VLSI Technology

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

Lecture - 15
Diffusion I – Theory of Diffusion and Fick's Laws

In our discussion about VLSI technology, so far we have discussed about the crystal structure, we have talked about crystal growth, we have talked about epitaxy and we have also talked about oxidation. So, now you see, physically you have your substrate material. By bulk crystal growth you have grown a bulk substrate material. If necessary, on top of that you can also grow an epitaxial layer and you have grown an oxide to cover the surface of that crystal. Logically, the next step is going to be open windows in that oxide by photolithography and then selectively dope certain regions. I will skip however, I will, permit me to be slightly illogical and I will discuss photolithography later in this course.

First, the topic I want to discuss today is the, how to introduce a selected amount of dopants into the already grown crystal, be it the bulk material or the epitaxial material. That is I will assume that by photolithography, windows have been opened in the oxide. How? The details we will discuss later, let us for the time being assume that we have opened windows in the oxide and we are going to introduce selected amount of dopant into this crystal by a technique called diffusion. So, I have more or less defined diffusion for you. Diffusion is a process by which controlled amount of impurity is introduced into the semiconductor.

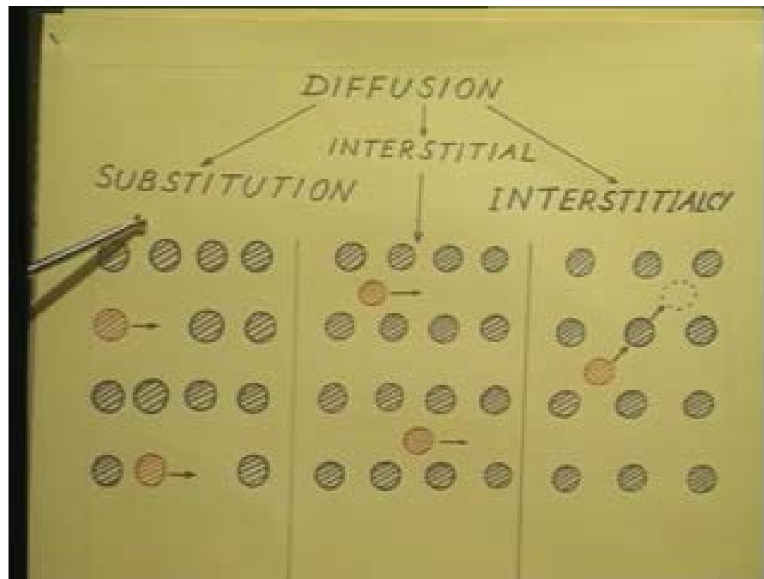
Physically speaking, what do we actually do? We subject the semiconductor at high temperature to an ambient containing the dopant impurity. Now, you know that a real crystal is not a perfect crystal. There are point defects in it. There are vacancies, there are interstitials. So, when this practical crystal is being subjected to an ambient containing the impurity atom, the impurity atoms will move inside the crystal, because a concentration gradient exists. I am harping on the point that the ambient containing the impurity atom, so the ambient surrounding the crystal, it contains the impurity. The concentration of the

impurity is high there and there are no impurities to begin with, inside the bulk crystal. So, the concentration of impurity is low there. Therefore, there exists a concentration gradient and under this concentration gradient impurity atoms are going to move inside the crystal.

This movement is going to be governed by the existence of the point defects that is vacancies and interstitials. Physically speaking, if the crystal has a lot of vacancies, it is very easy for the impurity atom to go and occupy these vacant sites, to give you just an example, right. So, the movement of the impurity atoms inside the semiconductor crystal will be governed by the presence of these point defects. So you see, I can view diffusion as atomic movement of the diffusing impurity, also called the diffusant. I can view it as the atomic movement of the diffusing impurity in the lattice through the vacancies or through the interstitial spaces, right.

Remember the interstitial spaces, silicon is a diamond crystal structure. In a diamond crystal structure you have seen that there are lots of empty space. Only about 30%, 34% of it is occupied by the lattice atoms and the rest 66% is free interstitial spaces. So, the impurity atoms can move in between the regular array of atoms or it can move in through the vacancies which will be available in the regular crystal lattice. So, this is diffusion - the atomic movement of the dopant impurity through the vacancies or through the interstitial spaces.

(Refer Slide Time: 6:32)

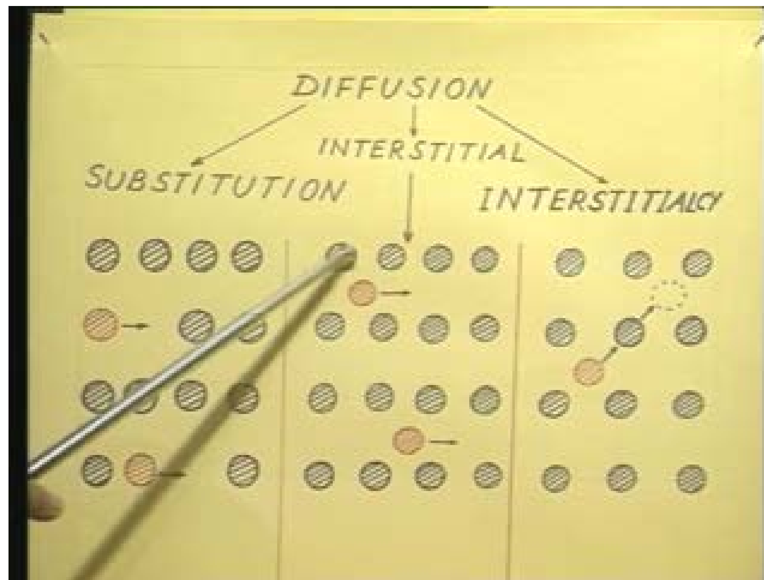


If we now look at the card, we find that I have divided the process of diffusion or let us say the process of movement of dopant atoms inside the crystal lattice in three categories: substitution, interstitial and interstitialcy. Looking at the substitutional movement of the dopant atom, what have I depicted here? These blue shaded circles are the regular silicon atoms in the crystal lattice and the red shaded circle is the dopant atom. So, let us assume that this is the direction of the bulk. That is on the left of the figure, I have the surface. So, at the surface I have a high concentration gradient. I can see a dopant atom sitting here and next to the dopant atom I have shown a vacant site. The vacancies exist in a real crystal.

Now, under this concentration gradient and because of the presence of this vacancy, this dopant atom can now move into the vacant site as depicted by the black arrow here. Something similar is happening here. I have a dopant atom and next to it there is an adjacent vacancy. So, under the concentration gradient, the dopant atom can move into the vacant site. So, this is called a substitutional impurity. That is the dopant atom is sitting on the lattice site normally occupied by the silicon atom. It is substituting for the host atom. So, this is called the substitutional impurity. Examples are almost all the common dopants in silicon.

For example, phosphorus, arsenic, antimony and remember, the substitutional impurities are going to be electronically active. That is if I put Group V elements, all of them, phosphorus, arsenic, antimony, they are all Group V elements; if I introduce these Group V elements, they are going to replace silicon in the host atom, sit in a substitutional site and because they have one extra electron, it can donate it. That fifth electron is free to move into the conduction band and therefore we have the donor, right. All Group V elements are donors for silicon and they take part in the electronic activity, they modulate the conductivity of the semiconductor material, agreed.

(Refer Slide Time: 10:07)



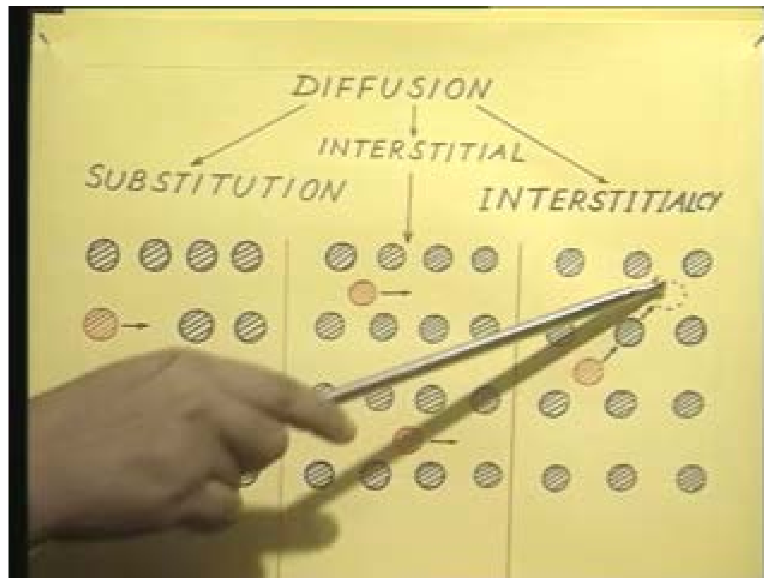
Let us look at the second possibility that is movement by interstitial. In this case what is going to happen? As I told you, in a diamond lattice there are lots of empty space in between the regular array of atom and the impurity is free to move through this interstitial space. You can well understand that the movement is going to be relatively unhampered. It does not have to wait for a vacant site; interstitial space is anyway available, it can move in through this interstitial space. So, movement is going to be much faster when the impurity is going to diffuse through the interstitial spaces. So, this is called diffusion by interstitial. All Group I and Group VIII elements that is lithium, sodium as well as helium and argon they move in silicon by this mechanism. Remember, lithium is actually a

notable exception. Remember, we said that lithium actually is an interstitial dopant. That is it moves through the interstitial spaces, but it still acts as a donor.

Now, a modification of this interstitial movement process will be a situation where the impurity atom can end up in both substitutional and interstitial sites. That is I may have a dopant which can sit either in a substitutional site or in an interstitial site. Why then am I classifying it with interstitial? Simply because, the movement of the dopant atom will be governed by its movement through the interstitial spaces. As I told you, if the dopant atom has to be a substitutional type, then its movement is greatly hampered. It depends on the availability of a vacancy.

On the other hand, if it is moving through the interstitial space, it can move very fast. So, if I have a dopant which can sit either in a substitutional site or in an interstitial site, then its movement will be governed only by the movement through the interstitial space. The substitutional atoms can be assumed to be more or less stationary. Their movement is going to be much less, right. So, the process of diffusion is going to be governed by the movement through the interstitial spaces. That is why I am classifying that as a sub class of the interstitial dopant, even though the dopant atom will be found both in the substitutional sites as well as in the interstitial sites and the example of such dopants are copper, nickel and gold. Particularly gold is very important, because you know gold is used sometimes as a deep impurity in silicon and its movement is like that, both substitutional and interstitial, but the movement is primarily governed by the interstitial movement and finally we come to the mechanism called interstitialcy.

(Refer Slide Time: 13:52)



This is very interesting. In interstitialcy, what happens? The dopant atom has started out the red hatched circle. It is originally in an interstitial site. Now, what does it do? It pushes a host atom out of its regular place in the lattice. It pushes it into this interstitial space and itself goes and occupies a substitutional site. So you see, it started out with being an interstitial, ended up by occupying a substitutional site. This is the movement of boron in silicon. Remember, boron is a very important dopant in silicon. It is the, almost the only p-type dopant we have in silicon and boron's movement is interstitialcy. That is boron is finally found in the substitutional site, but it, in the process it has dislodged a host atom pushed it into an interstitial site and itself has gone and occupied the host atom position.

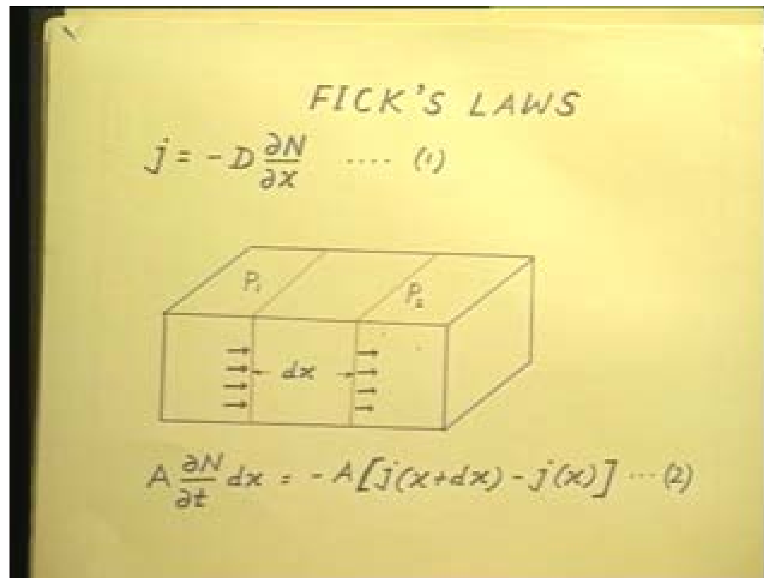
So, these are the three basic mechanisms by which an impurity atom can move inside silicon. Substitutional, that is the most straight forward. It substitutes for a host atom which is done by in case of phosphorus, arsenic, antimony. Then interstitial, that is also straight forward, movement is through the interstitial spaces and that is lithium, sodium, helium, argon. A modification of interstitial in which we have both substitutional and interstitial sites occupied by dopants as done by copper, gold and nickel and then a very interesting movement, interstitialcy in which the dopant atom starts out with being an

interstitial, pushes the host atom from its regular lattice site into the interstitial space and itself occupies the substitutional, which is boron in silicon. You see, boron also has to have electronic activity. It should also modulate the conductivity of the semiconductor, so it has to end up in a substitutional site. Even though it starts out with being an interstitial, it ends up in the substitutional site and of course, I can have any amount of combinations, but these are the three basic mechanisms. The combinations are relatively less. This is viewing the diffusion from a physical point of view. What exactly is happening?

After the atom, the dopant atom, moves inside the crystal what exactly is happening? How are they moving inside? Now, what are the important things we associate with doping or diffusion? We would like to know how far the dopant atom has gone, we would like to know what is the doping profile in silicon. That is as we move form surface to bulk, how is the doping concentration changing? We would like to know where the junction is. If it is a p-type semiconductor and we are introducing n-type impurity where the two doping concentrations become equal that is the junction. What is the junction depth and we will also like to know what the surface concentration is. All these things are important in order to fully characterize the diffusion profile or the doping profile of the impurity in silicon and in order to do that, we should carry out the mathematical analysis, right.

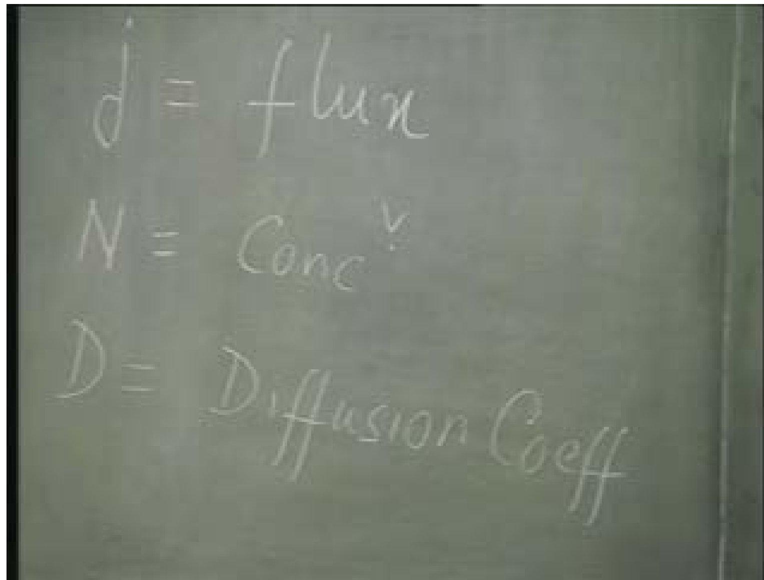
Now, the mathematical analysis of diffusion, if we view it only in one dimension that is just the movement of diffusion atom in this direction, one dimensional analysis is done by Fick’s law.

(Refer Slide Time: 18:36)



If we look at the second card, you will find that Fick's first law is given simply by j is equal to minus D del N del x. What are the individual terms? j of course stands for the flux. What is flux? Flux is rate of transfer of solute or the dopant per unit area. This is valid for a dilute solution. What do I mean by a dilute solution? That is the dopant concentration is going to be, amount of dopant atoms incorporated in silicon is going to be much less, than its, than the number of host atoms. So, that is the dilute solution, right, isn't it. For example you see, let me give you a physical example. Even a very highly doped silicon will have may be a surface concentration of 10 power 19 or 20, 19 or 20 atoms per centimeter square. So, that will work out to be, I mean, much, much less than its molar concentration, the number of silicon atoms per unit area. So, it is a very dilute solution, so to speak. That is the, only very few dopant atoms, actually physically may be 1 in a million or 1 in 10,000 is incorporated in the crystal.

(Refer Slide Time: 20:30)



So, this Fick's law is valid actually for this dilute solutions and in this we can say that j is actually the flux which is the rate of transfer of the dopant per unit area, N is the concentration of dopant, x is the direction in which the movement of the dopant atom is taking place, therefore del N del x is the concentration gradient and the movement is obviously from the high concentration side to the low concentration side. That is why I have a negative sign, a minus sign, because the movement is from the high concentration side to the low concentration side and D is the diffusion coefficient. In other words, what we are trying to say is that the flux is proportional to the concentration gradient of the dopants and D is the proportionality constant, the diffusion coefficient and the negative sign simply signifies the direction of the movement of dopant from the high concentration side to the low concentration side. That is Fick's first law, very simple.

Now, what actually is this D?

(Refer Slide Time: 22:18)

D is the diffusion coefficient and I may have already mentioned this that D is given usually by D 0 exponential minus E A by KT. This D 0 is considered to be independent of temperature. So, this is the temperature independent part, this is the temperature dependent part. Obviously, you can make out that as the temperature increases, the diffusion coefficient or the diffusivity is going to increase, right. Now, if I plot D versus 1 by T, then, D versus 1 by T in a log scale, then from the slope I can find what is E A, E suffix A that is the activation energy. This is called the activation energy of diffusion.

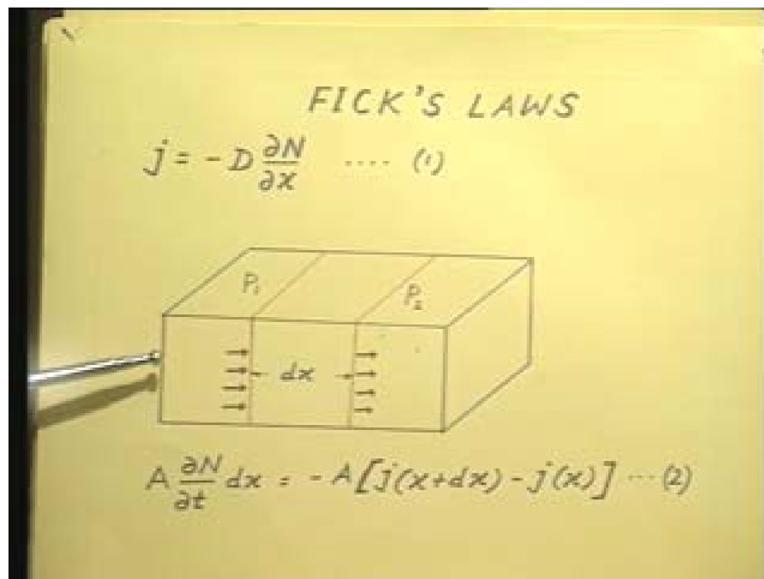
Now, this activation energy, the value of this activation energy is actually going to tell you how difficult or how easy it is for the dopant atoms to move inside the semiconductor. For example, if the dopant atom is substitutional, then the activation energy value is 3 to 4 electron volts. Compare it with the interstitial dopant in which case the activation energy is 0.6 to 1.2 electron volts, much less. In other words, E A for substitutional is much, much greater than E A for interstitial. What does that signify?

It simply signifies that the movement through interstitial spaces is a much faster process.
The diffusion coefficient is going to be much greater in this particular case for the interstitial movement, isn't it. So, what you have seen physically is now being given to

you mathematically. This is the significance of this mathematical representation that the diffusion coefficient for an interstitial dopant is much higher. That is the diffusion is much faster in case of a dopant which is interstitial. Physically, obviously you understand that that is because its movement is relatively unhampered, it can move through the interstitial spaces. On the other hand, the movement through the substitutional sites is relatively more difficult because, one, the dopant atom has to depend on the availability of the vacancies. So, that is why the activation energy associated with the substitutional impurity is much larger than the activation energy associated with the interstitial impurity.

Now Fick's second law.

(Refer Slide Time: 26:00)

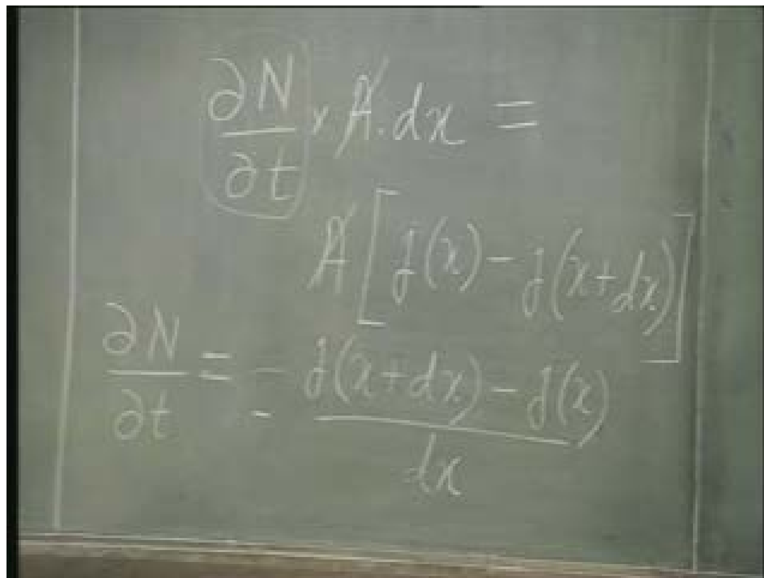


In order to understand Fick's second law, again we have to look at the movement of the dopant atoms through this parallelepiped. Imagine that this is the crystal extending in this direction. The dopant atoms are moving from the left to the right, agreed. Now, I have identified two planes in this parallelepiped, marked by red and called P 1 and P 2. So you see, physically the dopant atoms are moving in, crossing the plane P 1, moving towards

the right. Let us consider what is happening between these two planes P 1 and P 2 separated by an incremental distance dx.

Now, dopant atoms are moving in through this space and some of them are going to be staying in the space in between these two planes. That is what diffusion is all about, impurities are going to substitute for the host atoms or are going to stay at the interstitial sites, right. So, in this movement of dopant atoms, there is going to be some accumulation of the dopant species in between these two planes. As it is moving from surface to bulk, at every region in between two successive planes some dopant atoms are going to be caught in the lattice. Rate of accumulation of the dopant atoms in between P 1 and P 2, what is that? Let us see. I have already defined N to be the concentration of the dopant atoms. So, if dopant atoms are going to be accumulating, then there is going to be a change in this concentration, yes.

(Refer Slide Time: 28:43)



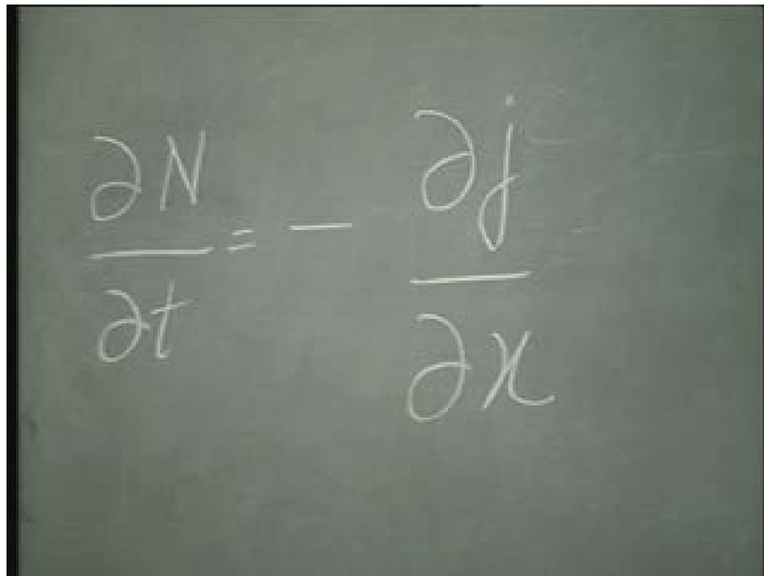
So del N del t is going to give me the rate of change in the dopant atom concentration. This is the rate at which the dopant atom concentration is changing and let us say that the cross sectional area of the planes P 1 and P 2 is A and in between P 1 and P 2, I have an incremental distance called dx. So, A dx is actually the incremental volume. So, this del

N del t multiplied by A dx is actually going to be the rate of accumulation of the dopant atoms in between these two planes.

We will take a very simple example. Let us say that at P 1, the concentration of the dopant atoms was 5 per unit volume, just some number, 5. Let us say, at P 2 the concentration is 4 per unit volume. Concentration has reduced obviously. So, how many atoms are taken in this? 5 minus 4 multiplied by the incremental volume A into dx. So, the rate of accumulation of the dopant atoms will be given by the rate of change in concentration multiplied by the incremental volume. So far, so good. Now, this is going to be equal to, I can equate it with the change in the flux density.

What is flux? Let me remind you once again, flux is rate of transfer of the dopants per unit area. So, at the plane P 1, there is going to be a flux. At the plane P 2, there is going to be a different flux. The difference of these two fluxes multiplied by the area is also going to be the rate of accumulation of the dopant atoms in between P 1 and P 2. So, I can say that this actually equal to A into ......, isn't it. I can therefore now rewrite it as ...., can I not?

(Refer Slide Time: 32:31)



When in the limit dx tending to zero, the right hand side can be simply written as ....., isn’t it.

(Refer Slide Time: 32:48)

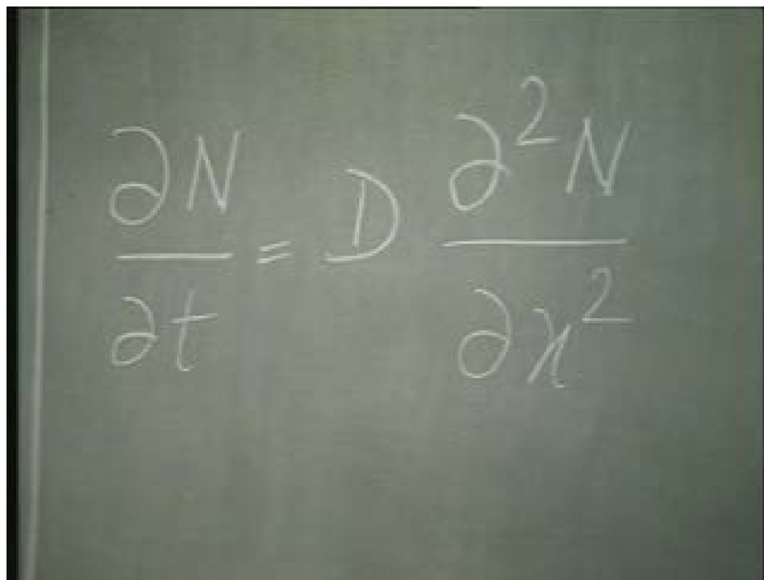
If I put ......, dx is after all an incremental distance between the two planes P 1 and P 2, then I can say del N del t is equal to minus del j del x.

(Refer Slide Time: 33:29)

Now, from the first law of Fick's, if I substitute for this j, then I can write that this is equal to ....; del N del t is equal del del x of D del N del x. This is Fick's second law of diffusion. The first law of diffusion is j equal to minus D del N del x. The second equation is del N del t is equal to del del x of D del N del x. These two equations govern the diffusion in one dimension. Diffusion of impurities in silicon in one dimension is governed by these two laws. They are general laws of diffusion, valid for all dilute solution cases and this is also used to find out the movement of dopant atoms in silicon. Now, Fick's second law actually is very interesting. You see, it says del N del t is del del x of d del N del x. Now the question is, what about D? Is D a constant or is it a variable? That is can I take this D out of the differentials? If D is a constant I could do that. I could simply take this D out. So, from here, therefore we branch; mathematically speaking, we are going to branch into two branches depending on whether the diffusion coefficient D is a constant or not.

Let us try the easier situation first. That is let us assume that the diffusion coefficient is a constant. Whether we are justified in making this assumption, we will discuss later when we come to the real life situation. But for the time being, let us simply take the case mathematically the easier option. That is let us make the diffusion coefficient a constant.

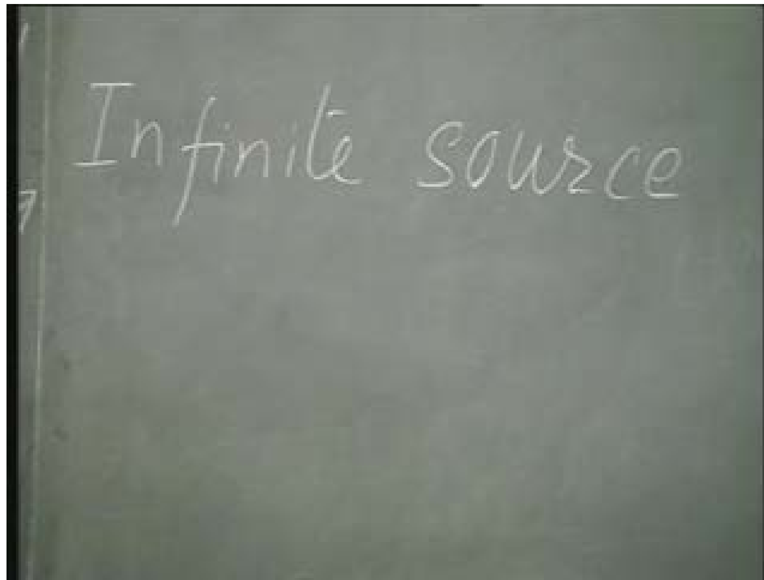
(Refer Slide Time: 36:27)



If I make the diffusion coefficient a constant, then this equation can be rewritten as ....
I have taken D out of the differentials. Therefore, I can rewrite this equation as del N del t is equal to D of, D into del 2 N del x 2. This is called Fick's simple law of diffusion, simple law of diffusion; simple, because we are taking the easier option that is the diffusion coefficient is a constant. If that is so, then let us investigate what about the doping profile? Now, we know that the diffusion is going to be governed by this equation. Assuming that the diffusion coefficient is a constant, the diffusion is going to be governed by this equation.

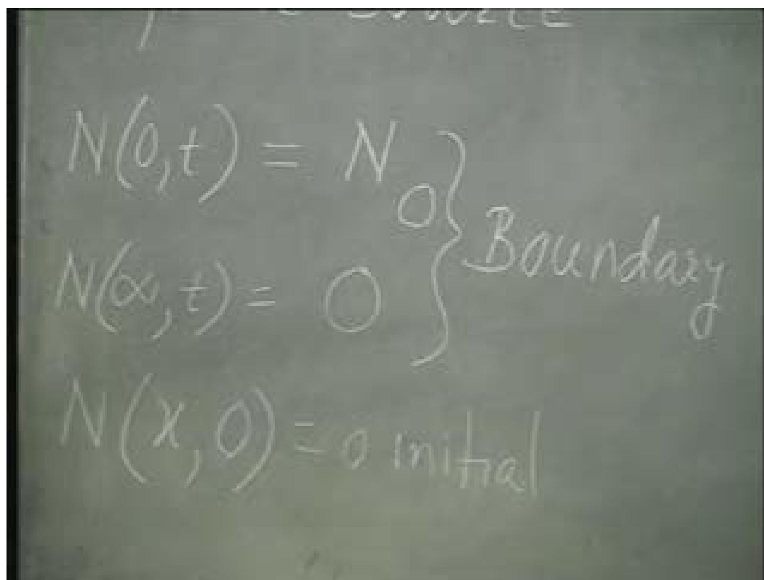
Now, what about the doping profile? In this case also, there are various possibilities. What are the possibilities? One possibility is I have an infinite source of diffusion, infinite source of diffusion. What does that mean? That means that the surface concentration of the dopant is always a constant. The surface concentration is given by its maximum possible value. That is the limit governed by solid solubility, right, solid solubility limit. When I have an infinite source, the surface concentration at most can be equal to the solid solubility limit of the dopant in the semiconductor. So, no matter for how long I carry out the diffusion, the infinite source is infinite, it is very big, so the surface concentration is never allowed to fall from this solid solubility limit. So, that is one situation, right. That is called the infinite source diffusion or sometimes the constant source diffusion; source is a constant, infinite.

(Refer Slide Time: 38:57)



So, from here therefore we can branch off to infinite source diffusion. In this infinite source diffusion, what are my conditions?

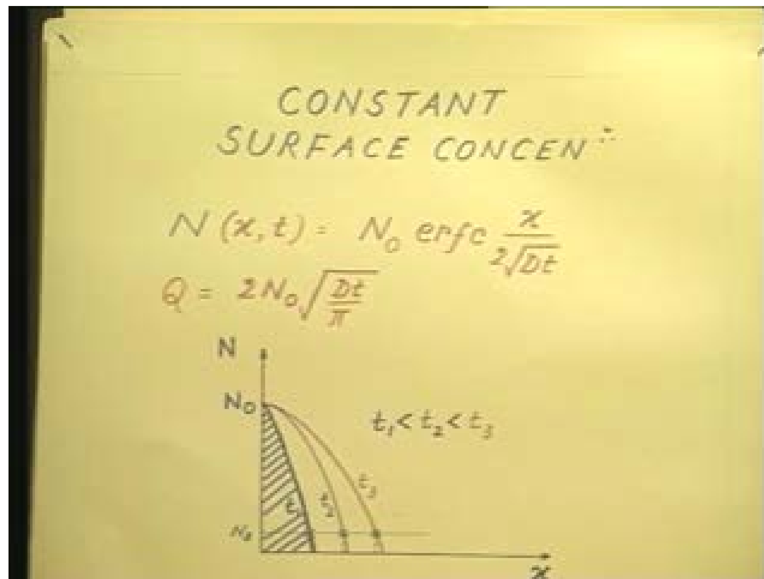
(Refer Slide Time: 39:21)



The first condition is as I told you that the concentration at the surface, at the surface x is equal to zero; at the surface, the concentration is given by the solid solubility limit. I am going to denote it by N 0. So, this is one boundary condition. At the surface, the doping

concentration is equal to the solid solubility limit. What is the other boundary condition? That is at an infinite distance from the source, the doping concentration is going to fall to zero. So, this is my other boundary condition and when I started the diffusion that is when t is equal to zero that is my initial condition, when t is equal to zero, then also the doping concentration was zero. These two are the boundary conditions and this is the initial condition.

(Refer Slide Time: 41:09)

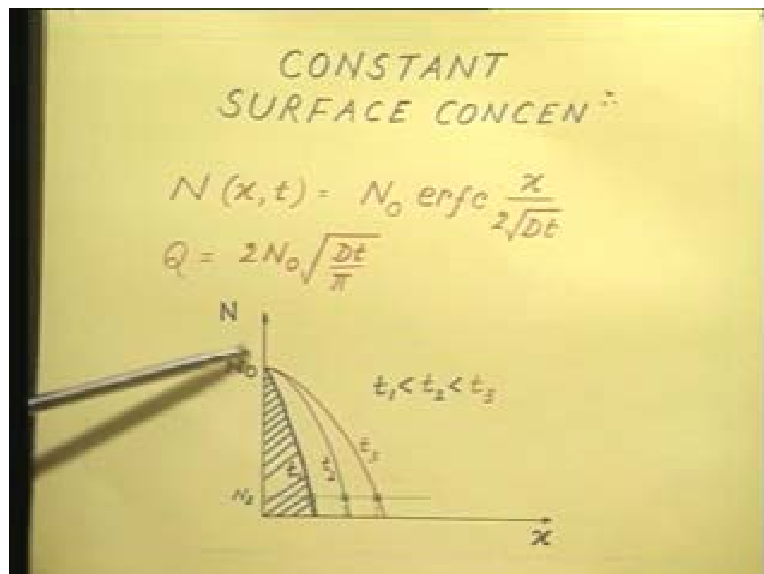


If we solve Fick's simple diffusion equation subject to these two, boundary and the initial condition, then we will see that the doping concentration in the crystal is given as shown in this third card as N, a function of x and t, is given as N 0 error function complement x by 2 root over D t, D being the diffusion coefficient, small t being the time. What is this complementary error function? Complementary error function is 1 minus error function.

(Refer Slide Time: 41:42)

That is error function complement of x is nothing but 1 minus error function of x. Now, what is error function of x? Error function of x is given as, error function of x is given by 2 by root pi, it is actually a definite integral, 2 by root pi integral with the limits zero to x exponential minus z square dz. That is error function of x. Complementary error function of x is 1 minus this error function of x and the nature of this error function curve is as shown here.

(Refer Slide Time: 43:02)



I have plotted N versus x, N is the concentration, concentration versus x with the time as a parameter. So, these three curves, the blue, the violet and the red, they are for three different times. t 1 being less than t 2, t 2 being less than t 3 and this is the shape or the nature of the complementary error function profile. Notice that all of these curves start from the same surface doping concentration given as N 0, no matter what the time is. For however long the diffusion is carried out, if it is an infinite source diffusion that is if the doping profile is governed by the complementary error function curve, then you see at x equal to zero, this term is going to be zero. Therefore, the complementary error function value is going to be 1 and N is always equal to N 0 at x equal to zero, no matter what the time is, agreed.

So, that is one important point about this infinite source diffusion, the surface concentration always remains constant. The surface concentration is given by the solid solubility limit N 0, in this case. So you see, we have found two important things regarding this doping profile, regarding the infinite source diffusion. One, that is the doping profile follows the complementary error function type, two - the surface concentration is given always by the solid solubility limit.

What are the other things that we want to know about this? Well, one very important thing is the total amount of impurity which has been introduced in the semiconductor during this diffusion process, the total amount of impurity and in order to obtain the total amount of impurity, what do we have to do? We have to integrate the area under the curve. The total amount of impurity is given by this hatched area, right. In this particular case, if I have carried the diffusion for t 1 time, then this is the doping profile and the area under that curve is actually the total amount of impurity.

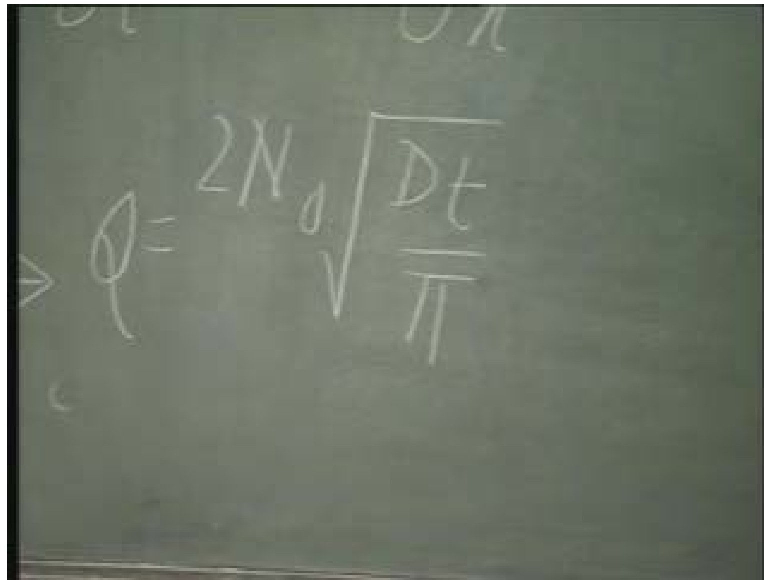
So, in order to obtain the total amount of impurity given by Q, we have to integrate the doping concentration profile.

(Refer Slide Time: 46:09)

If I do that, then we get Q is actually equal to integral zero to infinity or in other words, I can write it as integral zero to infinity N 0 error function complement of x by 2 root D t dx or I can write it as N 0, I have changed the variable, from x I have changed it to z; that is I have taken x by 2 root D t to be equal to z, so that dz is equal to dx by 2 root D t. So, I have replaced this dx by dz, taken 2 root D t outside, N 0 being a constant can be taken out of the integral. So this is what I obtain that Q is equal to N 0 into 2 root over D t integral zero to infinity error function complement of z dz.

Now, this error function complement of z dz integrated in the limit zero to infinity is a definite integral and its value is given by 1 upon root pi.

(Refer Slide Time: 48:03)



So, Q is given by twice N 0 root over D t by pi.

(Refer Slide Time: 48:28)

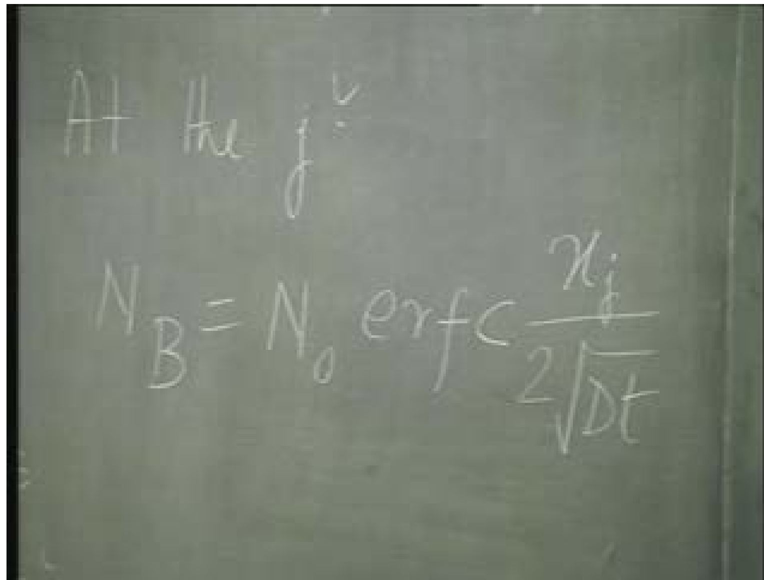
That is simply because this is equal to 1 upon root pi. So, I have found the third important parameter that is the total amount of impurity, the total amount of impurity that has been introduced during this diffusion. Notice that this is actually dependant on the time of diffusion, very rightly so. The more, the longer period you diffuse, the more impurity is

going to be introduced and this square root of D t is a very important factor as well, because this is going to, this is going to be a measure of how much the dopant has introduced, how far the dopant has gone inside the semiconductor. You see, D t actually has, square root of D t actually has the unit of length. It is often referred to as the diffusion length. So, this is actually a measure of the junction depth in a way, which brings me again to the fourth important factor and that is the junction depth. I know that square root of D t is going to be a measure of the junction depth.

How do I find the junction depth? Well, in order to the find the junction depth, I must first know the background doping concentration that is the substrate doping concentration. Before diffusion the substrate must have had some doping concentration. So, how do I define the junction? If the substrate had this doping concentration shown by the green line N B, uniform doping concentration, then at the point at which the complementary error function profile is going to cut it, intersect it, that is going to be my junction depth. In other words, that is the point where the two impurity concentrations are going to be same. So, this is the junction, right.

How do we function the junction depth? All I have to do is I know that at this point the doping concentration is equal to the background doping concentration. That is N is equal to N B.

(Refer Slide Time: 51:20)



So, I only have to solve this equation that is at the junction N is equal to N B. That is N B is equal to N 0 complementary error function x j by 2 root Dt. So solving this you can find out what the junction depth is going to be and again notice that junction depth is going to be actually some factor dependent on N B and N 0 and then square root of D t. So, square root of D t is a very important factor, do not forget square root of D t. Square root of D t has the units of length and your junction depth is going to be directly proportional to this square root of D t. So, if you want to have a deep junction, either carry out the diffusion for longer time or try to increase the diffusion coefficient by increasing the temperature. That is why in a practical diffusion process, these are the two parameters we play with, time and temperature. Let us stop here today.

In the next class I will decide the other, we will discuss about the other branch that is when the source is no longer infinite.