High Speed Devices and Circuits

Prof. K. N. Bhat

Department of Electrical engineering

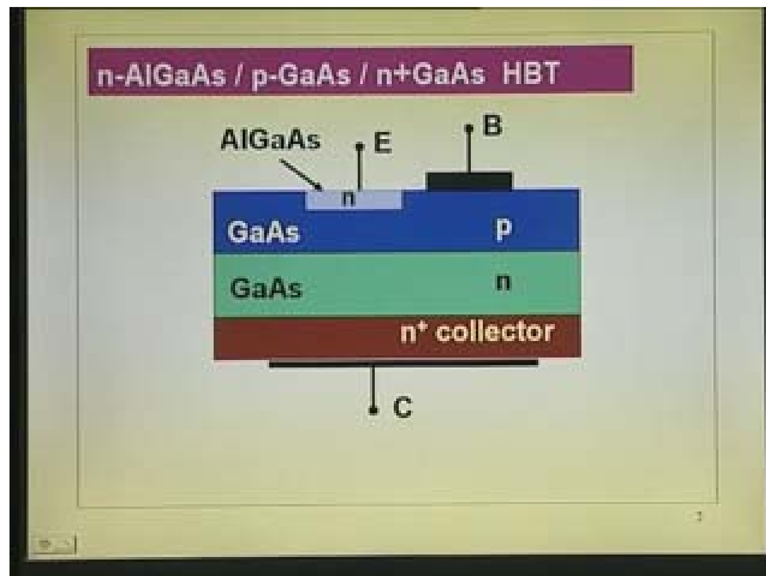
Indian Institute of Technology, Madras

Lecture - 39

Heterojunction Bipolar Transistors (HBT) - 2 (Continued)

We have been discussing the heterojunction bipolar transistor HBT and we will continue on that in our discussion.

(Refer Slide Time: 01:17)



In fact, what we have said is we have the same n p n transistor as in the conventional bipolar transistor. The difference is the emitter has a wider band gap compared to the base. So in the gallium arsenide based technology it will be n plus collector. In fact, this could be if it is just starting with n plus then we can grow epitaxial layer. Epitaxial p layer gallium arsenide, then we can grow epitaxial AlGaAs. Otherwise, if it is an integrated circuit, we can put this n plus layer on the semi insulating substrate and then actually take the contact from the top, just like in the case of bipolar in silicon. In a similar way, we can achieve n down this and take a contact on this on to that. We will see

some of the structures; so that contact is taken from the top. The most important thing is the current flow is vertical here; current flow is vertical, the entire area is conducting. This is one of major differences in all these FET's, except in the case of vertical D MOSFET, but even there the entire area is not conducting.

So the result is you get for a given current, we can realize the given current, with a smaller area; because, if it is a edge that is emitting, this area is wasted; fully edge will be emitting in the case of FET's. So that is why for a given current level or a given power level, the FET will occupy more space compared to bipolar. Of course, we are trying to get into higher velocities and higher speeds, higher cut off frequencies reducing the base width and we said that should be accompanied by higher doping concentration here. We can also ensure that the depletion layer is wide enough by taking the doping on the AlGaAs less than the doping in the base. This is not acceptable in homo junction transistors.

(Refer Slide Time: 03:30)

$$\beta = \frac{D_{nB}}{D_{pE}} \frac{N_{DTE}}{N_{ATB}} \frac{n_{iB}^2}{n_{iE}^2} = \frac{D_{nB}}{D_{pE}} \frac{N_{DTE}}{N_{ATB}} \frac{e^{-E_{gB}/kT}}{e^{-E_{gE}/kT}}$$

$$N_{DTE} = \int_{Emitter} N_D(x)\, dx$$

$$N_{ATB} = \int_{Base} N_A(x)\, dx$$

We have gone through this derivation, so we can see that the beta which is determined by the emitter efficiency, we are neglecting the total recombination in the base because that is just very small, 0.5 micron width of that order. So everything will be governed by the emitter efficiency. In fact, even the silicon transistors today by homo junction, it is the

emitter efficiency which governs the beta, because, base widths are small. Here what you got ultimately is deficient coefficient of electrons in the base divided by the deficient coefficient of holes in the emitter, multiplied by total doping in emitter divided by total acceptor concentration in the base. Normally, the homo junction cancels this, but now we will have some constant which cancels out practically and e to the power of minus E_{gB} band gap in the base divided by e to the power of minus this quantity.

(Refer Slide Time: 04:20)

$$\Delta E_g = \left(E_{gE} - E_{gB}\right)$$

$$\beta = \left(\frac{D_{nB}}{D_{pE}}\frac{N_{DTE}}{N_{ATB}}\right) e^{\Delta E_g / kT}$$

Typically, $\Delta E_g = 0.3\text{eV}$, $e^{\Delta E_g / kT} = 1.63 \times 10^5$

When, $\beta \approx \frac{N_{DTE}}{N_{ATB}} \cong \frac{1}{200}$ and $\frac{D_{nB}}{D_{pE}} = 2.5$

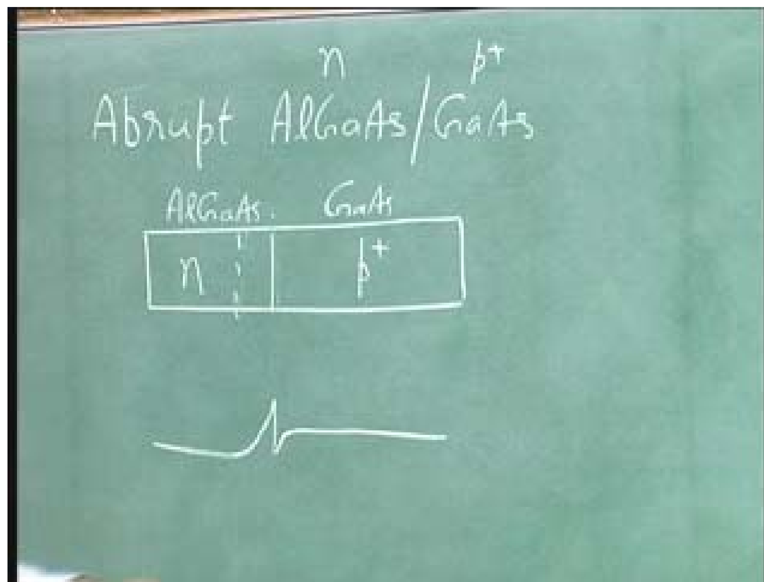
$$\beta = 2.5 \times \frac{1}{200} \times 1.63 \times 10^5 = 2038$$

This actually turns out to be that delta E_g band gap in the emitter minus band gap in the base; beta will actually be this quantity. This quantity will be the beta of the transistor homo junction. We have a multiplying factor which is e to the power delta E_g by kT. We have got another dimension to play with; we have got delta E_g if we make it something like 0.3 electron Volts, as we see in this example at room temperature that quantity is 10 to the power of 5. That is 10 to the power of 5, we can afford to make the total doping in the emitter less than total doping in the base, still this ratio will be large. In fact, put some numbers, I just arbitrarily some numbers which people sometimes use when the doping in the emitter - total doping, integrated doping in the emitter - that is to see this total doping in the emitter is integral over the entire emitter width, that ratio doping in the emitter and the base if it is 1 by 200, we deliberately did it to overcome some of the problems which we discussed in last lecture. D_{nB} by D_{pE} will be..; in fact, it will be much more than that,

because, the hole mobility in be the emitter of the gallium arsenide will not be very high; it will further lower, if the doping is higher, but here that doping if it is not there much, but still that is low compared to the D_n. I have just taken a pessimistic value of 2.5, maybe 3 or 4.

So with that I compute the width beta which has come to about 2000; 2000 with be the base doping is actually it is 200 times larger than that of emitter doping we get. Now, there are a couple of things that we have to discuss: one of the things that we discussed is - what is the practical value that is realized? Before we go to that let us take a look at the energy band diagram of this structure. We have so much talked about AlGaAs GaAs structure with a notch, with a discontinuity in the energy band diagram. Here also it will be present there, but we can overcome that if we do not like it.

(Refer Slide Time: 06:38)



I will just project two type of structures which are there: that is abrupt AlGaAs on gallium arsenide; this is n, this is p, n plus, p plus plus. We can just put it as… I will just put it as p plus, indicating that doping is higher than that. Now an abrupt junction, if I have it like this, n and p plus; I may be talking of 10 to the power of 18, I may be talking of 10 to the power of 17; that is, we call this as p plus. In fact, that diagram, the depletion layer width here will be small in this case, unlike in the case of the AlGaAs GaAs

HEMT; because, there this is high which contributes to the electrons here and the doping here is low. Here entire built-in potential practically will be on this side; that is one difference that we can see. So in an abrupt junction, when I say abrupt, doping is abrupt plus the gradation is abrupt from the gallium arsenide. This is gallium arsenide right up to that point and from here its AlGaAs.

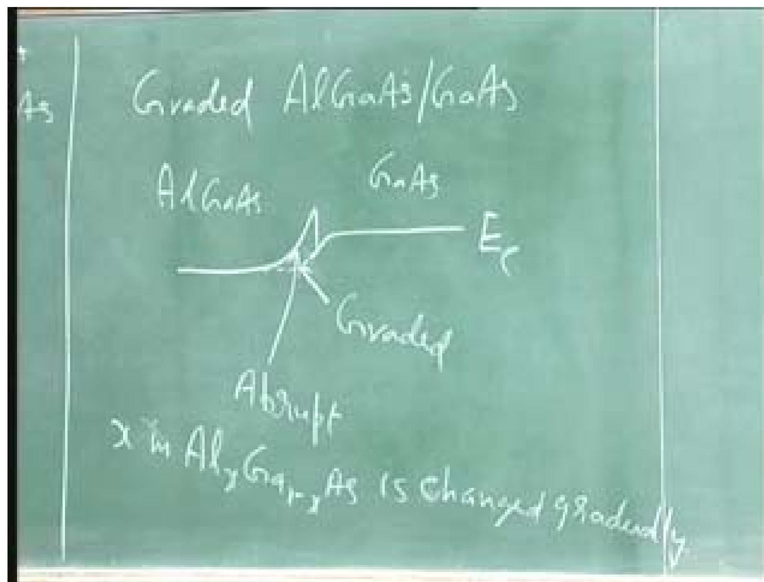
We will have a depletion layer like this. So energy band diagram now will be…. I am just drawing it in that fashion; come down, notice the difference. This is very narrow, the depletion layer width there and the rise in the potential there is very small, but still we have notch. We do not have to worry about this notch collecting electrons, because, that is very small notch there; the height is very small. Still to show I have put it, but we may say I have got this barrier for the electrons which are coming up from this side. When we inject if we have this extra height there, if it is bothering, it does hurt the injection of electrons to some extent. So to prevent that what is done is make it smooth. How to make it smooth? Very easy, on the board for me here; so I will just show that here. This is like that it is there. I will just draw that separately. So here we have this thing like this, like that; that is the discontinuity here, from there onwards very small. The most important thing one should notice here is this height and this height will be different on this side. (Refer Slide Time: 09:49)

In fact, because, of discontinuity here, here it is just going up smoothly there, like that; here it is there, then there is a discontinuity. Here what you see is the barrier for electrons here is smaller compared to the barrier for the holes here. These holes are injected from this side to this side; if the barrier rate is higher here, the hole injection from the base to the emitter is difficult. In fact, your expression there includes that effect; your expression there includes in terms of smaller n_i squared here. See the hole injection jp into that emitter is q D_p pn_0 divided by that width pn_0 is p n bar e to the power of V_{BE} by V_T p n bar is n_i squared by doping; n_i squared is smaller; so, the hole injection is smaller. Another way we are looking at is this barrier is higher for the holes.

This is the barrier and this is the discontinuity and sort of notch that is present here. Now graded, in fact people work out on that, that is the initial set of device that were fabricated

were like this only. If I do not change from gallium arsenide to AlGaAs abruptly; for example, if this is 0.3x and aluminum content 0.3, I do not change from here abruptly to 0.3 aluminum. What we do is change from gallium arsenide slowly to AlGaAs 0.01, 0.02 0.03 like that we keep on changing by the time we change about 100 angstroms, a small layer there, totally change it to AlGaAs.

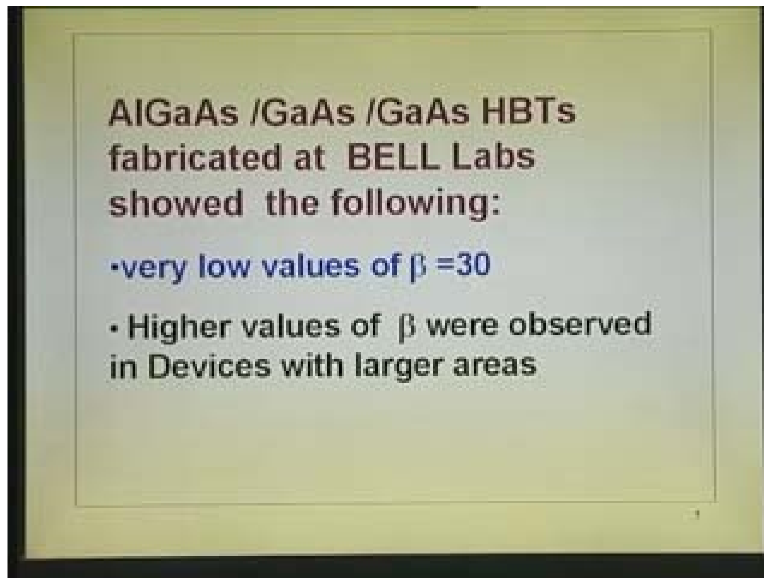
(Refer Slide Time: 11:47)



So that is graded AlGaAs, Gallium arsenide structure; that is emitter, this is base. So how will that be? So what we do is instead of taking up like this and then like that conduction band, what you do is do not change it abruptly, we will see the notch. If this were homo junction how it would have come? It has followed this path. Now actually, it will not follow the exact path, but slightly bring it down to that. So this is instead of having the notch up there, we change the aluminum content gradually, so we will not see the notch. An abrupt transit is not seen, but transit is slowly transferred into that till the band gap is high. We get a high band gap from there and so I am showing the conduction band. So this is gallium arsenide, this is aluminum gallium arsenide, this is graded (Refer Slide Time: 13:00) and this one is abrupt. Change aluminum content gradually, that is, x in is AlGaAs is varied or changed gradually; x is changed gradually.

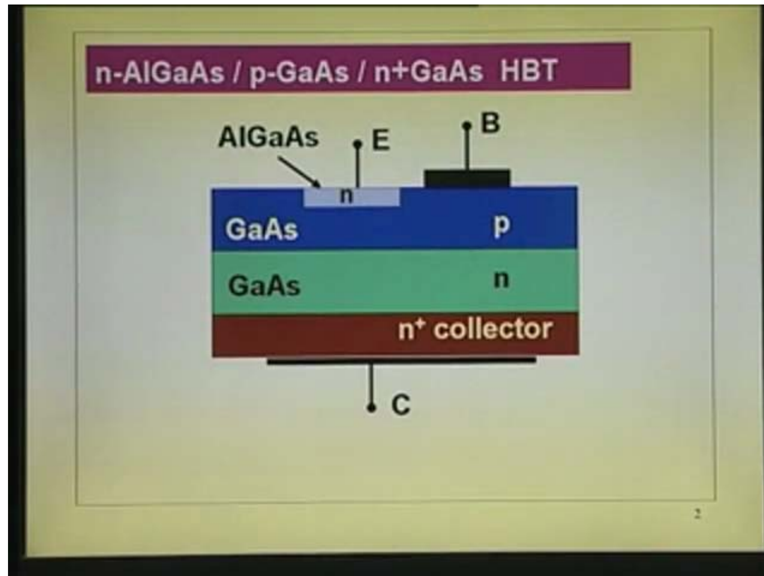
I think we may not be able to see it changed gradually here near the junction, so that this abrupt discontinuity and additional peak that we see for this transfer of electrons from here to here that is smoothened out. So in most of hetero junctions, we will see the smooth diagram like that. So I just put this across, so that we do not worry. In the case of HEMT we were discussing the notch; where has the notch gone? Notch is still there, but notice even when it is there this is because heavily doped it is the small thing, but it will provide some barrier to the electrons here, additional that is removed by smoothening out. So these are some of the changes, improvements, people have made subsequently, but the people did not do that smoothening out. They were not much worried about the beta being low when they made the device; when they made the device what they saw was the beta was about 30.

(Refer Slide Time: 14:58)



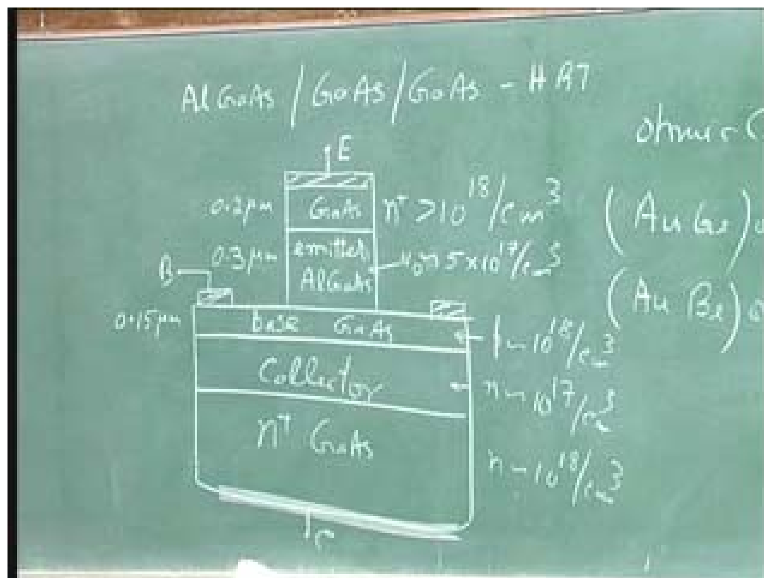
Aluminum gallium arsenide, gallium arsenide gallium on gallium arsenide, HEMT n p n HBT, the BELL lab, New Jersey; beta was low and the beta was higher, if the device has lager area. Now I just want to point out the structure that they had done, what they have put, just make it clear, because, this particular structure gives a wrong picture.

(Refer Slide Time: 15:29)



This is what we see in the conventional homo junctions. I just put it for continuity. Actually what we do is what they did was this structure (Refer Slide Time: 14:58). So since we had completed our discussion, I will just remove this.

(Refer Slide Time: 16:08)



The structure schematic of course is like this: AlGaAs, gallium arsenide, gallium arsenide - HBT. In fact this is like this. We have got contact on the top for the emitter and from

there we have got a gallium arsenide layer. We will explain each one of them; why those layers are. There are modifications of this structure, but this is typical structure that is done when we say discrete device. If it is integrated circuit, we have semi insulating layer at the bottom if we take contact from the collector from the top. This is n plus greater than 10 to the power of 18 per centimeter cubed. n plus gallium arsenide not AlGaAs. Then below that the emitter; this is the emitter; this is the emitter and that is aluminum, I will just put AlGaAs; it is understood that aluminum gallium arsenide layer and this is about 0.2 microns. This is I think the device that they made in the lab initially - 0.3 microns

Now you see there is talk of layers that are even thinner; the typical layer which we have; we just saw the AlGaAs. So that is actually the emitter - contact. So these are not done by diffusion and things like that; these are done by growing one layer over the epitaxial layer. So then we have the base region; so we have the base 0.15 micron. This is the number that in the first article they have, put this is the base contact, contact, gold beryllium; this is the gold germanium, gold beryllium or gold zinc. The p type contacts are made with gold. See here gold and germanium; germanium is n type and zinc is p type. Gold with little bit of zinc makes it p plus there, makes contact. So we can have gold zinc or gold beryllium alloy evaporated and made contact. So contacts are gold germanium on n plus gold beryllium on p base; gold germanium, germanium it makes n plus to make contact; gold beryllium is p type, second group, a good p type dopant. So what happens is the beryllium goes and oxides the gallium side; so that makes p type, the second group. This is actually the base gallium arsenide. We can have force contacts taken on either side and it can be circular contact, all round.

Below that we have got collector and this is actually p type doping about 10 to the power of 18 per centimeter cube, the typical number. This one, the doping here is N_D is equal to roughly about 5 times 10 to the power of 17. Then this collector is n approximately, I think 10 to the power of 17 or so; n minus and then n plus gallium arsenide with n is about 10 to the power of 18. This number let me call it 10 to the power of 17, so lightly doped here. So that the depletion layer extends on to that region; it does not move into that. Then this is about 10 to the power of 18 and we have got the contact there.

The entire key thing here is we can see there are so many layers present. This is n plus layer which is present in all HBT's to make the series resistance reduction and this layer is that layer is about 0.5 micron. Just hold the depletion layer; much larger compared to the 0.15. We have got n plus gallium arsenide, starting the way we do this device is the heavily doped n plus wafer, on that grow by MOCBD. A collector layer which is 10 to the power of 17 per centimeter cubed 0.5 microns. On the top of that grow a p type, all that we do is all these are gallium arsenide, gallium arsenide, gallium arsenide. It is not necessary that we can have gallium arsenide here, we can have AlGaAs, but in this device it is gallium arsenide lightly doped then about 0.15 microns base width gallium arsenide.

Then grow another layer on the top of that all the way; a layer which is aluminum gallium arsenide, we can have gradation here. The width over the gradation takes place is over in about 100 angstroms; that particular diagram which I drew, we see, instead of taking up like that we take it like this (Refer Slide Time: 24:21), that width becomes more because that depends up on doping level. So the transition from here to here can be taken place at a very small range of 100 angstroms; this is about 0.3 micron and doping is about 10 to the power of 17.

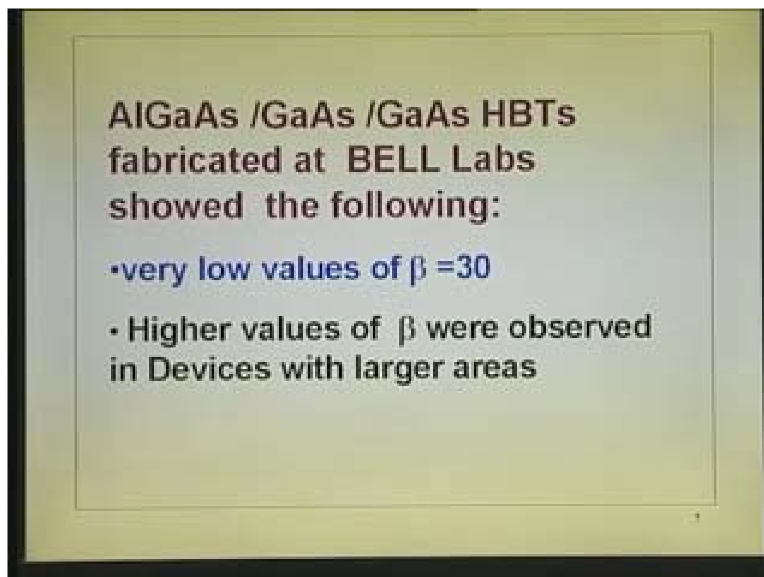
The depletion layer, emitter based depletion layer, will be expanding on to this layer rather than this. Notice doping is lower compared to the base by a factor of 2; not as much as high, but apparently they have taken the aluminum content much more, so that delta E_g is not 0.3 it is much more than that. If we have aluminum content x equal to say 0.3 or so, we will get about 0.5, 0.4 delta E_g that will give we e to the power delta E_g much more. So this is actually the emitter. What is this layer doing there - n plus gallium arsenide? There is only ohmic contact; the emitter region is over here. In fact, whatever is injected here does not reach up to that and whatever reaches here it is just as good as extending ohmic contact on to this. The wide band gap semiconductor and a narrow band gap semiconductor, we can make good ohmic contact over there.

If we do not have this layer, I can put the contact on to this it becomes difficult; lot of contact resistance comes into picture. So that is why when we deal with wide band gap

materials we will have a contact layer, which is the lower band gap. In fact, we saw if we recall the discussion that we are doing on the HEMT, there we saw a top contact layer over which the metal is put; ohmic contact is made. The top contact layer is a lower band gap material. So that is the purpose. So if we want to calculate the integrated doping that is the portion and this integrated doping here. Anyway, what they have done is that is they made this device, growing these layers and edge all the way up to this point; edge the AlGaAs and layer up to the gallium arsenide. In fact, we can have selective etch ends; the moment it sees it is gallium arsenide, it will stop. So then put a contact on to this base contact.

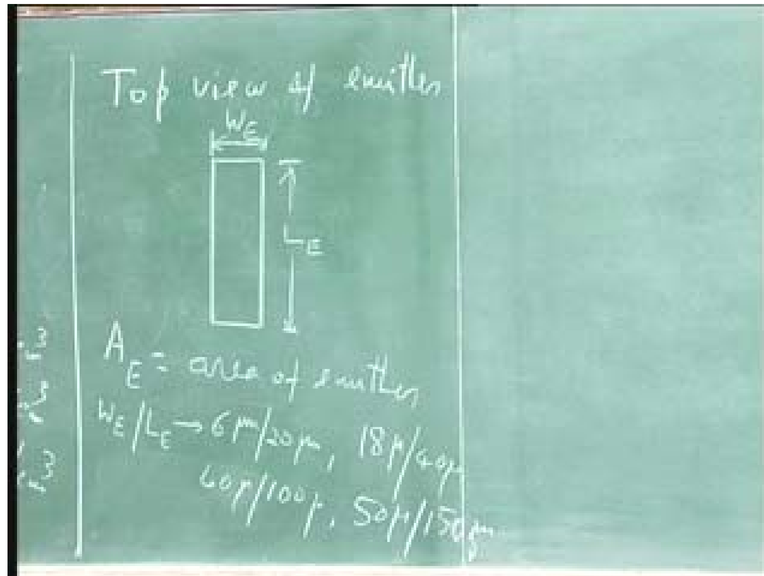
I put two stripes coming on both sides and of course, make contact on the emitter and the collector. So this is the type of structure that they have made. In fact, if we want to have a device integrated circuit what we will have will be this n plus layer will be grown on a semi insulating substrate; all the layers will be grown; then, we will have a step here; we will etch down all the way up to this point ohmic contact on to this here. Collector contact is taken from here; base from here; emitter from there; like that three layers. So contact is made from bottom to top. The semi insulating layer will be isolating layer. So this is the type of device that we had and to their surprise they got beta about 30; that is average very low beta.

(Refer Slide Time: 27:28)



They expected 2000; also, they saw, they made device areas which was larger beta was better than this; it is very low when the area is small. In fact, their structures were like this.

(Refer Slide Time: 27:51)



The device structures the top view of emitter was like this; that is, L_E like that and W_E emitter. Area of the emitter is W_E, L_E. What were the values that they have chosen? They have chosen values which are 6 micro meter by 20 micro meters. W_E by L_E ratios were 6 micron by 20 micron; that is smallest device. 20 micron we see the very small area emitter area. We talk of areas in millimeter or 500 microns of that order is very small. Then 18 micrometer divided by 40 micrometer; that is another type of device. 40 micrometers by 100 micrometer; also they had 50 micron by 150 micron. I am just pointing out this, because, they had made devices with varieties of areas and found this give poorest beta; better, better, better. So, what is the reason?

(Refer Slide Time: 30:06)

Emitter efficiency is very high.

β governed by recombination?

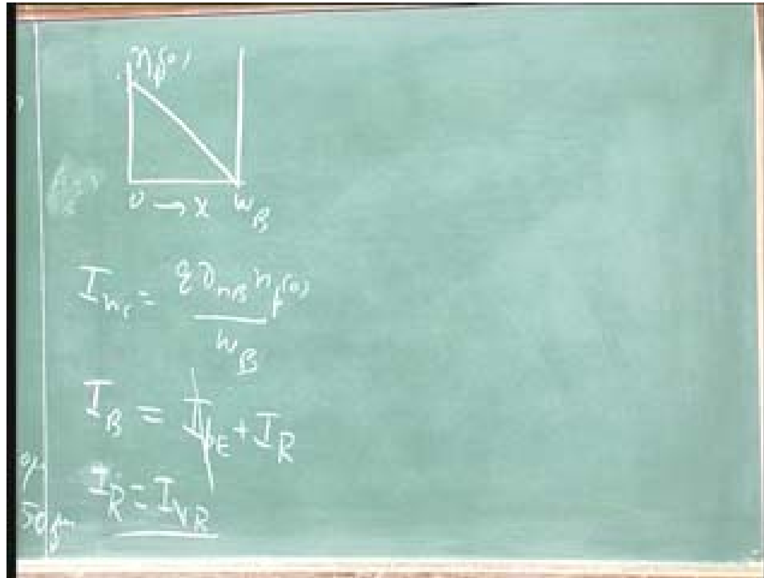
$$\beta_0 = \frac{I_{nC}}{I_{vR}} = \frac{qD_n n_p(0)}{W_B}\,\frac{2\tau_n}{qn_p(0)W_B} = \frac{2L_n^2}{W_B^2}$$

$$= \frac{2 \times 100 \times 10^{-7}}{10^{-10}} = 2 \times 10^5$$

This is too large to be concerned about

That means apart from emitter efficiency some other parameters is playing role. Let us see if it could be for this type of structure, whether it could be by recombination into the base region. If we have not noted down 6 micron by 20 micron, 18 micron by 40 micron, 40 micron by 100 micron, 50 micron by 150 micron, W_E/L_E and the area is product of the two. If the beta were governed by the volume recombination, would we get low beta? So that is why we just computed that here. Let us say that emitter efficiency is not the one that is limiting it. If it were emitter efficiency you must have got beta about 2000 or higher, about that. So here in the base region, if the beta is governed by the transport in the base region, collector current divided by the base current and base current I am now neglecting the whole current injected to the emitter, taking that it is totally due to volume recombination. That is, actually this quantity is the collector current.

(Refer Slide Time: 31:34)



What we are trying to point out is if I have the base region that is W_B, then I have n_p0 without the symbols do not have meaning whatever equation/ we write n_p0; so this I_{nc} per unit area of course; it is actually the current density that is A is equal to 1, A_e is equal to 1 if I take; so I_{nc} is actually equal to… this is the well know formula for D_{nB} n_p0 by W_B q D_n by d of x; n_p0 by W_B is actually the D_n by d of x, diffusion current. Volume recombination I_B is equal to I_{pE} plus I recombination in the base. Let me write it slightly bit better. I_B taking up all other unit area; otherwise, current densities. I_B is equal to the hole current injected to the emitter plus recombination current in the base.

We are telling that let us neglect initially we calculated beta by finding this divided by that; that should give high value. We now say let us neglect that and see if I take I_R how much will be the beta and what will be I_R? I_R is actually the area under this curve that is total charge divided by life time. What is the area under curve? n_p0 into W by 2 into q of course, q n_p0 into W_B by 2 is the area charge divided by life time tau_n; that is the volume recombination current. So I_R I am taking it as I_{VR} where I_{VR} is the volume recombination, current, recombination taking in the volume of the base, total volume; total charge in the volume divided by lifetime.

(Refer Slide Time 34:01)

Emitter efficiency is very high.

β governed by recombination?

$$\beta_0 = \frac{I_{nc}}{I_{vR}} = \frac{qD_n n_p(0)}{W_B} \frac{2\tau_n}{qn_p(0)W_B} = \frac{2L_n^2}{W_B^2}$$

$$= \frac{2 \times 100 \times 10^{-7}}{10^{-10}} = 2 \times 10^5$$

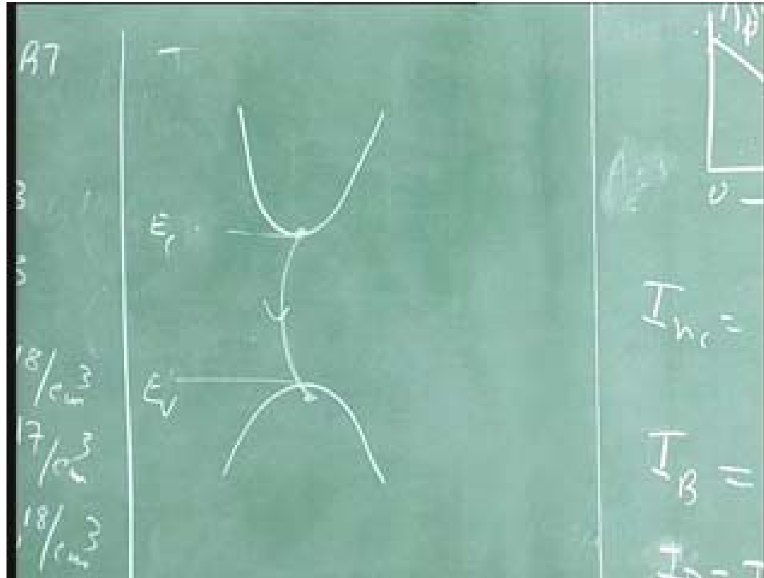
This is too large to be concerned about

That is what we have put here. This is the collector current, electron current injected to the collector which is the collector current qD_n, n_p0 by W_B area, of course, I have taken 1; otherwise, current density even otherwise it cancels out here when I take the ratio. This is the total charge q n_p0 W_B by two divided by tau_n. This is 1 by area; so that actually is 2 Ln square; q_n p gets cancelled D_n tau_n L_n squared into 2 divided by W_B squared; that is the beta. If we compute that for a base width of about 0.1 micron; in this example it is 0.15, but just taking the example for the particular device, for particular doping, particular base width 0.5 micron W_B square is 10 to the power of minus 7 centimeter cubed to Ln squared D_n tau_n. D_n into tau_n; D_n is 100. See for silicon D_n is 36 centimeter Volt per second; I have taken as 100; it is pessimistic value it is more than that. Even if the doping is high, I have taken this deliberately low, because doping in the base region is high 10 to the power of 18 of that order.

Then even it will be about 100, just about 2.5 times that of the silicon D_n; that is mobility is of 2.5 times mobility of silicon electrons of silicon 1500 into 2.5; that is about 400 or 4000 or so. That is why 100 I have taken D_n and lifetime it is 10 to the power of minus 7. We may say we have taken very bad life time 0.4 micro second. A silicon technologist will say that is very poor life time we have defects there. But in the case of gallium arsenide and direct band gap semiconductors lifetime should be small. The chance of

recombination is very good there, because, if I just recall what we have been discussing earlier.

(Refer Slide Time: 36:26)



If we see the band gap, that is the conduction band, that is the valence band. Here electrons can directly transfer; very easily recombination will take place by transfer of electrons, just by emitting some radiation light visible or invisible depending up on the band gap. 1.24 by E_g will be the wavelength that is associated with that. Because this chance of recombination is very good from here, it is enough if the holes and electrons are in the vicinity of each other, they will recombine; whereas, in silicon it is not just enough if the holes and electrons are in the vicinity there should be another particle which participates in the transfer of electrons from the conduction band to valence band.

So probability of recombination is lower if there are no defects. Here even if there are no defects, the probability of recombination is very good so long as they are in the vicinity of each other which will recombine.

(Refer Slide Time: 37:27)

Emitter efficiency is very high.

β governed by recombination?

$$\beta_0 = \frac{I_{nc}}{I_{vR}} = \frac{qD_n n_p(0)}{W_B} \frac{2\tau_n}{qn_p(0)W_B} = \frac{2L_n^2}{W_B^2}$$

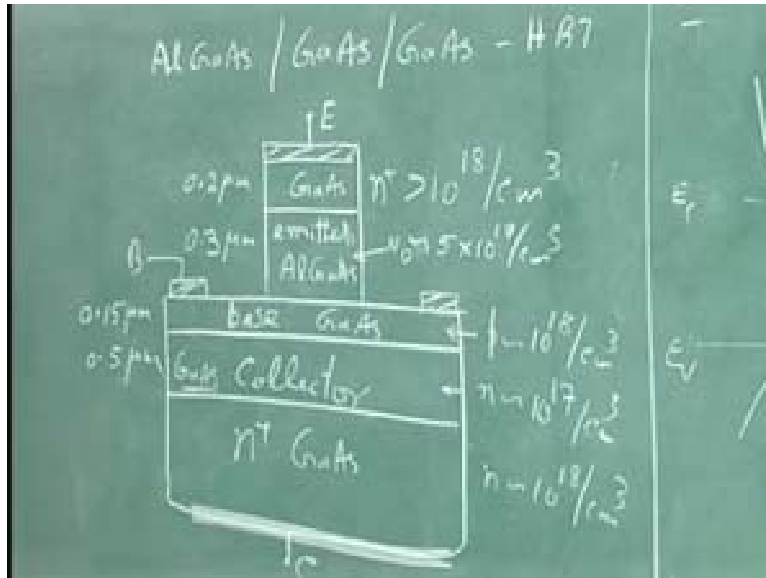
$$= \frac{2 \times 100 \times 10^{-7}}{10^{-10}} = 2 \times 10^5$$

This is too large to be concerned about

I have taken 10 to the power of minus 7; still it is beta associated by, if I compute the beta which is due to this transport of electrons or recombination in the electrons, then it is about 10 to the power of 5 beta, we would not get that. In fact, compared to this, the beta governed by the emitter efficiency will be slower we found about 2000. So if we take this and that together it will be 2000. Beta governed by volume recombination is 10 to the power of 5; beta governed by the emitter efficiency is 2000. So between the two, we can say approximately 1 by beta is 1 by beta 1 plus 1 by beta 2, by putting this two processes.

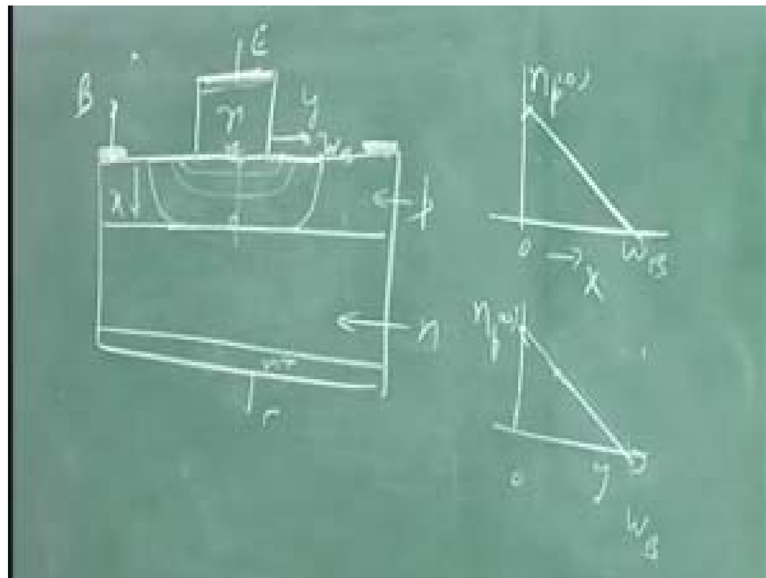
So it will…. We do not have to worry about this; it is not recombination. Even if we take the lifetime 1 or more magnitude smaller that means 10 to the power of 4, it will be just I have taken up 100 nanoseconds, 0.2 micro seconds. This is too large to be concerned about, we do not have to worry about it all, recombination in such small volume neither in silicon nor in gallium arsenide.

(Refer Slide Time: 38:35)



But now what we said in this derivation here, in this derivation is the electrons which are injected from here are recombining in this volume. Let me remove the whole thing now; just show only this portion, so that it is clear.

(Refer Slide Time: 39:02)



If I take a base region like that, that is the p base and this is the n and n plus, does not matter; we have the n plus region here collector and we have the emitter and the base it

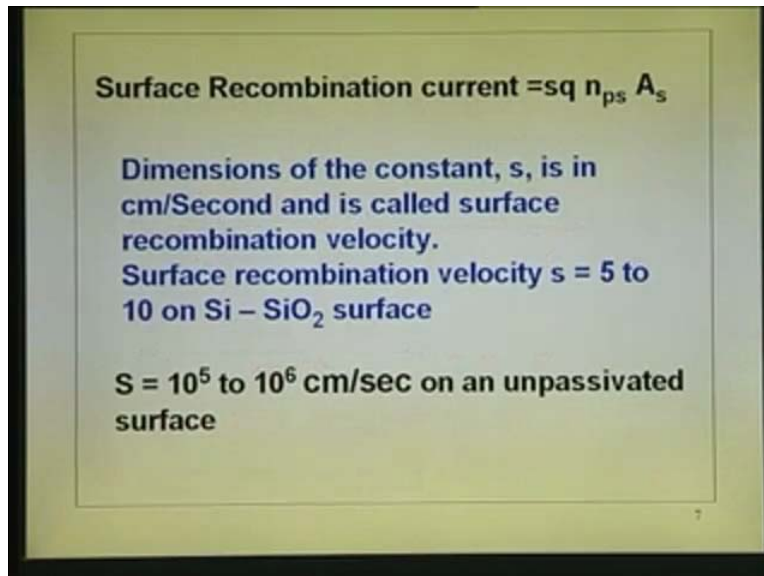
maybe on both sides. This is the n, this is p; emitter, base, collector; what we assumed when deriving the equation is that is the region that we took one-dimensional analysis current flow is entirely here. Actually what happens is we will have just spreading up current flow like this plus current flow like that, that does not bother much. How much it will spread in this direction? If this is 0.1 micron, this is not going to spread too much it may go slightly more than 0.1. That is we will have current flow like this, like that; that must increase base recombination slightly more; it will be a fraction of that. Instead of 10 to the power of 5, we may get beta of 0.9 into 10 to the power of 5, that sort of thing, but the killer is if we take a look at the carrier distribution, carrier concentration is n_p0 here. We all the time plot the carrier distribution like this n_p0, 0 to W_B; that is in this direction. What about in this direction y? How will the current distribution be? This is n_p0 here, boundary and this is 0. Why? Because it is reverse biased.

Carrier concentration 0 at the collector base junction, n_p0 here. If I take it as 1, this is 0 n_p0 is 1. I can draw for uniformly doped base region and carry distribution like that 1, like that, like that, ultimately like this. These are the carrier distributions. It is like if I have potential plot like this, we know if we have Poisson's equation d squared phi by dx squared is equal to 0, then the potential distribution will be governed by that; instead of d square phi by dx squared is equal to 0 we have got d squared n by dx squared is equal to 0. But del square is equal to 0, not 0, but del square n is equal to n by tau; that is the equation that we got towards the distribution here. What I am trying to point out is we will have n is equal to n_p0; this is 1 and here it is 0. With this line if I plot the distribution is like that - linear.

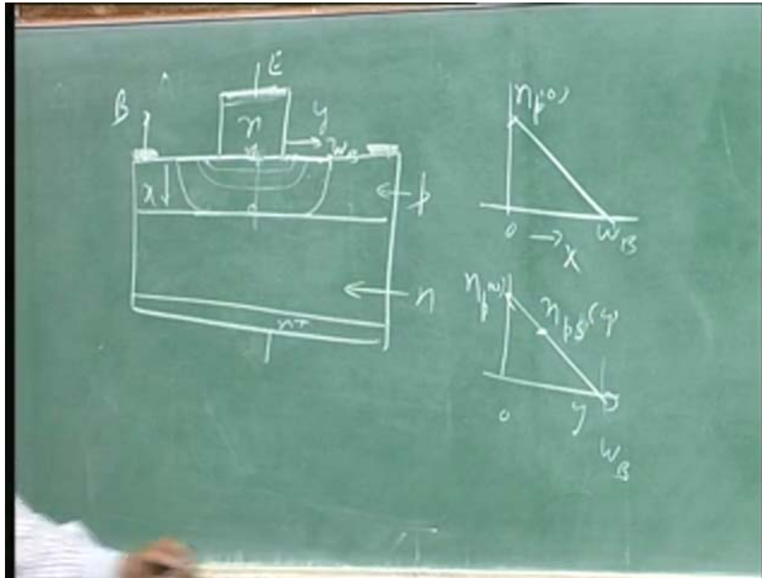
Similarly, if I plot from the surface this width will be actually practically equal to this width. Surface up to which the carried distribution will be governed that we solve by computer we will see this. We will solve this equation del square phi minus k phi is equal to 0; where phi is n_p0. We will get this sort of distribution for potential or carried distribution 10 to the power of 16, 0 and linear if we plot like this; let us plot here and here also y direction, also we will have practically n_p0 up to a width equal to W_B we will have distribution.

Why are we worrying about this distribution on this region? The difference between this region and this region is dramatic, drastic unless we do some treatment. This is the bulk region; the recombination is only due to the recombination, bulk recombination, which can be direct transition or due to some defects. This region may not be defective; still we get low lifetime, but that is not bothering. What about this region? Carrier concentration here is n_p0 and is falling down almost linearly up to a width equal to W_B. What will be this area? (Refer Slide Time: 43:43) This area depends upon the base width; this area depends upon width. Why should we worry about area? Let us see how it happens. In this region what happens? The carriers are present here. The carriers which are also present here will also get recombined; that is the key thing. That is recombination will take place on the surface; that is what we are trying to point out now.

(Refer Slide Time: 44:14 min)



(Refer Slide Time: 44:27)



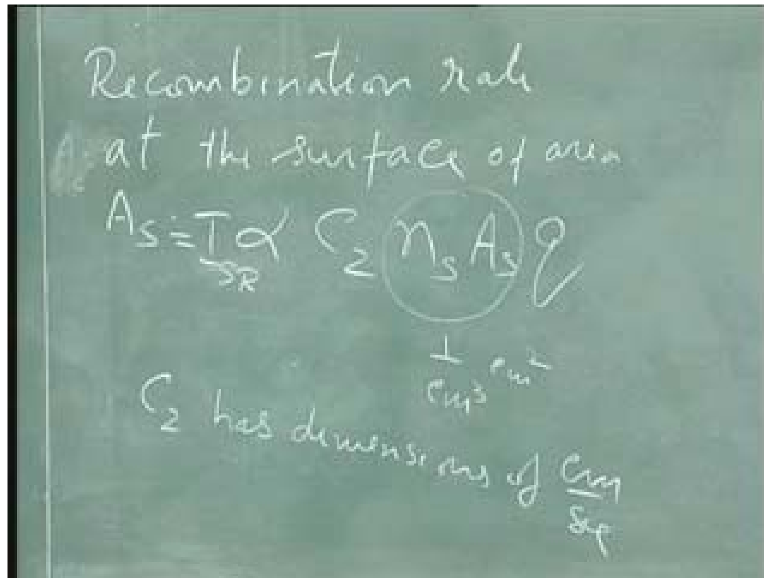
If at any point this is n_{ps} at any y or let me just keep it as it is n_p as function y. At any n_p there will be recombination and that is called as surface recombination. Let me put that down here. Hope this argument is clear, the distribution is linear here, it is almost linear on surface also. There are carrier concentration raised above thermal equilibrium on surface and those carriers which are there, whenever when we raise the carrier concentration above thermal equilibrium, there will be recombination. The recombination rate will be proportional to carrier concentration: n by tau, the bulk. Let us see that. In fact, I think I will introduce we this concept today.

(Refer Slide Time: 45:16)

See high recombination rate in the bulk that is volume is equal to carrier concentration divided by lifetime. In fact, let me put it in this way, is equal to some constant C_1 into n per unit volume into V; V is the volume. In the volume V is that recombination current into q current charge per second is the current, this is per centimeter cube; volume is centimeter cube. So this is dimensionless, this is charge so proportionality constant at a dimension equal to 1 by second. So that is why this is given as n_{qV} by tau; it has got dimensions of 1 by tau, I by time; so exactly the proportionality constant which we call it as lifetime in the bulk; charge divided by the time is the current in the bulk. So we call it as lifetime.

Surface also will have some equivalent constant; what is the current? what is the recombining on the surface? Let us take a look at that. Recombination rate at the surface.

(Refer Slide Time: 47:23)

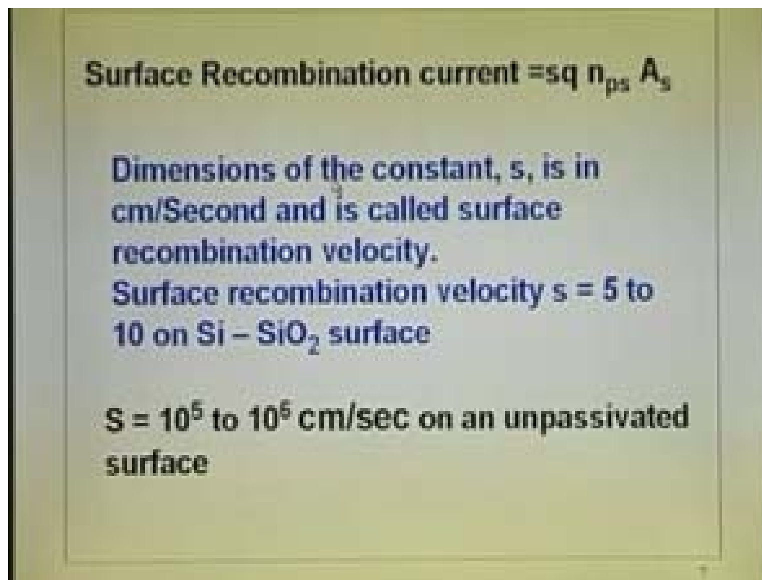


This is an important concept to understand that we will see how important to give a good treatment to the surface. We must actually take care of the surface very well; otherwise. it will give very bad performance. So recombination of the surface of area A_s, A surface is equal to is proportional to some C_2 in the carrier concentration there on the surface into area into q. In this case, we have got proportional to n multiply by volume to get the recombination of the entire volume. Here we do not have the volume, we are finding out how much is the recombination I taking place at area A_s; n_s is the carrier concentration per centimeter cubed from surface into area into q proportional to C_2.

What is the dimension of this? I_{SR}, this is I_{VR}; this is centimeter cubed. This is 1 by centimeter cube; this is centimeter square; this is charge. So this together has a dimension equal to 1 by centimeter cube into centimeter square. So together it has got 1 by centimeter. What is the dimension of this? This charge, of course, takes care of this. Finally this should be charge per second; that means, entire thing must have dimensions of 1 by second. Here it was very easy for us to give 1 by tau. Here 1 by cm is the dimension here (Refer Slide Time: 49:37). I must have this as 1 by second; so this must have dimensions of centimeter per second. So C_2 has dimensions of centimeter per second. What is that? That is velocity.

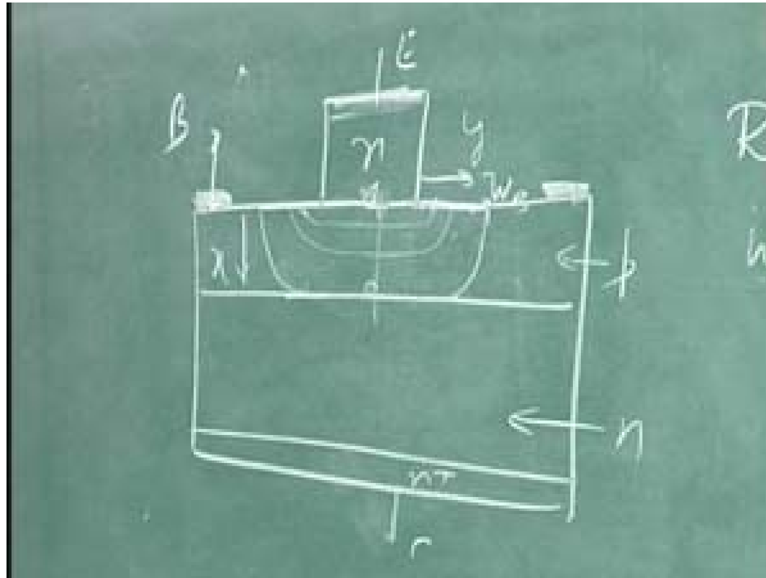
This is proportionality constant. Once we put proportionality constant is equal to…. So this is actually equal to s q_n A_s; where s is C_2; this is the standard symbol that is used; S is C_2 has dimensions of centimeter per second and is denoted as S, where S is actually the dimensions of centimeter per second. This is somehow they did not want to confuse with the velocity of the electrons there, though this has dimensions velocity, they do not want to confuse with the velocity of electron. This is actually a constant which we multiply by this quantity, it gives us recombination current; S is called the recombination velocity.

(Refer Slide Time: 51:12)



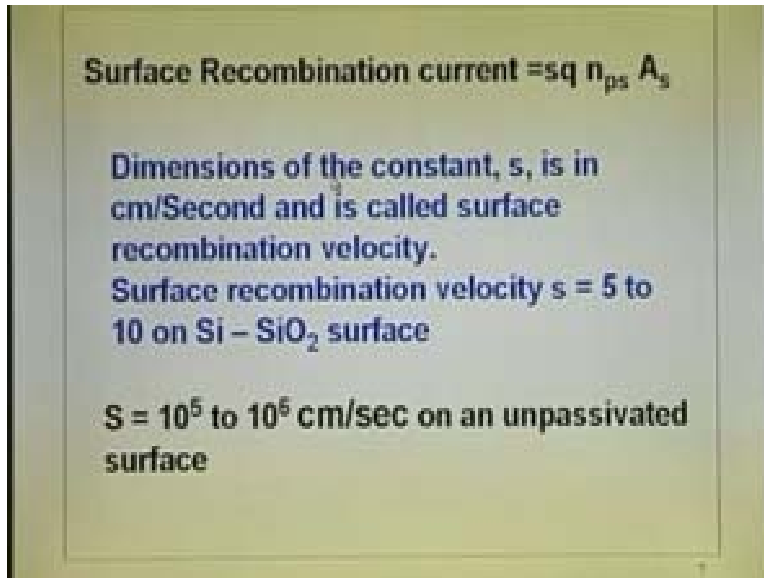
What you are trying to tell is the surface recombination current, how much is the current recombining? The surface depend up on that constant of proportionality that is recombination velocity, surface recombination velocity. This S is the surface recombination velocity; n_{ps} is the average carrier concentration over the surface; A_s is the area. When we took the bulk recombination, we took the integral on volume and we get q n_p0 W by tau for the current.

(Refer Slide Time: 51:55)



Here what is the average n_{ps} over the entire area emitter? From here to here (Refer Slide Time: 51:57). I can compute this area, if I know this area. What is the average carrier concentration here? This is n_p0 here; this is 0 here; linearly falling n_p0 by 2.

(Refer Slide Time: 52:17)



The surface recombination current in the surface will be q n_p0 by 2 into area into S. So now if we want to compute this, we have to compute what is the area? What is the

recombination velocity S? To give an idea when we compute the bulk recombination volume recombination I_{VR}, we talk of q n_p0 by lifetime; that lifetime is 10 to the power of minus 6, 10 to the power of minus 7, etc. Whereas, in the case of surface, recombination is very heavy. How heavy it depends up on how bad the surface is. Suppose, we have a surface which is left open like that, like an open wound, it is very highly reactive, the recombination rate is very high. When we say the recombination rate is high, what is the quantity which is high? Recombination velocity. When I say bulk recombination is very high, lifetime is very small; that is, high means lifetime is very small. If the surface recombination is very high means actually this proportionality constant is very high; actually it is very bad.

This is the characteristic of the surface recombination velocity. In silicon devices, silicon bipolar junction transistor, where we have the oxide, the recombination velocity is something like 50 or people even talk of 0, no recombination, but it cannot be 0, it is 550 centimeter per second, but in a bare gallium arsenide if we leave it open like this, that will be of order of 10 to the power of 6 centimeter per second; tremendous, because we have all the tangling modes there, which are serving as the defective centers, which are serving as recombination centers. So because of that we will have a lot of current going to that.

(Refer Slide Time 54:26)

Surface Recombination Current I_{SR}

$$I_{SR} = \frac{qn_p(0)\,s\,A_s}{2}$$

$$I_{VR} = \frac{qn_p(0)\,W_B\,A_E}{2\tau_n}$$

$$\frac{I_{SR}}{I_{VR}} = \frac{s\tau_n}{W_B}\frac{A_s}{A_E}$$

A_s is the effective area over which surface recom - bination takes place. A_E is the emitter area

Surface recombination effects dominate in small area devices

So the surface recombination velocity, in fact without going into the thing we can see now. What is the ratio between the surface recombination and this volume recombination for this case?

I will just finish this before we discuss before we go further discussion on the effect of beta with…. Surface recombination is q n_p0 S into A_s by 2 because q_n p n_p0 by 2 is the average carrier concentration, S is the proportionality constant, A is the surface area. Instead of q by tau we have got q into S. Volume recombination is… notice here this is the surface area, this is volume area, divided by the two we get this. How much is the ratio between the surface and the volume recombination depends upon now recombination velocity and ratio of surface to the volume surface area to ratio this area? How much this surface area depends upon that.

In fact, if we make this area smaller and smaller, we will get that ratio larger and larger. I think I will not get past this today, but all that I am pointing out is how much is the surface recombination depends very critically on this ratio and also the A_s. So when we make a device we must take care of this. So I will continue on this discussion in the next lecture, very interesting features of this.