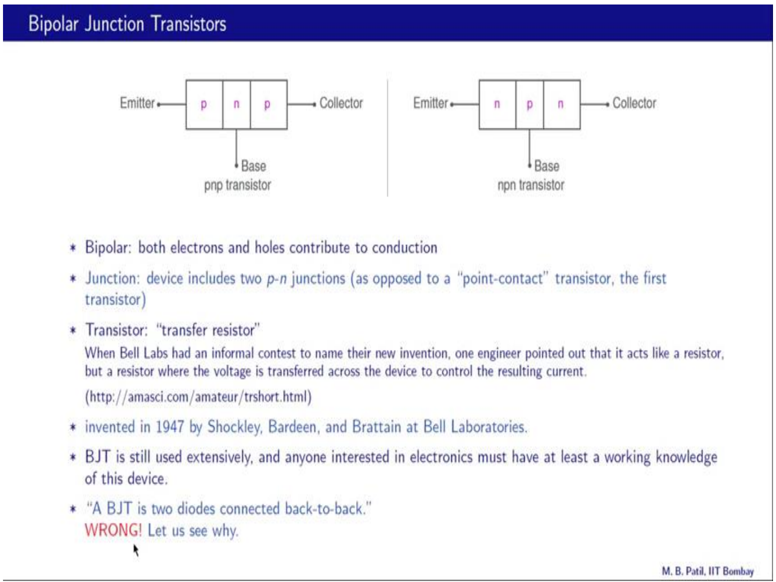
Basic Electronics

Prof. Mahesh Patil
Department of Electrical Engineering
Indian Institute of Technology, Bombay

Lecture - 22
Bipolar Junction Transistor

Welcome back to Basic Electronics. We will now start our discussion of the Bipolar Junction Transistor or the BJT in short. The BJT is much more complex than the diode. And in the next few classes, we will see how to represent its behavior for circuit analysis. In this class, we will start with the BJT symbols for the two types of BJT namely npn and pnp-BJTs. We will look at a simple model to represent the BJT in the so called active mode. So, let us get started.

(Refer Slide Time: 00:54)



Let us start our discussion of the bipolar junction transistors or the BJT in short. BJTs are of two kinds' pnp transistor and npn transistor. In the pnp transistor, we have a p-type semiconductor region followed by an n-type region and then a p-type region again. Here we have an n-type region then p-type region and again an n-type region. And in both cases, there are three contacts emitter, base and collector. And we should note that the schematic diagrams shown here are only representative. In reality the transistor structures are very different and in fact these two junctions this p n junction and this p n junction are also very different from each other, and that has significant implications on the performance of the BJT as we will see later.

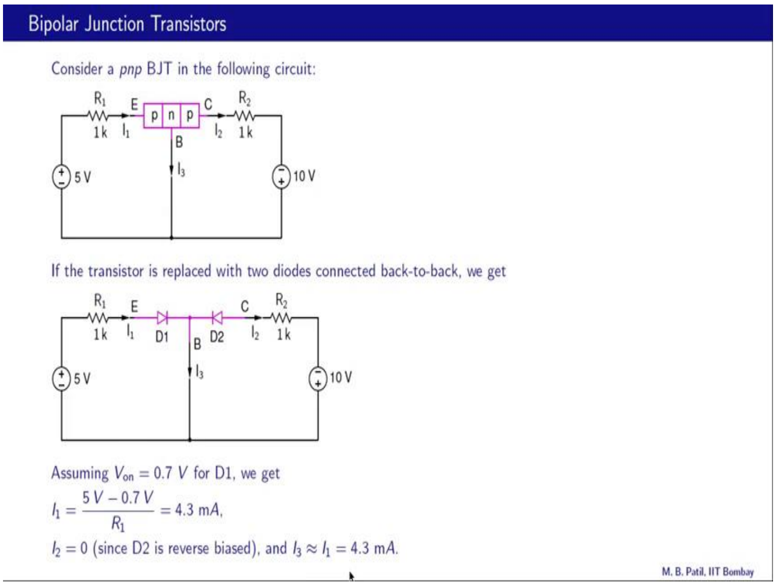
Let us now look at why BJT is called bipolar junction transistor. Bipolar because both electrons and holes contribute to conduction in these devices as opposed to for example, MOS transistors in which only one type of carrier contributes. Junction, because the device includes two pn junctions one here and one here and this should be contrasted with the point contact transistor which was the first transistor to be invented by Shockley and others in 1947. This point contact transistor is not used anymore. And finally, this word transistor stands for transfer register, and you can read more about that at this website.

As we mentioned the bipolar transistor was invented in 1947 by Shockley, Bardeen and Brattain at Bell labs. And the device they came up with was substantially different than the bipolar junction transistor that we use today. And over the years, the bipolar transistor technology has matured following several developments and innovations. And today the BJT is a very commonly used device both as a single transistor that is discrete transistor and also as part of many, many integrated circuits.

The bipolar transistor is still used extensively, and anyone who is interested in electronics must definitely have at least a working knowledge of this device; and a word of caution here because the BJT has been around for so many decades invented in 1947. It has now become somewhat fashionable to say that oh it is an old device which is not so exciting and so on, but that would be a really big mistake. In reality BJT is the workhorse for analog electronics; and it is used in a huge number of applications even today. It is true that the MOS transistor has replaced the BJT in digital chips such as processors or memory, but analog circuits is a completely different story, and BJTs are used in many, many analog circuits.

And if someone is really serious about electronics, he or she must definitely develop at least a working knowledge of the BJT. So, let us treat the BJT with the respect it deserves; and to begin with let us see how it works, here is a statement that one might encounter in textbooks and so on. A BJT is two-diodes connected back-to-back. It seems to make sense, we have a diode here another diode here connected back-to-back, a diode here another diode here again connected back-to-back, but there is something seriously wrong with this statement.

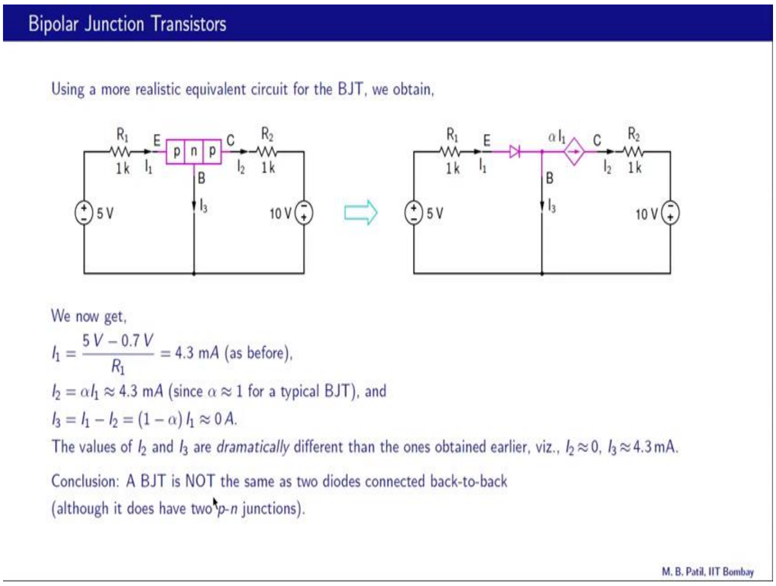
(Refer Slide Time: 05:39)



Let us see why. Let us look at this simple circuit involving a pnp transistor. Let us replace the transistor with two-diodes connected back-to-back like we were discussing. So, here is one diode pn that is D 1, here is another diode pn again D 2, and there n ends are connected together like here. Now, let us find I 1, I 2 and I 3 in this situation. We notice that the p-end of the diode D 1 is at a positive potential as compared to its n-end which is at a negative potential. So, therefore, we expect D 1 to be under forward bias.

And let us assume that V on for this diode is 0.7 volts. So, then we have I 1 is equal to this 5 volts minus 0.7 volts drop here that divided by R 1 which is 1 k. So, I 1 turns out to be 4.3 milliamps; and I 2 is negligibly small, because D 2 is under reverse bias, it is p n is negative as compared to the n terminal. And therefore, D 2 is under reversed bias, the current is negligibly small and so therefore, I 2 is 0. I 3 is then nearly equal to I 1, because this I 2 is 0 and therefore, I 3 is also about 4.3 milliamps.

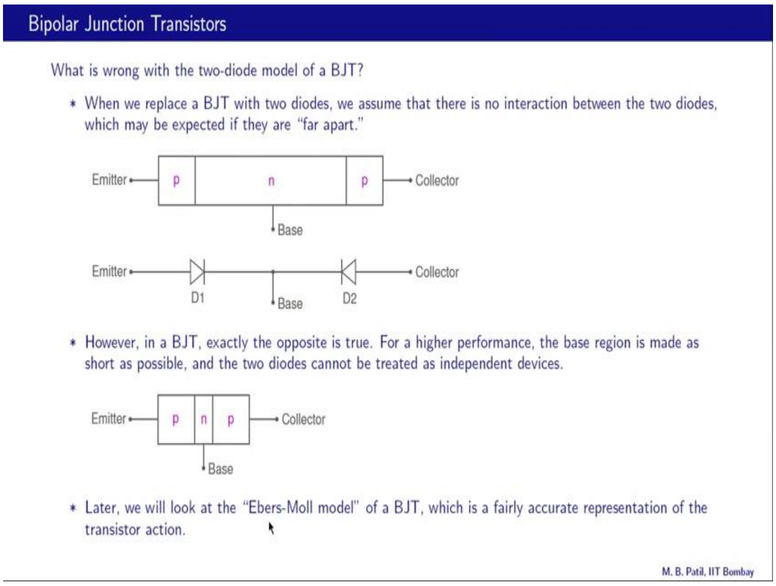
(Refer Slide Time: 07:32)



Let us now look at the same circuit, but with the transistor replaced with a more realistic BJT model, this one. In this model we have a diode between emitter and base as before, but between base and collector now we have a current controlled current source. And if this current entering the emitter is I 1 that current is alpha times I 1. So, there is a big difference between the model that we saw in the last slide and this one. Let us now look at I 1, I 2, I 3. As far as I 1 is concerned nothing has changed really, we have 5 volts here of which 0.7 will drop here and I 1 will then be 4.3 divided by 1 k, which is 4.3 milliamp.

What about I 2, I 2 is given by alpha times I 1 and alpha is nearly equal to 1 for a typical transistor. So, therefore, I 2 is also equal to about 4.3 milliamp. What about I 3, I 3 is I 1 minus I 2; now I 2 is alpha times I 1, so therefore, this is 1 minus alpha times I 1 and since alpha is close to 1, I 3 will turn out to be 0, very small. Now, these values of I 2 and I 3 are dramatically different than the once we have obtained earlier in the last slide that was I 2 equal to 0; and I 3 equal to 4.3 milliamp. Compared that I 2 with this I 2, and that I 3 with this I 3, they are very, very different. And in the conclusion, therefore, we can say that BJT is not the same as two-diodes connected back-to-back although it does have 2 pn junctions.

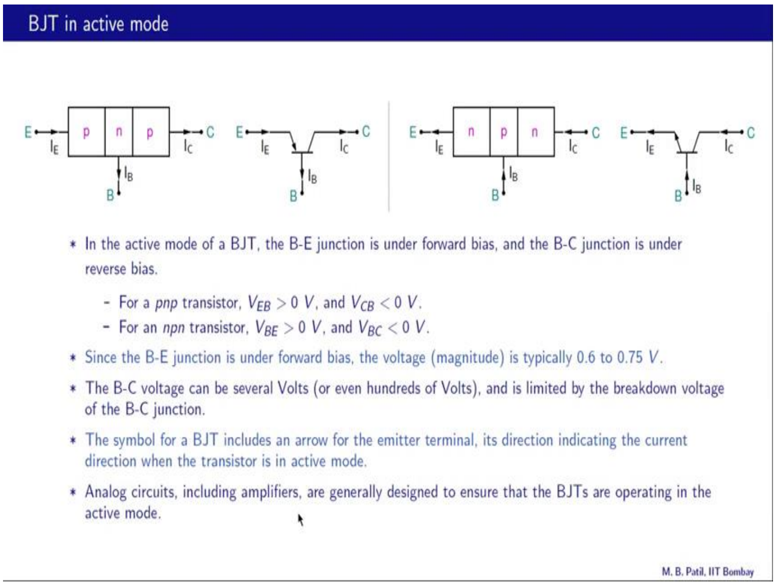
(Refer Slide Time: 09:36)



Let us now see what is wrong with the two-diode model of a BJT. Here is the two-diode model two-diodes connected back-to-back. What it implies is that this n region the base region is very long, and that these two junctions the emitter-base junction and the collector-base junction are far apart and that implies that there is no interaction between these two-diodes.

Now, in reality exactly the opposite is true. So, this region is in fact, made very narrow for higher performance; and the two-diodes cannot be treated as independent devices. And later we will look at the Ebers-Moll model of a BJT and understand the behavior better. This model is a fairly accurate representation of the transistor action, and it is still simple enough to understand.

(Refer Slide Time: 10:47)



Let us now consider the BJT in the active mode. What is the meaning of active mode, here is a pnp transistor, active mode means the emitter base junction is under forward bias, and base collector junction is under reverse bias. So, that means, this node the emitter node is at a higher potential than the base node and compared to base node the collector is at a lower potential. Similarly, for an npn BJT, the emitter base junction is under forward bias in the active mode that means, the base is at a higher potential than the emitter; and the base collector junction is under reverse bias that means, the collector is sitting at a higher potential compared to the base.

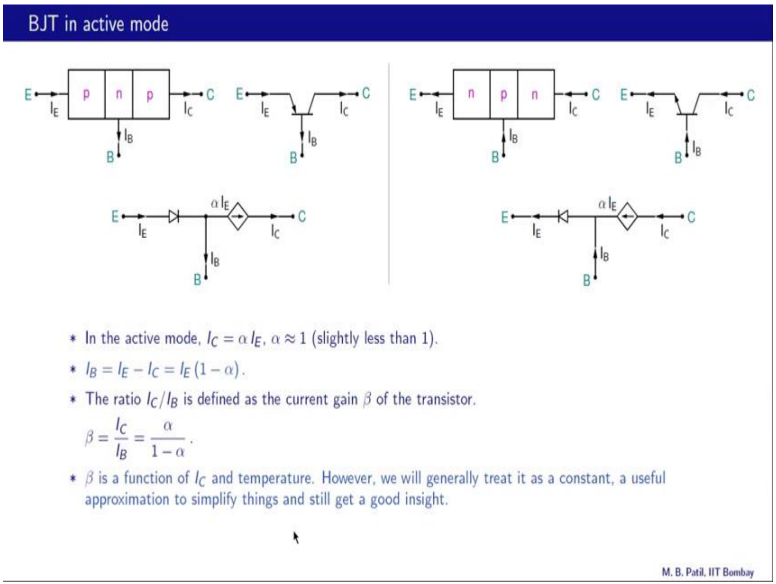
So, this is the summary of what we just said in the active mode of a BJT the base emitter junction is under forward bias and the base collector junction is under reverse bias. And for a pnp transistor it means V E B that is V E minus V B is greater than 0 volts; and V C B - V C minus V B is less than 0 volts. For an npn transistor, the opposite is true V B E is greater than 0. So, V B E - V B minus V E is greater than 0; and V B C - V B minus V C is less than 0.

Since the base emitter junction is under forward bias the voltage is typically in this range 0.6 to 0.75 volts. So, what it means is this difference V E minus V B for the pnp transistor is in that range; and for the npn transistor, V B minus V E is in that range, when the transistor is conducting. The base collector voltage can be several volts or even hundreds of volts, and is limited only by the breakdown voltage of the base collector junction. The base collector junction is under reverse bias and that reverse bias essentially can be a few volts or can be much larger depending on how the transistor is fabricated and it is basically limited by the breakdown voltage of the base collector junction. So, once again base emitter forward bias, base collectors reverse bias in the active mode; same for npn, base emitter forward bias base collector reverse bias.

Now, the symbol for a BJT includes an arrow for the emitter terminal. So, look at this symbol, this arrow here is always at the emitter; and the direction indicates the current direction when the transistor into the active mode. So, when the transistor is in the active mode, this junction is under forward bias; and the current would be in that direction and that is the same as the direction of the arrow.

Similarly, when this junction is under forward bias, the current would be in that direction and that is why this arrow here and diameter is going out from the device. And finally, in analog circuits including amplifiers, we generally have the BJT s operating in the active mode it is, they are designed like that. And of course, we will consider some analog circuits later and we will observe this particular point.

(Refer Slide Time: 15:23)



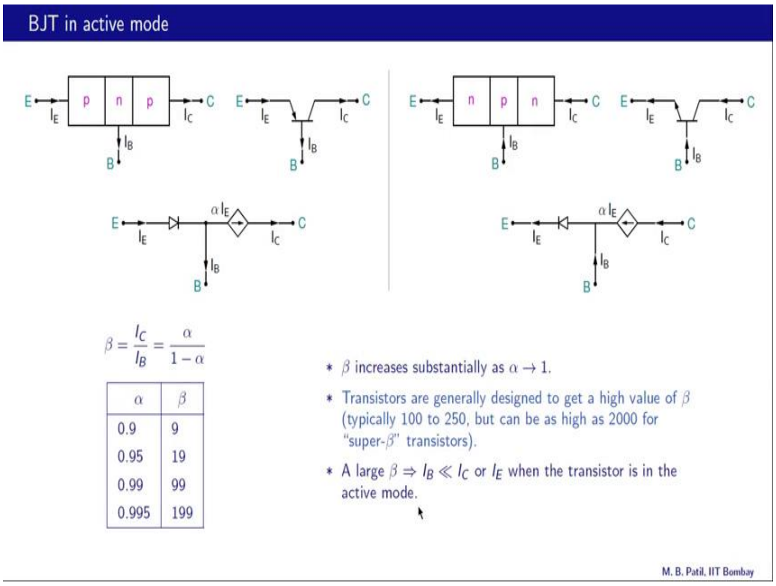
Here are the equivalent circuits for BJT in active mode. This circuit is for a pnp transistor, and this circuit here is for an npn transistor. So, what if we are here between the emitter, and base we have a pn junction that is represented by this diode here; between base and collector we have this current controlled current source if this current is I E that current is alpha times I E. And as we have commented earlier this alpha is close to 1 little less than 1. For the npn transistor, we have a similar situation there is a p n diode here between base and emitter that is a diode; and between the base and collector, we have this current controlled current source. Once again if this current is I E that current is alpha times I E.

Let us note that the direction of currents in the pnp transistor, and in the npn transistor are different. Here if the diode is p n it will conduct always in that direction and the emitter current is in the same direction. Same case here, this diode will conduct always in that direction and the emitter current is also in that same direction. I C is in the same direction as alpha I E in both cases; and I B is coming out of the device in the pnp case and I B is going into the device for the npn case. And an easy way to remember this is I E is equal to I B plus I C in both of these situations.

In the active mode, which is what we have shown here I C is alpha times I E as we have already seen and alpha is approximately 1 slightly less than 1 I B is I E minus I C and since I C is alpha times I E, I B is I E times 1 minus alpha. And we have chosen this current directions, so that this equation is valid for both the pnp and npn case without any change of sign. Now, the ratio I C by I B is a very important one and that is called the current gain beta of the transistor. So, beta is I C by I B, and we already have an expression for I B, I E times 1 minus alpha I C is alpha times I E. And therefore, when we take the ratio of these two that I E will cancel and we get alpha 1 minus alpha.

And in general beta is a function of I C and also temperature in our studies, we will generally treat that as a constant and it is a useful approximation to simplify things and still get a reasonably good insight in the circuit operation. So, this complication definitely should be taken into account when we design circuits for a white temperature range, for example, but in our examples we will say that the temperature is constant usually the room temperature and we will treat beta as a just a constant number.

(Refer Slide Time: 19:41)



Let us now look at the relationship between beta and alpha. Let us calculate beta for a few values of alpha for example, when alpha is 0.9, we have beta equal to 0.9 divided by 1 minus 0.9 that is 0.1. So, 0.9 divided by 0.1 or 9. When alpha is closer to 1, say 0.95 beta increases to 19, still closer to 1, beta increases further; from 0.99, if it becomes 0.995 beta increases even further. So, there is a strong relationship between beta and alpha and beta increases substantially as alpha approaches 1.

Now, most transistors are designed to get a higher value of beta typically 100 to 250 sometimes beta can be as high as 2000, and these transistors are called super beta transistors. We should mention one exception in which the beta is not quite as high as 100 or 250 or could be 50 or less than that and that is power transistors. In power transistors, the main concern is that the transistor should be able to conduct large currents and also it should be able to block large base to collector voltages.

So, all those constraints mean that the value of beta cannot be too large. In our applications, we are only going to consider electronic circuits. So, then our beta values would be in that range. One important implication of a large beta is this. What is beta is I C by I B. So, I B is I C divided by beta so that means, if beta is large, I B would be much smaller than I E or for that matter for I E when the transistor is in the active mode.

In summary, we have got started with a new electronic device in this class namely the bipolar junction transistor or the BJT. We have seen in what way a BJT differs from two-diodes connected back-to-back. We have seen the meaning of active mode of a BJT. In the next class, we will consider a simple BJT circuit, and see how to obtain the currents and voltages in that circuit using what we already know. So, see you next time.