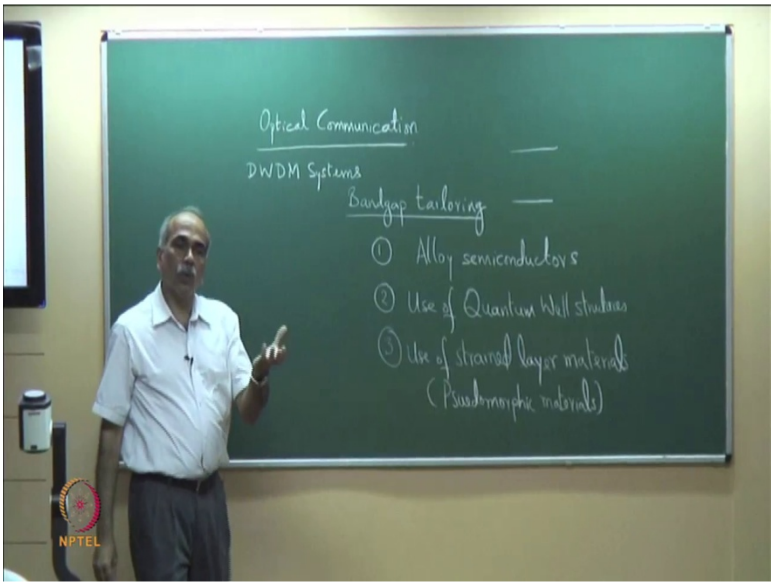
Semiconductor Optoelectronics
Prof. M. R. Shenoy
Department of Physics
Indian Institute of Technology, Delhi

Lecture - 13
Band gap Engineering

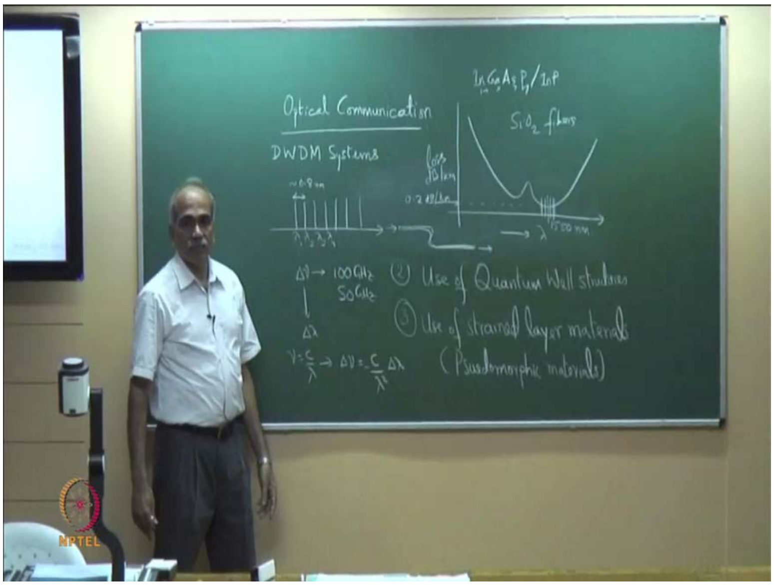
(Refer Slide Time: 00:38)



Let us continue with the lectures, and today we will discuss about band gap engineering. Band gap engineering refers to or it is also called sometimes as band gap tailoring band gap tailoring refers to the various techniques which are employed to achieve band gap modification. We have seen some of them. Some of the techniques which are used are; first we have already seen by alloying or use of alloy semiconductors use of alloy semiconductors. By alloying taking different compositions you can realize materials which have different band gaps.

The second technique which is used is use of quantum well structures, use of strained layer strain layer materials. These are sometimes referred to as Psuedomorphic materials Psuedomorphic materials Psuedomorphic materials. We have been discussing about quantum wells in the last class; we will continue the discussion in this context of band gap engineering or band gap tailoring. Why do we need band gap tailoring or band gap engineering, because we need to have materials with different band gaps. For example, for example, with respect respect to optical communication optical communication in the context of optical communication. Today we use D W D M systems D W D M systems which means dense wavelength division multiplexed systems; a large number of wavelengths in the low loss window of optical fiber.

(Refer Slide Time: 03:41)



We use a large number of wavelengths in the low loss window of the optical fiber; so those of you are not familiar. Todays communication system, optical communication system uses W D M or D W D M system. So, if you see wavelength versus loss in D B per kilometer is a standard graph which we see; fiber shows a low loss window. So, this is loss; low loss window centered around 1550 nanometer 1550 nanometer. The loss here is approximately 0.2 D V per kilometer; this loss here is 0.2 D V per kilometer; silica fiber. What I have plotted is the attenuation spectrum of silica fiber; silica based SiO2 fibers, optical fibers which are used for communication.

So, in this low loss window one can put large number of wavelengths in the vicinity here, in this within this window. So, if you zoom this a little bit then you can have around 1550 you can have large number of wavelengths; there are standard communication grid wavelengths, so lambda 1 lambda 2 lambda 3 lambda 4 and so on. These are separated; the channel separation between these is typically they depend on what is the channel separation which is used. In communication technology, the channel separation delta nu is the standard ones are 100 gigahertz or 50 gigahertz. So, corresponding to delta nu, you can calculate what is delta lambda. Because nu is equal to

C by lambda; therefore, delta nu this implies delta nu is equal to C divided by lambda square into delta lambda. So, around 1550 window for every delta nu you can find out what is delta lambda with a negative sign of course which means 100 gigahertz spacing here corresponds to approximately; please verify this, 0.8 nanometer separation.

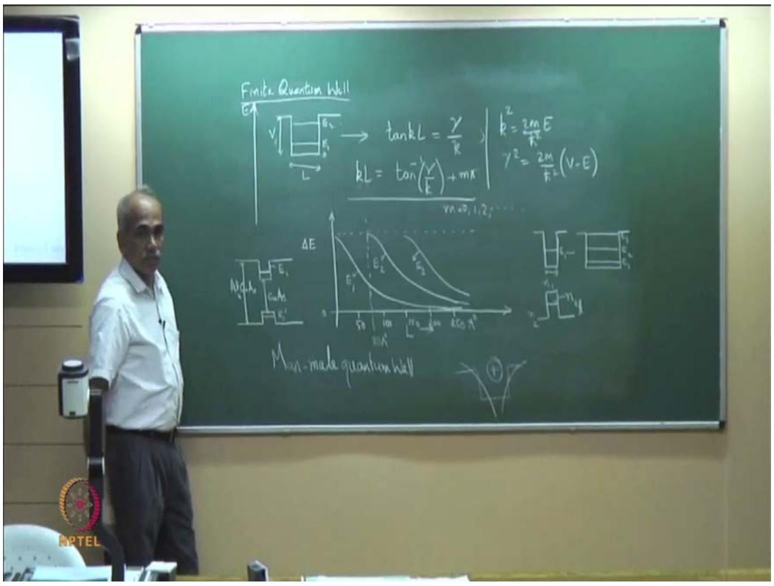
So, in this low loss window around 1550 you can insert large number of wavelengths with a separation of approximately 0.8 nanometers, in the same fiber. Which means one fiber one fiber, so the optical fiber here carries all these wavelengths at the input. The advantage is every wavelength has a certain channel capacity and you have large number of channels. So, the wavelengths are very closed; this 0.8 nanometer separation is quite closed and therefore, the word dense wavelength division multiplexing. Normally, one users W D M wavelength division multiplexing or frequency division multiplexing. Here it is the additional D is to say dense wavelength division multiplexing; very closely spaced channels.

To realize such a system you also need sources which is a laser, which is used as an optical source; you need sources which are very closely spaced. One can realize these lasers using the same material system, but while going for band gap tailoring a little bit of tailoring of the band gap can result in wavelength emission at very closely spaced lines. So, wavelength emissions which are very closely separated; so you do not have to change the material system. For example, in communication one generally uses the Indium Gallium Arsenide Phosphate Indium Gallium Arsenate Phosphate or an Indium Phosphate substrate. So, this is Gallium x Arsenic y Indium 1 minus x and P 1 minus y.

So, this is the material system which is used to realize these lasers in the 1550 window. Using the same composition it is possible by using the second or the third technique, same exactly same value of x and y, but using this second or third technique it is possible to realize lasers which emit at slightly different wavelengths. Of course, by changing x and y that is the first technique which is alloying, we can any how achieve band gap modification. But even keeping x and y constant, also it is possible to vary and thus why the word band gap tailoring. How does this happen we will discuss in detail.

So, the need for band gap engineering or band gap tailoring is to realize sources or even detectors or certain devices with a specific wavelength response. We will discuss more about this at a later stage when we will discuss about the devices and tunable laser diodes. But at this stage let us go into the techniques; what are the techniques that are used. So, the first technique I do not discuss any further because as we know by changing the composition x of ternary compounds or by changing the composition x and y of quaternary compounds, we can realize materials with different band gaps. Band gap is a function of the composition of the alloys. So, let us straightly straight away go to second technique that is use of quantum well structures. This will bring as in continuity with the last class on the discussion of quantum well structures.

(Refer Slide Time: 10:09)



So, if we take come back to the quantum well structure; if we take a finite quantum well finite quantum well. Finite quantum well refers to this finite barrier height potential V or V 0; this is the quantum well. We had used notation L for the width and v is the well height. If you take a finite quantum well, as we saw in the last class the energy I can value equation are given by tan kL is equal to gamma by k or you can write this as tan inverse form. So, kL is equal to where please recall that this k is k square is 2M by h cross square into E and gamma square was 2 M by h cross square into V minus E.

So, kL is equal to tan inverse gamma by kappa, gamma by k and plus m pi. If you want to get all solutions for a given L, you should first write in this form; because tan is a multi valued function. And if you write it in this form where m equal to 0, 1, 2 etcetera. For a given L we are continuing the discussion of the last class, but in the context now of band gap engineering, band gap tailoring. For a given L please see that gamma and k contain E, this will give energy Eigen values and therefore, for a given L you can find out what is the energy Eigen value. So, if you plot this as a function of well width L here L versus E that is. So, if I have here E 1 E 2 and so on, depending on the width it may support one solution or it may have several solutions. So, if you plot this value E which I actually want to use the word the notation delta E, because E is the total energy here axis.

Now, this E 1 is with respect to this that is this is 0. With respect to this, this is E 1 and with respect to this, this is E 2. And therefore, this is actually delta E 1 and delta E 2 this axis is our total energy E, which has E c E v and so on. So, this delta E here refers to the energy of the Eigen states. So, if you plot this then you will get solutions like this. Let me first draw the graph and then you see whether you understand whether this is alright or not. So, this is asymptotically. So, what I have plotted is this is for the first state E 1, this is second state E 2 and third state E 3. Variation of E 1 with the width of the value; if you have a small well the width of the well is very small, then this may support only one solution. Just like those who have studied optical wave guides if you take a single wave guide fiber. So, this is n 2, this is n 1 you have only one solution which is the n effective here; that is the n effective of the fundamental load.

If you have a wider core then it can support the second more third more it becomes a multimoded wave guide. Similarly, in quantum mechanics if you have a well of certain width L; then depending on the value of L it may support only one solution and if the well width becomes larger it may have many solutions. So, E 1, E 2, E 3 and so on. So, what I have plotted is as the width increases this comes down; this goes on dropping down E 1. Because from top a new solution will come E 2 and similarly, from top a new solution will come E 3; that is what is happening. As L increases E 1 is coming down, this is the energy value. So, this is 0 this is delta E and after some value of L the second solution start and E 2 varies like this.

So, the point here is if you locate for example this region, the numbers may be approximate numbers may be this is 50 Armstrong, this is 100 Armstrong, 150 typical numbers for a semiconductor say gallium arsenate 200, 250; so L in Armstrongs. So, what it means is if from 0 to 80 Armstrong, let say this is 80 Armstrong; am just giving an example to illustrate the point 80 Armstrong. For all well width from a small value up to 80 Armstrongs, there is only one allowed solution; there is only one E 1, there is no E

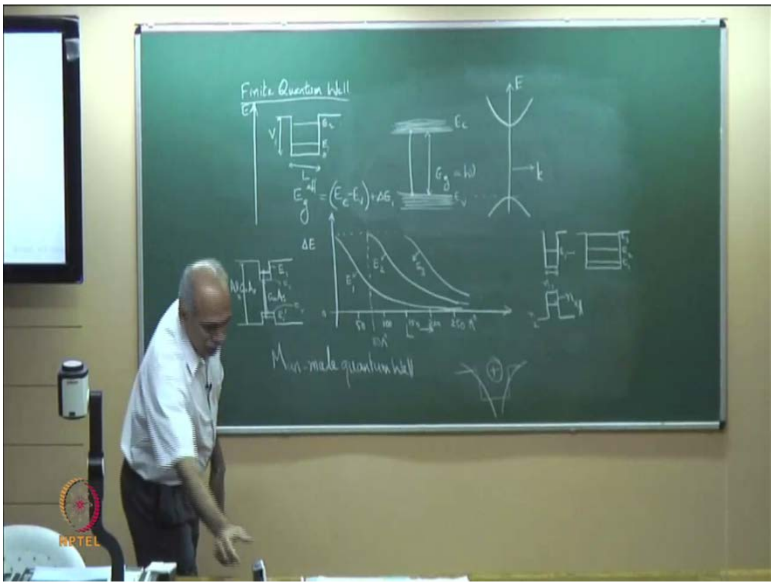
2, E 3. If you cross this then the second solution will come. So, what has been plotted is the variation of energy icon values with the well width.

So, depending on your well width you will have different values for the E 1. Let us focus for the time being at our well supports only one solution, which well are we talking of? We are talking of a quantum well which means we are looking at a double hetero structure in which the sand witched layer, low band gap layer has a very small band; has a smaller band gap. So, this is for example, gallium arsenate and this is aluminum arsenate; so Al x gallium arsenate. So, we are looking at this well; this is the quantum well. The potential energy well in the band structure or a band diagram, energy band diagram of a double hetero structure. When the dimension of the sand witched layer becomes small, it is a quantum well. So, it is this well that I have been drawing here.

These are also called interestingly man-made quantum wells. This is man-made, this structure is made by man; you have made it. So, it is man-made quantum. Particle in a potential box is a standard topic in quantum mechanics and all M.Sc physics students have studied in great detail and this existed for very long time. The particle in a box problem was basically an approximation to potential problems which one encountered. If you have an atom here, a positively charged atom. Then as we have discussed in our first class the potential energy varies like this. And we approximate this by a box to calculate the energy Eigen values. It was an approximation, simplification; so that you have analytical solutions. Today this is a real problem and you have real particle in a box; this box is real number. It is no more an hypothetical approximation; because these structures are real hetero structures; hence the name man-made quantum wells.

You can grow double hetero structures and you can realize quantum wells where the original results of particle in a box that which we used to study in quantum mechanics or almost directly applicable. So, we come back here that over this range if I vary the width, energy Eigen value changes. Which means come back to the well here, E 1 is here. And similarly, there is a well here it is a barrier for electrons or well for holes which also has energy Eigen values which I call E 1 dash; for the valence band. In the valence band also you have a potential well or potential barrier which also have discrete energy Eigen values. The lowest allowed energy Eigen value is E 1, the highest energy Eigen value for electrons or lowest energy for holes is E 1 dash. Please see when we had I have to erase some things let me rub these.

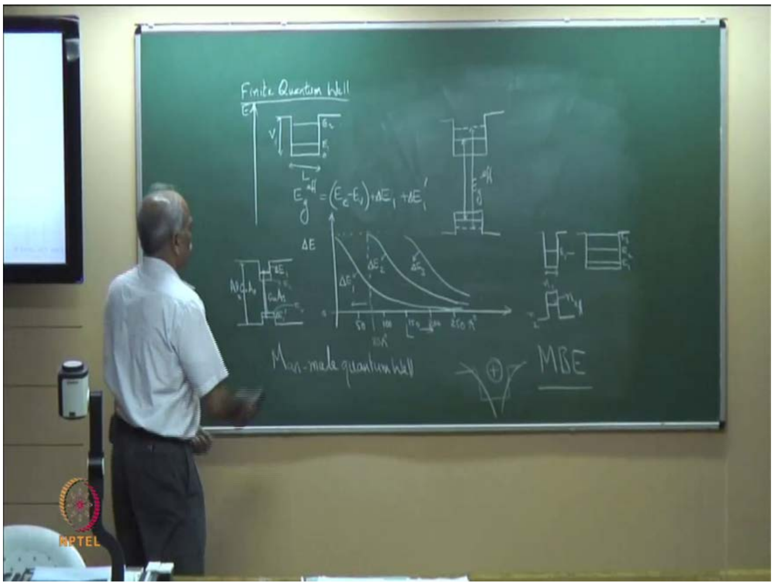
(Refer Slide Time: 21:05)



When we draw the energy band diagram E v and E c for a bulk semiconductor, there are we understand that there are large number of allowed states here and it forms a band. Correspondingly, the E k diagram if you see then we have a conduction band and the valence band; E k diagram for the band. So, these are equivalent. So, this is E and k. There are large number of states here; right from E c. And therefore, the energy band gap E g which is here and electron which is sitting right at E c could come down and recombine with a hole which is sitting at E v. And the energy of the photon which is emitted will be equal to E g equal to h nu.

Now, this is E c see this is E c of gallium arsenate and this is E v of gallium arsenate; bulk. But there is no state there, first allowed electron state is here and first allowed hole state is here; that is the lowest energy hole state. So, the electron can sit here, electron can occupy this state and this can recombine with a hole which is sitting here. Which means, now effectively the band gap is this, it is not E c minus E v of the bulk, but the band gap is this; effective band gap. So, E g effective E g effective is equal to E c minus E v; that is the original E g this plus delta E 1 delta E 1 that is E 1 actually which I call delta E. Let mere move this. So, plus E 1 if I call this is E 1, then this is E 1 plus E 1 dash for the lower band.

(Refer Slide Time: 23:18)



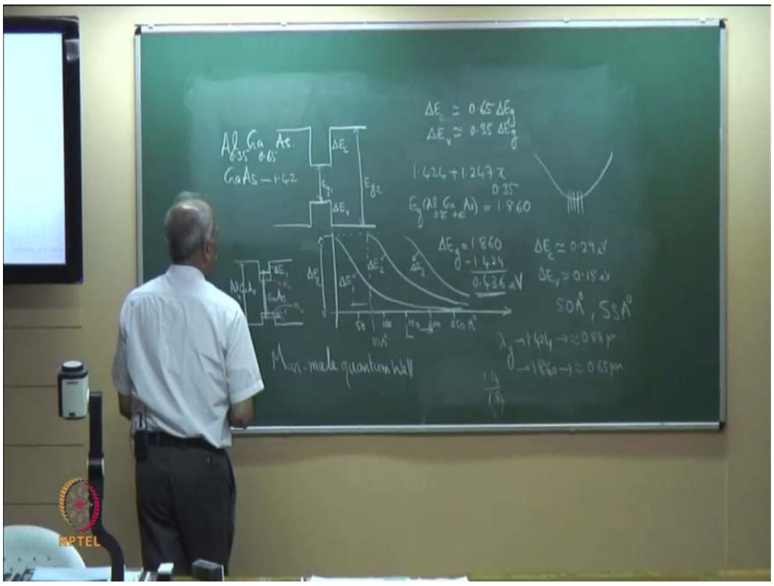
So, this was E g, now we have E 1 please see E 1, this E 1 is the difference E 1 or maybe it would be preferable that I please correct this it would be preferable that we write because we have used E 1 for an energy of electron in the valence band and E 2 for energy of electron in the conduction band. And therefore, it would be preferable to use this delta E 1, delta E 2, delta E 3; in which case this is delta E 1, delta E 2. Kindly make this correction; so that let me show this as delta E 1 dash and delta E 1. The delta referring to that these energies with respect to the band edge here.

So, this is the effective band gap which is larger than E g. So, this is one point that the effective band gap in quantum well structures is let me draw that again. Effective band gap in quantum well structure, if I have an energy level here and an energy level here this is E g effective. The interesting point is if you change the width of the gap, width of the sand witched layer that is width of this well, if you change the width of the well you are moving let us say you are decreasing the width. So, if you are decreasing the width then energy this is going higher. So, this comes to the new place here and this also goes to a new value and therefore, the effective band gap now has increased.

So, what have we got? We have tailored the band gap by simply changing the width of the quantum well or by width of the sand witched layer; this is tailoring. Of course, one can change composition, but if you see it is easier to say that we can change the composition, but when you really see the fabrication process of MBE or VPE or any one of the processes, it is very difficult to you can change the composition. But it is very difficult to control the composition to the required value. But controlling the thickness of the layer is very easy because in MBE technique; for example, today you can grow layer by layer mono layer. So, you can control thicknesses to one mono layer, one mono layer is 3 to 5 Armstrong thickness.

So, that kind of precise control is available in fabrication and therefore, by simply controlling the thickness we can change the effective band gap. If you change the effective band gap the emission wavelength from this source, when we realize a source using such structure the emission wavelength also changes. So, band gap tailoring by use of quantum well structures. You can take some numbers; how much, what kind of change is this? Put some number and see, let us see. Let me take an example, just to get a feel, how much can I tune really? How much, what is the amount that I can tune? See the maximum tuning possible of E 1 is this region; that is the width of the well from bottom of the well I can tune up to this that is the maximum variation that you can get.

(Refer Slide Time: 27:24)



Let us say you realize a quantum well, with a material let me take of familiar gallium arsenate and aluminum gallium arsenate. Let me put a composition say 0.35, gallium 0.65 and arsenate. Gallium arsenate is 1.42 E v, 1.42 and aluminum gallium arsenate you can calculate 1.424 plus 1.247 into x; where x is 0.35, you can substitute the numbers. So, you will have E g, E g of aluminum gallium arsenate 0.35, gallium 0.65 arsenate equal to you can calculate this I have a done a small calculation. You will see that this comes out to be 1.860. And therefore, delta E g delta E g delta E g is the band gap difference between this and this, delta E g is 1.860 minus 1.424 which is 6 3 and 4 0.436 E v is the delta E g. Very quick estimation that how much you will be able to tune, what is the wavelength range you are which you can tune? You will see that the number is really very big, you can tune over a relatively wide range for all (( )) from the practical point of view.

There is something here. This was of E g of gallium arsenate and this is E g of aluminum gallium arsenate. So, E g 1 and E g 2; delta E g is the difference. Where how much will appear here and how much will appear here; what is the discontinuity that will be? The point is I have shown that there is a discontinuity here and a discontinuity here. It could have been that these are aligned and only the discontinuity is here; but no… In practice, there is a discontinuity band gap discontinuity here and here; this is called as delta E c discontinuity and this is this discontinuity is called delta E v.

This has been determined by experiments that this delta E c is approximately 65 percent; that is 0.65 times delta E g. And delta E v the discontinuity at the valance band here is about 35 percent of the total delta E g and this is 65 percent. In the early 80s, this number was 85 and 15, 85 percent delta E c and 15. But subsequently advanced experiments have determined that this discontinuity is approximately 65 percent and 35 percent. Why this is this is a bigger issue?

(( ))

This the question is whether the 65 and 35 is a universal for all materials? No, it depends on the material system, but for most materials which are used in optoelectronics it is approximately in this region. Sometimes, you will see 33 and 67, people are using 33 and 67, 30 and 70, but it is around this value, but it depends on the material system.

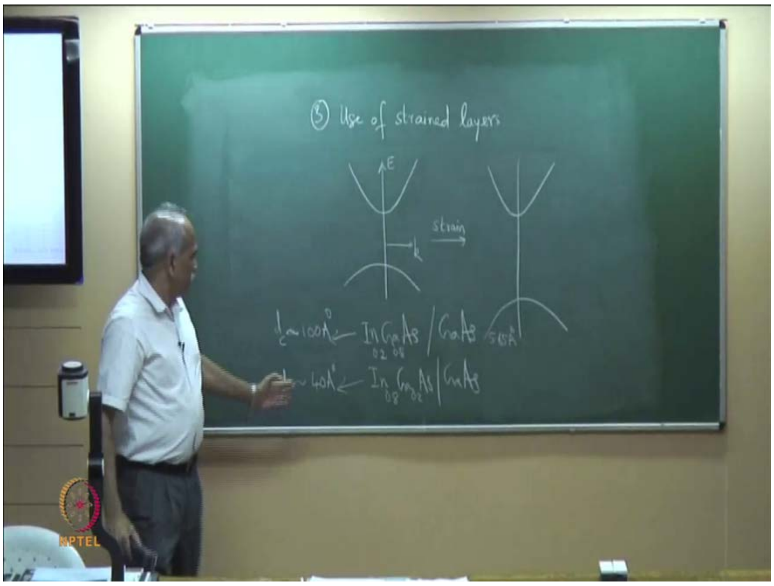
(( ))

Dependence on doping, the question is whether it depends on doping also? That that dependence on doping is much smaller; the primary dependence on the material itself. So, let us assume this number. So, if you assume 65 percent and 35 percent, please see our delta E g is here. Then this discontinuity that we have, why we are interested in this discontinuity? Because that is the maximum change that you can get; from here to here this is delta delta E c. The maximum variation of E 1 can be from the bottom of the well up to the top of the well. So, delta E c is 65 percent of this. So, you know delta E g. So, you can calculate what is delta E c, you can calculate what is delta E v; this is I think delta E c is approximately you can see some numbers 0.29 E v. And delta E v is approximately 0.15 E v approximate number; you can multiply and see.

Therefore, the total band gap total effective band gap can vary from 1.424 up to 1.860; the maximum variation possible. You can calculate the wavelength of emission that is lambda g corresponding to 1.424, you can calculate; check what is the number, we know immediately 1.42 divided by 1.424 will give you I think approximately 0.88 micro meter. If you put for lambda g corresponding to 1.860. So, this is 1.24 divided by 1.860, the number will be much smaller. So, this will come out to be how much will it come? 1.24 divided by 1.86 approximately one-third, two-third approximately two-third. If I drop this and this, it is 12 by 18 which is two-third; so it is approximately 0.65. The point is you can tune the emission from 880 nanometer to 650 nanometer; maximum tuning variation. But this is a when I talked about optical communication window, we need wavelength to be varied by 0.8 nano 0.8 nanometer; less than 1 nanometer is the variation that is that we have talking.

So, therefore, tuning that is changing the emission wavelength by controlling the thickness is the easiest way; because compositions, the gas mixture, gas storage in the fabrication process are kept constant. Simply, the thickness is controlled. In one case, you may have some 50 Armstrong and then next time for the next layer you may take 55 Armstrong, 60 Armstrong. So, you change the thickness by 5 Armstrong’s and the emission wavelength will change. So, this is band gap engineering or tailoring; it is tailoring, it is adjusting the dimension. So, that is why the name tailoring; band gap tailoring by use of quantum (( )). We will talk little bit more about these when we come to quantum well lasers, the devices quantum well lasers and strained quantum well lasers. The third point let me briefly talk about the third point; that is strained quantum well structures. These are the three important techniques which are used for band gap tailoring or band gap engineering.

(Refer Slide Time: 36:13)



So, the third point is use of strained layers. In the last class we talked about strained layer epitaxy, strained layers. When you introduced strain in a medium, so normally you have the EK diagram; here this is the band structure. In the presence of a strain, what strain are we talking of? At the interface there could be compressive strain or tensile strain. Tensile strain is where it is elongated in the compressive strain it is compressed and this is done automatically; the layer adjusted itself, so that bonds are complete. When bonds are complete, the energy of the system is minimum. That is why it tries to do that; it tries to do even if the layer is thick, but it is not able to do and then it leads to cracks in the system, in the cracks in the material.

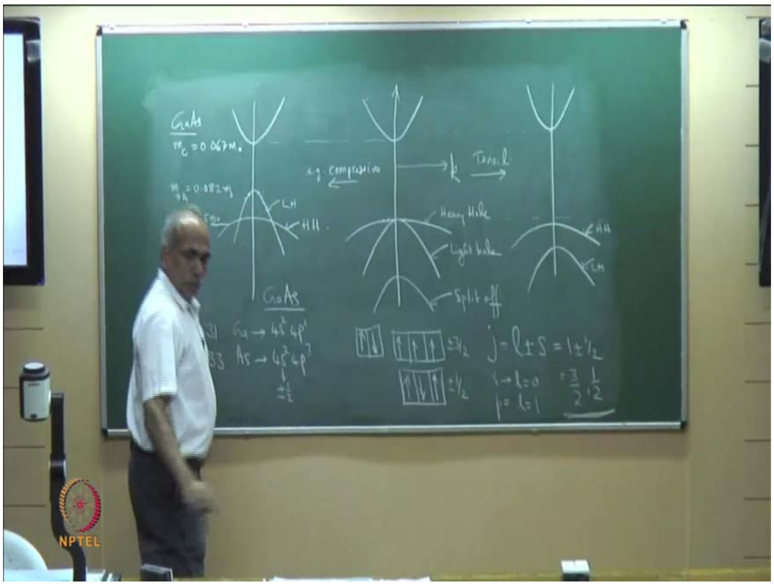
And, therefore, if the layer of the semiconductor which is being grown is less than the critical thickness then it is possible to grow strained epitaxial layers without any defects. Now, one of the important applications of this are consequences of this is change in so due to strain the band gap changes; first the simple picture. So, the band gap could change which means it could become larger or it could become smaller. Depending on in a given system, in a given material system if the strain is compressive and then if the band gap becomes larger, then if the strain is tensile then the band gap here becomes smaller.

So, by control of strain, how can you control strain? Strain can be controlled by choosing the lattice mismatch. So, you can choose for example, indium gallium arsenate on gallium arsenate, let us say indium gallium arsenate. You could choose 0.2 0.8 arsenate or you could choose indium 0.8, gallium 0.2 arsenate on gallium arsenate. The critical thickness in this case, which one do you expect critical thickness to be more? Please see gallium arsenate has 5.65 Armstrong as the lattice constant, indium arsenate has 5.87 has the lattice constant. Here only 0.2 fraction of gallium is replaced.

Therefore, this will have a lattice constant closer to this; this will have a lattice constant closer to indium arsenate means further from gallium arsenate. Indeed if you put some numbers you will see the d c, the critical thickness in this case is about approximately 100 Armstrong; approximate numbers of the order of 100 Armstrong. And in this case (( )) in this case d c because; so we expect it to be more or less? Critical thickness is expected to be less. So, d c is of the order of 40 Armstrong. Because the mismatch is higher. So, one can put the numbers and we have seen that you can calculate d c is equal to A s by 2 mod epsilon and see.

So, by changing the composition you can change the strain; by changing the strain you can change the band gap. So, one can choose a proper combination to change the band gap. Now, you could also argue, if you want to change the composition why go for strain, we could directly change the composition and change the band gap. Yes, strain has strain induced changes have many other advantages. A little bit more complication some amount of quantum mechanics has to be brought in here; but let me very briefly introduce. The valance band in a semiconductor is more complicated, it is not simple.

(Refer Slide Time: 41:13)



We have been always drawing a simple E k diagram in the valance band; the valance band is actually more complicated. It comprises of a light hole band, a heavy hole band and a split off band. So, I am now making it a little bit more complicated and a split off band. This is called this is a heavy hole band, this is light hole band and this is a split off band.

The same E k diagram; all the while earlier while I was drawing, I was drawing the upper most band. Because the electrons will get first field here or the holes which are interacting will be the holes which are here. And therefore, it is this band which is most important in normal circumstances. But when you introduced strain what happens in this material is, at k equal to 0 that is k; at k equal to 0 the degeneracy between light hole and heavy hole breaks and the two bands separate out. Depending on the strain it could be like this; so it could take. So, the light hole could come up and the heavy hole band could still remain there; heavy hole band and the light hole band. Depending on the type of strain whether it is tensile strain or compressive strain, you could have a picture which is like this. So, the light hole band is here, heavy hole band is here; I will tell you what is this heavy hole and light hole in a minute, but let me first talk about band gap tuning.

So, you can see that the effective band gap has been tuned and the holes which are now interacting with electrons are in this band; light hole band. The band gap has changed and the effective mass of holes M v has changed, because this band is varying rapidly.

So, if you find out for example, for gallium arsenate we know that m c is equal to 0.067 M 0. And maybe I have given you what is M v was 0.45 M 0; the light hole has a mass m v of light hole is approximately 0.082 times M 0. It is light compare to this; first point. Heavy hole it is heavy in mass, this is light; this is very simple level of talking light, but basically it is because of the band variation here.

So, now the carriers which are participating are electron here and holes in the light hole band. In my original diagram here it is the holes which have energy because they are coming first. So, they are the once which where interacting with this, now it is this. So, the property of the semiconductor changes very much; this as certain applications later we will discuss in realizing ultra low threshold semiconductor lasers. We will go bring those complications a little later; for the time being in this context the band gap has changed and the effective mass of the carriers have changed by strain. If you had, let us say for example, this is compressive strain. If you use tensile strain the opposite will have; that is this remains here, this changes a little bit and the degeneracy's broken. But and of course, forget the split of split of its quite deep. So, now, the light hole band is here and heavy hole band is here. So, if you use tensile strain for the same material system the band gap has increased, here the effective band gap has decreased.

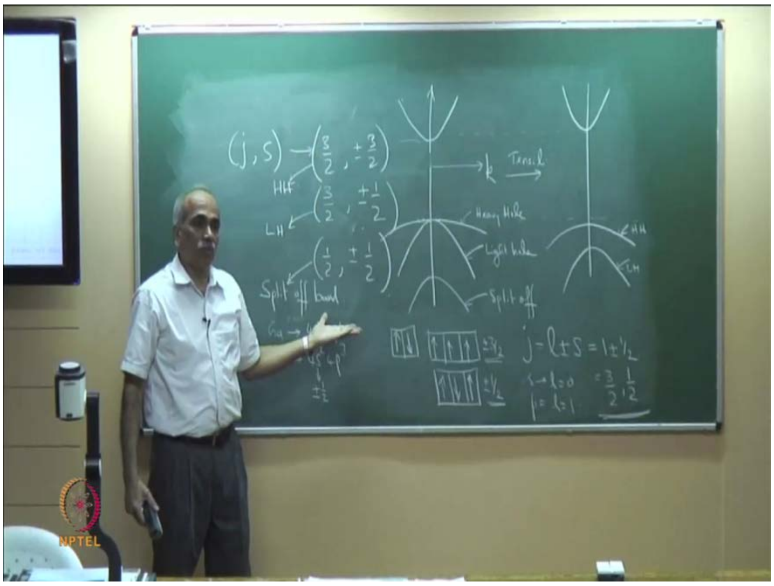
So, please remember that the purpose is not just do decrease or increase the band gap, there are additional benefits which will come because of the effective mass of holes; instead of heavy holes now light holes will be participating. This we will appreciate when we discuss about devices of because that has an implication on the movement of Fermi level under injection. And the movement of Fermi level in they are injection conditions determines the threshold; we will see this at a later state. But for the time being, by use of strain you have a band gap modification and change in the carrier mass, effective mass of the carriers.

Now, we have a few more minutes; what are these light holes and heavy holes? Very quickly, but you may have to refer to some more basic books. If we take gallium arsenate, every time I take gallium arsenate because one it is most widely used material; reason number 1. Reason number 2 is we become familiar with one particular material, so that for any other material you can always extend the arguments. We know that this has 4 S 2 4 P 1 gallium, is 31 atomic number and outer (( )) has 4 S 2 4 P 1 arsenate atomic number 33 has 4 S 2 4 P 3. S 2 means spin it has plus minus half spin; please see

S 2 means there can be two electrons. In p there are 3 shells; so the three electrons could be all like this or it could be.

So, the total spin in this case is 3 by 2 and in this case it is half, plus minus is possible; depending on whether all of them are down or all of them are up, this is the total spin. And the angular momentum j is given by l plus minus s; this is called l s coupling. l plus minus s; l is the orbital quantum number, s is the spin quantum number. l for s shells, l equal to 0, for p shells l equal to 1; this is the value of spin. So, l equal to 0 and l equal to 1; therefore, the total quantum number here will become you can take values. So, 1 1 plus minus half, the total spin; which means this can become 3 by 2 and half. The angular momentum can take values 3 by 2 and half very quickly and therefore, you have three combinations.

(Refer Slide Time: 49:55)



Now, please see s can take values 3 by 2 and plus minus half, j can take 3 by 2 and half. And therefore, we have combination of states which are s, which are (j comma s) the states are represented by (j comma s) which gives the combinations possible are 3 by 2, plus minus 3 by 2, 3 by 2, 3 by 2 here and plus minus 3 by 2, 3 by 2 plus minus half and the last one is half comma plus minus half. The first number is j, j takes values 3 by 2 and half, 3 by 2 and half; s takes values plus minus 3 by 2 and plus minus half. So, we the combinations which are possible are these. This is nothing but the heavy hole band, this corresponds to the light hole band and this corresponds to the split off band.

Those of you are interested you can read more about this, but I thought because I have drawn them, what do they really represent? These represent the allowed quantum states of electron. So, depending on the quantum state of an electron they may occupy this band or this band or this band. So, this is light hole and heavy hole. The concept of light hole and heavy hole is a quantum mechanical concept. Please refer to certain books if you wish to know more about this, but our objective has been to see the band gap variation due to strain, due to tailoring of quantum well structures and due to composition.

So, let me stop at this point, we have covered most of the material aspects. In the next class we will discuss about p-n junctions; you have studied p-n junctions. But my interest will be to see hetero structure p-n junctions. How would be a p-n junction or what is a energy band diagram of a hetero structure p-n junction. Because most of the devices that we will see are p-n hetero structures; it is not the simple p-n junction which we normally study in electronics. So, we will see how they are because that has lot of implications in the final device characteristics.