Fundamentals of Semiconductor Devices

Prof. Digbijoy N. Nath
Centre for Nano Science and Engineering
Indian Institute of Science, Bangalore

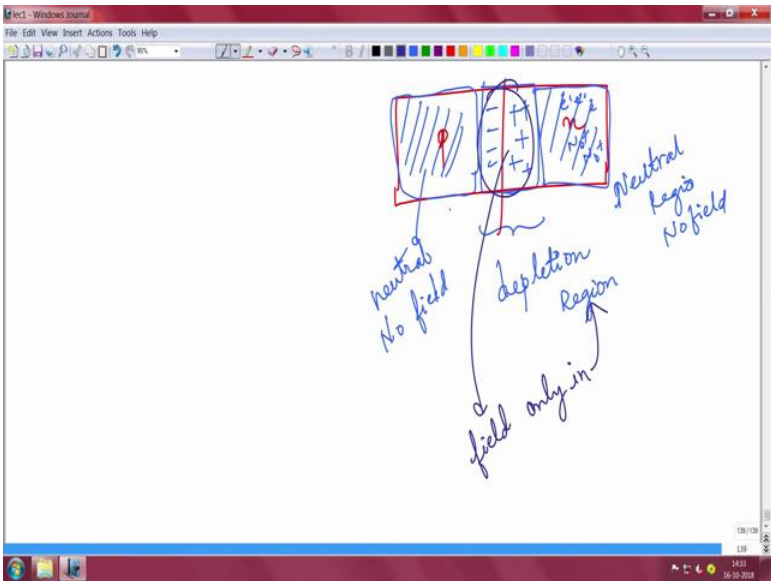
Lecture – 15
p-n junction under equilibrium

Welcome back. So, we ended the last lecture with the introduction of p-n junction; I told you the definition of depletion region. How a depletion region is formed and why is it called depletion, how electrons and holes will diffuse from one side to another side and create that depletion region ok. I also mentioned repeatedly that there is a field that will develop in the depletion region; we will see the in the today's class how that fields will be you know operational and all. But the field will oppose any further movement of electrons from n side and holes from p side, from diffusing to the other side, they will prevent that the, that field will prevent that.

So, today's lecture we will focus on the energy band diagram of the p-n junction and we shall introduce a few concepts like built in potential for example and try to understand what defines the extent of depletion on the region, you know depletion region how wide is that region. Is it 5 nanometer, is it 500 nanometer, is it 5000 nanometer whatever how do you know how much depletion has formed and what is built in potential.

We will only focus on static condition, which means no voltage is applied, no light is shining; its equilibrium; there is absolutely no electricity flowing through p-n junction that is called p-n junction in equilibrium and that is what we are studying till now ok no voltage is applied. So, we will come to the whiteboard now and we will continue from where we had left last time.

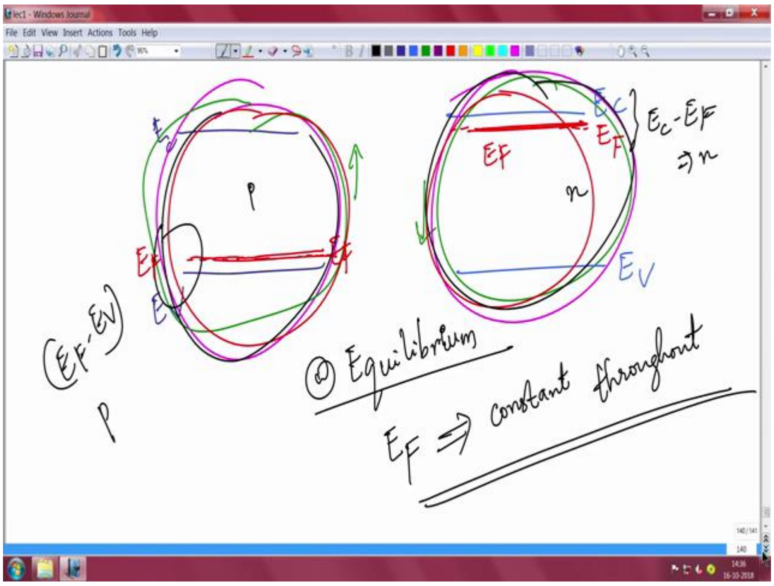
(Refer Slide Time: 01:49)



So, I told you that you have a p you have an n, because holes will try to diffuse from p side to n side electrons will try to come from n side to p side they will recombine near the junction and there is a depletion region that forms. The name depletion means that this region is depleted of carrier because electrons and holes have recombined; there are no mobile carriers here. What remains is negatively charged acceptor ions, positively charged donor ions; Of course, this region has lot of mobile holes and also ionized acceptors, this is called neutral region. This is also called neutral region, this is also called neutral region because they are very number of negatively charged electrons but also equal number of positively charged you know donor ions.

So, this region is called neutral region, this region is called also neutral region, there are no fields here, there will be no electric field a neutral region ok. No field, no field in the neutral region. The field only exist in this region; there is a field only in the depletion region, only in depletion region and the field prevents from further diffusion I told you. So, how will actually the energy band diagram of this semiconductor look like of this junction look like? Ok.

(Refer Slide Time: 03:23)



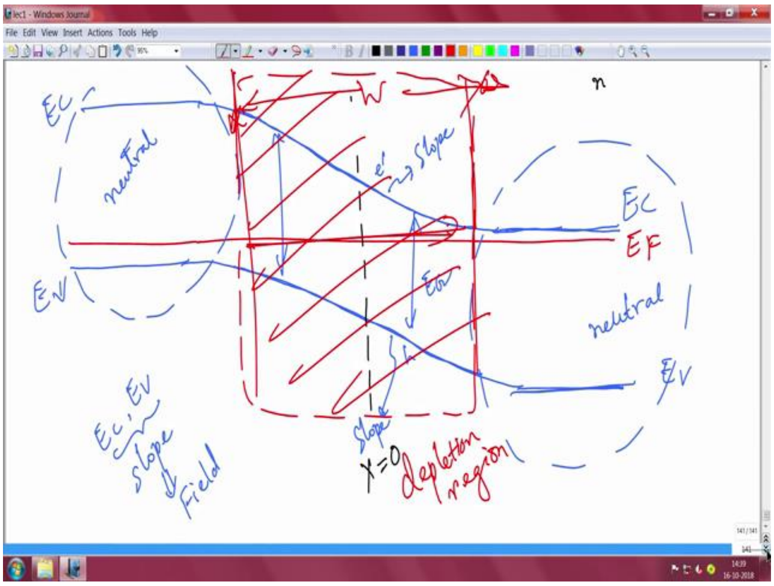
So, you have a p type semiconductor. So, suppose this is E_c, this is valence band E_v and suppose the Fermi level is the p type semiconductor. So, Fermi level is here this is your Fermi level ok. Now, this is your n type semiconductor, conduction band, valence band and this is a Fermi level. If you remember, these are all lightly doped semiconductors at room temperature, 100 percent ionized.

So, there is no incomplete ionization. Now these are before joining, this is p type, this is n type. Now when we join, the thing that we should keep in mind in equilibrium, at equilibrium there will be no current flow right. If there is no current flow then Fermi level E_F, Fermi level will have to be constant throughout.

If there is equilibrium and there is no net current flow, then Fermi level has to be constant throughout because any slope in Fermi level, any gradient in Fermi level means it is not equilibrium, current is flowing. So, now, basically what it means is that; what does it mean? It means that the Fermi level on this side has to basically become equal to the Fermi level on this side. In other words, this side will have to come up and this side will have to come slightly down so that this Fermi level aligns with this Fermi level.

So, I will bring this up, I will bring this down, until the Fermi levels align and that is equilibrium.

(Refer Slide Time: 05:18)



So, eventually the Fermi level has to be same throughout, eventually the Fermi level has to be same throughout; it has to be constant and same throughout. Suppose, this was your junction that has formed; this was x = 0 where on the left side p type is there, on this side p type is there on the right side you have n type and they have joined here.

So, the Fermi level has become equal. So, far away from the junction, if I go back to the previous slide, far away from the junction, this spacing of conduction band and Fermi level has to remain constant because the doping has not changed. Far away from the junction, this difference of Fermi level minus valence band that dictates your p type concentration by the way, this dictates your n type concentration by the way, that cannot change because the doping has not changed; only near the junction some spice is taking place you know some event is occurring.

Away from the junction, this should be preserved, this diagram should be preserved ok. So, away from the junction so, far away from the junction, this is the n side right. So, far away from the junction, your conduction band and Fermi level have to have the same difference. This is valance band, you know and far away from the junction on the p side also, your valance band, this is your E_v and this is your conduction band E_c this is same.

So, you see this particular region and this particular region is the same before and after joining, but the Fermi level has now aligned if you look carefully in the red colored E_F the Fermi level has aligned right. So now, what you do is you join this conduction band and valence band. So, conduction band will now be joined like this, sorry it will join here I am not very good at drawing here, but you see my point right. So, conduction band will like that.

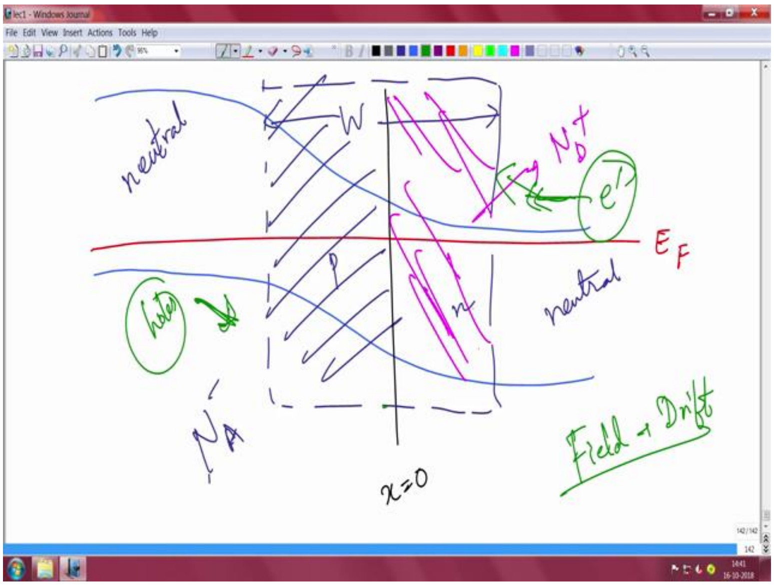
There is no discontinuity; its smooths, its smooth my drawing is little bit this thing; similarly valence band will have this, it is smooth, there is no discontinuity here, its smooth. So now, that is good. So, this is x = 0, which is the depletion region you know. The depletion region is the region where there is a field and field exist whenever conduction band or valence band, whenever they have a slope, whenever they have a slope, it means there is a field. The field need not be externally applied, the field can be internal also, as in this case I have not applied any external field.

The field exists internally without application of any external voltage. Any slope in conduction and valence band means that there is a field; is there a slope here? No there is no slope; is there slope here? Of course, there is no slope. So, this region is your neutral region that I had been talking about. There is no slope here no slope here. So, this is your neutral region that I had been talking about; the slope exists in this region.

If you look at this region carefully, if you look at this region carefully, there is a slope the conduction band is bending; that means, there is a field that is up to here at this point. Similarly, it starts sloping, there is a field here eventually it is up to here. So, of course, slope in conduction and valence band is the same because your band gap is the same your band gap is everywhere same you know. So, whatever slope you have here is also the same slope that you have in valence band. So, the field experienced by electrons here and the holes here is the same basically ok.

Because, there is a slope here, there is a field here and the field exist only in the depletion region. So, this region is your depletion region, actually that is your depletion region where there are no mobile carriers ok. You see the depletion region, the depletion region has a definite width and that is this W. I call it W, that extends from here to up to here this line; to this line this, is your depletion width. So, let me come to the next page, again I will draw the same thing.

(Refer Slide Time: 09:25)



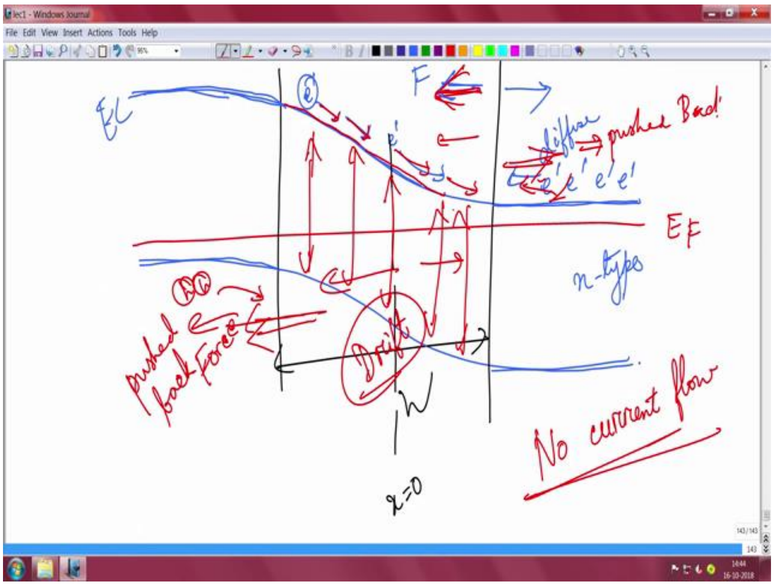
Now, I will keep drawing the conduction this energy band diagram because this is like a very important thing. So, you know this is here, this is here, suppose this was the x = 0 point the junction is formed. There is a field here, there is a depletion is extending up to here up to here. This is W the total depletion; some depletion is on the n side some depletion in as on the p side of course, this regions are neutral ok.

There is no field here and everything is neutral there, this is also neutral region only here it is as a field right; So, some depletion has extended to the n side, some depletion has extended to the p side. So, whatever has extended to the p side, this part has your negatively charged ionized acceptors. And whatever you have depleted on the n side, you have positively charged donor ions. So, essentially that is what is happened, this is a positively charged donor side and this is a negatively charged donor side.

Further electrons, further electrons from n side will be blocked by this field and further holes from the p side will also be blocked by the field. So, the field will prevent the hole from diffusing to the n side and prevent the electron from diffusing to the p side, which means and you know the field induced; the field related current is always called the drift current. So, essentially the drift component of the current, will precisely balance the diffusion component trying to go the other side here. So, drift and diffusion will cancel each other and the total current is 0 here.

Now, you might ask how is the field I mean which direction is the field, how is it blocking, how do you know the field is not blocking or blocking you know. So, again I will draw this band diagram because it is very easy to explain with the band diagram.

(Refer Slide Time: 11:10)



So, this is the Fermi level, again this is the p type, this is the n type, p type n type. So, let us focus on the depletion region which is say from here to up to here; this is the depletion region, this is your W of course this is your junction for example, x equal to 0. So, you know this is of course, this is flat. So, there is no field here, I keep telling you. Similarly there is a flat here, its flat and starts from here flat starts from here. So, there is no field here, the field only exists in the depletion region. Now I told you that look at this conduction band, this is conduction band right. So, look at this conduction band. If you think of it this as a cemented floor and if you place a rolling ball here, which direction will it roll, it will roll this side you know.

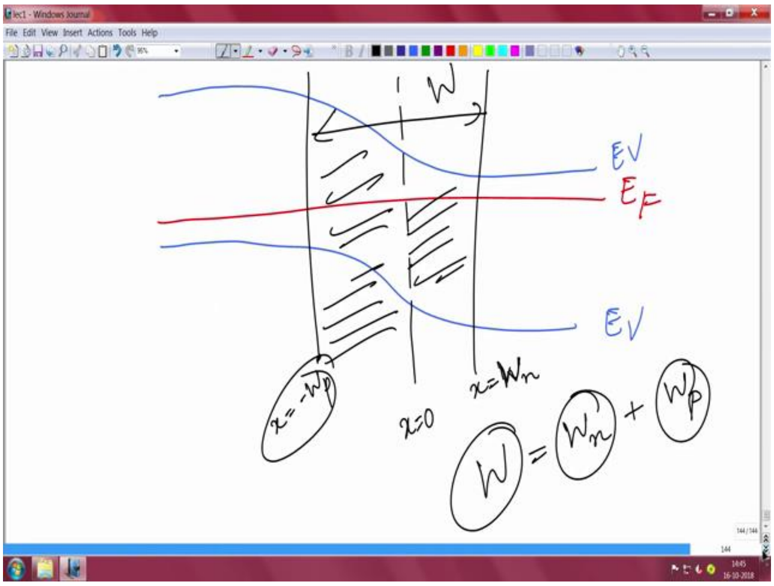
So, same thing; an electron here will basically try to move in this direction, which means the field is in this direction ok. If the field is in this direction, the force on electron will be in this direction because it is negatively charged. So, whatever electrons are here in the n type region, this is the n type region you know, there many many electrons here, they try to diffuse to the other side. They try to diffuse to the other side, but the field is in this direction, the field is in this direction, which means this electrons that are trying to diffuse will be pushed back in this direction right.

They will be pushed back, which means there will not be able to go this some on this barrier. Similarly, if the field is in this direction then the force on holes also will be in this direction because positively charged. So, whatever holes are here, if they try to diffuse on the other side of n, you know because its lower concentration, the field is in this direction. So, holes will be push back in this direction. So, holes also will not be able to actually go back to the other side go to the other side.

So, the drift that is happening in this depletion region, the drift that is happening in the depletion region, will exactly balance out the diffusion tendencies of electrons and holes from both side, the diffusion component ok. And that is why there will be no current that will flow, no current will eventually flow; it is all in equilibrium, it is all in equilibrium. That is why the field that is how you find out the field direction you just think of this is a slope as a cemented floor, put a ball here we will; which it will go in the direction which means exactly the electrons will also go in this direction which means the field is in this direction ok.

That is why you basically find out the direction of the field in a semiconductor and a field has to be a field will be only there when there is the slope in the conduction band or valence band. So, please keep that in mind, because the band gap is same everywhere, the slope is also same. But remember in certain semiconductor the band gap might keep changing with distance, those are hetro-structure sort of devices. If the band gap keep changing with distance, then the field on electron and hole might be also different; we will come to that in one of the far classes in the next you know sometime later.

(Refer Slide Time: 14:04)

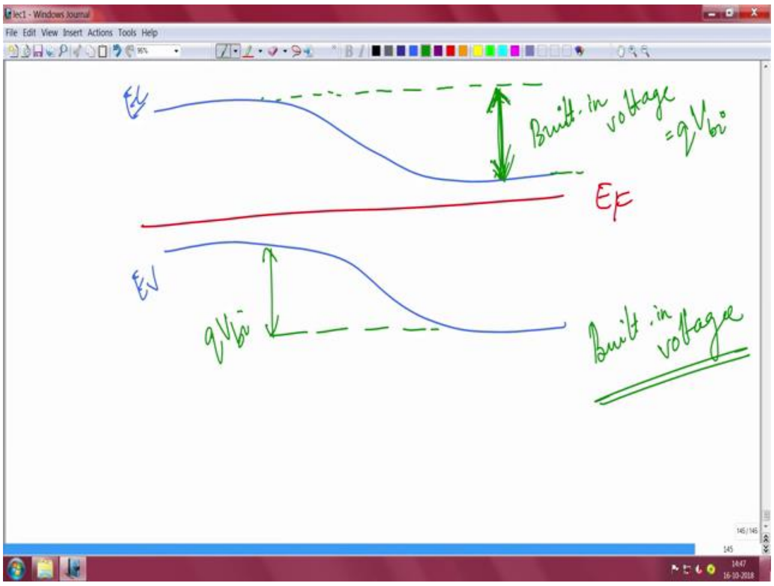


So, again I will draw this junction now, this is your Fermi level, this is your conduction band, this is your valence band, and this is a depletion that has formed. I told you this is suppose the 0, x = 0. So, the depletion has extended little bit on the n side. I call this distance as $x = W_n$; that is W_n is the depletion on the n side and on this side, I call this point as $x = -W_p$. So, depletion on the p side that has extended to the other side, this is x = 0. The total depletion is this that has formed, and total depletion is of course, equal to $W_n + W_p$ both are positive quantities that I am taking ok i.e.

$$W = W_n + W_p$$

We want to know how much is the depletion width, what is the depletion on the n side what is the depletion on the p side, those are very important terms that will actually decide how the device performs; they will decide how the device performs. That is also integrally connected with something else again I will draw it.

(Refer Slide Time: 15:05)

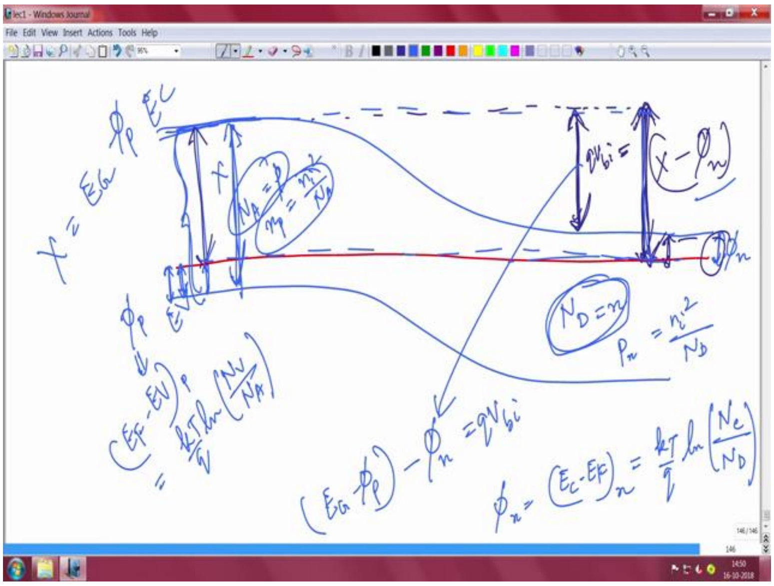


Suppose I have Fermi level here, I keep drawing this a that you know it becomes very very familiar all the time. So, it is band gap is same I am doing a mistake here. You see this is a conduction band here, if you look at the conduction band think again carefully you know I will take green color here for example, this conduction band and this side to the conduction band, they have a difference here in the energy.

This is called the built-in potential, built-in voltage or built-in potential same thing. We designate it as qV_{bi}. You see the conduction band on the both side and same thing with valence band of course, is the same thing away is the exactly the same thing because the same band gap. You see that that voltage that has built up because of that field, you know that that voltage is called built-in voltage you cannot measure it by the way you can only estimate or calculate it. You cannot measure it because, to measure you have to pass current, or you have to apply voltage; the moment you apply voltage or pass current, this equilibrium will be distorted and you cannot measure this again.

This is called built in voltage and we shall try to find out what is the value of this built in voltage and you know how is it important. Oh this built in voltage is very important in devices, extremely important.

(Refer Slide Time: 16:24)



So, how do you find out the built in voltage Fermi level keep drawing this conduction band, valence band, let me again not write this down because you know that this is conduction band this is valence band. So, this is your conduction band. So, this is the Fermi level and this is the conduction band on both side.

Now, this gap, this doping on n side is N_D and the doping on p side is N_A; this and you know N_D = n here because it is hundred percent ionize, this is equal to p here. You see this gap between conduction band and Fermi level, I will call that φ_n that is conduction band and Fermi level difference on the n side, this is the n side you know, it is called φ_n, I will call it φ_n generally. And this value is nothing but kT/q, you remember Maxwell Boltzmann approximation because this is not a high doping ln of N_C by n electron concentration which is N_D basically. n is only N_D here you know same thing.

Similarly this gap on the p side is phi p this is essentially your the Fermi level minus the valence band on the p side, on the p side and that value is called as is actually kT/q ln of N_V by p. Of course, instead of p, I can write N_A or the doping on the n side. Now by the way, if there is an electron concentration here of this side you know what is the minority hole here? That minority hole here, I will say p of n because on this side is n_i^2 by N_D.

That is the minority hole and it is very small in quantity, this is p here it is very high, electrons on this side, n p this is minority electron on p side will be n_i^2 by N_A, this is a very small quantity this is a very large quantity right. So, I know this is φ_p, I know this is φ_n, now if you look carefully again, this energy difference is actually if you recall very closely, this value qV_{bi} is actually equal to this hole value minus this value.

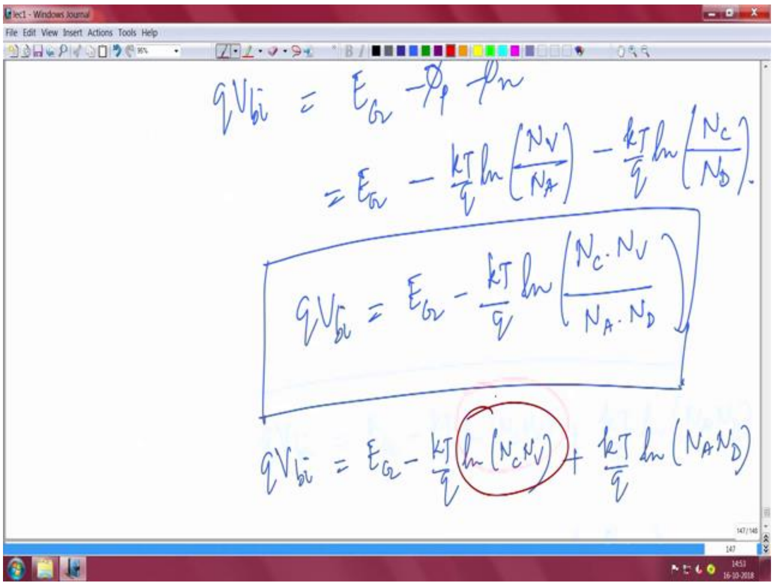
So, some value here which is say some unknown quantity x maybe, minus this φ_n that is equal to qV_{bi} this is this, right. This value is actually equal to this minus this all agreed yeah. This value by the way this value x that I am putting here, if you come this side, you will see that this is this value actually, of course, the band you know, the blue curve here is constant everywhere. So, this is actually x, this is, this is your x, right. You look at the Fermi level right, and this x is nothing but this whole thing which is the band gap, minus φ_n ; φ_p this the hole is the band gap.

Now, this entire thing is the band gap remember this is band gap conduction band this is valence band. So, this is whole thing is band gap. So, this is band gap minus this gap this φ_p, band gap minus φ_p is actually your x, this unknown quantity x. So, qV_{bi} is actually given by x which is nothing but band gap minus φ_p minus this φ_n, this is your qV_{bi}.

$qV_{bi} = E_G - \varphi_p - \varphi_n$

Let me go to the next page again.

(Refer Slide Time: 19:55)



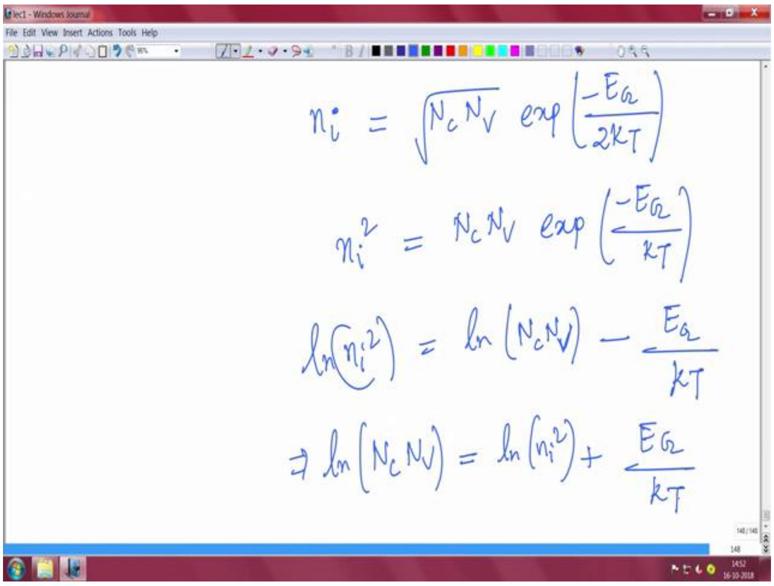
So, qV_{bi} is band gap minus φ_p that is the, this φ_p basically you recall is your this gap, and φ_n is your this gap in the neutral region right. So, φ_p - φ_n. E_G the band gap of the semiconductor minus φ_p is nothing but kT by q ln N_V by N_A remember from the last page this is minus kT by q ln N_C by N D i.e.

$$qV_{bi} = E_G - \frac{kT}{q}\ln(\frac{N_v}{N_A}) - \frac{kT}{q}\ln(\frac{N_C}{N_D})$$

So, I can write qV_{bi} equal to E_G minus kT by q ln N_C times N_V these are conduction band effective density of stage that are provided in the question always, divided by N_A into N_D. This N_A is the p side doping remember and this N_D is the n side doping remember that. So, this is your equation for your built-in potential; can you simplify it further we can actually simplify it further. I will tell you how. Let us simplify it even further.

$$qV_{bi} = E_G - \frac{kT}{q}\ln(\frac{N_C N_v}{N_D N_A})$$

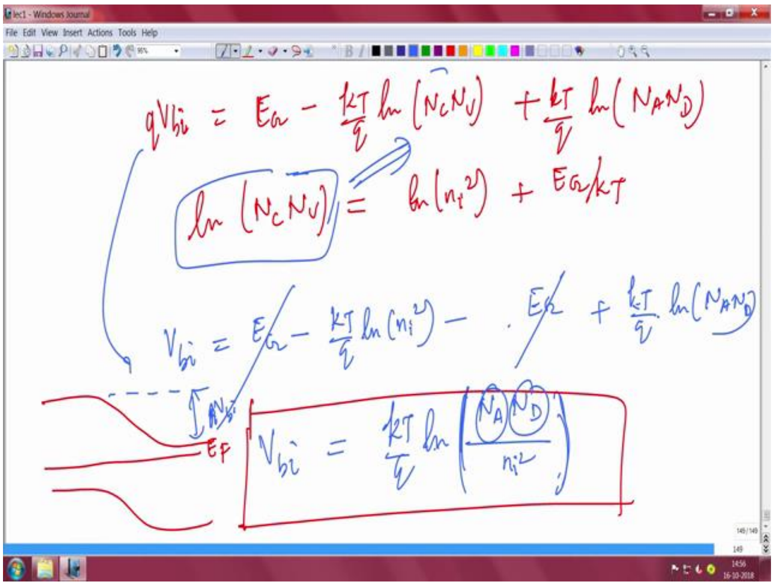
(Refer Slide Time: 21:04)



You remember the intrinsic carrier concentration of a semiconductor n_i is given by N_c N_v square root exponential of the band gap by 2kT. So, if I take a square here, n_i^2 becomes N_C N_V exponential minus E_G by 2kT, if I take a log here, ln of n_i^2 is equal to ln of $N_C N_V$ plus log of exponential which is minus this that will be minus E_G by this is actually not to put if this is kT because there is a square term kT. N_C N_V minus E_G by kT.

Shall we do this shall we put this equation? So, basically what it means is that ln of $N_C N_V$ is equal to ln of n_i^2 plus E_G by kT. Remember this equation, we will put it in the equation that you saw here. qV_{bi} equal to E_G minus kT by q log of N_C N_V plus because it is a minus minus here, kT by q log of N_A N_D. So, this quantity, that quantity we know different equation there right. You remember here, this is this is the equation right. So, I will write this the other previous equation and I will put it, I will put this equation there. What will happen then?

(Refer Slide Time: 22:51)



So, I will say qV_{bi}, again let me write it, is equal to E_G minus kT by q ln of N_C N_V plus k T by q ln of N_A N_D. And the second equation was that this ln of N_C N_V is equal to ln of n_i^2 plus E_G by k T. So, I just need to put this into this into this; This quantity is here. So, how do I do that? I will write this again qV_{bi} equal to E_G minus kT by q ln of this is this. So, ln of n_i^2 right, minus kT by q into E_G by kT, of course, this is kT by q only. I mean there is will be an extra q there.

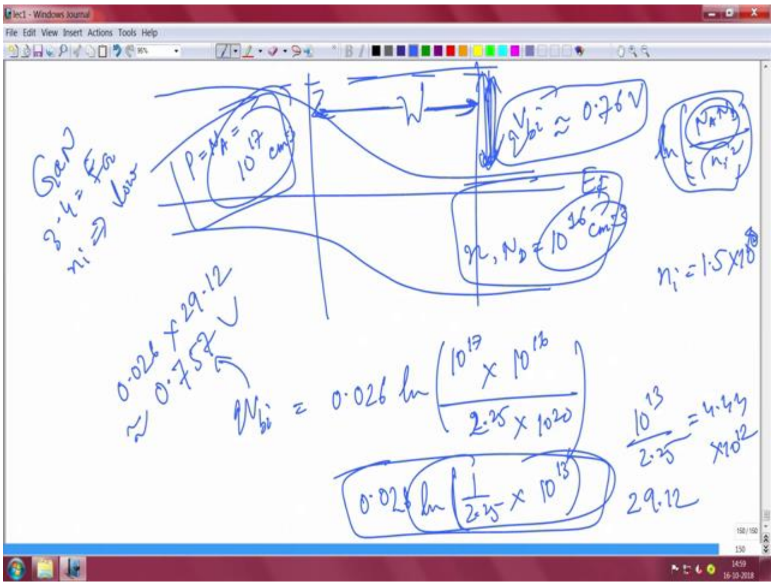
So, that of course, cancels out here and then plus kT by q ln N_A N_D. This this cancels out. So, V_{bi} actually this will be V_{bi} only because there is a q already here. So, V_{bi} will be equal to kT by q ln of N_A, doping on the n side, doping on the p side by n_i^2 i.e.

$$V_{bi} = \frac{kT}{q}\ln(\frac{N_A N_D}{{n_i}^2})$$

This is your expression for the built in potential; the built in potential that has formed in a p-n junction; you remember this is a Fermi level, this a conduction band, this is a valence band, this built in potential that has formed qV_{bi} is this actually. I mean I can remove the q here of course, this is a same thing.

So, it depends as kT by q ln N_A N_D by n_i^2. So, N_A is the doping on the p side N_D is the doping on the n side. So, the product of the doping divided by n_i^2, that log of that value times kT will give it a built-in voltage. So, for example, let us take an example here, we will take an example here.

(Refer Slide Time: 25:18)



So, suppose this is your Fermi level, this is the conduction band, suppose this is n type doped $10^{16}/cm^3$. This is p type doped say $10^{17}/cm^3$, then this built-in voltage will be given by kT which is 0.026 at room temperature ln N_A into N_D which is 10^{17} into 10^{16}, the doping by n_i^2. What is n_i for silicon if you remember its 1.5 X 10^{10}. So, this is 2.25 square will be 2.25 X 10^{20}. So, this will be 0.026 log of 1 by 2.2**5** into 10^{13}. So, you can just take log of this 10^{13} divided by 2.25. So, if I use the calculator here, I am going to use a calculator and I will quickly let you know what the value is.

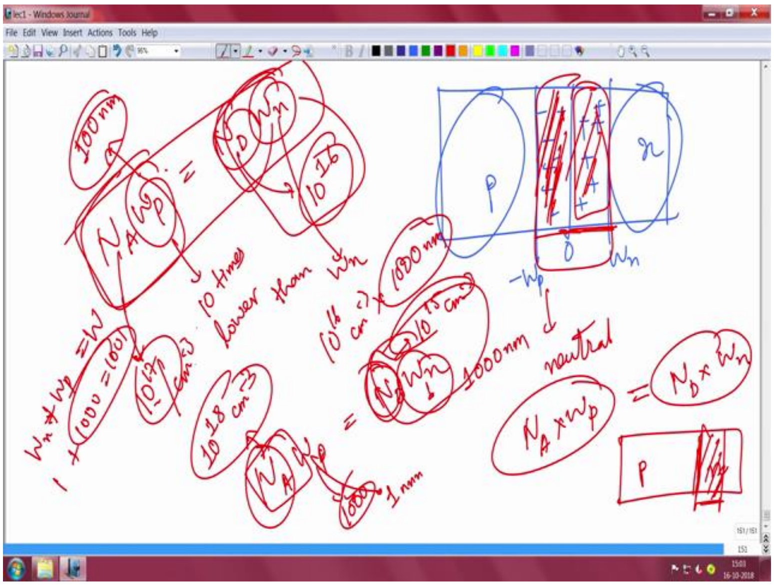
So, for example, I will you can do a calculator math actually, but the exact value will be how much? So, for example, I take 10^{13}. So, one you know, e to the power 13. So, that will be 10^{13}, that will be 10^{13} divided by 2.25 is equal to approximately 4.44 X 10^{12}. So, log of that quantity you know how much will the log of that quantity be; so, ln of 4.44 X 10^{12}. So, that is into 10^{12} will be equal to, this quantity will be equal to 29.12. So, the built-in potential will be kT, which is 0.0026 into 29.12 that will be almost equal to, how much, 0.757 volt.

So, you see if you have a doping of 10^{17} on one side, doping of 10^{16} on one side, the built-in potential will be almost equal to 0.76 volt. So, that is the voltage that is dropping across the p and n, this is the built-in field here, the built-in potential that is here and this potential cannot be measured, but this can be calculated. So, this built in potential of 0.76 volt has drop. You see this depends on the product of ln N_A N_D by n_i^2. So, if your doping on both side increases, then this quantity also increases you know, product and so, the built-in potential also increases as doping increases.

If on the other hand, if you take a narrow band gap semiconductor, your n_i will be very large, then a built-in potential will be very small. If you take a wide band gap semiconductor, then your if the if the band gap is large, say you know you take gallium nitride for example, the band gap is 3.4 eV right. So, your ni will be very very low, very very low; if ni is very low then this, your kT by q times this quantity, the built in voltage will be large.

So, built-in voltage will be large if your band gap is also large; so, a larger band gap essentially corresponds to a larger built-in voltage and a larger doping also corresponds to a larger built-in voltage. And built-in voltage is very integrally connected with this length of this depletion region. The length of this depletion region or this width you can say, is very integrally connected to the built-in voltage. So, if you recall actually the block diagram.

(Refer Slide Time: 29:09)



So, this is your p type n side. I told you the depletion extends on both sides little bit, this is the minus W_p this is 0, this is W_n. So, it has extended this side, extended this side, this is all negatively charged acceptor ions, this is all positively charged acceptor ions. The charge is neutral here I told you, the charge is neutral here I told you. The charge also has to be neutral in this, and for the charge to be neutral, the aerial density of this and this has to be equal. So, this is nothing but N_A^-. So, $N_A^-$is equal to N_A only times this W_p. This is the actually the area under this, you know like the aerial density. that has to be equal to this charge, which is N_D the doping that you are doing of course, into W_n this one right, that.

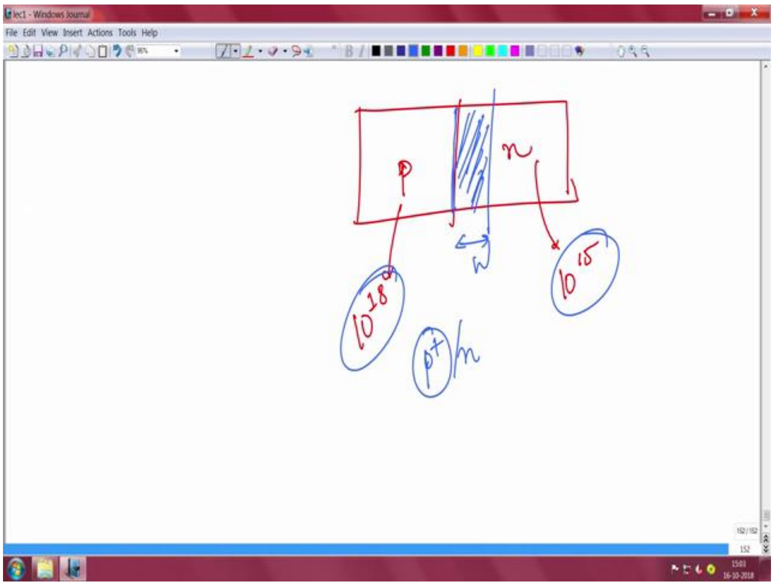
So, you see $N_A W_p$ must be equal to $N_D W_n$, the total charge on either side has to be same i.e.$N_A W_P = N_D W_n$

This means that if doping on one side is higher than the other, the depletion on that side is lower. So, suppose I am doping p type is 10^{17} and n type is 10^{16}. So, p type doping is 10 times higher and n type doping, which means W_p has to be 10 times lower than W_n. Only then the product will be same know, if W_p is suppose 100 nanometer, then $10^{17}/cm^3$ into 100 nanometer will be equal to $10^{16}/cm^3$, into thousand nanometer. So, then this will be 10 times larger. So, basically there in the inverse ratio; if the doping is high, then the depletion on that side is low, if the doping is low 10^{16}, the depletion on that side is high.

Only then the product will be same know. Which means if I take for example, N_AW_p equal to N_DW_n, if I take this as very high doping, say 10^{18} and suppose this is very low doping say 10^{15}, this is thousand times lower doping than this, which means this will be thousand times lower than this. So, if this quantity is say thousand nanometer, then this quantity W_p will be only one nanometer, which means in a realistic diagram when you draw, this is p this is n, you know this will be almost negligible, only one nanometer, on the other hand this will be thousand nanometer.

So, total depletion width will be $W_n + W_p = W$; 1 nanometer plus thousand nanometer is equal to thousand one nanometer, which is almost the same as thousand. What I am trying to say is the doping on one side is extremely high compared to the doping on the other side, then the entire depletion region extends to the region where the doping is low, the depletion on the doping and the region where the doping is high can be neglected when the doping ratio is very high.

(Refer Slide Time: 32:08)



So, to reemphasize and reiterate, you know. If I have a p n junction and this is doped very high, say 10^{18}, this is doped very low, say 10^{15}, then I can say that depletion is all the depletion is almost entirely in this side, there is negligible depletion on this side, because the depletion ratio is in inverse portion to the doping. So, this is the 1000 times more than this, which means this is thousand times lower thickness then this. So, all the depletion actually is on the n side. This kind of junction is call p plus n, p plus is very highly doped n is n. Now, it is reversed n is very highly doped when you call it n plus p junction, then you will call it as n plus p junction.

So, we will end the lecture here today ok. So, we have finished of the derivation of p n junction built in voltage, we have you know discussed the depletion width, how the depletion width actually splits in the ratio the inverse ratio of the doping, and how the depletion how the built in potential actually is proportional to the doping concentration the product of them, the band gap because, ni square also comes in there.

So, these things we have discussed; the field, how the field opposes, the band diagram, we have drawn many many times. So, now, hopefully you are aware of how the band diagram looks like, this is all assuming room temperature, 100 percent ionization of carriers and it is moderate doping. So, you know there is no Joyce-Dixon sort of thing here and finally, now in the next class, we will discuss about the relation between the depletion region and the built-in voltage. For that we will have to solve the Poisson equation. That we will do in the next class.

Thank you for your time and we will continue in the next class.