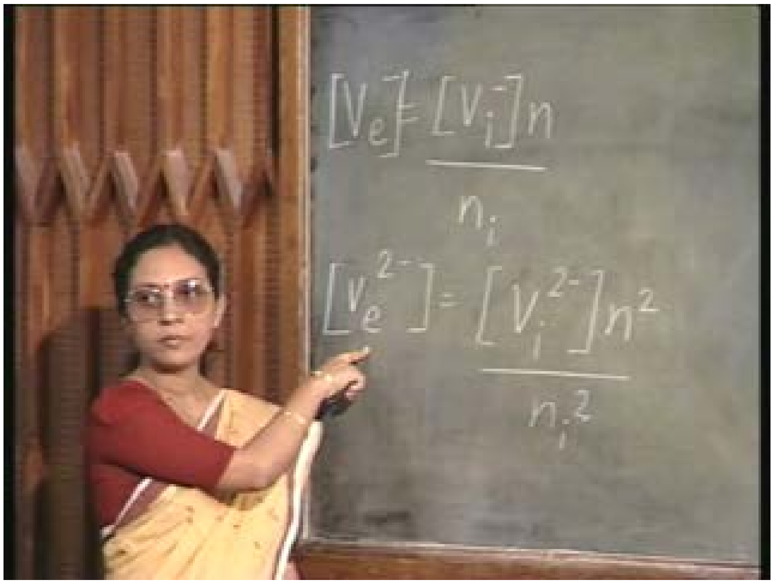


$$K = \frac{[V_i^+]}{[V_i^0]n_i} = \frac{[V_e^+]}{[V_i^0]p}$$

$$[V_e^+] = \frac{[V_i^+]p}{n_i}$$

So, in both cases I can replace it by and therefore I could write that ....., can I not?

(Refer Slide Time: 9:06)

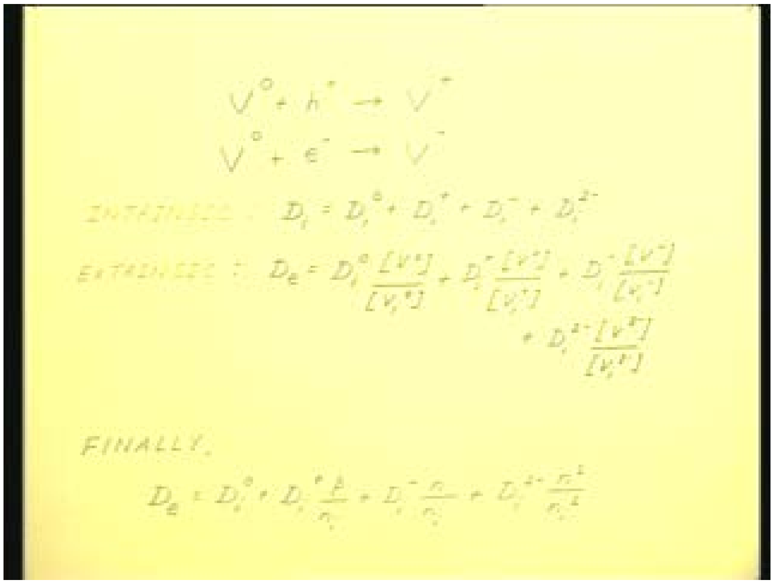


$$[V_e^-] = \frac{[V_i^-]\,n}{n_i}$$

$$[V_e^{2-}] = \frac{[V_i^{2-}]\,n^2}{n_i^2}$$

By analogy therefore, I could also write that, by analogy I could therefore write that V e minus is given by V i minus into n by n i and concentration of V e 2 minus is given as concentration of V i 2 minus multiplied by n square by n i square. In other words, I have derived a relationship between the concentration of the charged vacancies in the intrinsic and the extrinsic case, right.

(Refer Slide Time: 10:49)



$$V^0 + h^+ \rightarrow V^+$$

$$V^0 + e^- \rightarrow V^-$$

INTRINSIC : $D_i = D_i^0 + D_i^+ + D_i^- + D_i^{2-}$

EXTRINSIC : $D_e = D_i^0 \frac{[V^0]}{[V_i^0]} + D_i^+ \frac{[V^+]}{[V_i^+]} + D_i^- \frac{[V^-]}{[V_i^-]} + D_i^{2-} \frac{[V^{2-}]}{[V_i^{2-}]}$

FINALLY,

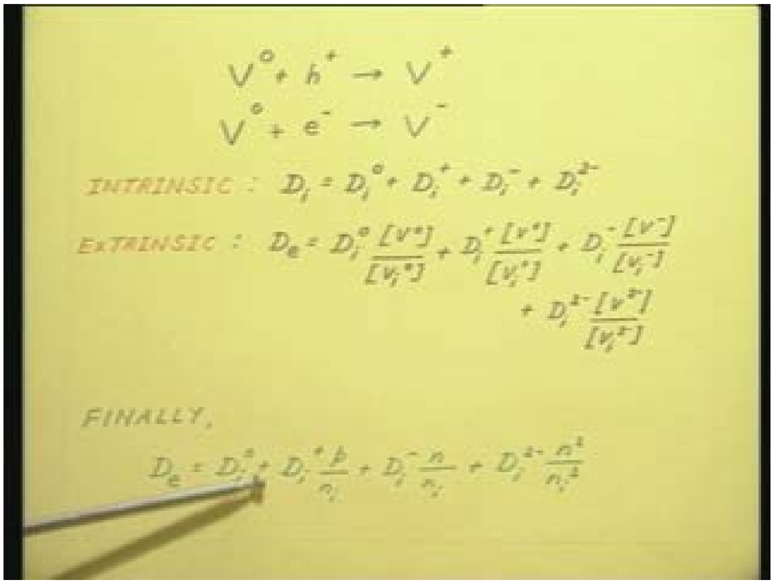
$$D_e = D_i^0 + D_i^+ \frac{p}{n_i} + D_i^- \frac{n}{n_i} + D_i^{2-} \frac{n^2}{n_i^2}$$

Now, if we consider that the interaction of the impurity with the vacancies in various states, neutral and various charge states, they are independent of each other, then I could write the total effective diffusion coefficient in the intrinsic case D i as D i 0. That is this is the diffusion coefficient where the impurity is reacting with the neutral vacancies plus D i plus that is this is the diffusion coefficient related to the interaction of the impurity with the positively charged vacancies plus D i minus, in which case it is related to the interaction between the impurity and singly negatively charged vacancies and D i 2 minus where it is related to the interaction of the impurity with the doubly negatively charged vacancies, right.

So you see, in the extrinsic case also I can analogously define the extrinsic diffusion coefficient D e as D e 0 plus D e plus plus D e minus plus D e double minus and considering that this is how they are related, I could express them in terms of the intrinsic diffusion coefficient as D i 0, concentration of neutral vacancies in the extrinsic and in the intrinsic case. In this case also, concentration of the positively charged vacancies in the extrinsic and intrinsic case, concentration of negatively charged vacancies in the extrinsic and intrinsic case and concentration of doubly negatively charged vacancies in the extrinsic and the intrinsic case, yes and now if I combine this with these expressions that is the relative concentration of the vacancies in various charge states between the extrinsic and intrinsic case, then what is concentration of neutral vacancies in intrinsic and extrinsic case?

The ratio is 1 as I told you, because the number of neutral vacancies is only going to depend on the temperature, not on anything else. So, this term becomes 1, simply I get D e. The first term is D i 0. In the second term, what happens? This is my extrinsic positively charged vacancies, this is my intrinsic positively charged vacancies and their ratio is given by P by n i. So, I can replace it by P by n i. In this third term this ratio is going to be n by n i and in the fourth term, it is going to be n square by n i square. So, I can therefore arrive at the expression for the extrinsic diffusion coefficient in terms of the intrinsic diffusion coefficient and the concentrations, right.

(Refer Slide Time: 14:25)



$$V^0 + h^+ \rightarrow V^+$$

$$V^0 + e^- \rightarrow V^-$$

INTRINSIC : $D_i = D_i^0 + D_i^+ + D_i^- + D_i^{2-}$

EXTRINSIC : $D_e = D_i^0 \frac{[V^0]}{[V_i^0]} + D_i^+ \frac{[V^+]}{[V_i^+]} + D_i^- \frac{[V^-]}{[V_i^-]} + D_i^{2-} \frac{[V^{2-}]}{[V_i^{2-}]}$

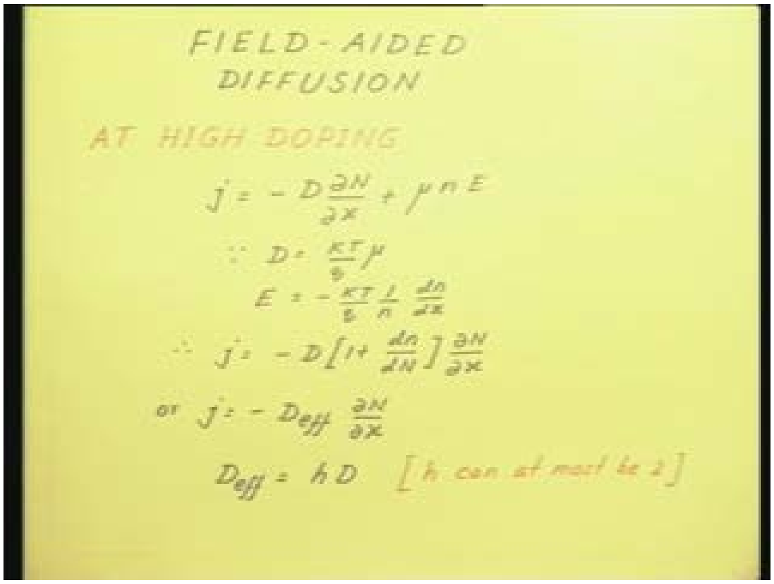
FINALLY,

$$D_e = D_i^0 + D_i^+ \frac{p}{n_i} + D_i^- \frac{n}{n_i} + D_i^{2-} \frac{n^2}{n_i^2}$$

Now, please focus your attention on this last line. So, for a general case, what do you see? You see that the effective diffusion coefficient, extrinsic effective diffusion coefficient, contains a number of terms and depending on which term is going to dominate we have either D is equal to constant, the first term which is not dependent on the concentration or D is proportional to the concentration depending on whether you have an acceptor like impurity or a donor like impurity, either of these two terms will come into picture, right and if you have very high concentration and an n-type impurity, then this third term may also dominate the situation, so in which case you will have D proportional to n square. So, in a real life situation everything is possible. You can have, for very high concentration D proportional to n square, for moderately high concentration D proportional to the concentration, either n or p depending on whether it is donor type or acceptor type and you can also have D is a constant in which case we say that the intrinsic condition is dominating. The dopant concentration is much lower compared to the intrinsic concentration at that temperature, right.

So, in a real life this is the situation. But, this is still not the whole of the story. In a real situation, the diffusion coefficient can get further modified if there is very high doping concentration. What happens? When there is a very high doping concentration, the impurity may be ionized. **A**t the time of entering the semiconductor itself it may be ionized. There will be an electron concentration associated with this ionization, ionized that is let us say I have positively charged impurity ion and the electron. Now, electrons will move faster. Obviously, they are much lighter than your positively charged ions, so electron will move faster. In the process there will be a space charge region, because this electron is separated from the positively charged ions, right and this will result in an electric field and the direction of the field will be such as to aid the movement of the diffusing impurity. Electrons are moving ahead, dopant, ionized dopant atoms are separated. The action of the field is from the positive to the negative, so it is actually aiding the movement of the dopant ions. This is an extra factor coming into picture, because of the ionization of the impurity atoms. If there was no ionization, this effect of field, this beneficial effect of field will not be there.

(Refer Slide Time: 18:12)



FIELD - AIDED DIFFUSION

AT HIGH DOPING

$$j = -D\frac{\partial N}{\partial x} + \mu n E$$

$$\because D = \frac{KT}{q}\mu$$

$$E = -\frac{KT}{q}\frac{1}{n}\frac{\partial n}{\partial x}$$

$$\therefore j = -D\left[1 + \frac{\partial n}{\partial N}\right]\frac{\partial N}{\partial x}$$

$$\text{or } j = -D_{eff}\frac{\partial N}{\partial x}$$

$$D_{eff} = hD \quad [h \text{ can at most be } 2]$$

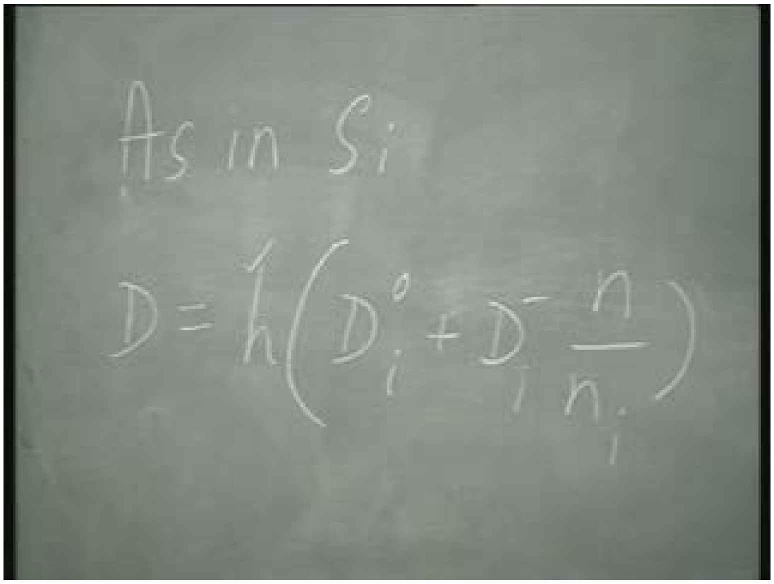
So, this is called field aided diffusion, field aided diffusion. In this case what happens? You see, so far when we have considered the diffusion process, we have considered the flux to be only proportional to the concentration gradient. That is only the first term in this equation was considered, right; j is equal to minus D del N del x. But now, in addition to this concentration gradient, I have the beneficial effect of the electric field and that is the second term, right, mobility times electron concentration times the electric field that is the drift component of the flux. Your first component is the diffusion component, second component is the drift component.

Now you all know Einstein's relationship that is D is proportional to mobility and this electric field can be expressed in terms of the electron concentration and the electron concentration gradient, so that eventually if I plug all these values into the first equation, I get, j is given by minus D into 1 plus d small n d capital N del N del x. Small n stands for the electron concentration, capital N stands for the impurity concentration. So, commonsense is going to tell you that this term can at most be equal to 1, right, at most all the impurity atoms can get ionized. This term can at most be equal to 1. Therefore, I can maximum have an enhancement in the diffusion coefficient by a factor of 2, right.

If I take this field aided motion of the impurity ions, then the flux will be given by minus D effective del N del x, where D effective is h D. h can at most be equal to 2 and this h is called the field aided factor, field enhancement factor or field aided factor. So you see, in a practical situation the diffusion coefficient is not necessarily a constant. It can be a constant, it can be a function of the dopant concentration; it can be proportional to the dopant concentration or even proportional to the square of the dopant concentration. In addition to that, I can have a field enhancement factor called h, which can at most be equal to 2. That is effectively I can have double the diffusion coefficient.

Let us now look at some of the actual dopants in silicon. What are the actual dopants in silicon? p-type dopant you know, there is only boron; n-type dopant, arsenic is there, phosphorus is there. Let us first take the case of arsenic diffusion in silicon, see how its diffusion coefficient behaves in the light of our discussion so far.

(Refer Slide Time: 22:18)

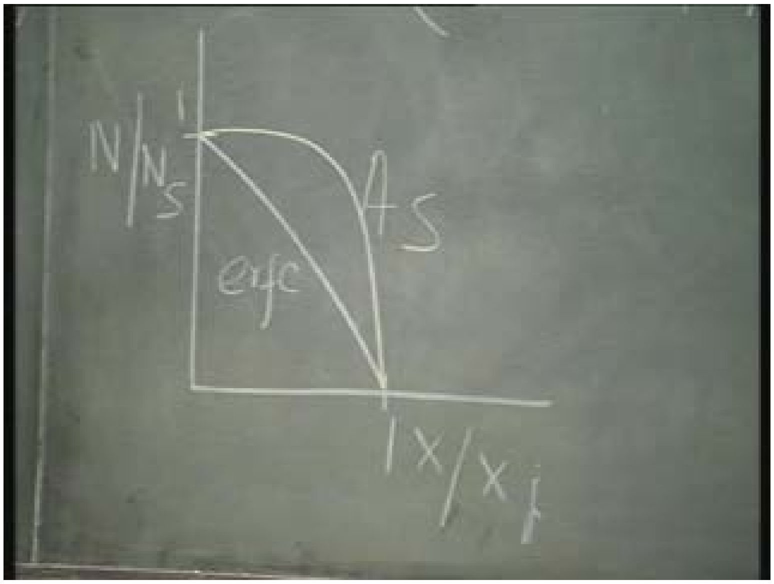


As in Si

$$D = h\left(D_i^0 + D_i^- \frac{n}{n_i}\right)$$

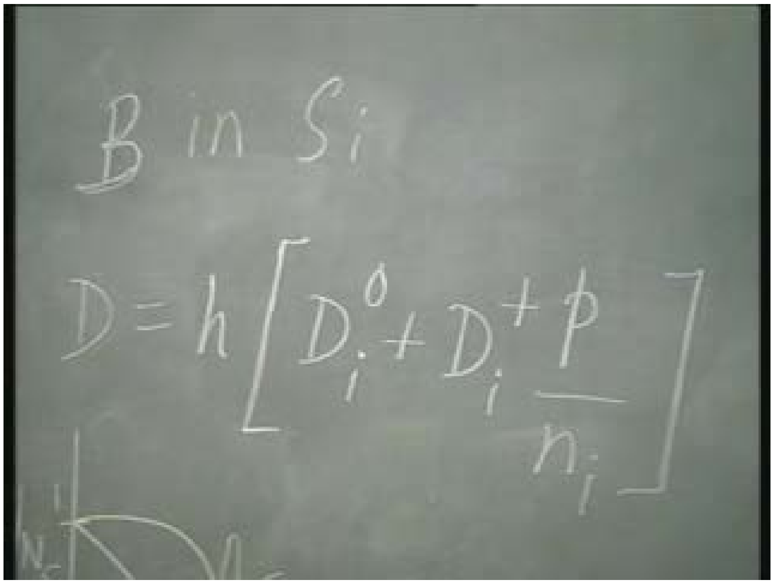
So, in case of arsenic in silicon, in case of arsenic doping in silicon, the effective diffusion coefficient D is given by, there is the field aided factor h and then you see, I have two terms in the diffusion coefficient. The first term is independent of the dopant concentration and the second term is proportional to the dopant concentration. That is this is determined by the interaction of arsenic with the singly negatively charged vacancies. This can be fitted by a numerical situation, by a numerical solution.

(Refer Slide Time: 23:54)



Now, if I want to look at the doping profile of this, if I want to plot say, normalized doping concentration N by N S, normalized to the surface concentration and if I want to normalize it with respect to junction depth X by X j, then you know that if this is the complementary error function profile, let us say this is the complementary error function profile, this value is 1 here and this value is also 1 here. Surface concentration is N S, so N by N S is equal to 1 and at the junction depth, X is equal to X j, so X by X j is also 1. So, if this is the error function, complementary error function profile and if this is what the arsenic diffusivity in silicon is, then it is going to follow a shape like, right. This is going to be more abrupt. What do I mean by more abrupt? That is the concentration will be closer to the surface concentration over a larger length, larger depth and then it will fall steeply at the junction. So, this what the profile is going to look like, arsenic profile in silicon. So, this is arsenic profile in silicon, right.

(Refer Slide Time: 26:06)

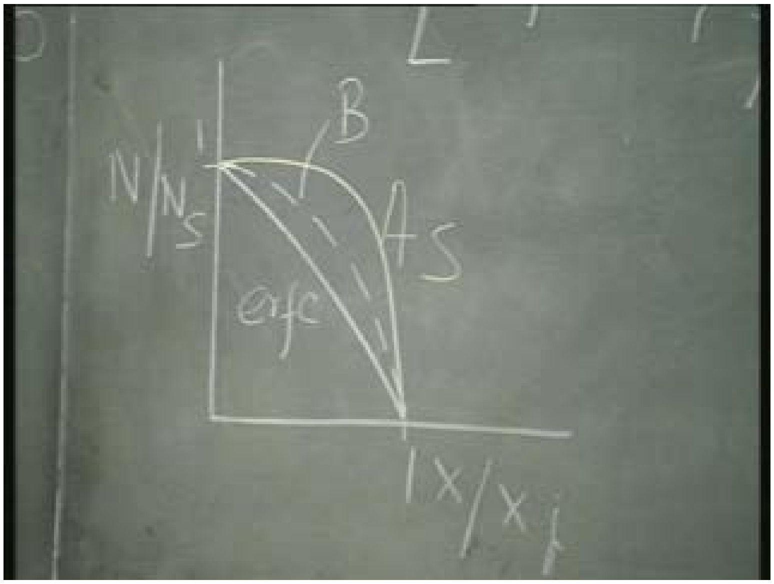


B in Si

$$D = h\left[D_i^0 + D_i^+ \frac{p}{n_i}\right]$$

Let us now consider boron in silicon. Boron in silicon also has a diffusivity. Remember, boron is an acceptor. So, you have the interaction with the positively charged vacancies. This is the hole concentration. So, this pattern looks quite similar to that of arsenic. However it is not as steep as that of arsenic.

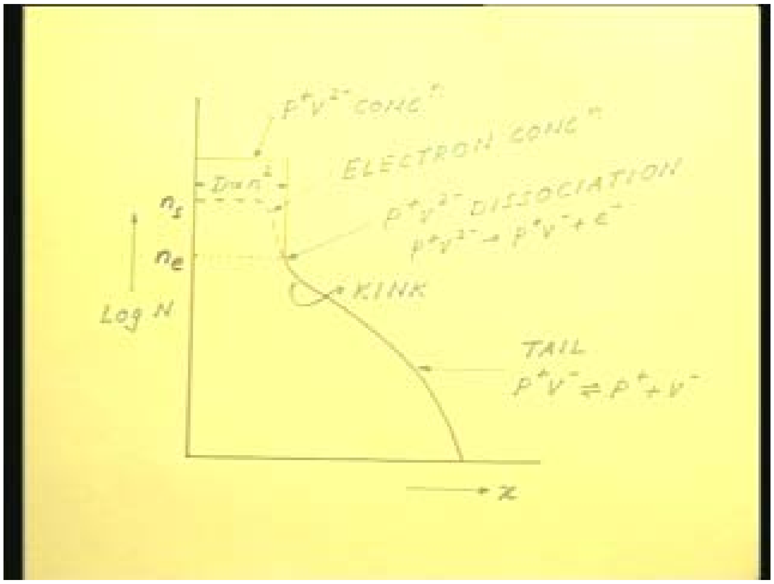
(Refer Slide Time: 27:11)



That is to say, the relative dominance of this term D i plus is comparatively lower, so that you are coming closer to the D equal to constant case and that profile is given by something like this, in between the complementary error function profile and the arsenic profile. So, this is profile of boron, so the profile of boron and arsenic in silicon can be explained by taking into account the possibility of interaction of the impurity atoms with the charged vacancies. In both cases it results in steeper, more abrupt profile than that predicted by the complementary error function for infinite source diffusion obviously, because D is not really a constant here, but there is a term which is proportional to N. In D, there is a term which is proportional to the hole or electron concentration.

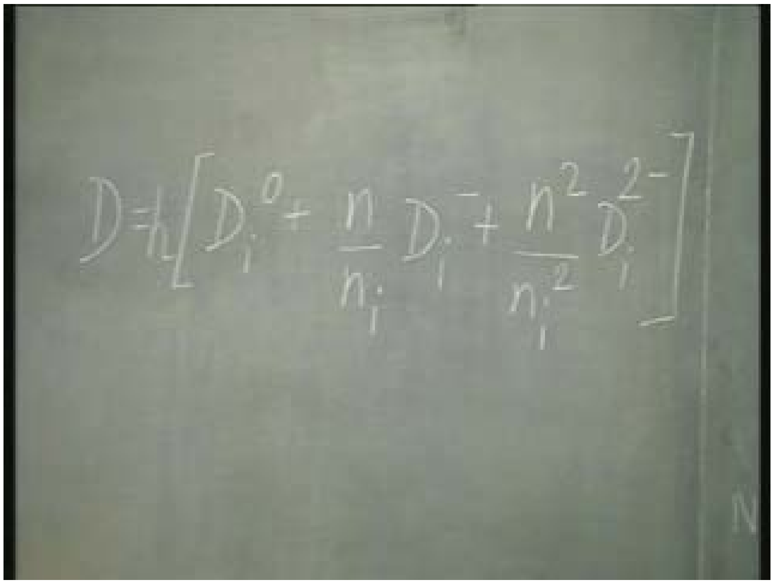
Let us now discuss the situation of phosphorus in silicon, but phosphorus in silicon cannot be given by such a simple doping profile.

(Refer Slide Time: 28:43)



Let us first look at the phosphorus profile in silicon. This is how it is going to look. I have plotted the Log concentration versus the depth and this is how the phosphorus profile is going to look.

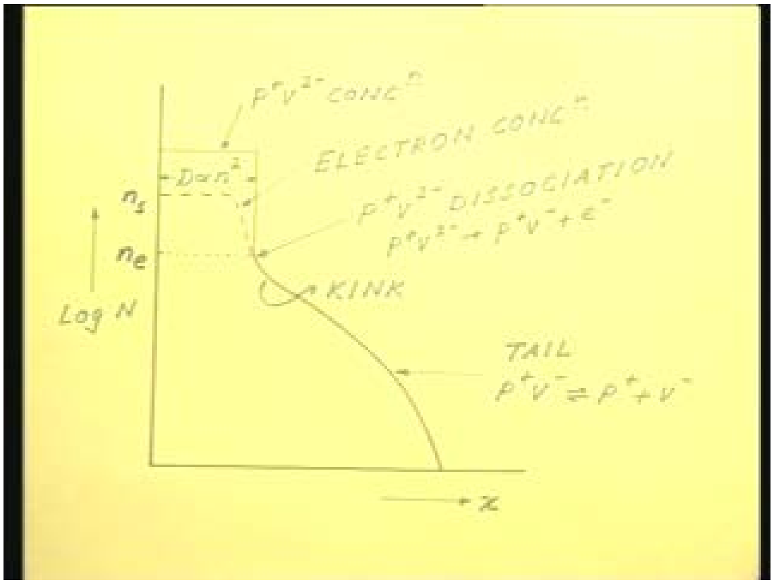
(Refer Slide Time: 29:27)



$$D = h\left[D_i^0 + \frac{n}{n_i} D_i^- + \frac{n^2}{n_i^2} D_i^{2-}\right]$$

In case of phosphorus, the diffusion coefficient is given by this. That is h, the field enhancement factor; then I have D i 0, then I have D i minus multiplied by n by n i and then I have D i 2 minus multiplied by n square by n i square. Therefore, near the surface where the dopant concentration is very high, the dominating term is going to be this. That is what I have also shown here, that in this flat topped region D is proportional to n square and since D is proportional to n square, this is going to be really very abrupt. Over this depth there is hardly going to be a change from the surface concentration, it is going to be truly box like and then it starts to fall abruptly, very steeply, right. As the dopant concentration starts to fall very steeply, the position of the fermi level changes, right. You know, for a very highly doped semiconductor fermi level will be very close to the conduction and then for a moderately doped or for an intrinsic semiconductor, let us say, fermi level will lie at the mid gap, near the mid gap. So, as the doping concentration changes fermi level will move from close to the conduction band to the mid gap point.

(Refer Slide Time: 31:54)



Now, in this flat topped portion you see, D is proportional to n square. That is this is the region, where the phosphorus impurity has interacted with the doubly negatively charged vacancies and that is given by this - P plus V 2 minus. As the fermi level position changes, as it falls 0.11electron volt below the conduction band, this pair starts to dissociate. Very simply speaking, it is something like this; doubly ionized vacancy, doubly negatively charged vacancies that means this vacancy had trapped 2 electrons, neutral vacancy traps 2 electron, that is how it became doubly negatively charged.

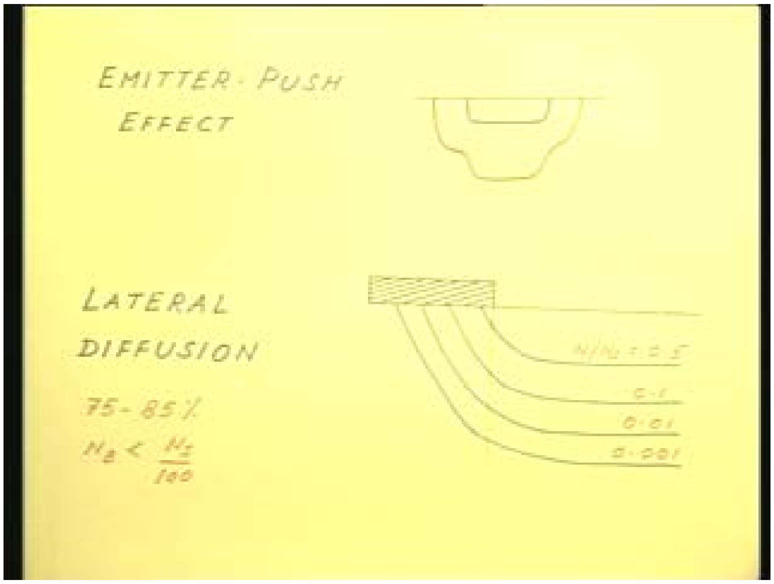
As the fermi level sweeps below 0.11electron volt, below the conduction band, some of the levels must be empty. So, it is not possible for these vacancies to retain all the electrons, right. The levels above the fermi level must be unoccupied. So, then this is the significance of this that then this doubly ionized vacancy, they begin to release electrons and it dissociates. See, P plus V 2 minus, it dissociates into P plus V minus that is the singly negatively ionized vacancy and it releases one electron, because the fermi level is now sweeping by. So, this is where we have a kink, because now the diffusion coefficient is no longer proportional to n square.

Now, these other terms start to dominate, so that the profile is no longer so abrupt, but it tends to be more flat and then as the concentration is lower and lower we enter into what is known as the tail region, for the phosphorus doping profile. In this tail region there is further dissociation of this P plus V minus into the positively charged phosphorus ions and the vacancies. So, in this process, there are more vacancies released. There is a dissociation, there are more vacancies released. That is because the binding energy of this P plus V minus is comparatively less. So, it dissociates more easily and in the process we have a lot of vacancies.

Now, because of these vacancies, there is an enhanced diffusion. You know, diffusion is going to be primarily dependent on the availability of the vacancies. So this, in the tail region a number of vacancies are created, because of the dissociation. So, the diffusion gets enhanced. That is why you have an extended tail into the bulk material. So, in the phosphorus doping profile, you will see that in the different region, different terms dominate. Closer to the surface where the doping concentration is very high, there the diffusion coefficient is proportional to n square. Then, as it falls below a critical level n e at which the fermi level is 0.11 electron volt below the conduction band, at that point this P plus V 2 minus dissociates giving rise to P plus V minus.

Therefore, here the diffusion coefficient is dominated by proportional to n factor and then as we further move into the semiconductor, this P plus V minus also dissociates giving rise to a number of vacancies and the diffusion is more because, diffusion is enhanced because, of the availability of these vacancies. So, the phosphorus profile in silicon is characterized by a kink and tail. A kink here where there is a switchover from D proportional to n square to D proportional to n or D is equal to constant case and then there is the tail because of the enhanced vacancies, right.

(Refer Slide Time: 37:50)

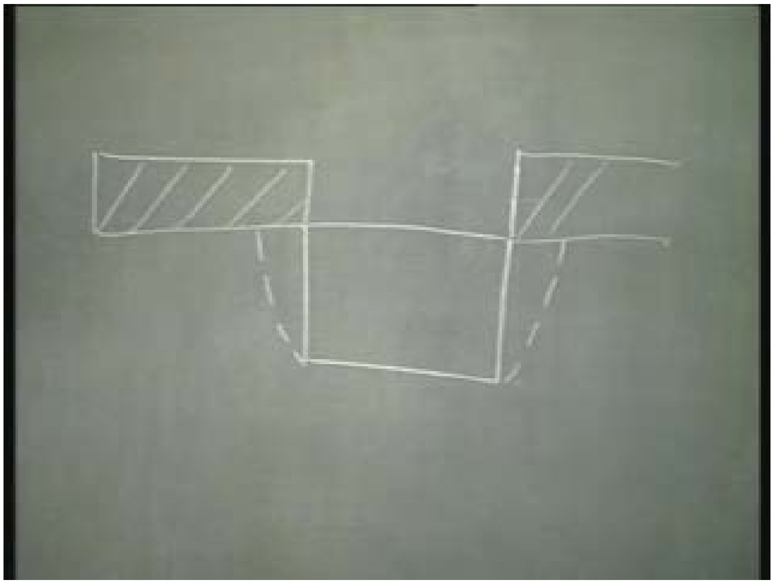


There is fallout of this phosphorus doping profile in silicon, which we will observe in the next slide. Consider an npn transistor. Let us say npn, therefore the base is boron doped and let us say that the emitter is doped with phosphorus. So, you know that after the base diffusion, we have to carry out the emitter diffusion. So, when we are doing the emitter diffusion, the base should also get further deeper diffused, but this deep diffusion should be uniform, right. That is to say, what we expect is this green boundary is my emitter and the base should have been given by this red solid and then the red dotted line, right. That is it should have got deeper diffused all round the emitter, isn't it. But in actual case, what we find is this. We find that there is an enhanced diffusion just under the emitter, just under the emitter. This line is just under the emitter. Just under the emitter there is an enhanced diffusion of the base. It is as if the emitter has physically pushed the base in. That is why it is called the emitter push effect.

Why should this happen? The answer lies in the phosphorus doping profile itself. You have already noticed that further away from the surface, when you are diffusing phosphorus further away from the surface, there is an enhanced diffusion, because of the creation of the vacancies, right. So you see, if this is my emitter, then this is the region where all these vacancy creations will take place, all the enhanced vacancies will be there. These vacancies are here at the emitter base boundary. What will happen to the base dopants there? They will also get diffused. Their diffusion also will get more enhanced, because of the presence of these vacancies because you know vacancies can migrate.

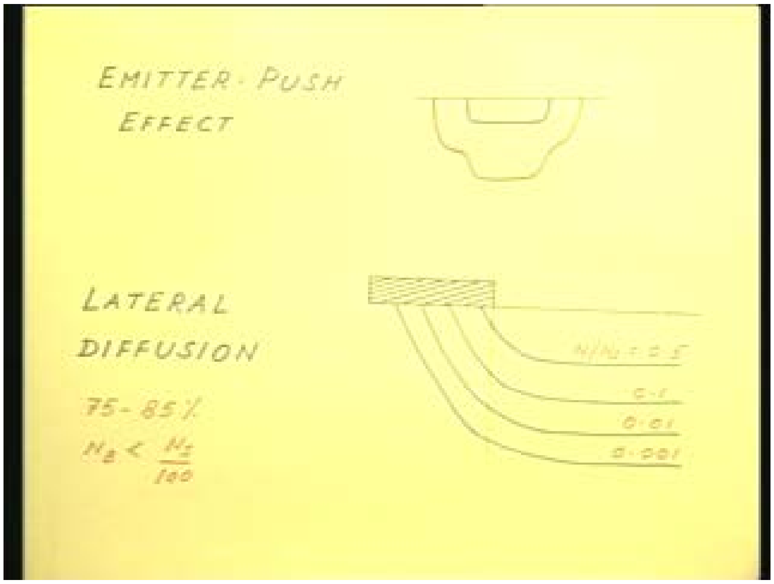
So, just under the emitter where the vacancy formation has taken place, because of the phosphorus diffusion in silicon and the dissociation of P plus V minus where all the enhanced vacancy formation has taken place that is where we will find the base has moved in deeper. If you do not want to have this emitter push effect, do not use phosphorus, use arsenic. Arsenic will not give rise to this emitter push effect. There is another problem associated with the diffusion process. See, under normal circumstances you will always diffuse through a window in the mask. You will open a window in the oxide and you will try to diffuse through that. So, it is going to be something like this.

(Refer Slide Time: 42:05)



You have silicon. Some part of it is masked by oxide and you are carrying out the diffusion. So, ideally what you want is the diffused region will be like this, right. But in practice, it will never be so. It will always move a bit inside and this is called the lateral diffusion.

(Refer Slide Time: 43:08)



EMITTER - PUSH EFFECT

LATERAL DIFFUSION

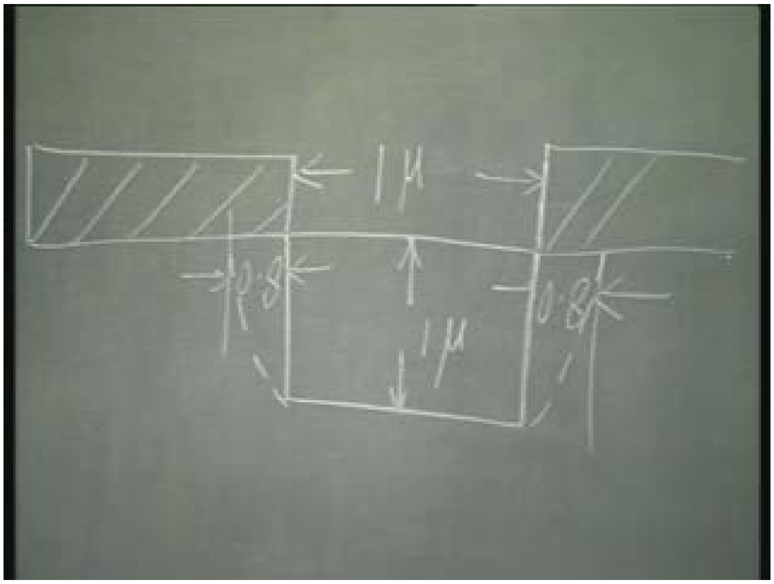
75 - 85%

$N_B < \frac{N_S}{100}$

So, in this case I have just shown one side of the mask window. This is the one side of the mask window and these green lines are the lines of equal concentration. That is here at the surface, the concentration is N S. Now, I am moving in deeper, so that the concentration is falling. This contour is for N by N S equal to 0.5 that is where it has fallen to half the surface doping concentration. Similarly in this case, this has fallen to one tenth of the surface doping concentration, one hundredth of the surface doping concentration and one thousandth of the surface doping concentration.

Notice that the lines are progressively moving under the oxide window. In a practical situation, your substrate doping concentration will probably be say, at least 100 times less than your doping concentration; background doping concentration will be at least 100 times less. In such cases when N B is 100 times, at least 100 times lower than the surface doping concentration that is we are taking about this contour line or lower than that that is the junction depth is here, junction is here, in that case it is found that the extent of lateral diffusion is 75 to 85% of the vertical junction depth.

(Refer Slide Time: 45:18)



That is to say if you want to have a junction depth of 1 micron and you have a window, let us say this window size is 1micron and you want to have a junction depth of, also of 1 micron, then actually this will be 0.8 on this side, 0.8 on this side, so that you will have one point instead. Even though your window size is 1 micron, you will actually have it as 2.6. So, you should take care of this in your device design. Otherwise you have two diffused regions. You are thinking that they are separated but, they might, just might not be separated; they might get shorted if you do not take into account this lateral diffusion situation and obviously the deeper the junction depth is the more will be the spread. If instead of 1 micron if it is 10 micron, this 1 micron window does not have any meaning at all, because your diffused region will be something like 17 micron. If you want to have deeper junction, they must also be widely spaced.

So, so far we have talked about the general process of diffusion. That is what the diffusion profile is going to look like for various dopants - arsenic, boron, phosphorus in silicon. Now, we come to the, next point to be discussed is the how do we carry out this diffusion? Essentially, again this is very simple. All you have to do is open the window in oxide, subject the wafer samples at high temperature to the dopant ambient. Just like oxidation, even here you use a fused quartz tube**,** because that gives rise to the least contamination and the temperature of diffusion is also approximately the same as the range of temperature for oxidation - 800 to 1200 degree centigrade. The question is the choice of dopant sources. Dopant sources can be solid, it can be liquid, it can be gaseous. They all have the various merits and anti-merits, but one thing is common. That is in most cases the basic relation, basic reaction is between the oxide of the dopant and the silicon.

In some cases, you directly use the oxide of the dopant, in some cases you use another product and then in an oxidizing ambient, you convert it to the oxide of the dopant and then allow it to react with silicon. Now you know, silicon has a great affinity for oxygen. Therefore, the reaction product will be the release of the dopant atom and silicon dioxide. So, quite often this diffusion is carried out in the oxidizing ambient; in most cases it is carried out in the oxidizing ambient. That will also serve the purpose of not allowing the out diffusion to take place. At the time of diffusion itself some amount of oxide will be grown on silicon and this will not allow the outward movement of the dopant atom, no out diffusion will be permitted. So, essentially all the diffusions in silicon is based on this principle: you take an oxide of the dopant, at high temperature let it react with silicon. It forms silicon dioxide and releases the dopant impurity and the dopant gets incorporated in the silicon. This is the infinite source diffusion and then also called the predeposition and then after the specified time, you shut off the source.

Now, see this is how the oxidizing ambient comes into the picture. You shut off the source and then you subject it to further high temperature processing called the drive-in. So, if there was no oxide on the silicon, there is a possibility that some of the dopants may get lost by out diffusion. Because of the presence of oxide, they will not get lost. Now that the source is shut off, the outside concentration is also low. But now, they will not get out diffused, they will only move deeper into silicon. So, these actual diffusion systems we will discuss in the next class and then we will move onto the more sophisticated technique that is ion implantation and we will see why, where the ion implantation scores over diffusion.