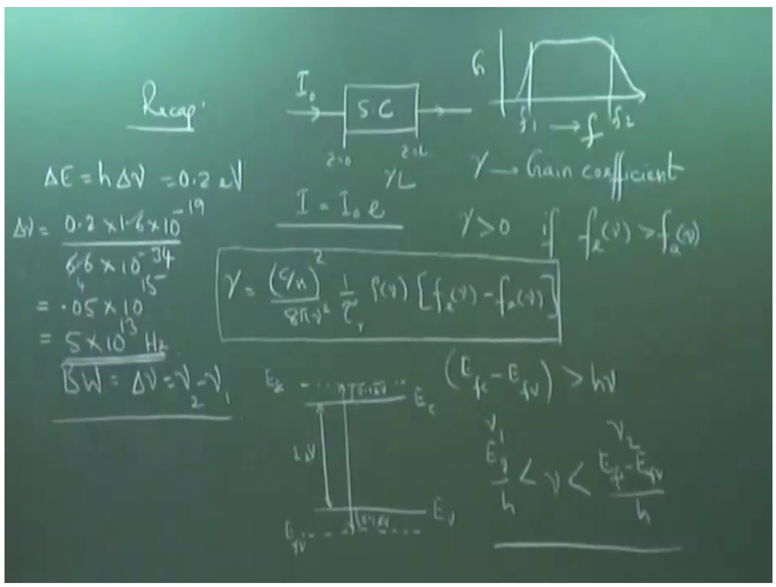
Semiconductor Optoelectronics
Prof. M. R. Shenoy
Department of Physics
India Institute of Technology - Delhi

Lecture - 21
The Semiconductor (Laser) Amplifier

In the last class, we saw the condition for Amplification by Stimulated Emission, and today we will take it further and discuss about Semiconductor Amplifier. I have written laser in brackets, normally semiconductor amplifier here refers to semiconductor laser amplifier, but amplifier itself the device we will discuss in detail a little later.

(Refer Slide Time: 00:52)



So recall if we pass a beam of radiation of intensity I0 at the input of a semiconductor, then if I0 is the intensity at z=0 then at z=L we have I=I0*e to the power gamma L, where gamma is the gain coefficient, gamma is >0 or gain if probability for emission is >probability of absorption. The gain coefficient gamma is given by c/n whole square square/8 pi nu square 1/tau r rho of nu*fe of nu-fa of nu, rho nu here is the optical joint density of states.

We can substitute this here so gamma is >0 if this is positive, if this is negative then gamma is < 0 and we will have absorption coefficient and we will see the absorption spectrum a little later, today so this is the expression for gain coefficient. Today, we want to know and we also have seen that this is positive if Efc-Efv is>h nu for all frequencies for which Efc-Efv >h nu we have gain or amplification.

And from a simple band diagram here so if Ec is here Ev, then if Efc and Efv happens to be in the bands Efc and Efv then for all frequencies which correspond to the range between this and this here that is for frequency for which Eg/h<nu< Efc-Efv/h we have amplification, there is amplification for all frequencies in this band. So this determines the amplification bandwidth, so if I call this frequency as nu 1 and this frequency as nu 2.

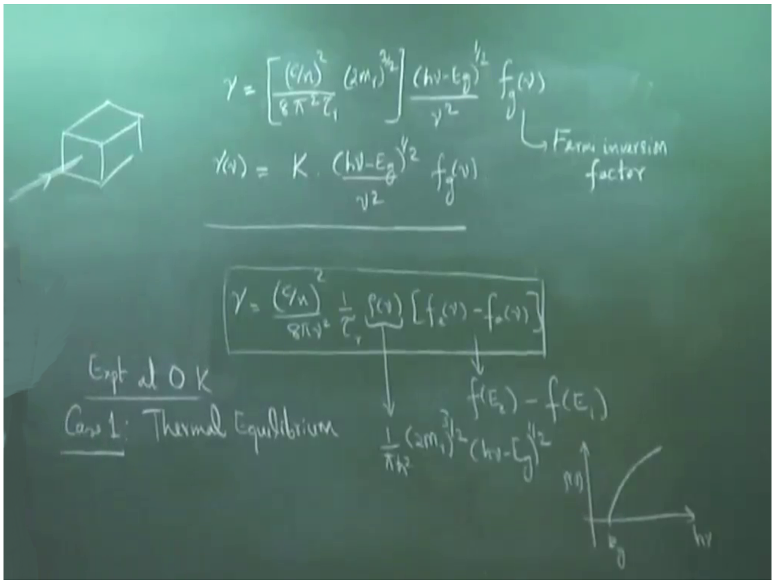
Then we know that the amplification bandwidth, so bandwidth is =delta nu=nu 2-nu 1, typically if this gap is say about 1.4 or 1.35 if you take a indium gallium arsenide phosphide amplifier or any amplifier, if this Eg is let us say 1 Ev, and if I say that this separation here is just to get an idea 0.1 Ev and this is also 0.1 Ev, then what would be the bandwidth? Bandwidth will correspond to delta E an additional delta E of 0.2.

So delta E=h*delta nu=0.2 Ev, the energy difference here corresponds to a frequency range and that is here h*delta nu=0.2 Ev, so delta nu here =0.2 so I have to convert it into joules because h is in joules so 10 to the power of-19/6.6*10 to the power of-34, so you see how much this will be if this is approximately 4 times and this is 0.2 so 0.05 so 0.05*this is 10 power 15, so the bandwidth here approximately 5*10 to the power of 13 hertz.

Indeed, semiconductor amplifiers have a bandwidth which is of the order of 10 to the power of 13 hertz. What we have got is an expression for gain coefficient, and an expression for bandwidth, what we would also like to know is the gain profile, how is the frequency if you take any amplifiers normal electronic amplifiers you would like to know the frequency response you generally plot f versus gain and maybe the amplifier has a gain curve like this.

And you have the 2 cutoff frequencies here f1 and f2 and delta is this, the gain profile is also very important there are application where you need very flat gain profiles. So we would like to see what is the gain profile of this amplifier? So we have got this number, cutoff frequency but we want to see the gain profile. So let us see the gain profile, so how to get the gain profile we have to know the variation of gain with frequency.

(Refer Slide Time: 07:37)



To begin with we carry out a thought experiment, a thought experiment is this an experiment at 0 K, why we are using 0 K you can guess because there are fermi functions here fe of nu and fa of nu, fe of nu and fa of nu this is =f of E2-f of E1, we already substituted you can substitute for the functions here and you get f of E2-f of E1, so we want to perform this experiment here at 0 K. Now at 0 K consider case 1 a semiconductor in thermal equilibrium.

We know that there will be no gain but let us see what we get, so thermal equilibrium we have already seen that quasi fermi levels the separation between quasi fermi levels have to be greater and we cannot achieve that at thermal equilibrium, but let us see what would be the profile. So what we now have is we have a semiconductor, so let us say this is a piece of semiconductor and incident radiation is passing through this.

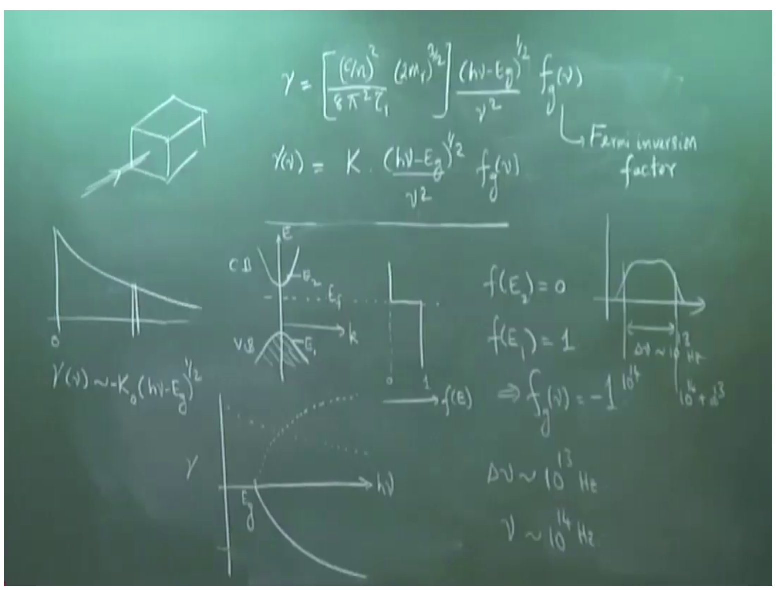
Let me rewrite this expression gamma=, because I want to substitute for rho of nu this is the optical joint density of states we have derived this expression one by1/h cross square twice mr to the power 3/2*h nu-Eg to the power 1/2, recall the density of states and we had a plot if you remember h nu and the from Eg was going like this, because h nu-Eg to the power 1/2 is the variation, so this is Eg and this is rho of nu, so substitute this expression here.

So I have c/n square 8 pi nu square there is 1 pi coming from here so 8 pi square and h cross square so c/n/8 pi h cross square, nu square is there I want to keep nu square outside because I am interested in finding the frequency dependence of gamma so nu do not want to get into this, so this pi h cross, let me write as it is let me not combine, so c/n whole square 8 pi square 1 pi I have taken there*tau r radiative recombination lifetime*2 mr to the power 3/2*this term is there.

So *h nu-Eg to the power /2 and there is a frequency dependence here nu square this term nu square, because this is independent of frequency this part so h nu-Eg to the power 1/2 by square*this term that is fe of nu-fa of nu we denote it as fg of nu where this is called the fermi inversion factor, why we will see why it is called inversion factor in a minute. So this difference I am denoting as fg of nu, this is a constant gamma of nu here.

So gamma of nu=some constant K*h nu-Eg to the power 1/2 by nu square*fg of nu. Let us look at the e-k diagram of the semiconductor, so thermal equilibrium a thought experiment at 0 absolute 0.

(Refer Slide Time: 12:49)



If we see the e-k diagram let us see the fermi level is somewhere here, I do not know somewhere there this is k this is=E, Ef is somewhere here, at 0 K the fermi distribution is given by a step function. So what I have now plotted here is 0, this is 1, and this is f of E, which means the conduction band is completely full all levels below the fermi level are completely full, I am sorry in the valence band is completely full so this is valence band this is conduction band and conduction band is completely empty.

So if you take any pair of states that is a value of energy E2 here, and a value of energy E1 here, f of E2=0 and f of E1=1 fermi function here behaves like a step function at 0 K, so f of E2-f of E1 so fg of nu this implies fg of nu=-1, so this factor is-1 at 0 K a semiconductor in thermal equilibrium the fermi function is somewhere here, you can take wherever you want the fermi function you can take the degenerate semiconductor also and can see that you will get fg of nu=-1.

And therefore the gain coefficient here gamma=-K*this, so how would this look like. So let us plot nu what is nu? nu is the frequency of radiation we have seen that the typical bandwidth is delta nu is of the order of 10 to the power of 13 hertz, and nu is of the order of 10 to the power of 14 hertz the frequency 2*10 power 14, 3*10 to the power 14, 4*10 to the power 14 this is the kind of number that we have. Therefore, our interest is to find the amplification response.

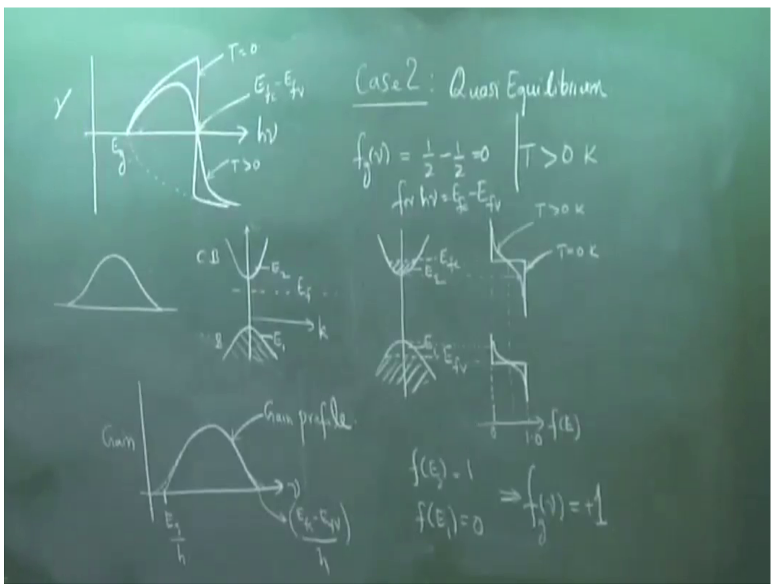
So the bandwidth here is approximately in our case delta nu is of the order of 10 to the power of 13 hertz, so in the range of interest here the absolute frequency itself varies a little that means for example you see this, this is 10 to the power of 14 hertz means this and this is 10 to the power of 14+10 to the power of 13 a small number, if this is 1 this is 1.1, so the variation of nu square over the interval that is actually if I plot nu square that will have a variation it is actually 1/x square please see it is 1/x square.

However, the range of interest where x varies very little and therefore, I can either assume it has almost a constant over the range of interest this is 1/nu square variation is almost constant, because 1/x square so starting from 0 if you take 1/x square it drops like this as you know 1/x square. However, we are considering a small range this is if you take from 0, our frequencies are just 10 to the power of 14+-10 to the power of 13, which means we are looking at a small variation here.

Therefore, 1/nu square varies very little and therefore, the variation is primarily determined by this, so at 0 K we have gamma of nu=a constant K 0*fg of nu is-1 so -K 0*h nu-Eg to the power 1/2, please see that 1/nu square dependence is very small because our range of interest is very small, if you wish you can keep and plot then also it does not make much difference, this starts from Eg, this is h nu I am plotting the gain coefficient gamma fg, let me plot, because it is negative I need to plot negative as well.

So it starts from here Eg and this would h nu-Eg to the power 1/2 would vary like this, therefore, -nu-Eg to the power 1/2 would vary like this, this is gamma this gamma is 0, gamma positive, gamma negative. We already know this at thermal equilibrium gamma is negative, because fg of nu is -1, I want to keep this graph here and let us see now let us go over to quasi equilibrium.

(Refer Slide Time: 19:16)



So case 2 quasi equilibrium, the same e-k diagram let me draw in parallel, but now we have pumped this semiconductor and assume that Efc is already in the band Efc and Efv, because we know that we get gain only if the difference is >h nu therefore, Efc is here, Efv is here. What does this mean Efc therefore, the variation here would be, and the variation here would be, so on this axis I have plotted this is 1, 1.0, this end is 0.

Corresponding to 2 fermi functions which describe the occupation probability of the 2 bands at 0 K the valence band is completely full up to Efv and the conduction band is full up to Efc.

Therefore, if you now consider a level E2 here and a level E1 here, note that f of E2=1 and f of E1=0 implies fg of nu=+1, so you see the fermi factor has been inverted earlier it was -1 at thermal equilibrium, and in quasi equilibrium fg of nu is +1.

So what we expect the gain profile to be now this one, so the gain now in this case the gain goes up like this, how far will it go this is h nu the photon energy, so let me just for differentiation let me show this like this, so this is for case 1, this is for case 2. How far will it go h nu is the photon energy increasing from Eg, it is increasing from this gap this gap is Eg so it is in increasing, when it comes to Efc and Efv are beyond this, this is 0 and this is 1.

So the factor is inverted at Efc-Efv, so along this line at Efc-Efv at the point Efc-Efv the factors suddenly becomes from +1 to -1, when energy h nu comes here that it becomes more than this difference when it comes here then this is 0 this is 1 and therefore, the factor in reverse and therefore, the gain profile drops down like this and continues on this line. So let redraw this here so h nu beyond Eg it starts increasing as h nu-Eg to the power 1/2 drops down this and continues here.

So this value here is Eg and this value is Efc-Efv, what is this on the axis is gain profile gamma, so gamma versus h nu, so this is the frequency band in which amplification takes place and the gain profile looks like this, this is at T =0. What will happen for T >0 in normal temperature, first thing you see is the fermi function gets smeared, so the fermi function does this so it 0.5 it has to cross at 0.5, so this is the fermi function this is for T=0 K, and this is T >0 K this also smears so we have this variation.

Whatever be the temperature at Efc and at Efv the probability is ½, so it is 0.5 here and 0.5 here, which means f of Efc-f of Efv is 1/2-1/2 that is 0, so fg of nu =1/2-1/2=0 for h nu=Efc-Efv, fg=0, fg=0 means what? Where is the gain expression? So it was multiplied by fg of nu here fg of nu, fg of nu=0 implies gain=0. And therefore, first point is at Efc-Efv gain is 0 even T>0 the gain is 0. Second, if I take a frequency close to this that is close to Eg, which means close to the band gap here, you see f of E2 is here.

I have actually shown like this but in practice as you know that the smearing is very little or f of Efc is close to 1 but <1 it is close to 1 but <1, and f of Efv here that is if I take at Eg so please see this h nu corresponds to Eg means E2 is here, so the corresponding of f of E1 is close to 1 but <1, corresponding to E1 here f of E1 is close to 0 but >0, so just if I say that if this is 0.99 and this is 0.01 the difference is 0.098, as you increase this becomes smaller and this becomes lesser and lesser it is going towards 0.5 and this is also going towards 0.5.

Therefore, the difference drops down from 0.1 to 0.0, to 0.1 to 0 which means here near Eg it is almost the same but the difference increases and finally, here it comes down to 0, let me draw this and then you, this is for T=0 this is for T>0, the difference near Eg the fermi for inversion factor fg is close to 1 but <1, and therefore, the multiplying factor is slightly <1, so you see it starts but as h nu increases the multiplying factor becomes smaller and smaller.

And therefore, the difference between T=0 and T>0 changes, and we know that whatever be the temperature the 0 crossing is this, so if you actually put the numbers and see that you will get a gain curve which is like this, this is the gain profile. So if I want to independently plot the gain profile, then so if I plot me then you will have then gain taking a shape something like this, so this is gain profile, let us say a room temperature T>0,

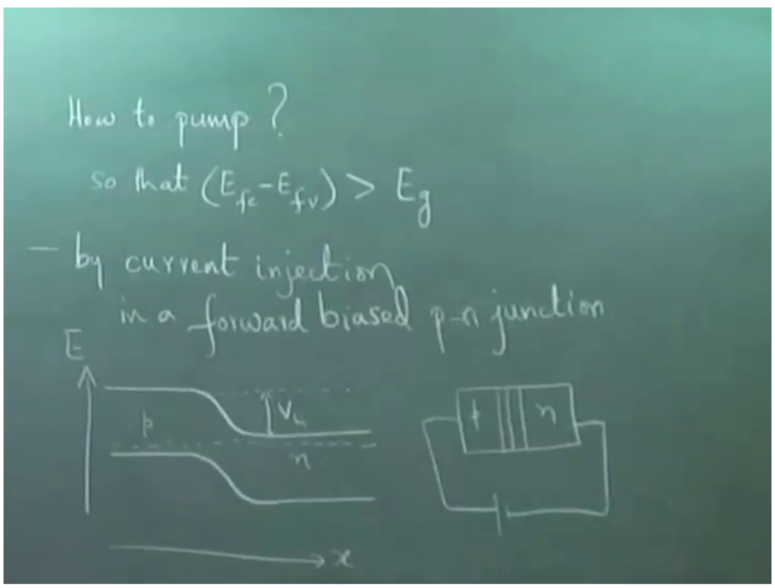
This value here is if this is frequency like you typically plot the frequency response, then this value here is Eg/h and this value here is Efc-Efv/h, I have not shown beyond that because beyond that it is negative, I am just plotting the gain profile, and typical bandwidth is 10 to the power of 13 hertz, these are practical numbers of semiconductor optical amplifiers. In theory it looks as if the gain starts right at Eg and at Efc it as if is it starting here and likes 0, but in practice it is slightly different.

The shape is the same but the ends are slightly different, because of band tail states these are highly doped semiconductor, and therefore, there are band tail states and this tapers down in practice like this. And similarly, it tapers to this, it is not an abrupt band this tapers down here, and it looks like it is not flat the first thing that you see that semiconductor optical amplifiers do not have a flat again profile.

And this is one of the reasons why in all WDM communications we do not use semiconductor amplifiers but we use erbium doped fiber amplifiers which have a flat gain profile. But the bandwidth in this case is quite large. So we know why this shape of gain profile that we have. So the summary is that if you pump a semiconductor and maintain the difference between the quasi fermi levels > the frequency h nu at which you need gain, then it is possible to have gain.

How do we pump the semiconductor? how to maintain so the next question would obviously be how to have Efc-Efv >Eg, it is not going to be easy, we can have quasi fermi levels all p-n diodes when you forward bias you have quasi fermi levels, but you are asking for too much that is Efc-Efv>Eg.

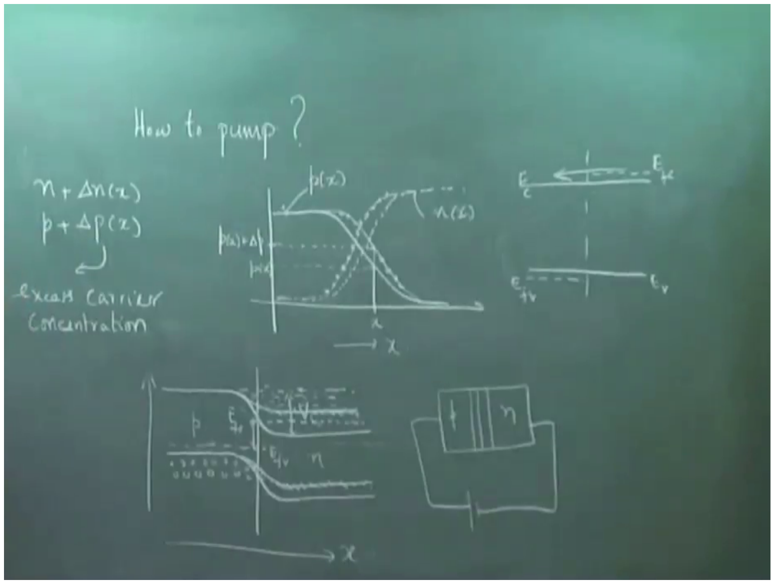
(Refer Slide Time: 32:36)



So how to pump? Any amplifier requires pump how to pump, so that the question is incomplete there how to pump so that Efc-Efv is >Eg >h nu for the frequency nu, but minimum is Eg therefore, I have written how to make it >Eg. The easiest way is by current injection I am writing the answer first and then we will discuss, by current injection in a forward biased p-n junction, so what do you mean by this? So you take a p-n junction we already drawn the band diagram.

So I will directly draw the band diagram, so what we have is this the p side, this is the n side, and we have fermi level here this is axis energy this is the distance x before pumping, and there is a built-in voltage built in potential so this is V built-in, E*V built-in is the energy. So when you forward biased the diode in our basic elementary picture, so we have forward bias so this is p, this is n, originally there was a depletion layer here and you forward bias.

(Refer Slide Time: 35:03)



I can also plot the carrier profile across the junction, very quickly call off the basics that we have already studied. So if you plot the carrier concentration, so this is p and this is n, so p,n what I have plotted is hole concentration p and this is n, and this is the same x across the junction, so on this side there are very little electron this is not fermi level, this is the concentration n of x so what I have plotted is n of x and p of x across the junction.

On the p side you have large number of holes, and on the n side you have, what happens when you pump? When you pump forward bias this goes up we already discussed that this goes up, which means the band now comes let us say here, and this band here this is the new position, let me draw them with the solid line itself and differentiate by putting crosses, because so let me draw in this fashion just to differentiate the second case that is after forward biasing.

The fermi level has to separate out because the fermi level has to remain, so the fermi level this one is here and the next one has moved, so Efc and Efv so what we have is Efv and Efc, the fermi levels have separated out good, you pump harder this going to a further up. How much will it go? When you pump harder you see this is full of electrons here, electrons are coming here so when you forward bias the career profile now becomes like this.

So far away from the junction there is very little change, but at the junction now the career profile has changed, please see this was before biasing how do I show okay let me put dots just to differentiate this is after forward biasing, and p profile has also moved like this so this is after forward bias. So what has happened is at a given value of x you see that both n of x and p of x have increased that is what you see here.

If I had taken this value of x here then earlier the number of electrons here was much less, but now we have large number of electrons. Similarly, holes which were here have also moved to this side now, and that is why this has also moved to that side. This was n of x please see at any value of x if I take this as the x, then this value here is p of x, after forward biasing this is so this is p of x, after forward biasing this is the new p of x, which means this value now represents p of x+ delta p of x.

Similarly, you have n of x and delta n of x, the delta n of x we are looking at the junction region, please remember we are looking at the junction region because the changes occur in the junction reaction because of forward biasing delta n of x and delta p of x we have got in the junction region delta n of x and delta p of x, originally it was just n before forward biasing, now it is this, this delta n of x and delta p of x are called excess carrier concentration.

So this is excess carrier concentration, n and p were called carrier concentration, delta n of x and delta p of x is called excess carrier concentration, and the point that we see is the fermi level has separated out. If you pump harder the level goes further up, at best the 2 levels let us say got equal is I have lifted this, the 2 levels have become equal this is the p-n junction original junction, now the level has raised, where is the fermi function for this is here, where is the function for this far away from the junction please see this.

This is the fermi function of the p side, this was the fermi function of the n side, so it is still <Ec that is this is now Efc, this is now Efv therefore, and this is Ec and this is Ev, what do you see

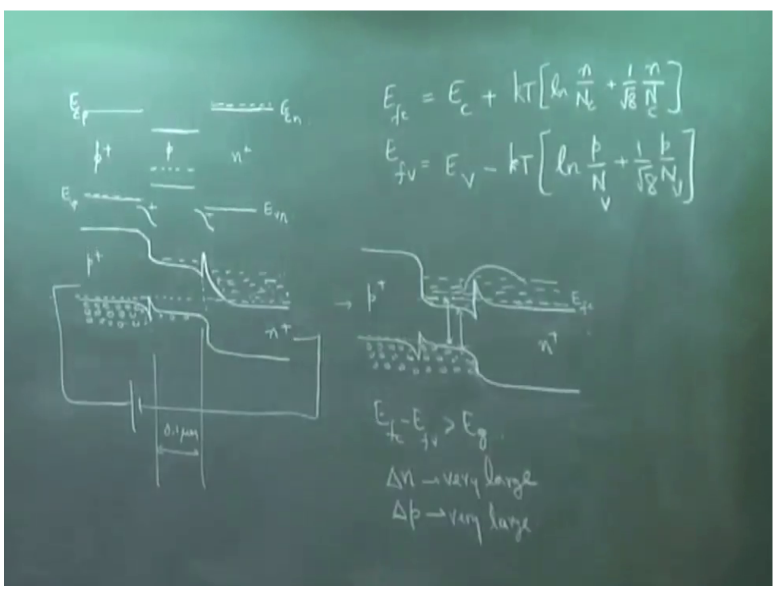
Efc-Efv is still <Eg, this is I said this is at best why am I saying this is at best, by the time you forward bias it is so much the current through the device is so high that the junction will burn up by this time, there is no barrier any electron injected here is simply going through the device.

This is very, very high forward biasing and the junction will simply burn out, yet you have not reached the condition Efc-Efv >Eg. So how to achieve this condition Efc-Efv>Eg, can we think of something? Yes, one of the ways is we start with highly doped p and n, so that degenerate p side and degenerate n side, if you start with the degenerate p semiconductor which means Ef is already inside here, and Ef is already into the conduction band.

In that case I can have Ef sitting here and Efc sitting there, please see if I start with a degenerate semiconductor there is a possibility of making Efc-Efv >Eg so Efv-Efc>Eg. So first point is we have to use it is necessary to use highly doped p and n materials to realize a p-n junction for a source which can act as an amplifier, even this is not a very practical solution, it is alright, in theory it is possible.

But as I said to achieve this kind of forward bias means the current through the device is so high that it will damage the junction, it more practical and correct way of getting this is by the use of double heterostructures, and which brought in as I said Nobel Prize for that discovery.

(Refer Slide Time: 43:34)



So use of double heterostructures that is why we discussed double heterostructures in detail and trying to draw the band diagram of double heterostructures, yes the method is by current injection but in double heterojunction double heterostructures p-n junctions, which form double heterojunction. So we recall the band diagram very quickly and now you see what happens that leads to carrier confinement, and the separation between Efc and Efv can be easily >Eg.

So recall the technique that we have for drawing the energy band diagram, so we have a double heterostructures with a low band gap material sandwiched between 2 high bandgap materials, highly doped p structure so let me draw the p here or inside the band okay. Let me for simplicity let me draw it here that is alright. And then we have another structure here which could be slightly p doped or n doped or intrinsic, and I take a highway doped n side here.

So this is n, n+, this is p, this is p+, this could be aluminium gallium arsenide, gallium arsenide and aluminium gallium arsenide. So we join the let us take the band diagram, so one of the as I said you could start anywhere, let me take the middle one here and show the fermi energy here. And then we know that before forward biasing the fermi energy should remain constant, and it is p to p+ therefore, the band will start bending like this here.

So the band show let us so the band will bend here, and then there is a discontinuity the fermi function is here therefore, E so this is Ec of the p side, this is Ec of the n side, Ev of the p side and Ev of the n side. So very quickly so the band starts bending like this, then it meets a discontinuity so I have a discontinuity and then the band continues, so the band continues. When it reaches this junction, this is the p-n junction.

So the band has to so here also we have + a potential energy variation like this, here also we have a potential energy variation like this, so the band starts bending further but it meets an upward discontinuity, so we take this up this discontinuity here. And then this continues further and because finally the fermi level has to be here. Similarly, this end there is this end starts from here, comes starts bending but there is an upward discontinuity here.

So this bends upward and then continues on this, comes here this has the downward discontinuity, so here there is a downward discontinuity this discontinuity downward discontinuity, and then we have the band continuity. I have shown a little actually the angle should be more because it is n+ so rapid drop, this is the advantage of chalk and board I can wipe whatever mistakes I make.

So this is the structure recall that this is there are plenty of electrons here, because this is a degenerate semiconductor, plenty of holes here, I just have to draw 1 diagram of forward biasing and then we will stop more of discussion we will do later, is the diagram alright? Clear? So this is the p side p+, this is the n+ side. We have not forward biased yet, so we are forward biasing now, so we apply positive voltage to this end and negative to this end.

As usual this will start going up so this starts going up, so let me draw the diagram separately here rather than showing it there so very quickly, so this discontinuity then we have this then we have this and then I am now showing the forward biased right, so originally I had this here. Similarly, from here I have this an upward discontinuity continuing like this, and then from here there is a downward discontinuity and let me show like this.

When I forward bias this end is raised, and therefore, we have the new diagram which is coming, so everywhere this portion gets raised here, and this comes here, so originally it was here now this has come to this. And therefore, the fermi level here or fermi level was inside the band, so the fermi level is here, so Efc do not worry you just have to see I have raised this and therefore, Efc is here Efv was here.

So Efv so this also came up now this got raised this got raised discontinuous because the band gap has to remain the same, and this was here. The layer which is here that is the sandwiched layer is of thickness approximately 0.1 micrometer, it is <the junction width in normal p-n junctions, and you have electrons because of forward biasing electrons completely filling here, because this has been raised, and because this level went up you have holes after forward biasing.

What has happened look at the junction region here, this was the p+ region, and this is the n+ region, this is the junction region. In the junction region Efc is already here, why Efc is here? Because the number of electrons are so large that the fermi function has moved here into the band, and number of holes are so large that here the fermi function has come into the band, now this difference you see this is the difference which is >Eg of the sandwiched layer.

The difference between Efc and Efv is >Eg of this layer, the layer which was sandwiched the low bandgap material. Second, point first therefore, you can clearly see that Efc-Efv can be >Eg, second point the carriers the number of carriers here number of carriers in a small volume is so large that the excess carrier concentration delta n and delta p become very large, we will put some numbers and quantify this what is this very large become extremely large.

Because these numbers become very large you know that you recall Efc and Efv by Joyce-Dixon approximation you recall this Efv=Ev-kT*ln p/Nv+1/square root of 8*p/Nv this is Joyce-Dixon formula which we have discussed earlier, the empirical formula ln n/Nc+ square root of 1/8*n/Nc. The carrier concentration n and p have become very large because of the pumping, this p is the original p+ delta p, delta p is due to current injection.

So this p has become very large p and n. And therefore, Efc and Efv, Efc becomes >Ec, and Efv becomes <Ec, because this is now a larger number a positive number and you can clearly see that Efc-Efv becomes >Eg. So Efc-Efv becomes >Ec-Ev, because both the quantities here are positive large quantities, if this small n is <Nc then this part will become negative log of negative fraction, but now when n has become very large it becomes positive.

We will put some numbers or you will put some numbers and see, but this is why double heterostructures are used in all laser devices, to achieve Efc-Efv >Eg in a practical way by passing very small currents, as we put numbers we will see that just by passing milliamperes, we do not have to pass 100s of milliamperes through the p-n junction or fear of burning the junction. So we will stop here and continue in the next class.