Digital Integrated Circuits
Dr. Amitava Dasgupta
Department of Electrical Engineering
Indian Institute of Technology, Madras
Lecture – 16
Schottky Transistor Logic; Heterojunction I2L

We shall continue our discussion on I squared L. We have seen that I squared L although it has lot of attractive features like low power dissipation, high packing density, the main drawback was speed and there has been a lot of effort in that direction, how to improve this speed of I squared L and last class we saw how improvement in technology helped to reduce the speed by having a different geometry where you reduce the areas, where excess charge is stored etc. Now there has been another direction in which work is gone on to reduce the propagation delay or improve this speed and that is by making certain modifications in the circuitry itself, small modifications and we shall take up a few case now.

So one of the variations I should say I squared L has been the short key transistor logic. This is the particular logic family Schottky transistor logic or STL and this is variation of the well-known I squared L. The basic circuitry, basic cell of this short key transistor logic looks like this. So you have diode here, short key diodes and this is the transistor. You can well recognize that this is short key transistor in fact. It is the bipolar transistor with a short key diode between the base and collector. I have drawn it out including instead of just drawing the symbol of a short key transistor because we would require this for some calculations. Here you have the input, here is the V_{cc}, you have a resistance here so this is another difference from the normal I squared L that is instead of a PNP transistor acting as a current source, you have a resistance.

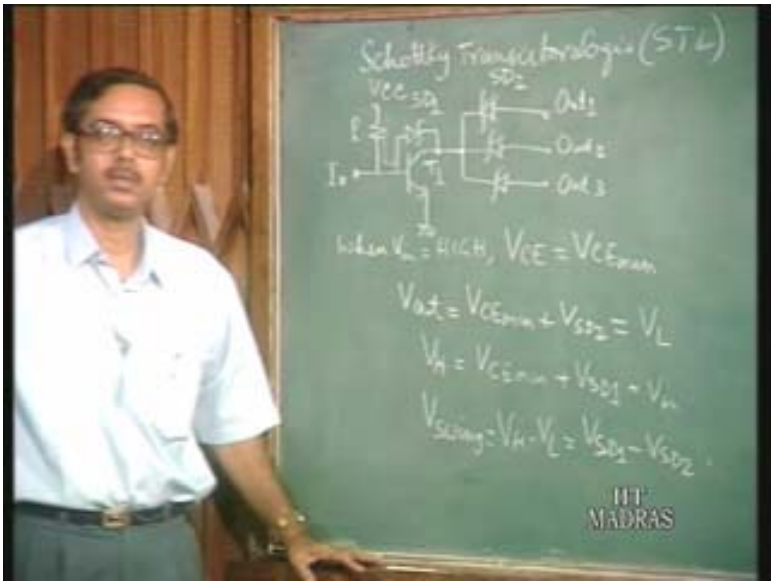
So in fact you will require resistance in this circuit and each cell should have a separate resistance because you cannot have a common resistance in which case all the input points should be basically shorted. So that is the major difference. So this is out one, out two and out three say the output is taken through short key diodes.
So common collector point and then you have short key diodes because just like you have multiple collectors in I squared L, it is a necessity so you must have multiple outputs in this case and this is how it is taken in this particular logic. So this short key diode let me call it SD one and these short key diodes let me call SD_2. Now let us see how this works. The principle of operation is almost the same as an I squared L. So if this input is low, what happens this out transistor is going to be cut off. The current actually flows this way into the input. Basically you have a saturated transistor at the input, this transistor is cut off and the output is going to be high.

On the other hand if the input is high that is at this point, you have that is the output of an I squared L cell where the transistor is cut off. So this input is high that means this current will flow into the base of this transistor. This transistor goes to saturation and the output will be low. So that is how it works, basically it's the same as the I squared L. The small difference

here is that you have short key diodes of the output which is going to give you an additional voltage drop. Now whether it is going to work that is what we have to see. What is the logic swing that is what we want to see first? Logic swing means voltage, the difference between the logic high and logic low. Now when the input is high say so we have to see what is the voltage levels, logic high and logic low. When the input is high say when V_{in} is high recall it T_1 does not go to saturation because of the short key diode. I should say V_{CEmin} that is the collector emitter voltage of this transistor goes to V_{CEmin} when the input is high. What is the output voltage? V output is equal to V_{CEmin} plus the drop across this short key diodes, V_{SD2} and this is actually logic low because when this transistor is conducting, the output is to be low.

When this transistor is off, the output is going to be high. So this is the logic low when this goes to the minimum value, you see in a short key transistor when its starts conducting the output voltage is going to fall but because of the short key diode across the base and collector, it can go only up to a point it cannot go below that which actually prevents it from going to saturation. So you know that, so this is V_L and now what is V_H that is logic high? Now in this case when this transistor is conducting the input voltage is actually logic high because this side the transistor is cut off. So whatever is the input voltage here is regarded as logic high.

(Refer Slide Time 10:15 min)



So what is V_H? V_H is again V_{CEmin} plus V_{SD1}, this is V_{CEmin} plus this drop that is the input voltage. So this is the logic high that is at the input. So this is actually V_{in}. The logic swing that is V_{swing} is V_H - V_L is equal to V SD one minus V SD two. You see the logic swing has been reduced even further, when you compare with an normal I squared L because of the fact that this V_{CE} which was normally going to 0.2 volts when the transistor goes to saturations, since this transistor does not go to saturation this is going to be higher plus you are having an additional drop because of this.

So the logic low itself is going to be much higher than the normal I squared L and the logic high itself cannot be more than the V_{BE} here. So the logic swing is actually in this case you see it is $V_{SD1} - V_{SD2}$. That is the difference in the drops of the two short key diodes. You know what is the short key diode. A short key diode is a junction between a metal and a semiconductor and the interesting thing is that the cut in voltage that is when the diode conducts, the voltage at which it starts conducting that cut in voltage actually depends on the material concern. You can adjust the cut in voltage by using different metals. so in fact this short key diode one, what they have used in this logic; short key diode one is usually platinum silicide one silicon and short key diode two is an alloy of titanium and tungsten on silicon and this gives a difference of the cut in volt V_{SD1} and V_{SD2} of around 200 milli volts. $V_{SD1} - V_{SD2}$ that is the logic swing is around 200 milli volts, 0.2 volts basically for these two combinations. So that is how this circuit works. Now so that is about the circuit. So it is basically like an I squared L. One advantage you can see directly here, it's a short key transistor so it is a non-saturated logic. It is going to be faster in the sense that the charge storage problems would be reduced.

Of course another advantage also is from the low logic swing, you know that from the initial classes what we did this logic swing also has an effect on the power delay product. The power delay product is directly proportional to the logic swing that is basically because of the fact, the logic swing means the difference between low and high. So when the output voltage changes from one level to the other, from low to high the output capacitance has to charge through only a small 0.2 volts. So it is going to be faster. If it has to charge through a higher voltage, it would take more time. So for any high speed circuits one has to sacrifice margin and the logic swing has to be reduced. So that is also another point in favour of I mean if you are looking at it from a high speed circuit point of view.

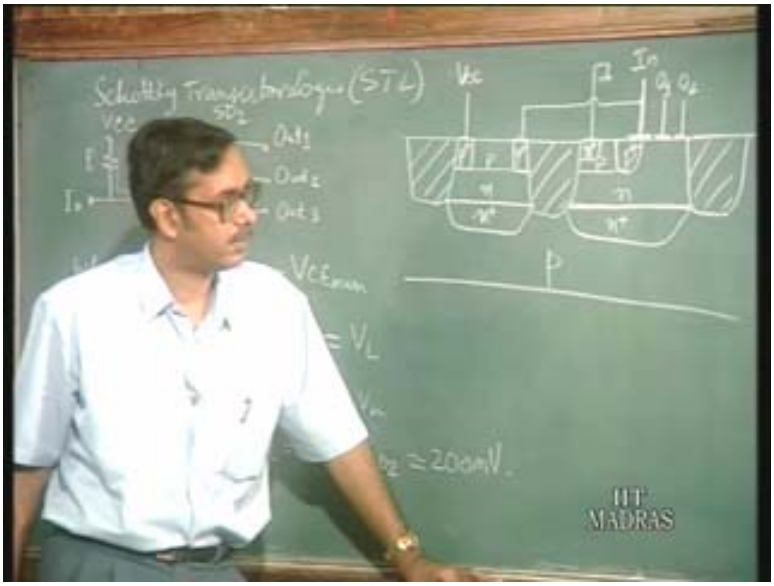
The other advantage we shall see, if you look at the circuit, the cross section of this cell is how you really fabricate it. I shall draw the cross section, it is slightly more complicated when compared to the standard I squared L cell. So what you have is again, it is like a bipolar junction transistor, the starting wafer. You have a P type substrate as you have seen in a normal bipolar junction transistor, this is the berid collector and this is n. So that is the normal bipolar junction transistor and this is actually silicon dioxide and that is used for isolation. So like a bipolar transistor you start with the P type substrate, make the berid collector then you grow epitaxial n region where you actually make the bipolar transistor.

So then you make the oxide isolation, when you have the base region P, so this is the base P region and for the P contacts you make P plus. So here you also have a P plus, this is P plus for contact and then you have the n plus here. So this is n plus P n. Now you have the transistor, if you look at it NPN transistor this is the emitter n plus PN. Now I will draw the contact, this is one contact then I draw two outputs. This is the emitter of the NPN which is actually grounded. This between here and here forms the resistance. This is the resistance, this is the input pointer (Refer Slide Time: 16:10), this is the output I will just put O here, O_1 and O_2. These are the outputs say, this is the input, this goes to V_{cc}.

So now if you go back to the circuit V_{cc}, this is the resistance, this P region forms the resistance which is the same as the base diffusion of the NPN transistor that is done simultaneously. So this is the resistance which goes to the input point which forms the base of the transistor. Now in a short key transistor how do you make the short key? So this is the collector region, so the base contact, there is an over hang on to the collector regions. So basically this extension of the base contact over the collector makes it a short key transistor. So this is the NPN transistor, this is the base contact and this is the short key transistor and this is the collector region and these are short key diodes. This metal on to the semi-conductor, this particular metal which you are putting here, so this is one type of metal which you have for the base contact and there is another type of metal which you are putting for the output because this is another short key diode. So this is one short key diode to the collector. So you have multiple short key diode connected to the collectors, this is n region. So common n region collector for the NPN transistor and you take the outputs through short key diodes. So this is the structure of the circuit, this is the cross section view of the circuit.

Now you can see another advantage, if you look at this figure compared to the normal I squared L is the fact that this particular NPN transistor is a normal NPN transistor. In I squared L you have the emitter down structure. Here is a normal emitter up structure where emitter is on top and the advantage is that this is heavily doped emitter, the charge storage will be less. In the emitter down structure, the emitter area was very large, lightly doped emitter where in the result there is a lot of charge storage. This is a normal NPN transistor which is a much faster device compared to the I squared L NPN emitter down NPN transistor. So these are the major advantages from the point of view of high speed device. Number one it is a short key transistor, so charge storage problems will be less, transistor not going to saturation, the logic swing is less with the result it is going to be faster. I know logic swing less means it is obviously faster and number three because the NPN transistor is a emitter up structure normal BJT structure, you have less problems of charge storage in the emitter as you have in the normal I squared L structure where we have a lightly doped emitter at the bottom, emitter of the bottom is larger area, so larger charge storage.

(Refer Slide Time 20:11 min)



So all these factors make it a high speed circuit and delays of less than one nanosecond can be very easily achieved using this particular structure. So that is the advantage of the short key transistor logic. Of course the disadvantage is you can see that compared to I squared L the technology is the bit more complicated. The simplicity is not there, you have to do I mean some more complicated technology but there is an advantage in terms of speed. I will just give you a reference in case if you want to do a further reading on this. This is I mean just D.J.Roulston and M.Depey electronic letters. So this particular structure was proposed in this paper volume 19, pages 21-22, January 1983. So this is the reference for this particular logic circuit. So that is about short key transistor logic which is considered to be a high speed version of the family of I squared L circuits. So that is about one of the ways, people have tackled the propagation delay problem of I squared L.

The other area where people have done a lot of work and in fact lot of work is also going on presently is by using hetero junction devices and you have what is called hetero structure I squared L (H I squared L). So before going into that circuit itself, I think I should say a few words about this hetero structure. What is hetero structure, how it improves the performance I think it is necessary. Now as the name implies the hetero structure is different from a homo structure, the normal devices which you have are called homo junction devices because the entire devices is made up of the same material, maybe silicon or mostly it is silicon. So the entire device is made out of silicon but if you have different regions of the device made out of different materials, for example if emitter is one material and the base is a different material, you have a hetero structure, this is the two different materials. How do you gain from having a hetero structure? That is the thing which we have to see. Now if you have a homo structure that is the same material the band gap of the entire structure in the entire device has to be the same because it is the same material.

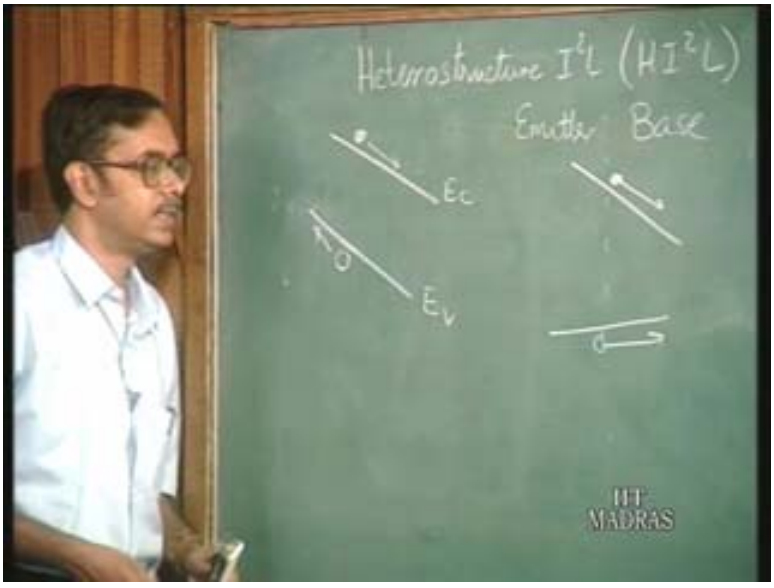
Now if you have a homo structure, the band will always be like this. If you draw the conduction band and valance band and it is changing. The conduction band has always to be

parallel with the valance band, I think it is clear. That is because the band gap is the same everywhere, so it must be parallel which means that if you have an electron in the conduction band and the hole say in the valance band, the forces acting on the electron and the hole must always be equal and opposite because the forces is proportional to the slope.

So the electron will move this way, the hole will move this way and the forces are always equal and opposite. So the current actually will depend on the number you have for different regions but the forces are equal and opposite but now if you can do some, what we call band gap engineering this particular, that is can change the band gap. You can modify the forces acting on the electrons and holes. So in the same region if you have electron and holes you can have different forces acting on that. For example if you have a band structure like this and if you have an electron here and the hole here, the force on the electron is going to be this side and the hole also it is going to be this side because of the slope. Now which means that you can modify the forces acting on the electrons and holes.

Now you can see for yourself that suppose you have an emitter base junction and this side is the emitter and this side is the base and is an NPN transistor, what you are basically doing by this structure is you have the electrons are going from the emitter to the base and the holes which constitute the base current, if you have a band like this you are preventing it from going from the base to the emitter. There is a force acting in the holes which is in a deduction opposite. No, it is force which is preventing the holes going from the base to the emitter. Normal structure it would have gone from base to emitter. So what are you doing basically?

(Refer Slide Time 27:18 min)

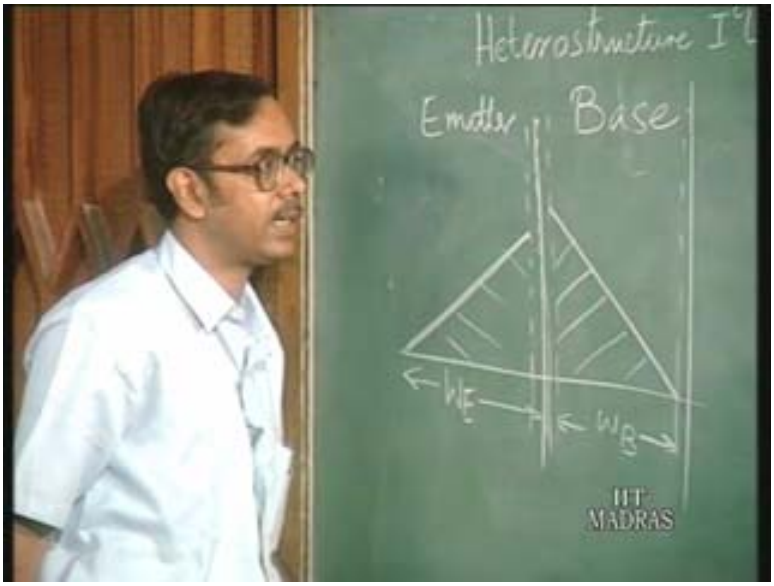


You are going to increase the beta of the transistor because the collector current which is basically due to electron moving from emitter to base that is going to get increased whereas the holes, the base current is going to get reduced because the force acting on the holes is preventing it from going to the emitter. So that is in essence that what you do in a hetero junction bipolar transistor, you have a wide band gap emitter, the emitter material is made up a wider band gap material and the base of a narrow band gap material.

So if you have a structure like this, you can increase the gain of the bipolar transistor by having a wider band gap emitter. So I shall again go into slightly more quantitative analysis. If we take emitter base junction again NPN transistors say emitter base junction, this side is emitter, this side is base and if I plot the charges on the emitter side so this is the contact say you have the excess charges. I am just plotting the excess charges. So this is say the collector, so the charges may be like this. So these are minority carrier stored charges on the emitter and the base.

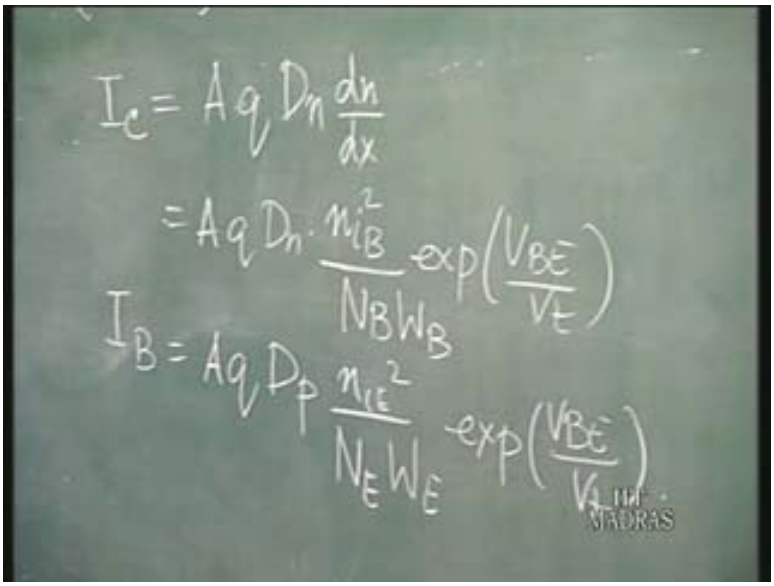
Now if we neglect the recombination in the base, we can write an approximate relation for the beta of the transistor. So this is say the emitter width from here to here and this is the base width WB, WE is the emitter width, WB is the base width. Now what is the collector current approximately? Say I am writing the collector current, I will write I_C is equal to A that is the area, q D_n that is the diffusion coefficient and the slope of the electron distribution in the base. So that is the diffusion current, this is well-known. You have already seen that. So what is dn dx that is what is of interest to us. The dn dx is the slope of this base profile.

(Refer Slide Time 29:46 min)



So the slope is nothing but the height of this triangle by the base. The height of the triangle is going to be given by Aq D_n, so the height of the triangle is n_i squared, I write B here because it is in the base region. That is n_i squared by the doping in the base. What is n_i squared by N_B is the minority carrier concentration at thermal equilibrium into exponential V_{BE} by V_t. That is the value at the depletion region edge, on the emitter base junction depletion edge. So this is the height of the triangle and the base divided by W_B which is the base of the triangle so that is dn dx. Similarly if I want to write I_B and I assume that the base current is primarily due to the injection of holes from the base to the emitter.

(Refer Slide Time 31:31 min)

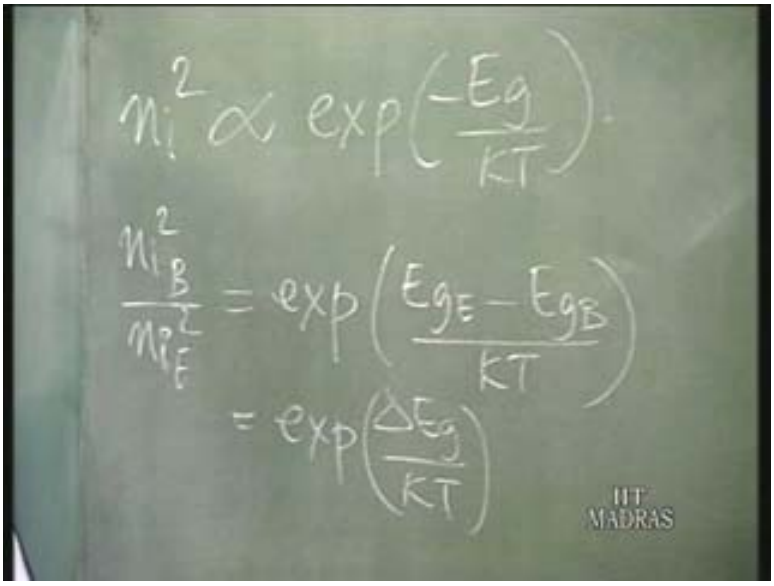


This slope will give us the base current and so by similar analysis I_B will be equal to Aq instead of dn you have Dp because it is diffusion coefficient of holes, n_i E squared by N_E that is the emitter. N_E is the doping concentration in the emitter, W_E is the emitter width, again exponential V_{BE} by V_t. Of course I neglect the recombination in the base which is going to be true especially if you have very short base width which is in fact true mostly in the case of hetero junction bipolar transistors.

Now so what do you have now is beta is I_C by I_B... So you have dn by dp then you have n_i B squared by n_i E squared N_E by N_B W_E by W_B. So it is almost the same relation as for homo junction except for that fact that we have written it in terms of n_i B square and n_i E squared. Why we have written is normally in a homo junction, the n_i is the intensive carrier concentration. In the emitter and the base would be the same but in a hetero junction they are going to different because of the difference in band gap. Now you know that n_i square that is the intrinsic carrier concentration depends very much on the band gap. If you have a larger band gap material, the intrinsic carrier concentration is less that is already we have seen that in the first few classes.

I think in fact n_i squared is proportional to exponential minus E_g by KT, n_i is proportional to minus exponential minus E_g by twice KT, so n_i squared is proportional to exponential minus E_g by KT. So what do you have is n_i squared B by n_i E square, this is going to be equal to exponential E_{gE} of the emitter minus E_{gB} of the base by KT and this is sometimes referred to as delta E_g that is delta E_g is the difference of the band gap of the emitter material and the base material, exponential delta E_g by KT.

(Refer Slide Time 34:39 min)



So you see that if you look back into this expression now, beta is equal to I just rewrite it again Dn by Dp N_E by N_B W_E by W_B exponential delta E_g by KT. Now what is this value of delta

E_g? If you take a normal system HPT system with the most commonly used hetero junction materials, if which is in use is aluminum gallium arsenide on gallium arsenide. Traditionally has been used as the hetero junction HPT material, of course lot of other materials which have been used later.

So if you take aluminum gallium arsenide that is aluminum 0.3, gallium 0.7 arsenide, this is the material on gallium arsenide. So gallium arsenide, for each gallium atom you have one arsenic atom that is the compound. Now what do you do is algasis, 30% of the gallium atoms have been replaced by aluminum in gallium arsenide. So for each arsenic atom or you can say for every ten arsenic atoms you have seven gallium atoms and three aluminum atoms. So it is written like this aluminum 0.3, gallium 0.7 arsenic. So it's a ternary compound of three materials and this material algas for this, the band gap difference delta E_g is 0.374 electron volts. Now if you take exponential delta E_g by KT for delta E_g as 0.374, at room temperature the value is almost 10 to the power 6 because the KT is just 26 milli volts. So this is an enormous number 10 to the power 6. So you see that by just changing the hetero junction material, the emitter material and having a wider band gap material and having a variation in band gap, you can have very large increase in beta.

So in fact I mean you cannot really have such high betas because once you have in this relation, what is going to happen is then the beta will be determined. Basically what we have been talking of its emitter efficiency, if this ratio becomes so large beta will be determined by the other factor that is the base transport factor that is the recombination in the base. So the beta is no longer going to be determined by this and so it is no longer advantageous to have such a high value of this factor.

So what you can do is now suppose in a homo junction, we have N_E by N_B that is the doping concentration of the emitter to the base is made very large. The emitter to the base doping concentration in a homo junction is made very large that is necessary to achieve high betas, we have already discussed that. Now for a homo junction transistors say N_E by N_B is 10 to the power 2 that is 100 and you have a corresponding hetero junction transistor say where N_E by N_B is equal to 10 to the power minus 2. That is the base doping concentration is more than the emitter doping concentration by a factor of 100, even then because in a hetero junction you have an extra factor of 10 to the power 6.

(Refer Slide Time 40:34 min)

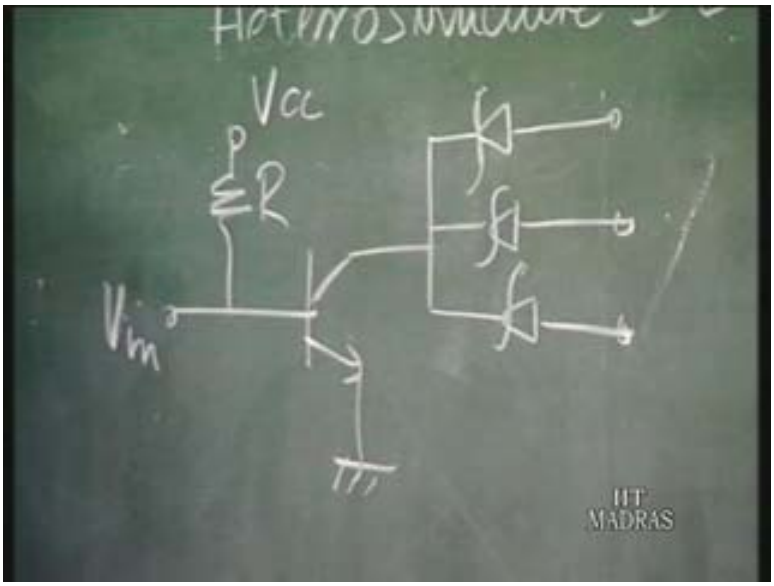
The beta if you just consider it to be determined by emitter efficiency is still going to be hundred times more than the corresponding homo junction. So in a hetero junction bipolar transistor because of the large gain which is achieved due to the difference in band gap, usually the base doping concentration is made much larger than the emitter doping concentration that is quite normal in a hetero junction bipolar transistor. It has got its lot of advantages, I need not go into all that because that takes a lot of time. Some of the advantages we can think for (Not understandable) the base resistance is reduced which is very important in high speed transistors because you have a higher doping concentration in the base, you can go for shorter base widths and you don't have to worry about base punch through an other effects which you would normally have to worry, if you have a shorter base width.

So you can actually have very short base width if you have high doping concentration in the base because if the doping concentration is high, the depletion width is going to be less and we don't have base points through. So there are lots of other advantages. So in a normal hetero junction bipolar transistor, you have a wider band gap emitter and a lightly doped emitter and the base doping concentration is much higher. So you can very well imagine that in an I squared L where the problem was that you have a lightly doped emitter and you had a lot of charge storage there, it is no longer a problem because of the fact that you now have the wider band gap emitter material and you know that basically the charge storage is proportional to n_i squared by N that is n_i square by N where N is the doping concentration in that region.

The problem was that n was less, the charge storage is proportional to n_i square by N where N is the doping concentration any region. The problem was because the doping concentration is less, the minority carrier concentration would be more but here n_i squared itself is going to be very much less because of the wider band gap material. So the charge storage problem is not of so much concern anymore because that is normally what is the expected in a hetero junction bipolar transistor, you have a lightly doped emitter. Also you have a heavily doped base region, so which also reduces the minority stored charge in the base. So that is the advantage of a hydro junction bipolar transistor.

Basically the structure is like this, you have heavily doped base and you have a lightly doped wider band gap emitter. So this hetero junction has been used in I squared L to improve the speed performance of I squared L circuits. In fact lot of recent works on I squared L which is come out has been in this area of hetero structure I squared L and you can see that why you should have a better performance if you use hetero junction. I think it is clear now. So the hetero structure I squared L has a structure which is quite similar to the short key transistor logic circuits which we had already seen and it consists of resistance and this is the BJT and you have the short key diodes here. In fact this is not a short key transistor, this is the normal transistor that is the only difference between a short key transistor logic circuit and this hetero junction I squared L circuit. So normal transistor you have short diodes here and so this is how it is going to work.

(Refer Slide Time 45:40 min)



I think in the next class, we shall take up the circuit or I think if I have some time I can just continue. So this is V_{in}, this is V_{cc}, this is R. So for saturation we have to ensure that basically this is a saturated logic and you can imagine that if this transistor goes to saturation there is no problem, this is point two. Now we have to ensure that this transistor actually goes to saturation. So for saturation beta times I_B must be greater than I_C this is the well-known

relation. Now what is I_B? I_B is equal to V_{cc} minus V_{BE} by R. What is I_C? This is going to individual bases, so you have V_{cc} minus V_{CE} plus V_{SD} by R into the number of fan out, the number of collectors. So the each of them is going to conduct, when this transistor goes to saturation this is connected to the base, as input to the next stage say. So current will flow this way for the next stage. So what is the current which is flowing? V_{cc} minus this voltage here which is V_{CE} plus V_{SD} divided by R. So the number of fan out here in this case you have 3, so you have three times that current. So that is the total current which is going to flow. Now if you put it here, so beta must be greater than I_C by I_B. You have to ensure that the beta of the transistor is greater than V_{cc} minus V_{CE} plus V_{SD} by V_{cc} minus V_{BE} into fan out.

(Refer Slide Time 49:32 min)

So once this is ensured, the transistor actually goes to saturation. What are the different voltages? Usually in hetero junctions as I said, you have aluminum gallium arsenide and gallium arsenide, the band gap of gallium arsenide is 1.42 electron volts compared to 1.1 for silicon. Aluminum gallium arsenide is much larger, its almost 1.8 electron volts. So these voltages V_{CE}, V_{BE} and even V_{SD} that is the short key diode on gallium arsenide, the cut in voltages of any junction it depends on the band gap of the materials. So these are going to be much larger compared to that of silicon. So in fact we shall continue with this discussion in the next class and we shall look at the structure and see some of the results on these hetero structures I squared L. In fact there has been also very resent results, some people are doing lot of work in this and maybe this is the area where I squared L is really going to make a mark in future.