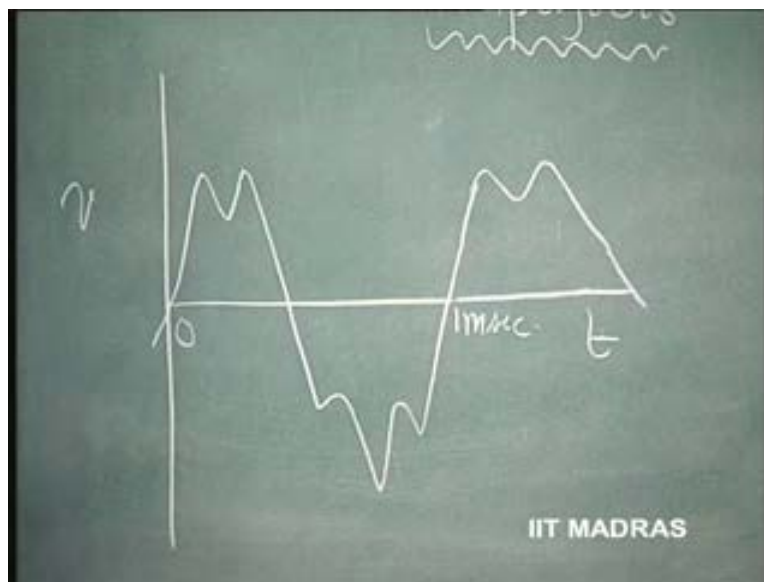


Electronics for Analog Signal Processing - I
Prof. K. Radhakrishna Rao
Department of Electrical Engineering
Indian Institute of Technology – Madras

Lecture - 19
Distortion in Amplifiers

Today, we will continue with our discussion about distortion in amplifiers. Signal distortion; that is what we mean by... For example, let us say, the input signal is something like this. Let us consider a periodic signal like this; and this has to be exactly reproduced. Let us say, this is 1 millisecond, time period.

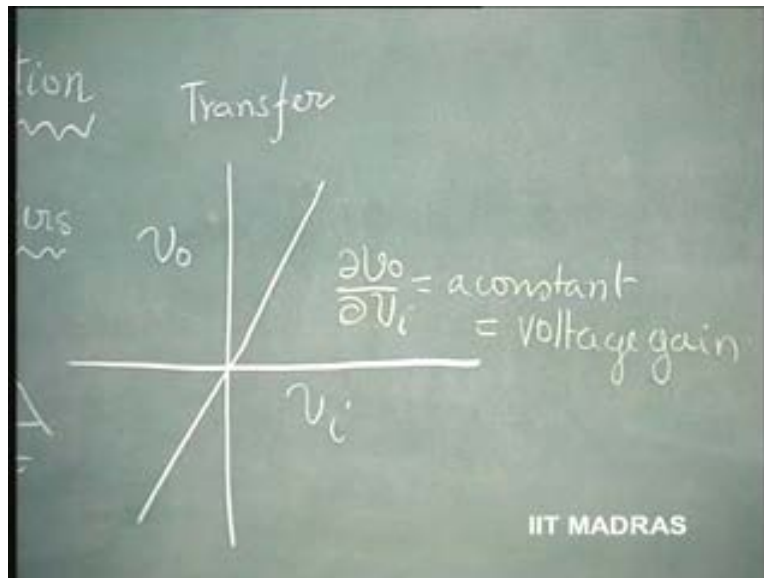
(Refer Slide Time: 02:30)



This is the voltage input. So, this has to be exactly reproduced at the output, following the same shape. This should not be in anyway distorted. Only thing is, it should be amplified; let us say, ten times or hundred times. That is the purpose of the amplifier. If it is, say voltage, it is a voltage amplifier; if it is a current, it is a current amplifier. It could be a transconductance amplifier or transresistance amplifier. You know all these types.

Now, we said that if the amplifier has a characteristic which is, let us say, V_o versus V_i ; this is called transfer characteristic, which is linear; then, we can see that the slope, ΔV_o by ΔV_i is a constant. It is a constant; throughout, this remains a constant. This is what is called as the voltage gain of the amplifier.

(Refer Slide Time: 04:01)



This slope, or this, is a constant throughout. It remains constant at all points of time, at all voltages, voltage levels. Then, it is a linear network. But, this is only possible theoretically. In practice, no amplifier can be that linear as depicted here. Whether it is V_o versus V_i or I_o versus I_i in the case of current amplifiers; or, I_o versus V_i in the case of transconductance amplifiers; or V_o versus I_i in the case of transresistance amplifiers, this slope has to be linear. For most of the amplifiers, we would like to make it as linear as possible by designing the circuit appropriately.

This is done, as I told you, by what is called biasing; proper operation of the device in the proper region, so that, the linearity is as much as possible for as wide a dynamic range. What it means is, it is operable over a large range of voltages; dynamic range. So, this is the purpose of designing amplifiers; to make it as linear as possible. But, the first thing that happens is, in order to bias this we need what is called as DC bias voltages. So, these

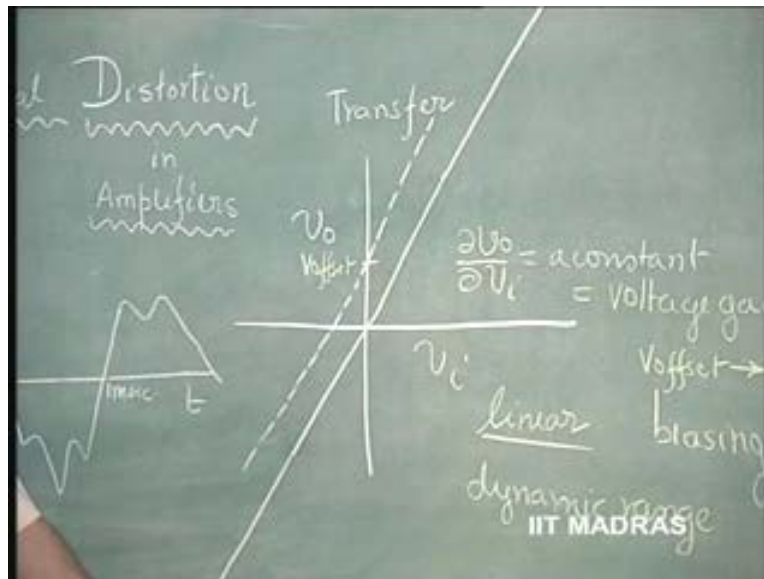
bias voltages will limit the amplitude of output, ultimately. So, these will be, let us say, plus V s and minus V s for dual supply.

(Refer Slide Time: 06:30)



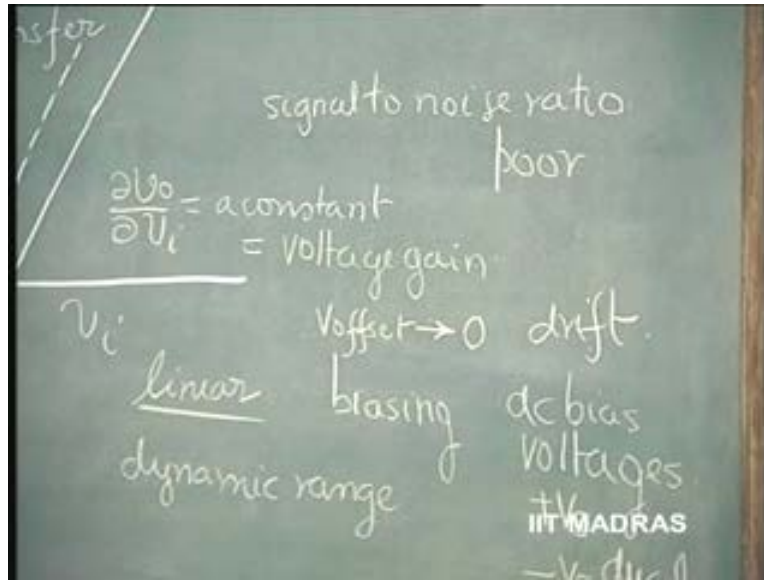
Why do we need dual supply? - so that, the voltage can be both negative as well as positive. And, the offset; this is another important thing. If it passes through the origin, there is no offset. If this characteristic goes like this, the gain is the same; it is linear, perfectly linear. But, there is said to be an offset voltage. So, this is having the same gain as this; but there is an offset. This offset can be positive or negative; so V offset is made very close to zero.

(Refer Slide Time: 07:25)



Why it should be made very close to zero? Because, this offset may not be constant with respect to temperature. You see, the offset voltage may not be constant with respect to temperature. It might be changing depending upon the time and temperature. This causes what is called as drift in offset voltage. So, this drift in offset voltage results in the signal, low frequency signal, appearing such that, we cannot distinguish it from the actual signal, time varying signal that we are using. So, this offset causes drift. This drift causes signal to noise ratio in the circuit to become poor.

(Refer Slide Time: 08:34)



We do not want our amplifier circuit to introduce noise which appears like signal. We do not know. We do not know what the signal is; and that signal is varying with respect to time; that, we are trying to amplify. Apart from that signal that we want to amplify, the amplifier itself is generating a signal which is getting mixed with our signal; and that is the noise.

So, the signal to noise ratio becomes poor because of the drift. This is particularly so in very low frequency amplifiers. In high frequency amplifiers, the drift being a low frequency component, we can isolate it using decoupling capacitors; whereas, in low frequency amplifier, since the signal itself is of low frequency, it mixes with the... what cause, what is caused by drift, and results in signal to noise ratio getting deteriorated within the amplifier.

So, this offset has to be necessarily made zero and kept around zero. The design of an amplifier therefore involves making the offset as close to zero as well as keeping it close to zero at all times, preventing any drift. So, these amplifiers are called low drift amplifiers. These are specially designed DC amplifiers.

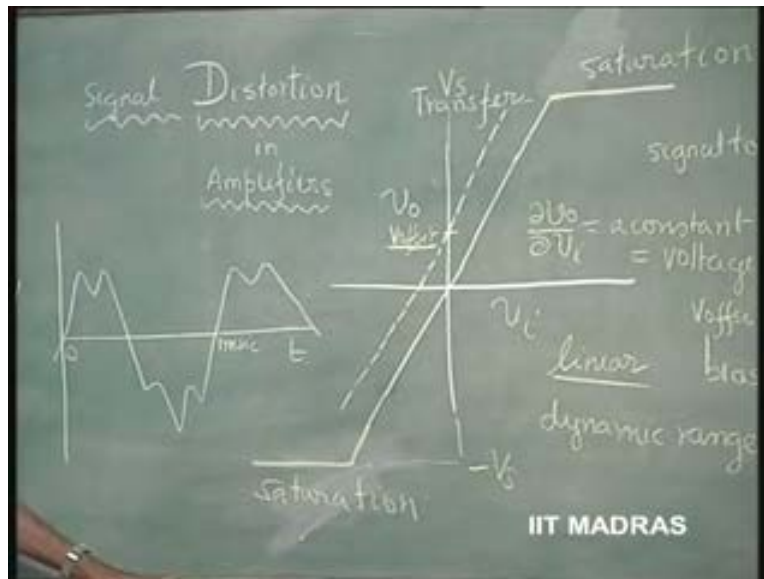
(Refer Slide Time: 10:07)



So, please remember that when we are given the task of designing amplifiers, for a specific type of amplifier, we have to concentrate on specific aspects of design. When it comes to low frequency **low frequency** amplifiers like biomedical amplifiers, etcetera, we are interested in designing low drift amplifiers. These are called DC amplifiers. So, in that case, V offset is an important factor.

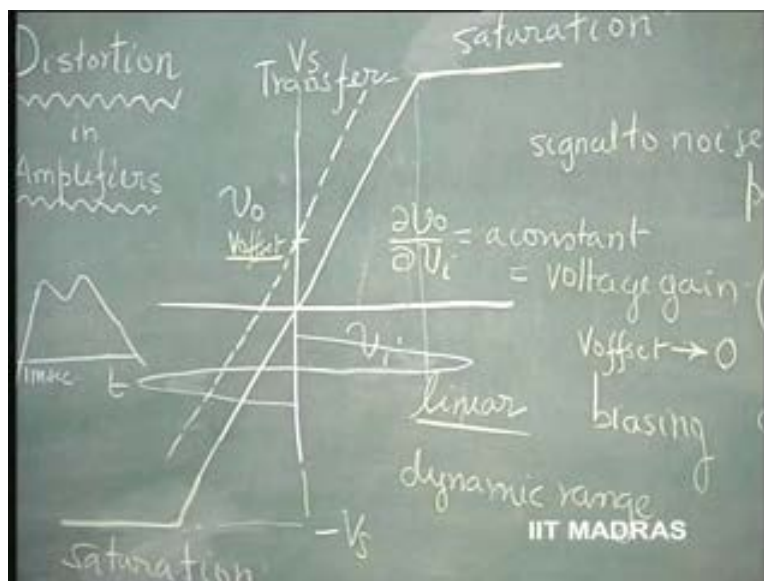
Next, this particular thing may be getting limited. Because of the supply that we are using, it is not going up to this. This is the maximum supply voltage. This is the minimum. In the negative direction, let us say, it is minus V s; this is positive direction; it is plus V s - dual supply. Above that, it is not capable of giving, because that is the bias supply voltage itself that is used for the active device biasing. So, this is called limitation due to saturation.

(Refer Slide Time: 11:41)



So, if I apply a signal which is having an amplitude greater than this, then, output is not going to increase further. It is going to remain constant. So, suppose I apply a sine wave here as the input; this sine wave, as long as it is within this, is going to be reproduced exactly at the output; no problem, as a sine wave. But, if this amplitude is exceeded, this amplitude is going to get limited here, chopped off here.

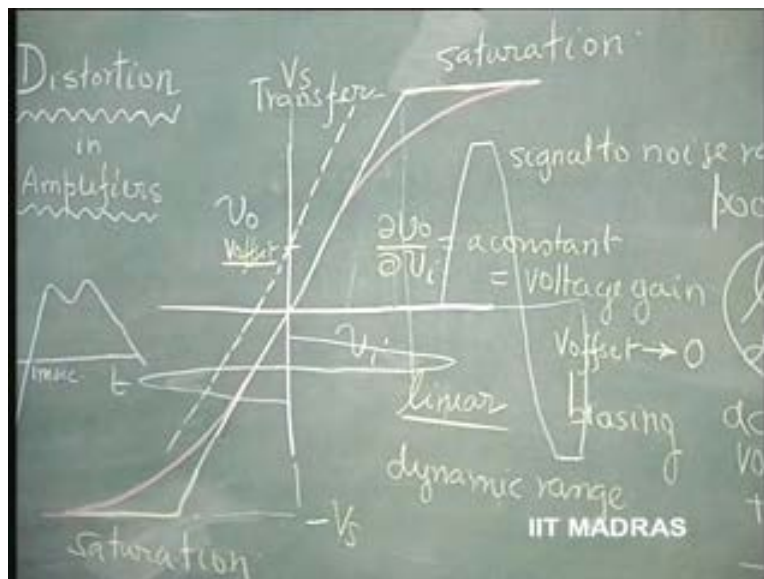
(Refer Slide Time: 12:33)



So, the output amplitude will be just getting limited. So, this is because, up to this input, amplitude, output is going to be reproduced exactly, proportionately. Thereafter, it is going to remain constant. Again, so is the case with the lower limit. So, this kind of chopping off is due to saturation. This occurs in all amplifiers. So, if you see anything like chopping off the head of your sine wave form, you can immediately conclude that you are applying a signal which is too large and therefore output is getting saturated.

So, other type of nonlinearity that we can discuss will be what is called as... this is no longer changing abruptly like this. So, the slope is maximum at this point and going on decreasing; ΔV_o by ΔV_i which is the amplification factor, this is different at different points. This is what happens because of what is called nonlinearity. So, this nonlinearity causes it to keep on decreasing and ultimately go to zero. At saturation, ΔV_o by ΔV_i is zero. That means, it is going from a maximum value to a zero value. So, this is what happens in practice, in most of the amplifiers, The nonlinearities may differ; it may not be exactly this nonlinearity.

(Refer Slide Time: 14:20)



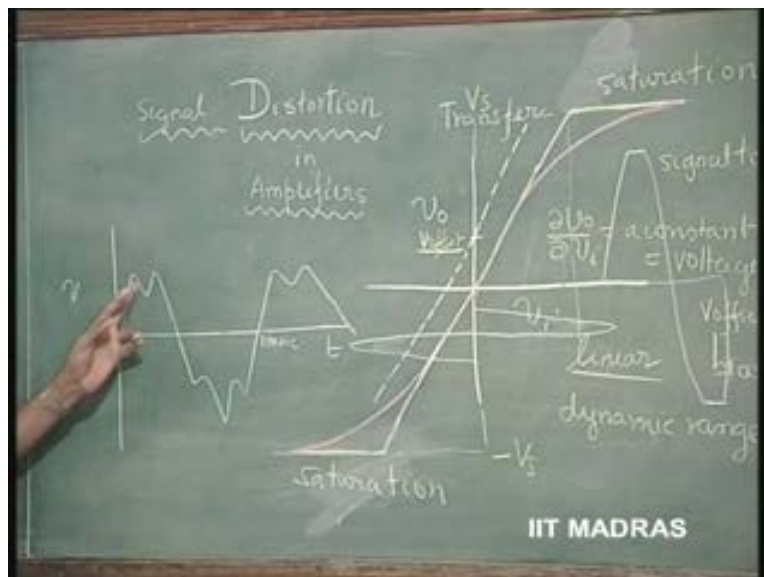
This is a symmetric nonlinearity **nonlinearity**. So, this nonlinearity is something that causes us to fix the operating point very carefully at the middle here, in this case, so that

the dynamic range is maximum. Operating point, because of offset, can shift to this. Even then, it is amplifying. But the dynamic range gets curtailed; here on this side, it can go only up to this point; on the other side, it can go this much.

Similarly, if the offset is here, then again, the range, dynamic range, of the amplifier gets curtailed. So, it is important to fix the operating point properly; and the operating point should be stable so that if the operating point drifts again, the dynamic range for your amplifier gets curtailed in one direction. So, operating point stability is also important in your circuit.

So, in the design of amplifier, we see to it that the operating point does not drift much. Particularly, we bother about this in the case of large signal amplifiers, not small signal. In the case of small signal amplifier, even if there is offset, since the signal itself is going to be very small, it is not going to matter much. So, the amplification factor may be still very high. So, this is what you have to bear in mind. So, dispersion comes about because of all these nonlinearities. So, let us therefore see what happens to this kind of signal if it is applied here.

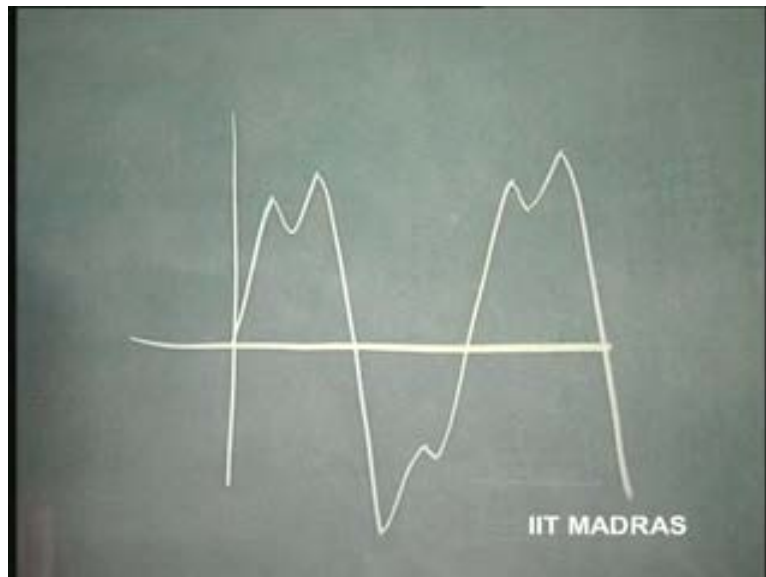
(Refer Slide Time: 16:11)



This is not going to be retaining the shape exactly here. Maybe, this portion of it may get chopped off; and this portion may be amplified properly. So, that is what is called distortion. How to measure distortion is what we are going to discuss today. See, the causes for distortion are very clear. This is caused by nonlinearity. I have told you the other type of distortion that is caused by the frequency dependence of the amplifier gain. That is called magnitude and phase distortion resulting in distortion of the signal. That is different from the distortion that is occurring because of nonlinearity.

So today, we will discuss about nonlinearity distortion. Normally, for testing the amplifiers, it is customary to use sinusoidal signal. Why? Because, sinusoidal signal can be sort of a generalized signal which... a periodic signal, for example, like this, can always be represented as a summation of sinusoidal. So, a test of an amplifier is normally done by using sinusoidal signals. If you know how it behaves for sinusoidal signals, you know how it can behave for a periodic signal; or, for that matter, you can also then conclude how it can behave for a periodic signal.

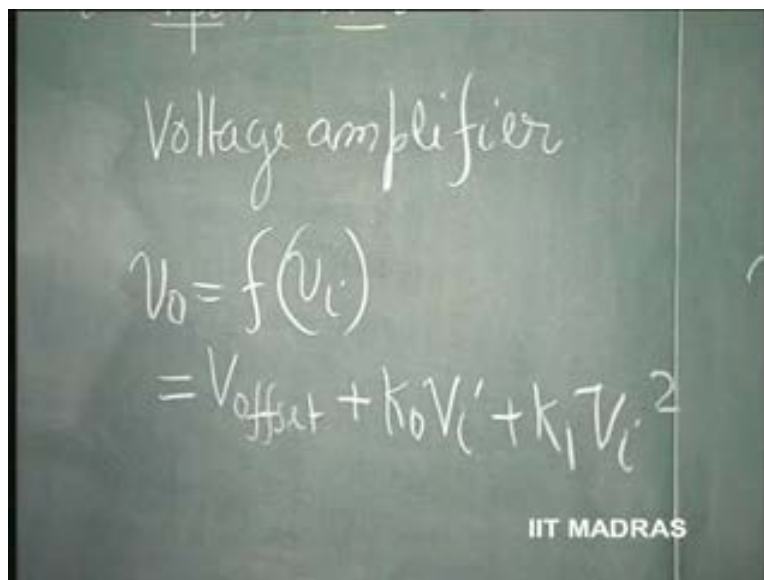
(Refer Slide Time: 17:54)



So, all these things you have learned in your networks course. Let us therefore assume that V_i is $V_p \sin \omega t$; ω is the frequency of the signal, V_p is the peak voltage of the input signal.

Now, if this is applied as an input to an amplifier, what happens because of nonlinearity? Let us assume that we are discussing a voltage amplifier; which means that, V_o is a function of V_i . I told you, this is going to be represented as V_{offset} , which is independent of V_i , plus K_0 into V_i plus K_1 into V_i^2 . I am considering now only nonlinearity due to this square term. The rest of the nonlinearity, I am ignoring. This is to illustrate as an example.

(Refer Slide Time: 19:23)



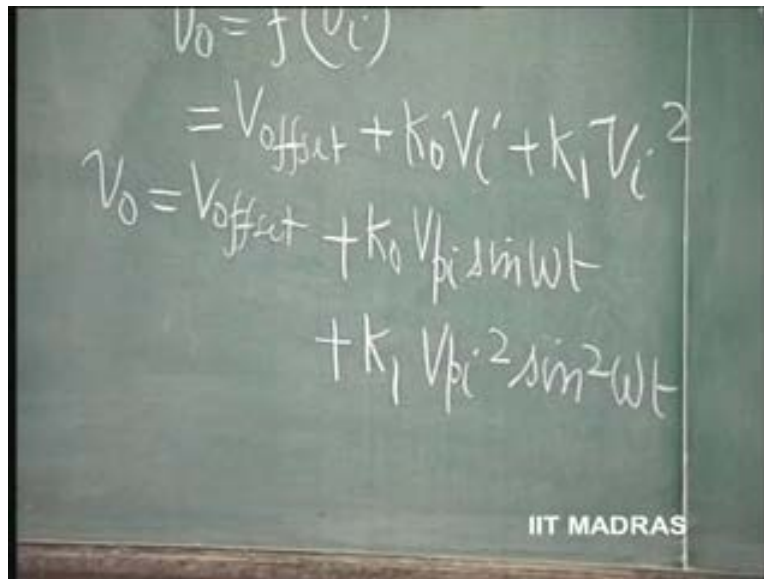
Voltage amplifier

$$V_o = f(V_i)$$
$$= V_{offset} + K_0 V_i + K_1 V_i^2$$

IIT MADRAS

So, this is the amplifier that we are going to discuss. So, if this is the nonlinearity that is present in an amplifier, then, what is going to be the amount of distortion? Normally, V_o therefore is going to be equal to V_{offset} ; this is independent of input. So, this remains unaltered; plus K_0 times $V_p \sin \omega t$. So, this is coming out beautifully. K_0 into V_p amplifies it. Next, plus K_1 into V_i^2 ; so that is coming as $V_p^2 \sin^2 \omega t$.

(Refer Slide Time: 20:20)



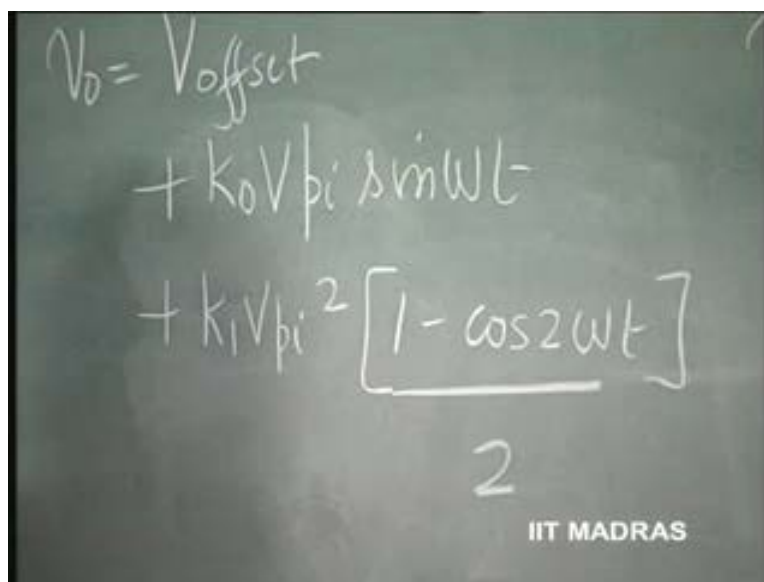
The image shows a chalkboard with the following handwritten equations:

$$v_o = f(v_i)$$
$$= v_{\text{offset}} + k_0 v_i + k_1 v_i^2$$
$$v_o = v_{\text{offset}} + k_0 v_{pi} \sin \omega t + k_1 v_{pi}^2 \sin^2 \omega t$$

IIT MADRAS

That therefore is going to be resulting in some unwanted component. What is that? We will see. We have therefore v_o equal to v_{offset} plus $k_0 v_{pi} \sin \omega t$ which is a wanted component; plus $k_1 v_{pi}^2 \sin^2 \omega t$ can be written as $1 - \cos 2 \omega t$ by 2.

(Refer Slide Time: 20:56)



The image shows a chalkboard with the following handwritten equations:

$$v_o = v_{\text{offset}} + k_0 v_{pi} \sin \omega t + k_1 v_{pi}^2 \left[\frac{1 - \cos 2 \omega t}{2} \right]$$

IIT MADRAS

So, please note; there are two things that are happening. It has now produced a 2ω component. That is, this harmonic generation occurs. This frequency component was not existing in the original input signal. So, this harmonic generation is something that we have to be careful about. This is causing distortion. The amplitude of this, if it is considerable, is going to cause considerable amount of distortion. It depends upon K_1 relative to K_0 . So, what is the extent of nonlinearity that is present?

So, $K_1 V_p \cos^2 \omega t$ is the second harmonic, divided by 2, is the second harmonic distortion. Now suppose, I want to know something about distortion. We know that the rms value of this is $K_0 V_p / \sqrt{2}$; the rms value of the distorted, that is, second harmonic, is $K_1 V_p^2 / (2\sqrt{2})$. So, percentage distortion equal to $K_1 V_p^2 / (2\sqrt{2})$ divided by $K_0 V_p / \sqrt{2}$, into 100. This is going to be the percentage distortion.

This is what we wanted. $K_0 V_p$ is what we wanted; but in addition, we have this distorted term appearing. So, $K_1 V_p^2 / 2$ divided by... so this is an important... this thing. The percentage distortion itself depends upon the magnitude of voltage V_p that you are applying. That will always be the case in the magnitude distortion occurring due to nonlinearities here.

(Refer Slide Time: 23:15)

The image shows a chalkboard with handwritten mathematical expressions. At the top left, the output voltage is given as $V_o = V_{offset} + k_0 V_{pi} \sin \omega t + k_1 V_{pi}^2 [1 - \cos 2\omega t]$. The input voltage is $V_i = V_{pi}$. The term $k_1 V_{pi}^2 [1 - \cos 2\omega t]$ is labeled as 'harmonic voltage generation'. Below this, the percentage of distortion is calculated as $\% \text{ distortion} = \frac{k_1 V_{pi}^2 / 2}{k_0 V_{pi}} \times 100$. On the right side, there are additional notes: $V_o =$, $= V$, and $V_o = V_{offset}$. At the bottom right, the IIT MADRAS logo is visible.

Now, how do you really get this value? $k_1 V_{pi}^2$ you have to get. It is very easy to get this. You do not have to really see the components of these two independently. This is going to mix with this. It is very difficult to isolate this. Whereas, we can see here that the DC offset originally was V_{offset} . Now, it is going to be shifted to $V_{offset} + k_1 V_{pi}^2 / 2$.

So, just by measuring the offset voltage without signal and offset voltage with signal... this is always the easiest way of measuring distortion. Offset voltage without signal is simply V_{offset} . Offset voltage with signal is $V_{offset} + k_1 V_{pi}^2 / 2$. You subtract this, you will get $k_1 V_{pi}^2 / 2$.

(Refer Slide Time: 24:25)

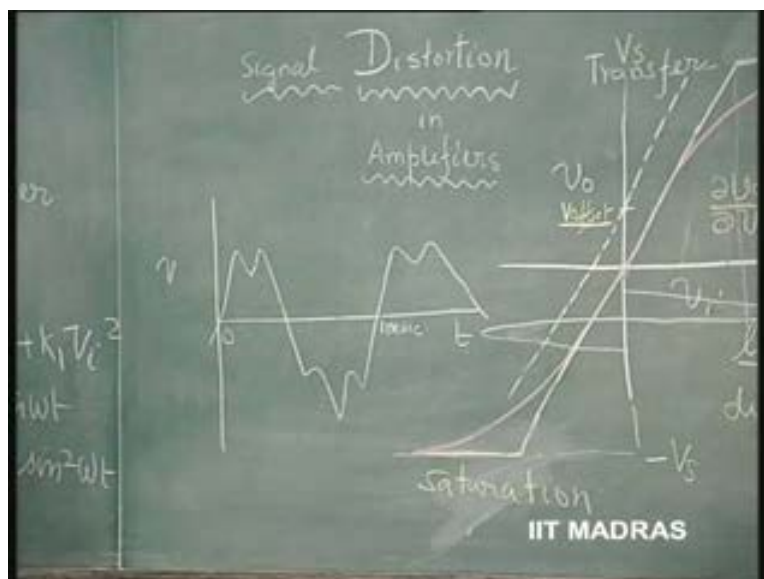
$$V_o = V_{\text{offset}} + k_o V_{pi} \sin \omega t + \frac{k_1 V_{pi}^2}{2} + k_1 V_{pi}^2 [1 - \cos 2\omega t]$$

$$\% \text{ distortion} = \frac{k_1 V_{pi}^2 / 2}{k_o V_{pi}} \times 100$$

IIT MADRAS

That means, the operating point shifts because of distortion; and the extent of distortion will be measured easily by finding out the extent of shift in the quiescent point or offset voltage. So, this is an easy way of practically measuring any amount of distortion. This may be due to second harmonic here, but it may have $K^2 V_i^3$; that does not matter. Even $K^2 V_i^3$, sort of, can be sort of giving you certain amount of distortion.

(Refer Slide Time: 25:06)



Then $K_3 V_i$ to the power 4... all these things will cause certain amount of distortion and all these distortions can be fairly easily measured by seeing the amount of offset shift. Predominantly, when we go to FET like structures, we have seen that it is a square law nonlinearity. Whether it is JFET or MOSFET, the output current is proportional to square of input voltage and this kind of distortion is quite common; and second harmonic distortion is good enough measure of the extent of distortion in the signal. So, please remember this. And, when it comes to transistors, bipolar transistors, again, it is an exponential non linearity. Again there will be the square law; this nonlinearity present in that.

(Refer Slide Time: 26:12)

nonlinear Voltage amplifier generation

$$v_o = f(v_i)$$

$$= v_{\text{offset}} + k_0 v_i + k_1 v_i^2$$

$$v_o = v_{\text{offset}} + k_0 v_{\pi} \sin \omega t + k_1 v_{\pi}^2 \sin^2 \omega t$$

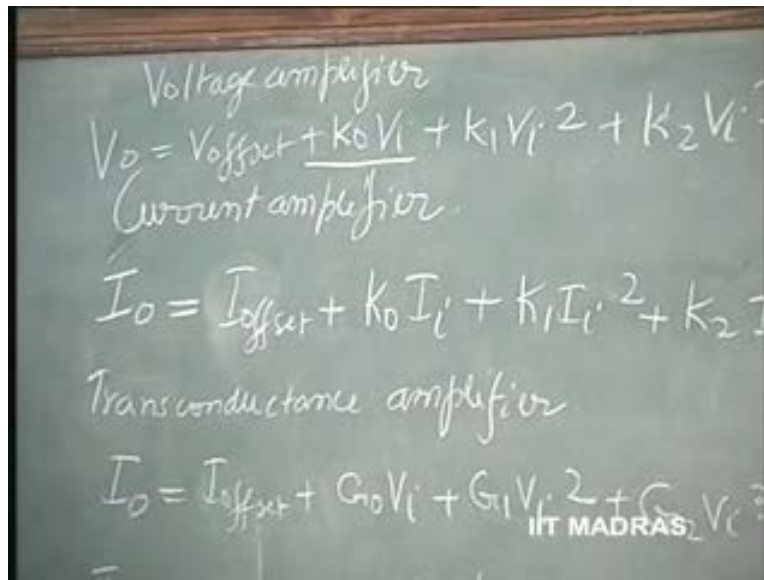
IIT MADRAS

This square law nonlinearity is absent only in a differential structure. There, the first nonlinearity is cubic. Square law is absent. It gets cancelled. Therefore, we will see that arranging these different devices, BJT or FET itself will cause us to get rid of certain dominant nonlinearity. So, this is also part of the design. In the present day design of integrated circuits, it is seen to it that this arrangement of devices is such that the major nonlinearities get mutually cancelled.

So, once again, the percentage distortion is an important measure of an amplifier. It tells us the effectiveness of our design such that this is minimized. Now, let us summarize the practical amplifier situation belonging to the four amplifier categories.

Voltage amplifier, that we had just discussed; V_{offset} is equal to V_{offset} plus K_{offset} V_i ; this is the important term that we should retain.

(Refer Slide Time: 27:42)



Rest of the things should be made as close to zero as possible. V_{offset} , K_1 , K_2 should go towards zero. That is how we are going to design a voltage amplifier. That means the operating point of the voltage amplifier should be done in such a way that V_{offset} is zero and K_1 and K_2 go towards zero.

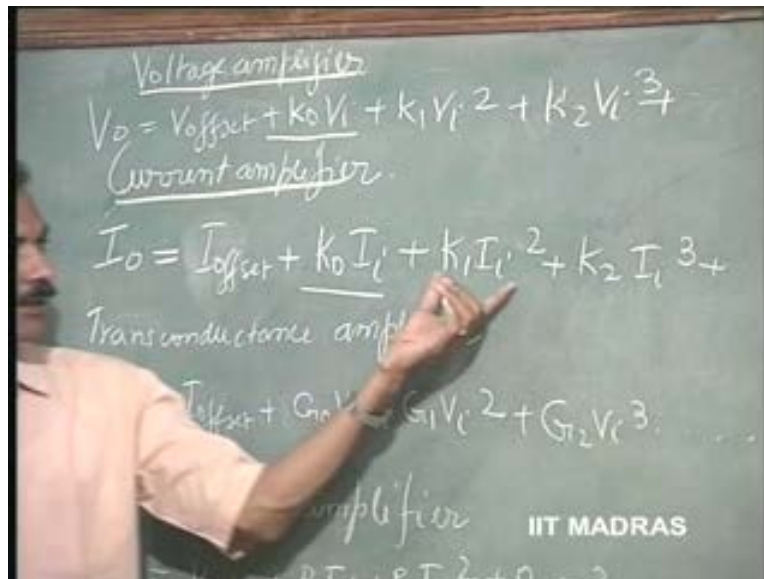
Next, how to design a current amplifier. I_{offset} is going to be I_{offset} . This is important. In the case of a current amplifier, there can exist an output current which is independent of the input current. I_{offset} plus K_{offset} I_i plus... this is what is important, linear term.

(Refer Slide Time: 28:32)



These are the nonlinear terms which should be made as close to zero as possible.

(Refer Slide Time: 28:33)



K_1 and K_2 should go towards zero. That means something that is going to be defined as current amplifier should inherently have, if possible, K_1 and K_2 very small. There is no point in therefore using this kind of characteristics for designing this kind of amplifier. We must have this K_1 , K_2 , basically inherently small, for this type of amplifier. Such a

configuration you should select; or, such a nonlinearity should be present in the active device.

The transconductance type amplifier - this is becoming quite popular these days in IC designs. I_o is equal to I_{offset} plus G_0 into V_i , linear term; and then, G_1 into V_i square, G_2 into V_i cube, etcetera. Your transistor amplifier can be put under this category; BJT or FET can be put under this category of transconductance amplifier.

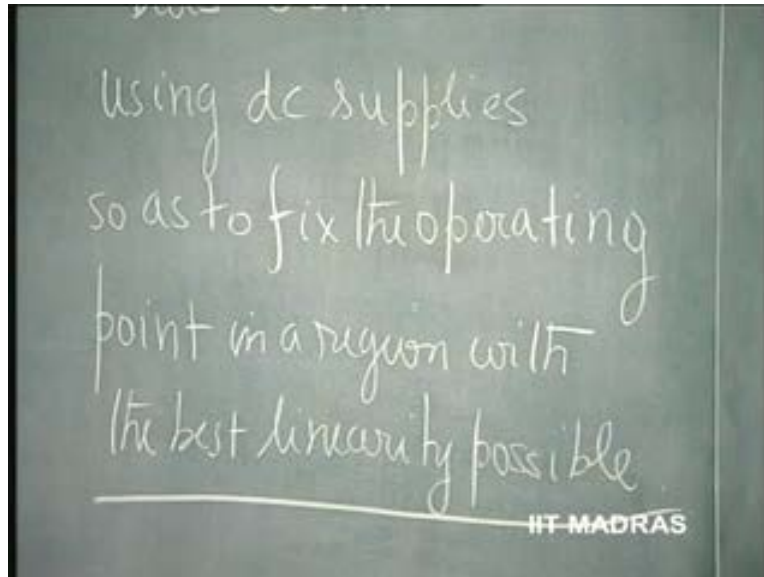
(Refer Slide Time: 29:40)

The image shows a chalkboard with three types of amplifier models written in white chalk. The first is a current amplifier with the equation $I_o = I_{offset} + k_0 I_i + k_1 I_i^2 + k_2 I_i^3 + \dots$. The second is a transconductance amplifier with the equation $I_o = I_{offset} + G_0 V_i + G_1 V_i^2 + G_2 V_i^3 + \dots$. The third is a transresistance amplifier with the equation $V_o = V_{offset} + R_0 I_i + R_1 I_i^2 + R_2 I_i^3 + \dots$. The text 'IIT MADRAS' is visible in the bottom right corner of the chalkboard.

Transresistance amplifier - V_o is equal to V_{offset} plus R_0 into I_i plus R_1 into I_i square plus R_2 into I_i cube... not yet been able to find out a device which is well suited for this kind of amplifier. For this, BJT and FET can very well do the design. Here, using BJTs, we can still obtain something which is called a current amplifier. Once again, here, this can be designed using a FET.

So, the essence of the design therefore is to bias these active devices using dc supplies so as to fix the operating point in a region with the best linearity possible.

(Refer Slide Time: 31:41)



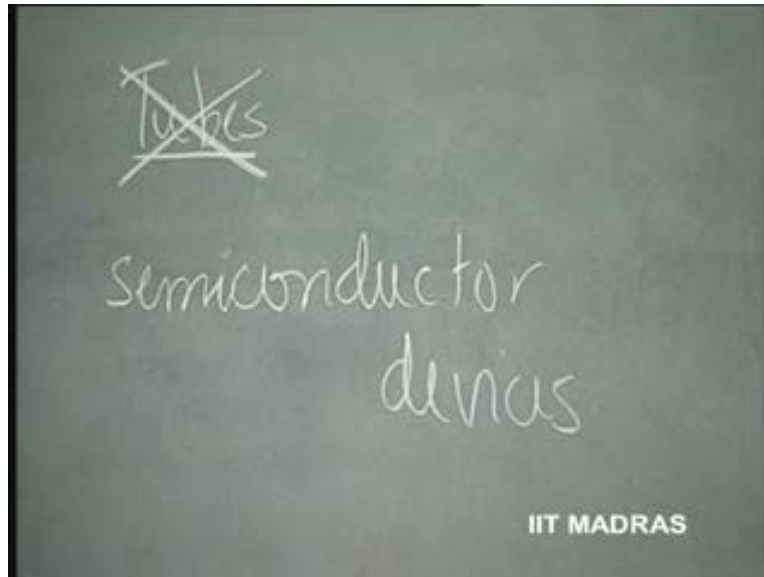
So, the next few lectures will be devoted especially to take examples of various devices like, bipolar transistors, field effect transistors, single transistors, pairs, combination of these transistors, etcetera, to achieve this purpose.

So, the basic idea is ultimately we should achieve one of the four category of amplifiers, with best linearity possible, with minimum amount of offset voltage, maximum bandwidth possible.

This is the idea of a design. So, we will now consider the basic elements which can be used for this purpose.

So, the elements that were being used were tubes earlier, electron tubes; these are all replaced now by semiconductor devices.

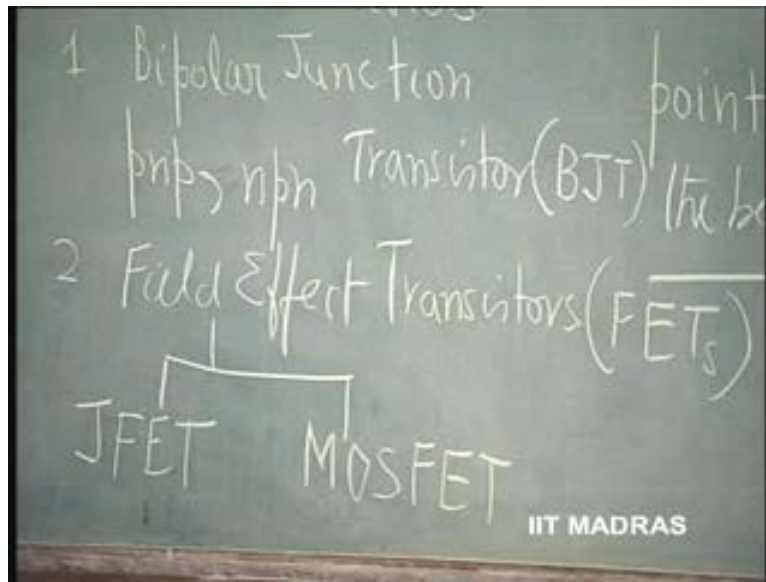
(Refer Slide Time: 33:02)



Like Bipolar Junction Transistor, popularly called BJT. There are two types in this: one is called p n p; another is called n p n.

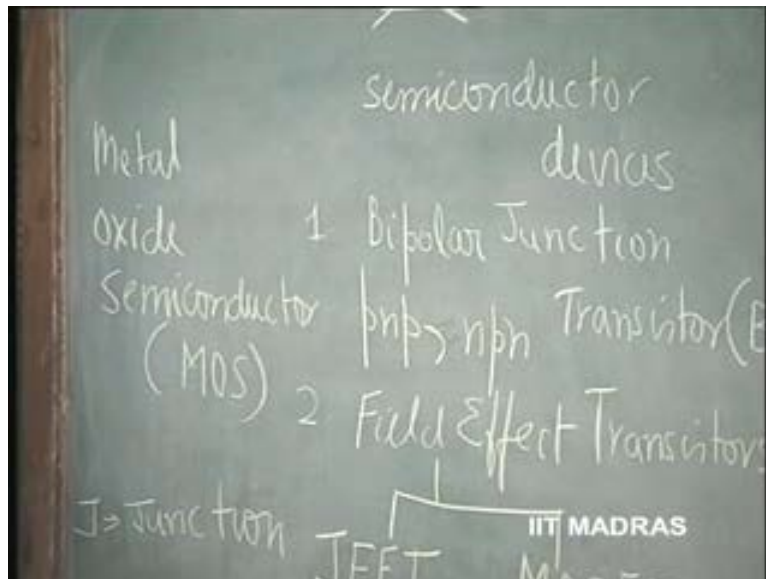
Next, we have what are called as Field Effect Transistors. These are popularly called FETs. There are two types in this: this is called Junction Field Effect Transistor; another is called Metal Oxide Semiconductor Field Effect Transistor.

(Refer Slide Time: 34:30)



So, J means Junction, J; and metal oxide semiconductor which is called MOSFET.

(Refer Slide Time: 34:55)



In these, in turn, we have two types. So, we have JFET- 2 types: n channel and p channel.

(Refer Slide Time: 35:23)



This JFET, because of the mode of operation of this is called depletion type; and therefore, in MOSFETs, we have two categories which are basically based on again the mode of operation – n-channel, p-channel. And then, we have two categories under this: this is called enhancement type; this is peculiar to only MOSFET and this is called depletion type. Similar thing, we have enhancement type of p-channel and we have depletion type. So, you can see now the variety of active devices available in semiconductor.

(Refer Slide Time: 37:00)

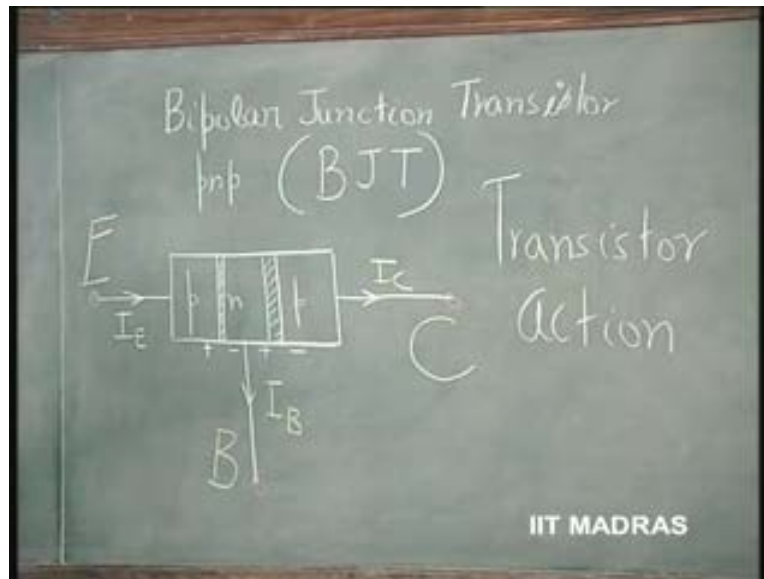


So, basically, these many devices are available for you to do the design of amplifiers. So, we have to only understand the characteristic of these: the input and output characteristic. We do not want to really bother about how this characteristic get obtained. This is the area of a Devices course.

We know that these are the various devices available and these have this kind of input output characteristics. So, we will understand in our further discussions about the operation, the biasing, the characteristics, of all these variety of devices that are available.

So, we are now getting introduced to these active devices which are working for us to make the amplifier design a success. So, let us therefore be first introduced to an important active device; that is a PNP bipolar junction transistor. So, I have now taken PNP; now, only qualitatively, we will now understand an important function in BJT. That is called Transistor action; how to make it active. So, this is important.

(Refer Slide Time: 38:36)



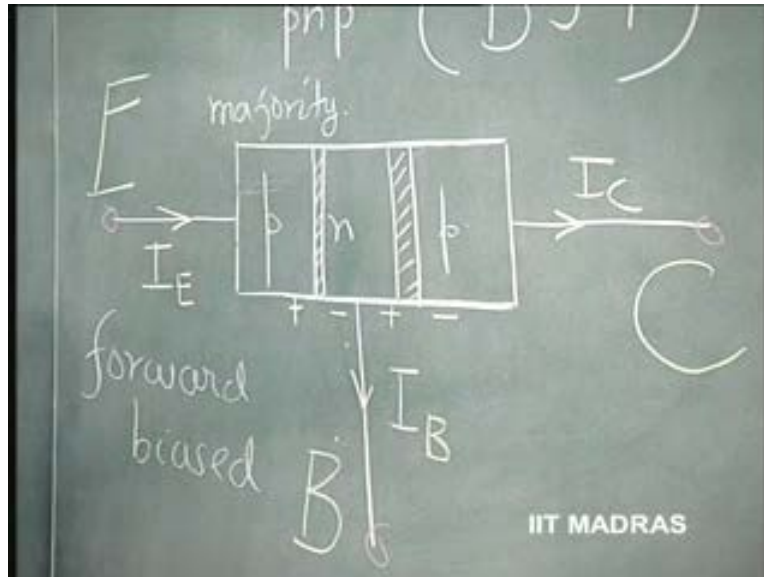
A PNP transistor built like this, let us say, is comprising two junctions. Compare it with what we discussed already in a diode. It comprised one junction, one p n junction; and we have understood its characteristic. It is exponentially related; that is, current and voltages are exponentially related.

So here, we have two such junctions; but it is not just equivalent to two diodes in series. This junction is very close to this junction. This is important in a transistor. In order to make a good transistor, I must have this junction very close to this junction. Now, what happens because of this? This junction, which is called emitter base junction, this junction is called emitter base junction; this is normally forward biased. For transistor action, this emitter base junction is normally forward biased.

It means, this p is made positive and n is made negative. The moment a junction is formed, you know that there is what is called a depletion layer which is devoid of all mobile charge carriers, electrons or holes. This layer is called depletion layer. This, we have talked about, again, in diode. The barrier, there is a barrier, potential barrier, which is of the order of point 5 to point 7 volts, in the case of silicon. That is why our diode starts conducting only after that cut-in voltage of about point 5 to point 7 volts. So, this

potential hill is there for the majority carriers. In this case, the majority carriers are, majority carriers are: holes on p side; n side, they are electrons.

(Refer Slide Time: 40:55)



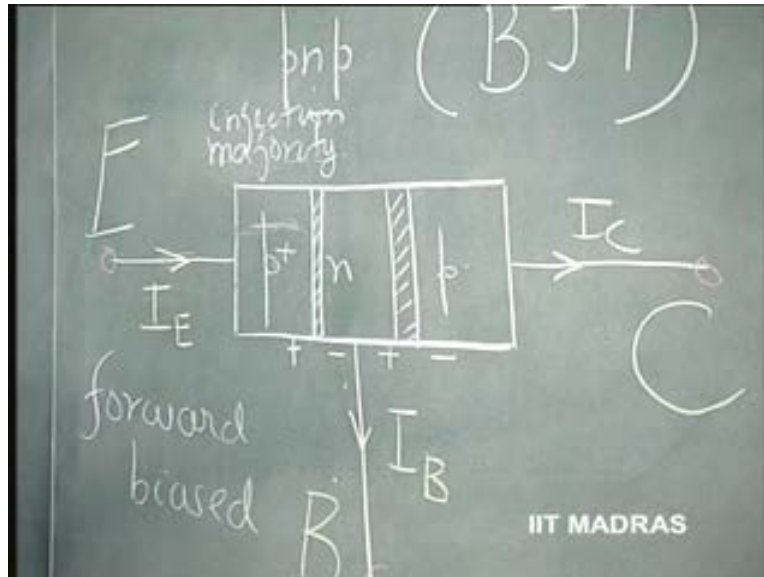
So, what is done in a transistor is the total current, normally is made up of... current flows due to both holes as well as electrons. In order that transistor action is good, I want to make this total current equal to mostly due to majority carriers in this. So, this particular region is therefore made heavily doped. So, this is really speaking, p plus, it is called. That means it is heavily doped. And therefore, the current here is going to be primarily made up of majority carriers that are holes.

So, this I_E is predominantly governed by hole currents in this region. So, these holes which are going this way constitute the current, because it is forward biased. These holes will be moving towards this. Electrons will be moving towards this easily; so because, the barrier height is reduced by forward biasing. So, a copious amount of current can flow because this junction is forward biased.

That means, injection of majority carriers will occur from emitter. That is why it is called emitter. This can now emit a large number of majority carriers from this point to this side.

So, we call this terminal as emitter. So, the design is done in such a manner that the total current is primarily made out of majority carriers.

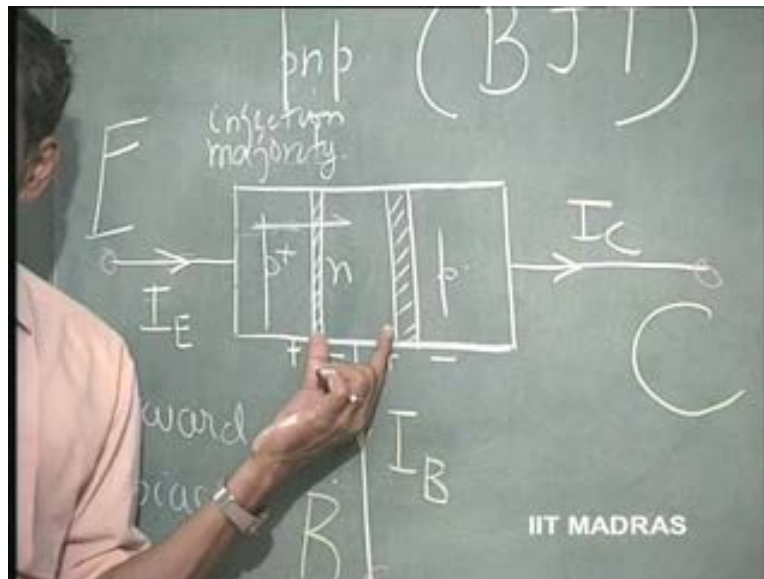
(Refer Slide Time: 42:49)



So, once these holes which are constituting primarily the current get injected on to this side... imagine them. Here, these holes in this region are minority carriers. So, they are likely to join up with electrons. They are going to be wasted. We do not want this to happen in a transistor. So, what do we do?

We make this region as thin as possible. That is, the construction of a transistor should be such that this region should be made p plus and this region should be made as thin as possible.

(Refer Slide Time: 43:29)



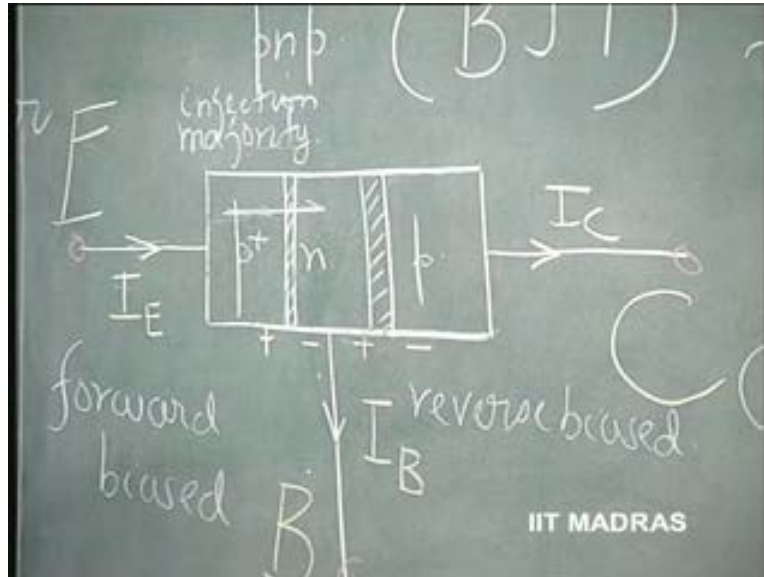
Therefore, we do not lose any of these injected carriers in this base region. This region is called the base region.

So, after they come over here, without getting lost, without recombining... this is called recombination factor. This recombination factor should be kept close to 1. So, nothing is lost in this region. So, these injected holes will entirely come to this junction; and here, you may... this junction is reverse biased. Why is that reverse biased? This reverse bias is going to cause this potential hill, as far as these majority carriers from this side is concerned, to increase; so, these majority carriers here cannot flow this way. But, the minority carriers here can now easily come to this side. But these minority carriers now coming to this side are holes, which are also injected from this side. So, most of these holes here find an easy path for this to get collected here.

So, this region is called collector. So, this is called collector. This is called emitter. So, what is the aim of transistor action? Transistor action is to make as many of the holes as possible on the emitter side to constitute the total current. That is done by making the emitter region heavily doped. Then, force most of these holes to get to come over here without recombining. That is done by making the base thin, very thin. Then, collecting

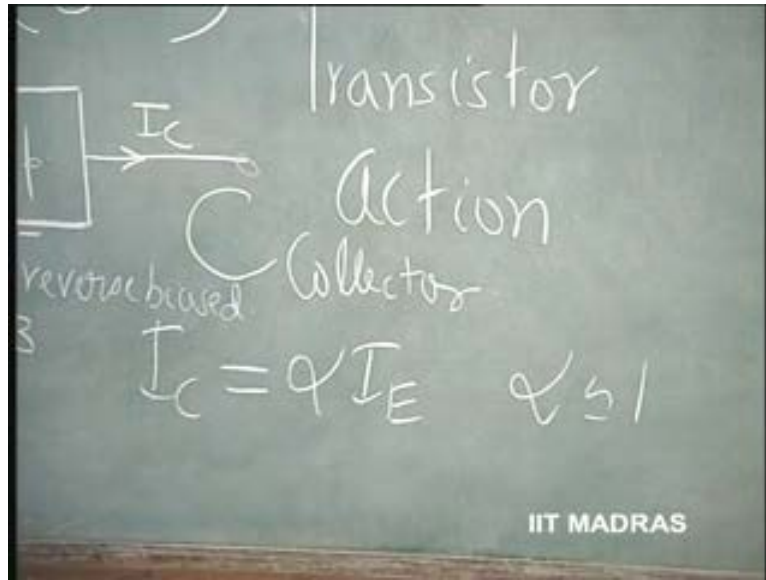
almost all the holes here. That is done by reverse biasing this junction. So, this potential hill is not a hill at all. It is an aiding field for these injected carriers, so most of them get collected here.

(Refer Slide Time: 45:54)



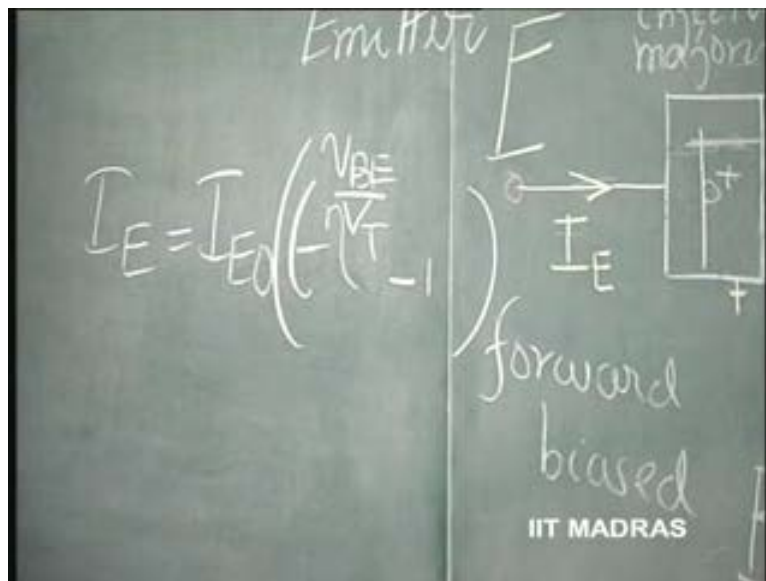
So, the transistor action lies in power biasing emitter base junction, reverse biasing collector base junction, such as to make I_C equal to I_E . That means, if it is not equal, that factor by which it is not equal, is called alpha. This alpha should be very close to 1.

(Refer Slide Time: 46:23)



That is a good transistor and this much should be sufficient for us for using this in its characteristic. Now, as far as the input is concerned, please remember; the input, emitter base junction voltage and emitter current, these are going to be related in the same fashion, I_E is equal to $I_{E0} \exp(V_{BE} / V_T)$. This η factor can also be there.

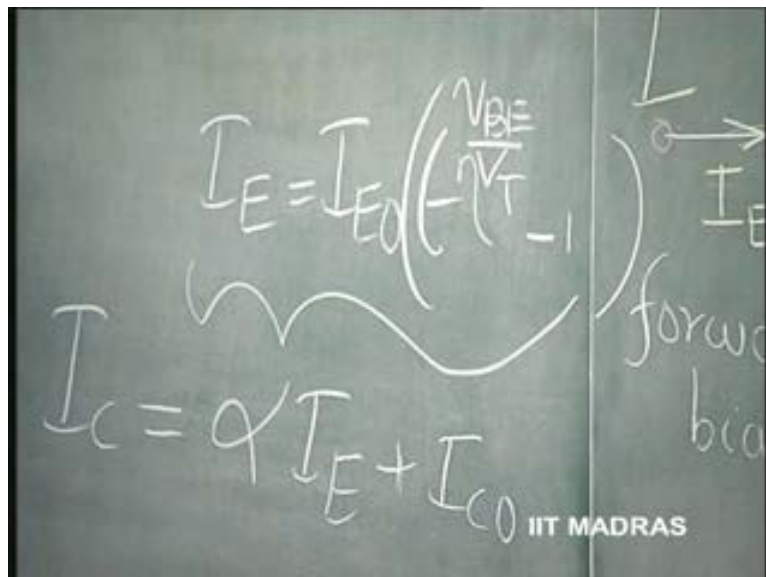
(Refer Slide Time: 47:05)



The same relationship that we had used for the diode; because, that is forward biased; emitter based junction is forward biased. And, the emitter current therefore is going to be I_E naught; this is the reverse saturation current, to exponent V_{BE} by ηV_T minus 1. So, this is something that we already know about, exponentially.

And, I_C is equal to α times I_E plus... strictly speaking, this is reverse biased. If this is totally reverse biased, then, there will be a reverse biased current flowing through this, which corresponds to I_C naught, which is the reverse saturation current of that particular junction.

(Refer Slide Time: 47:58)



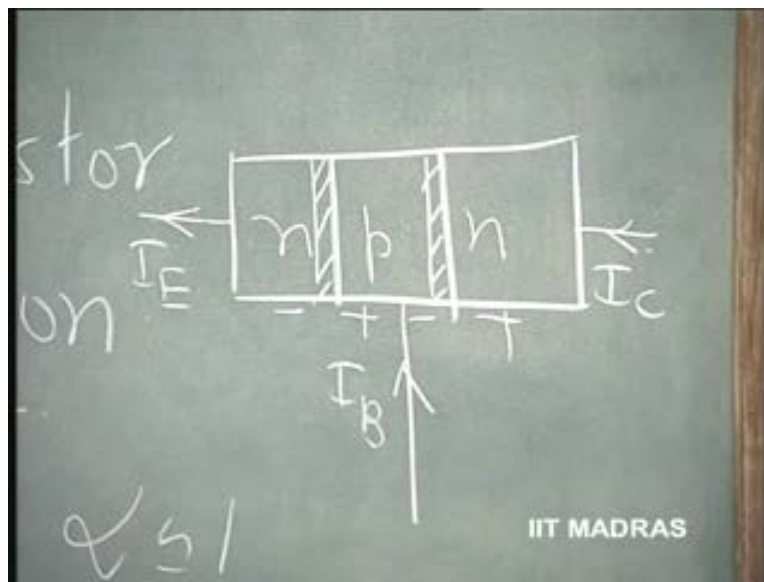
So, please remember. When a junction is reverse biased, if V_{BE} is reverse biased for example, there would have been I_E naught flowing. So, same way, there will be a reverse biased junction here; and therefore; there is this I_C naught flowing in addition to what has been injected from the emitter. So, if you understand these two equations, it should be sufficient for us to really use them in most of our design.

So, this is valid, whether the transistor is PNP or NPN. It does not make any difference. In the case of NPN transistor, this will be N, this will be P, this will be N; and therefore,

this is minus, this is plus, this is minus, this is plus. And then, this will be forward biased, that will be reverse biased again. Only the current directions will be, actual direction of the current will be, different. They will be this way, this way and this way. So, that is the only difference in the case of an NPN transistor.

So, let us put down that on this side. So, instead of this going in, these are the actual directions of the current, as against this here.

(Refer Slide Time: 49:55)



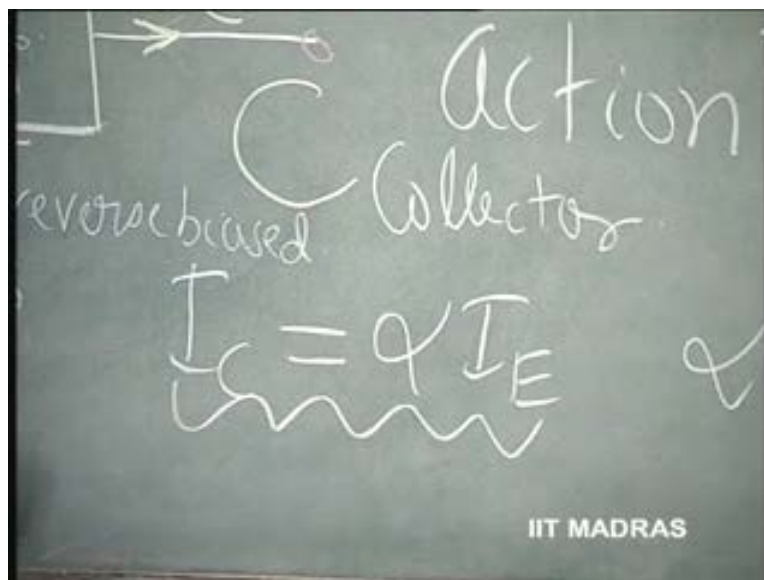
So, the biasing of the transistor, in order to make it work in the active region, is what is important. So, if you understand this, then, we have got what is called as a current source, which is totally independent of the voltage that is applied here. The reverse bias is not determining what the value of the current should be. That current is totally determined by emitter base voltage that you are applying here; and, I_E is going to be fixed by emitter base voltage at the input. And, this I_E is now flowing at the output and you can connect it to any load. So, this is a current controlled current source.

If I have 1 ampere here, this will be 1 ampere. This 1 ampere is flowing into almost a short circuit; and this 1 ampere can flow into any load. So, you have therefore forced an

output current which is dependent upon an input current. So, current controlled current source.

So, you can therefore see that the emitter base junction voltage, which is the input voltage, and input current; they are related exponentially. So, that is the nonlinearity. Emitter base voltage and input current are related exponentially. Output current and input current are related linearly. This is absolutely linear. This linearity is valid over a wide variation in current. So, this is a great advantage associated with this device; which is a bipolar junction transistor device. The current relationship is absolutely linear. This is important. So basically, this is a current amplifier.

(Refer Slide Time: 51:59)

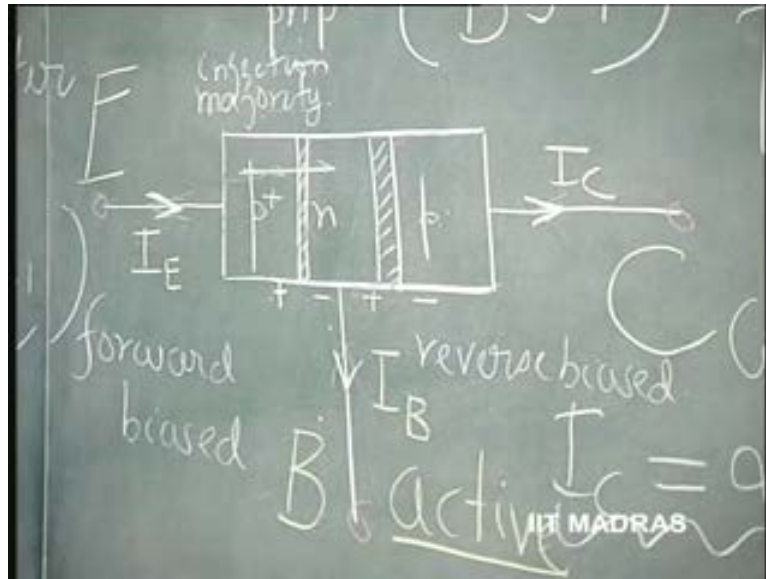


In the next class we will get introduced to the field effect transistors. How to have transistor action in field effect transistors. Just like this. How to make them work in a region where they can give amplification.

So, if you now reverse bias this and reverse bias this, both the junctions will be off. That is called cut-off region. This is reverse biased; this diode is reverse biased; this diode also is reverse biased. The entire thing is an open circuit. So, this is called cut-off region.

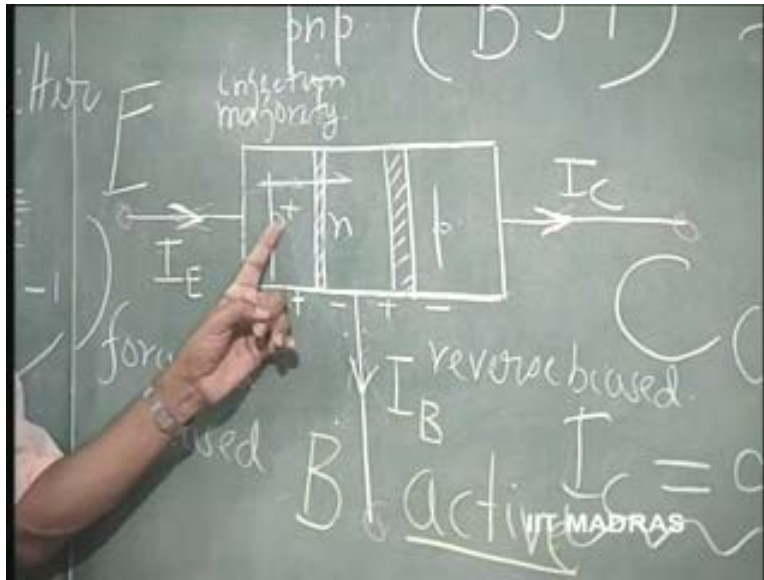
Transistor is said to be cut-off. This is the region in which transistor is active. Where one junction is forward biased, the other junction is reverse biased, the transistor is said to be active. Both junctions reverse biased, the transistor is cut-off; it is open. It is equivalent to open circuit here and open circuit here.

(Refer Slide Time: 53:00)



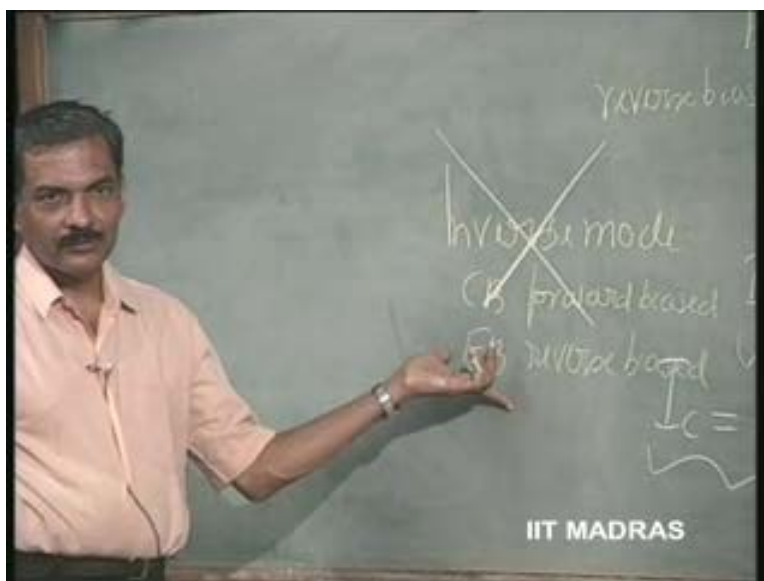
Now, collector base junction - forward biased, emitter base junction - reverse biased. This is called inverse mode of operation. You are not supposed to operate a normal transistor this way because, we have designed this transistor with this junction heavily doped.

(Refer Slide Time: 53:36)



This is not so heavily doped. So, its efficiency is poor. So, its alpha is going to be not so close to 1, as this one is. So, inverse mode of operation, collector base junction - forward biased, emitter based junction - reverse biased. Please remember that this is also an active region of operation; but alpha is not so close to 1. So, this mode of operation is not normally used.

(Refer Slide Time: 54:23)



Then, both junctions forward biased; this junction forward biased, this junction also is forward biased. That means both junctions are short circuits. So, the entire transistor is a short circuit; or, transistor is said to be in saturation. Both junctions forward biased - transistor in saturation.

(Refer Slide Time: 55:09)

