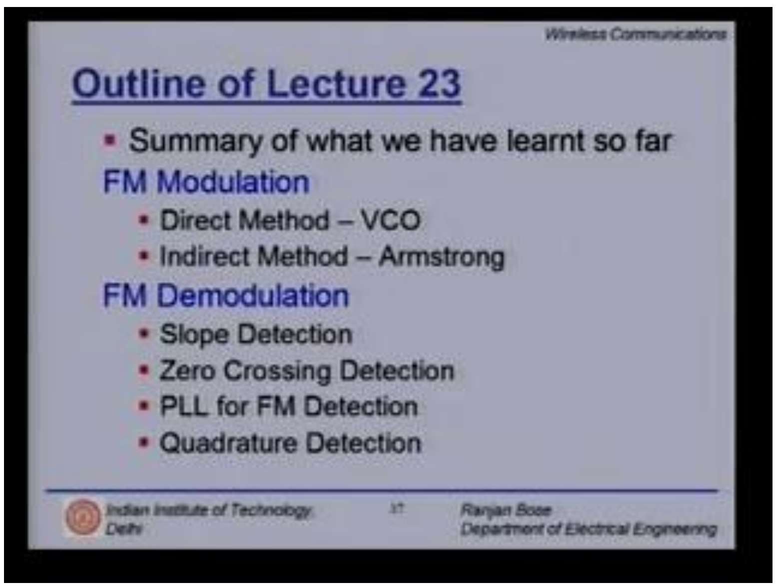
Wireless Communications
Dr. Ranjan Bose
Department of Electrical Engineering
Indian Institute of Technology, Delhi
Lecture No. # 23
Modulation Techniques for Mobile Communications (Continued)

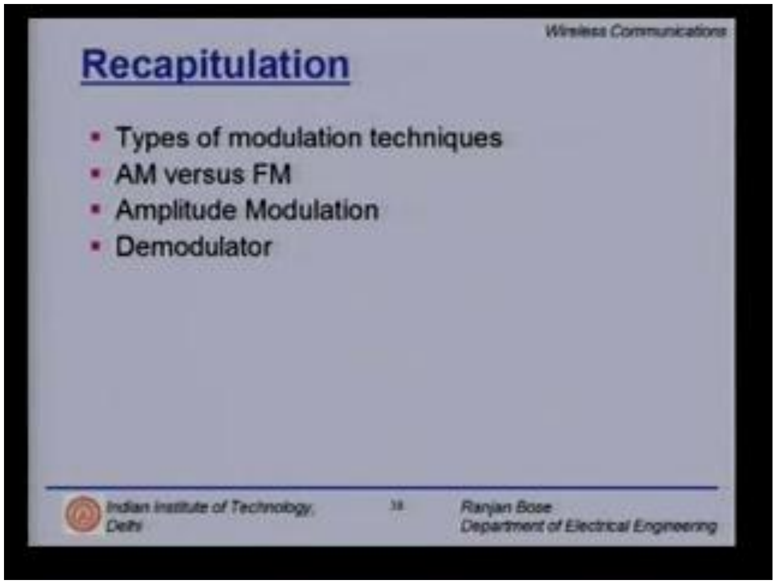
Welcome to the next lecture on modulation techniques for mobile communications. The outline of today's talk is as follows.

(Refer Slide Time: 00:01:24 min)



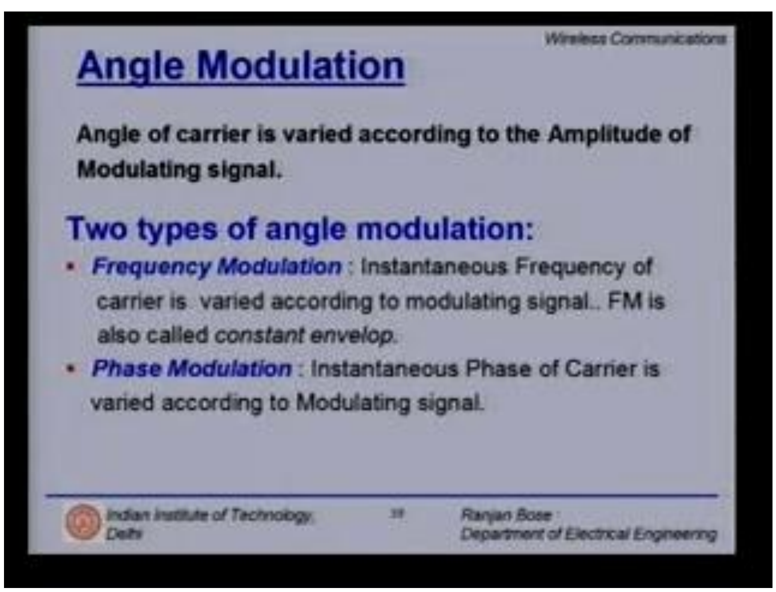
We will first summarize briefly what we learnt so far. We will then focus on FM modulation. We will look at the direct method and the indirect method for frequency modulation that is, generation of FM signals. Of course the other important thing is FM demodulation. We will look at the following four techniques. The slope detection technique, the zero crossing method, the PLL for FM detection and quadrature detection. So this will be the brief outline of today's talk.

(Refer Slide Time: 00:02:02 min)



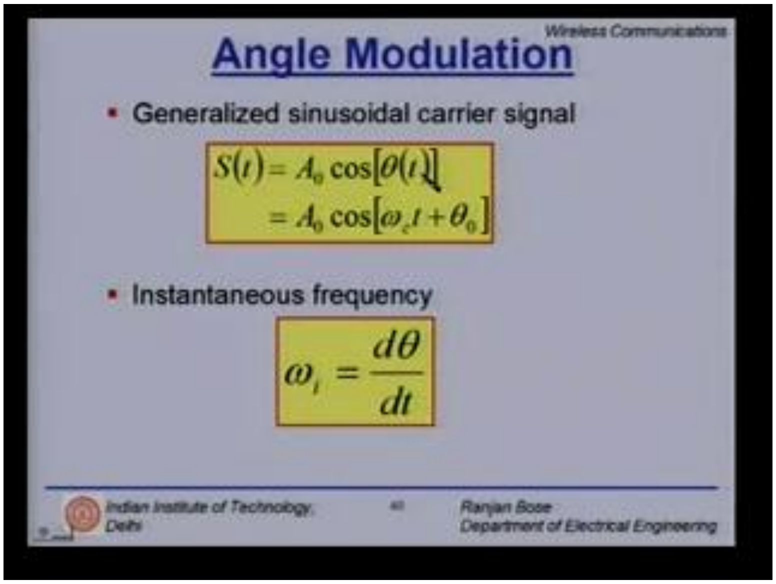
let us first summarize what we learnt last time. We looked at the different kinds of modulation techniques popularly used in mobile communications. Specifically, we looked at the analog modulation techniques and we talked about 'AM' amplitude modulation versus frequency modulation 'FM'. Then we discussed amplitude modulation in detail including generation of AM signals as well as demodulation of AM signals.

(Refer Slide Time: 00:02:36 min)



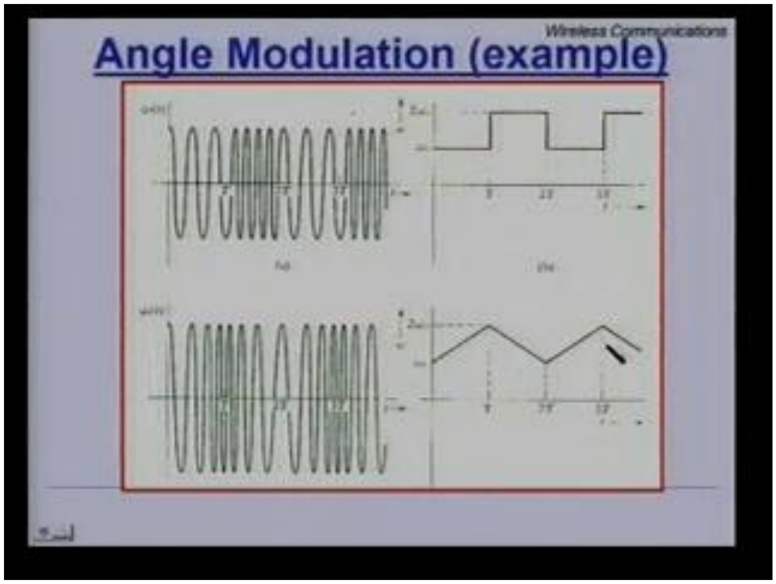
Now today even though our focus is on FM or frequency modulation, it is a sub-class of a general class of modulation techniques called angle modulation. What is angle modulation? The angle of the carrier is varied according to the amplitude of the modulating signal. There are two most important kinds of angle modulation. One is frequency modulation or FM and the other is phase modulation or PM. in frequency modulation, the instantaneous frequency of the carrier is varied according to a modulating signal. FM does not touch the amplitude of the carrier and hence it is sometimes called the constant envelop method phase modulation. On the other hand, it is placed with the instantaneous phase of the carrier and is varied according to the modulating signal.

(Refer Slide Time: 00:03:44 min)



Now if you look at the generalized sinusoidal carrier signal S(t) is given by A $_0$ cos theta t and here we can write it as A $_0$ cos omega ct + theta $_0$. Omega c is 2 pi fc, fc being the carrier frequency and theta $_0$ is the phase. The instantaneous phase in this case is given by omega $_i$ equal to the derivative of theta, d theta by dt. Now we have to either modulate the frequency part as per the input waveform or the phase part so as to either obtain frequency modulation or phase modulation. In general both these techniques are called angle modulation because of this thing. So we are changing theta which is the angle of cosine.

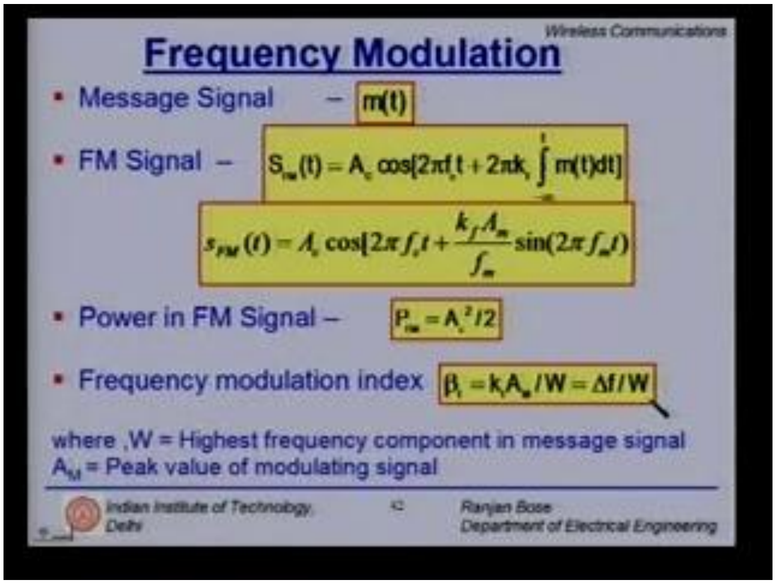
(Refer Slide Time: 00:04:41 min)



Let us look at an example of angle modulation. First consider the figure on this top right end. Here is a simple square wave. So for example for duration from zero to t we have ah signal which is frequency f1 and for a higher amplitude here, the squarewave I have a higher frequency and then again there is a trough. I again have a lower frequency and a higher frequency. So simply I am changing the frequency and depicting what could be the possible input m(t).

Slightly more complicated case when the amplitude of the message is varying continuously. So the frequency must vary continuously. So as the amplitude is low here, the frequency is low. When the amplitude is increasing, frequency is increasing. When the amplitude peaks, frequency peaks too and gradually goes down and the frequency decreases. We are just conveying the amplitude of the message in terms of the frequency.

(Refer Slide Time: 00:05:57 min)

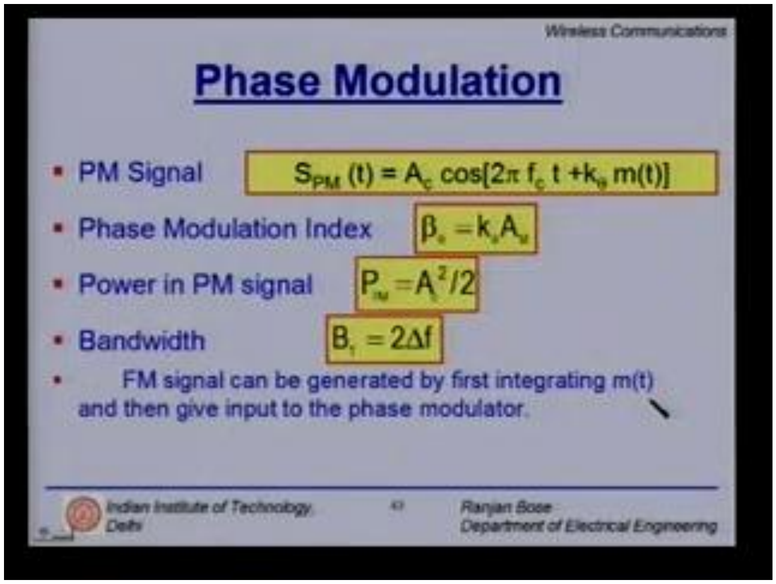


So let us define certain parameters. We have the message signal m(t). The FM signal is given by S_{FM} = Ac cosine two pi fct carrier + m(t) integrated with respect to time. so SFM can be expressed like follows the power in the FM signal is given by A_c squared by 2 and the frequency modulation index is defined as beta $_f$ $=k_f A_M$ over W where W is the highest frequency component in the message signal. We will use these parameters later today to characterize the spectrum of the frequency modulated signal. A_M is the peak value of the modulating signal. So A_M and W both figure in the calculation of frequency modulation index.

Student: so sir, does index depend on the proportion of highest embedded to the highest frequency?

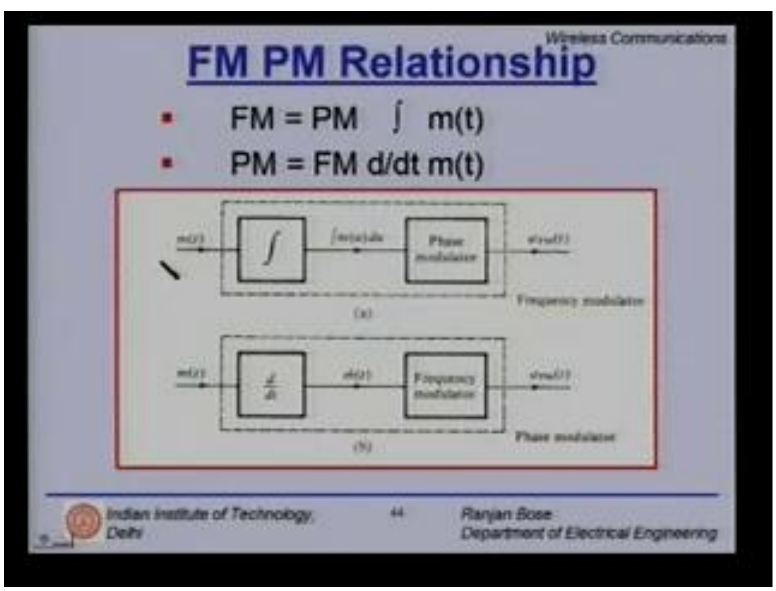
Professor: Yes. If you look at beta $_f$ physically, it is delta f over W. so it is actually the variation in the frequency. Clearly the frequency has to be changed as per the input amplitude. But there has to be a range. So the modulation index or specifically the frequency modulation index is a measure of how much is the variation in frequency delta $_f$ normalized by the maximum or the highest frequency component in the message signal. So there is a very physical meaning attached to it.

(Refer Slide Time: 00:07:58 min)



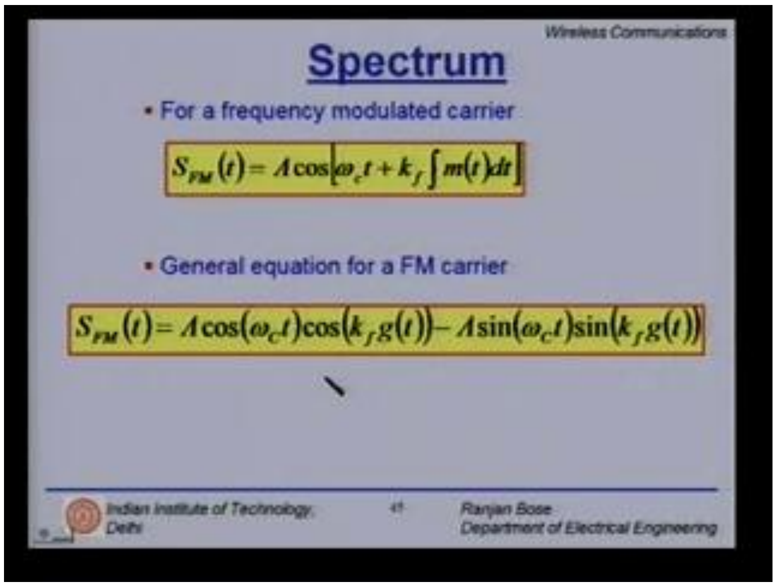
Let us look at the other brother which is the phase modulation. Phase modulation signal S $_{PM}$ (t) can be given as A_c, the amplitude of the carrier cosine 2 pi fct + k $_0$ m(t) directly the message signal is being used to alter the phase. The phase modulation index here is B $_{theta}$ is equal to k $_{theta}$ A_M. The power in the PM signal is again A $_c$ squared by 2. The bandwidth is 2 times delta f. so clearly if you compare the S $_{PM}$(t) with the previous case S_{FM} (t), then we observe that the FM signal can simply be generated by first integrated m(t) and then give the input to the phase modulator.

(Refer Slide Time: 00:09:00 min)



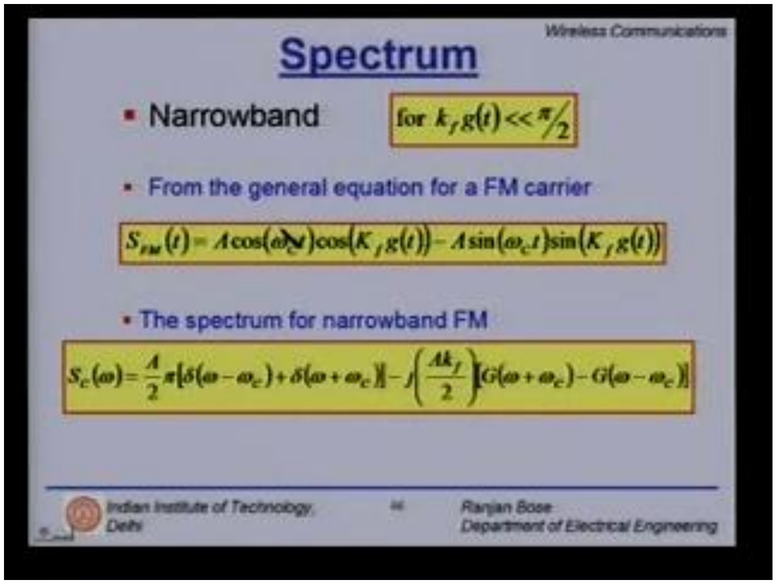
So what does this mean? PM when given the input of m(t) integrated is equal to the frequency modulation and on the other side, if you give to the FM, the derivative with respect to time of the input m(t), you get the phase modulation. Here graphically it is given by this m(t) coming in, passing through an integrator, then through a phase modulator gives you frequency modulation. On the other hand, the message signal derivative with respect to time then passing through the frequency modulator gives you the phase modulator output.

(Refer Slide Time: 00:09:42 min)



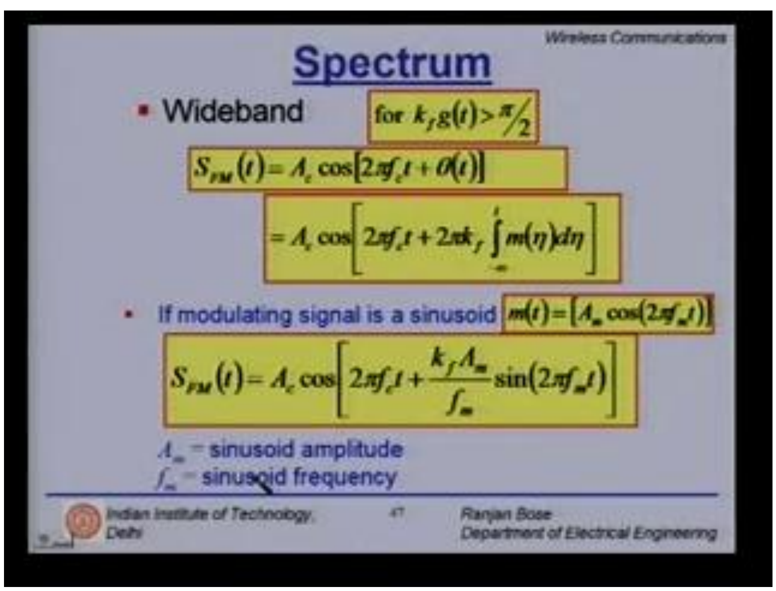
Let us now look at the spectral aspects. Now for a frequency modulated carrier S $_{FM}$, it is given by A cosine omega $_c$ t + kf integrated m(t) dt. The general equation for a FM carrier can be given as follows. So you can just expand this out and you will get S_{FM}t as A cosine omega ct carrier multiplied by cosine k_fg (t) - A sine omega $_c$ t sine k_f g(t), coming from this first equation. We will use this general equation of FM carrier to do further analysis on FM. also if you understand that an FM carrier can be represented as follows. We can think of efficient modulation techniques for FM.

(Refer Slide Time: 00:10:47 min)



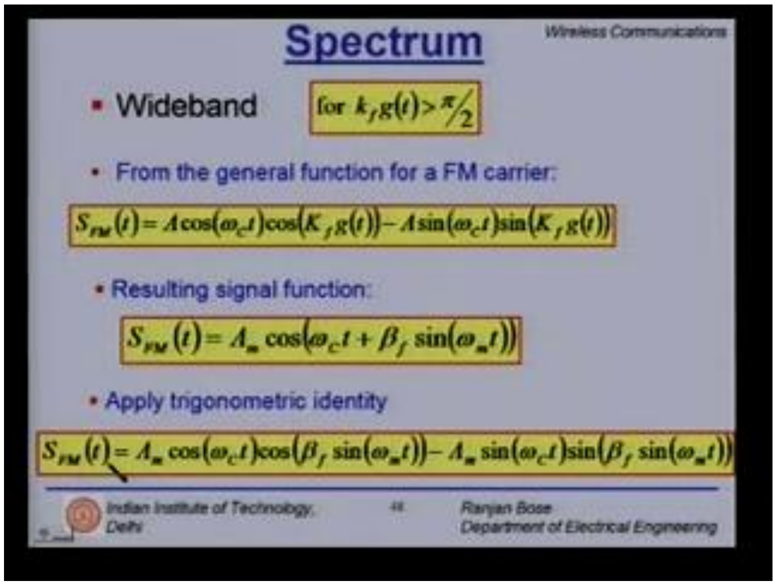
Now first consider the narrowband scenario. Let us define narrowband as the case when $k_f g(t)$ is much much less than pi by two. In this scenario from the general equation of an FM carrier which is this (Refer Slide Time: 11:07), the spectrum of the narrowband FM can simply be obtained by taking the Fourier transform. So for narrowband scenario, the spectrum of FM is given by this equation.

Refer Slide Time: 00:11:27 min)



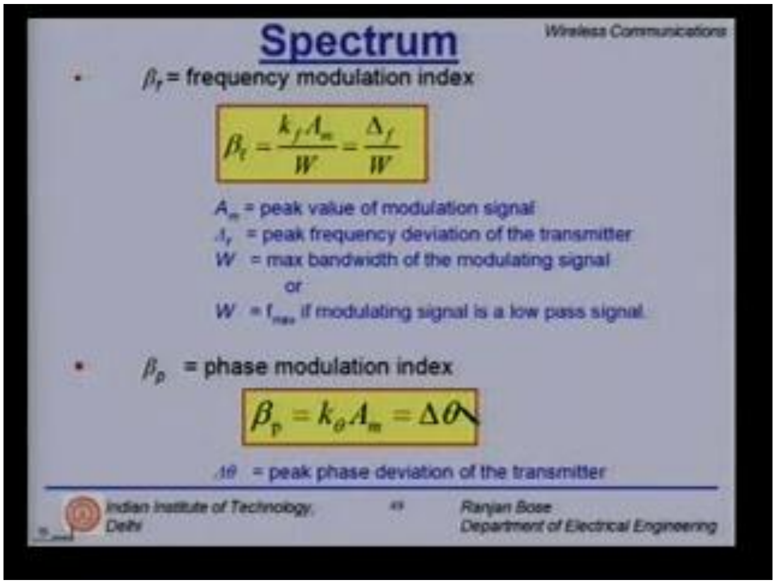
Now for the wideband case, we have the condition k $_f$g(t) greater than pi by 2 and in that case if you write S_{FM} in that formula A_c cos two pi f_c(t) + 2pi k_f integration of m(t) dt, if the modulating signal is a sinusoid, that is, m(t) is actually a cosine function, then if you solve these two, you get the SFMt as Ac cos 2 pi f_ct + $k_f A_M$ over f_m sin 2 pi f_m(t). what is AM? A_M is the amplitude of the sinusoid f_M is a frequency.

(Refer Slide Time: 00:12:21 min)



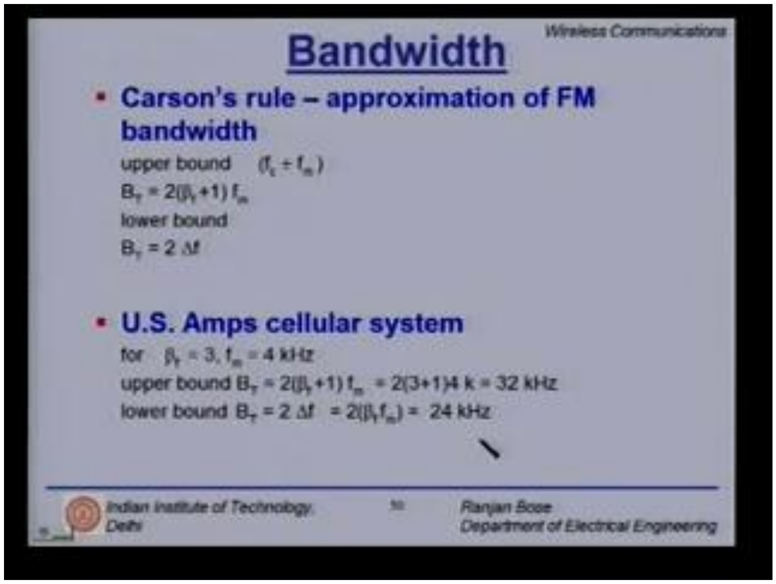
Now continuing with the wideband case, let us start with the general function of the FM carrier. As we realized, it’s a product of cosines - product of sines. We have the carrier frequency here and the message information here. so the resulting signal function, if you do the basic trigonometry is given by S $_{FM}$(t) is A_m cosine omega $_c$t + an interesting term, B $_f$ sin omega $_m$t. if you can apply a trigonometric identity, you can get the following expansion. So you have been able to express the time domain frequency modulated signal in terms of A_M beta $_f$ omega $_m$.

(Refer Slide Time: 00:13:28 min)



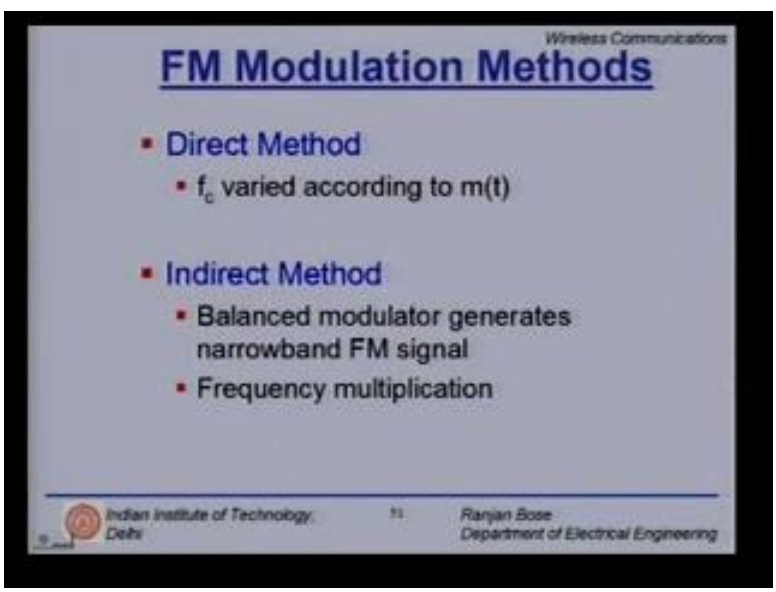
So let us talk about the frequency modulation index. Frequency modulation index beta $_f$ is defined as k_f A_M over W. as we saw, it tells you the maximum frequency deviation possible divided by the maximum bandwidth of the modulating signal. For the phase modulation index, beta $_p$ is defined as k $_{theta}$ A_m as nothing but the maximum deviation in the phase. While designing a transmitter, these two, beta $_f$ and beta $_p$ must be taken into consideration. Both give us a notion of how much maximum frequency deviation in the case of FM and phase deviation in the case of PM is permissible. If I have to have amplifiers, I should have them working linearly in this domain.

(Refer Slide Time: 00:14:43 min)



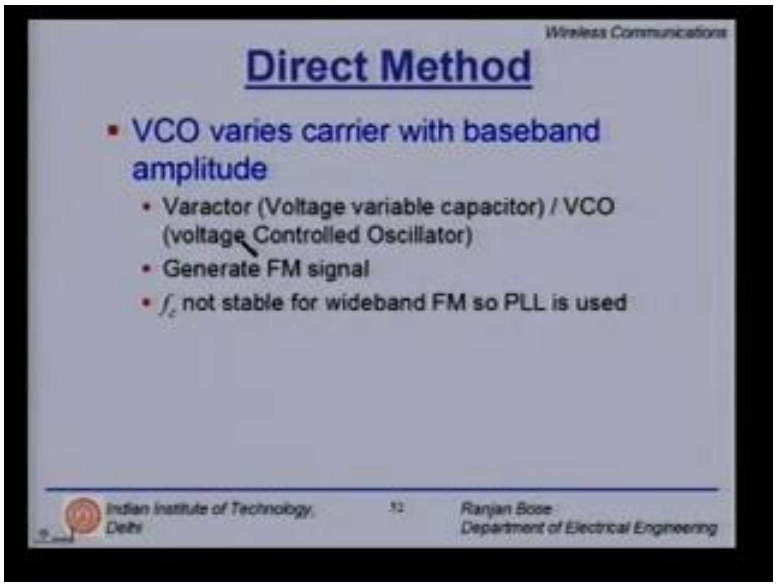
Let us talk about the approximate bandwidth of FM. we have the Carson's rule. The upper bound f_c +/ - f_m is given by B_t is equal to 2(beta $_f$ + 1) f_m. beta $_m$ comes from this frequency modulation index. the lower bound is given by B_t is equal 2 delta f, where delta f is coming from the product of kf and Am, Am being the peak value of the modulation signal. Just a small example. If you look at the first generation mobile communication systems, the U.S. AMPS cellular system which used FM for beta $_f$ =3, f_m = 4KHz. if you use this Carson's rule, you get the upper bound B_t by plugging in the formula 32 KHz and lower bound 24 KHz.

(Refer Slide Time: 00:15:55 min)



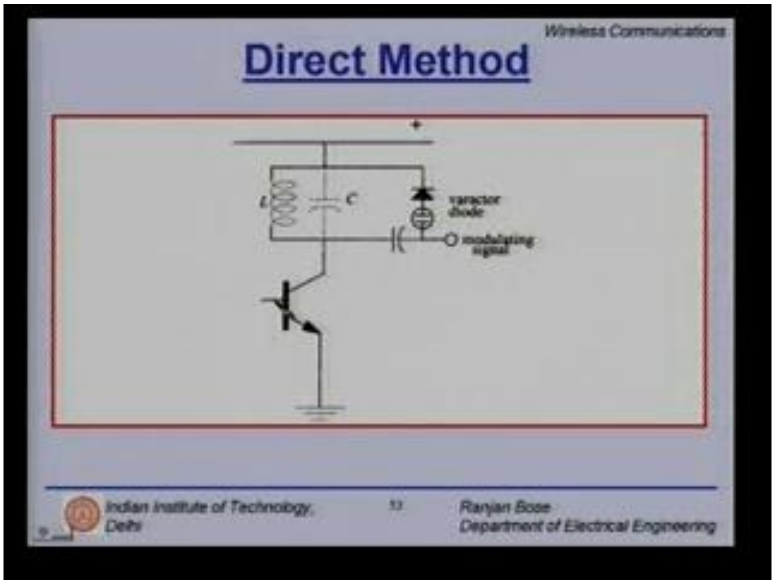
Now let us spend some time looking at FM modulation techniques. There are two broad categories in which we can classify FM modulation methods. The first one is called the direct method which, as the name suggest, simply varies the fc- carrier frequency according to m(t) -the message. The indirect method uses two kinds. One could be a balanced modulator. We generate a narrow band FM signal or frequency multiplication techniques. Let us now look at these methods.

(Refer Slide Time: 00:16:43 min)



In direct method, the voltage controlled oscillator or VCO varies carrier with the baseband amplitude. A varactor which is nothing but a voltage variable capacitor or a voltage control oscillator is used to simply control the frequency of the carrier with respect to the input m(t) thereby generate FM signals. But the problem is the moment you go from narrowband to wideband FM, fc is no longer stable. If fc is not stable, then we introduce inherent distortions. So a PLL a phased locked looped must be used to stabilize the fc. This is the direct method. What is it? Simply change the frequency of the carrier proportionate to the input amplitude.

(Refer Slide Time: 00:17:48 min)

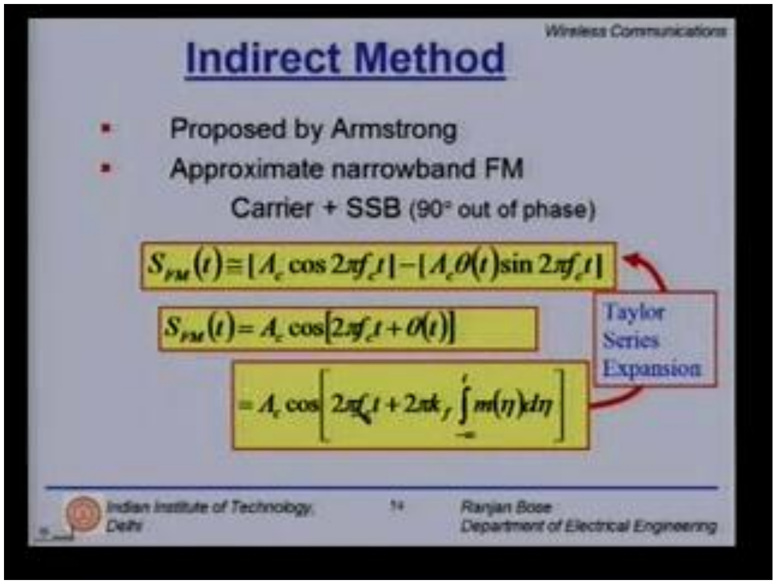


This is a simple realization of the direct method where this is the varactor and here the modulating signal is there. So simply by changing the input here, you can obtain a modulated signal. So it is extremely easy to use except that it works only for narrowband scenarios. Please note as you change the input amplitude here, you change the current being drawn here and as the voltage across this changes, the varactor will have a different frequency output corresponding to the voltage across it.

Conversation between student and professor: the question is: what does the PLL do if I have to stabilize it?

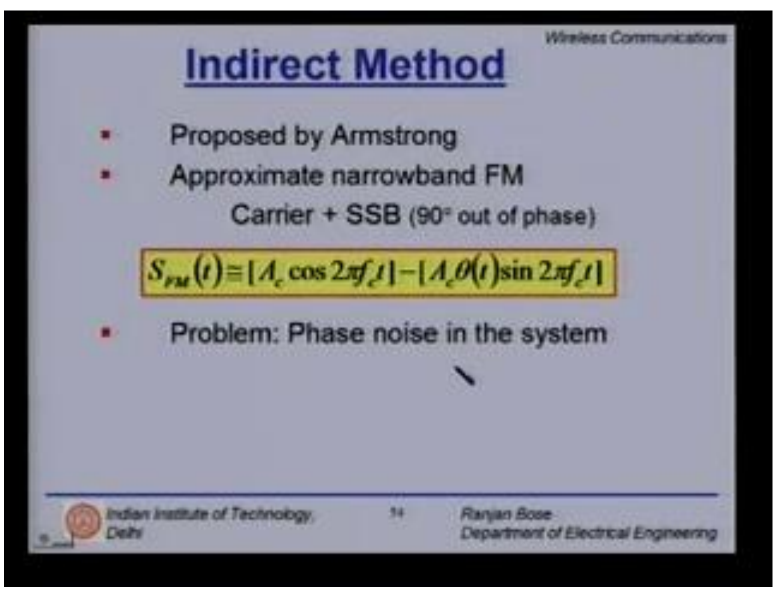
The PLL will ensure that the carrier frequency fc is locked. So it doesn't deviate much even though we move across the wideband.

(Refer Slide Time: 00:19:01 min)

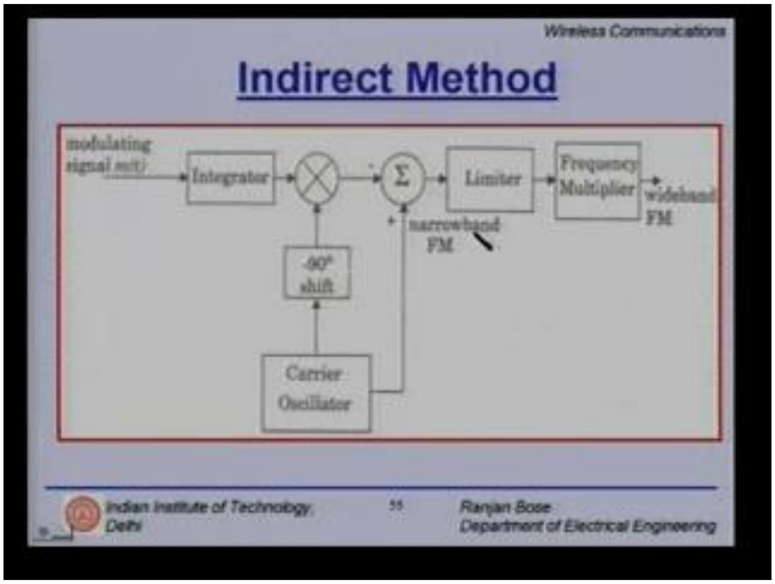


Now let us look at the indirect method it was proposed by Armstrong. The approximate narrowband FM is basically carrier + SSB (90 degrees out of phase). From where does this figure arise? See, if you remember the general formula $S_{FM}(t)$, if you do a Taylor series expansion on this equation, you can get an approximate value of S_{FM}. the problem with this method is phase noise in the system. So the phase noise kills the system

(Refer Slide Time: 00:19:58 min)

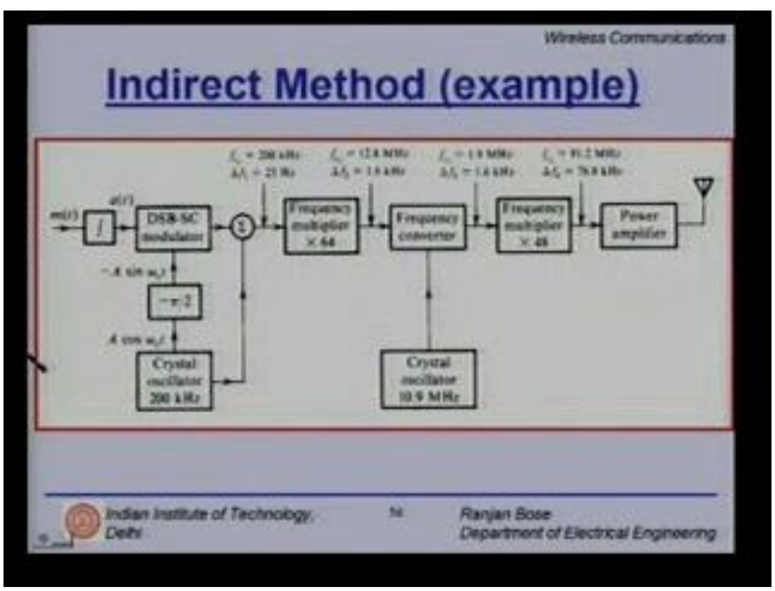


(Refer Slide Time: 00:20:10 min)



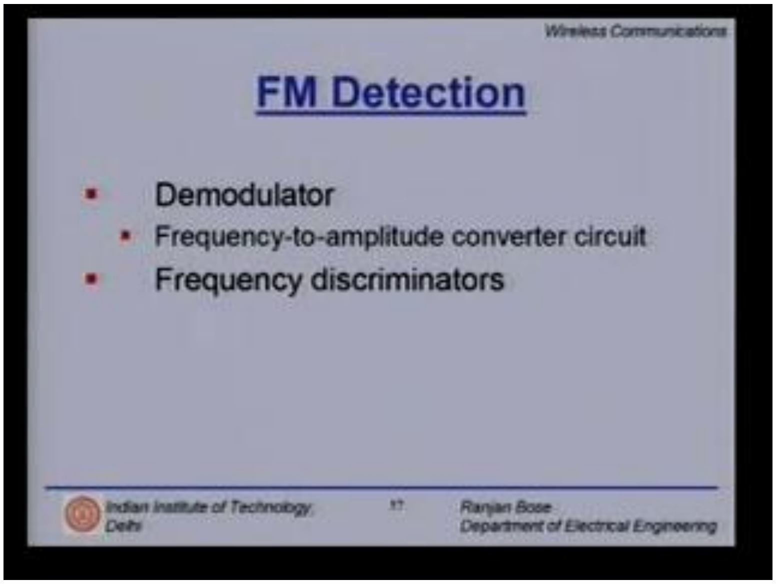
Let us look at an indirect method. Here you have the modulating signal. It passes through an integrator, -90 degree phase shift and here you add them up with a negative sign and you obtain the narrowband FM. it is nothing but a realization of this equation. It's a carrier, 90 degree phase shifted sin, here cosine right and here theta t (Refer Slide Time: 20:47). It's just the realization of this equation in this method. If you have a frequency multiplier, you can obtain a wideband FM.

(Refer Slide Time: 00:21:03 min)



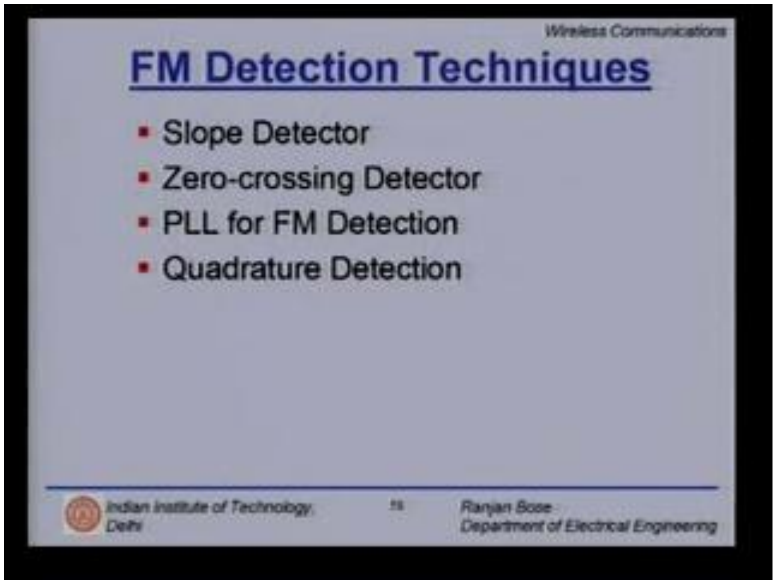
Let’s look at an example. your m(t) comes in. It is just that example of the indirect method. Then DSB-SC. here you have A sin omega ct coming in and with a - pi by 2 phase shift, you have A cosine omega ct summation. Here you get the FM. you do a frequency multiplier. The problem with this method is as you increase the frequency multiplier number ‘n’, the phase noise increases. That’s the problem. Frequency converter, power amplifier and radiator. So this is an indirect method to generate FM. It’s just the realization of the Armstrong method.

(Refer Slide Time: 00:21:57 min)



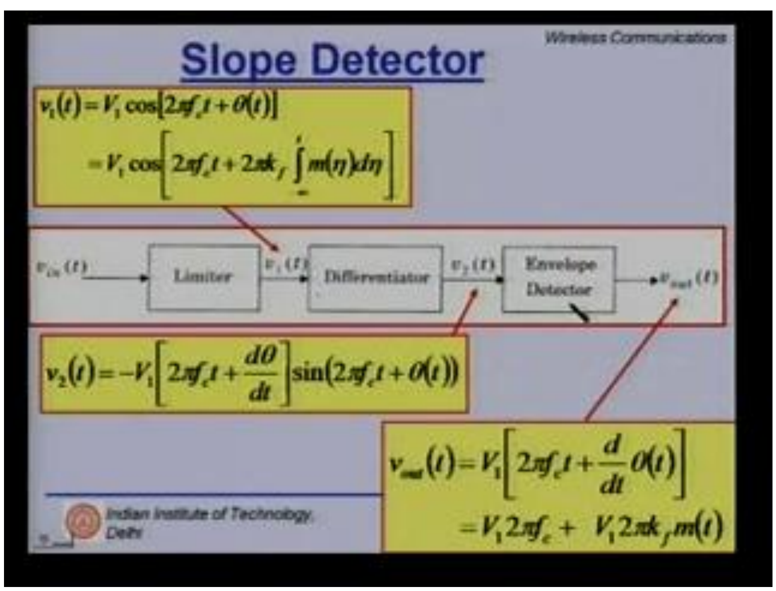
Now let us spend some time talking about the FM detection. So the demodulator can be either a simple frequency to amplitude convertor circuit. That is, find out the frequency and then convert it to the amplitude or you can have frequency discriminator.

(Refer Slide Time: 00:22:25 min)



Now FM detection techniques can be classified as slope detector, zero crossing detector, PLL for FM detection and quadrature detection. Let us talk about all of these one by one.

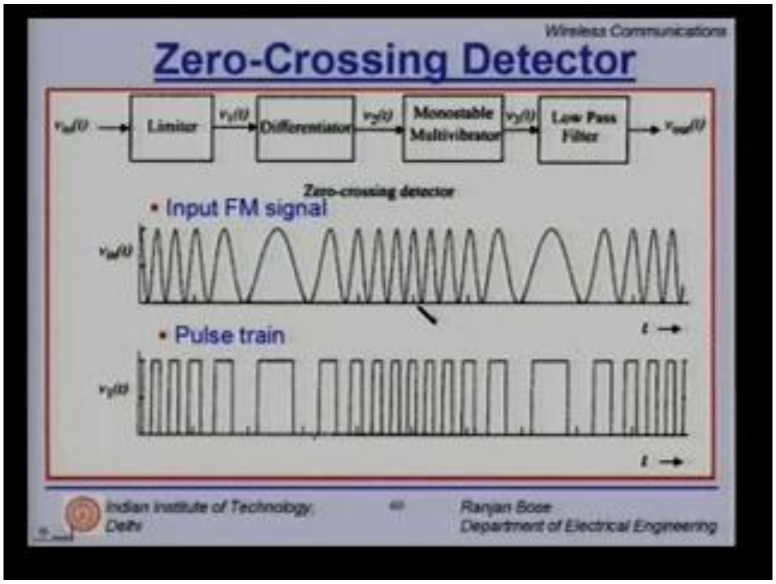
(Refer Slide Time: 00:22:49 min)



Let’s talk about the slope detector circuit. In fact, the only job of this is to figure out what is the instantaneous frequency. So here is the input to the limiter. then there is v_1 (t) which is given by v_1 cosine 2pi f_c(t) + theta t. please remember right now, we are trying to demodulate the FM. so the signal that we already have is a frequency modulated signal.

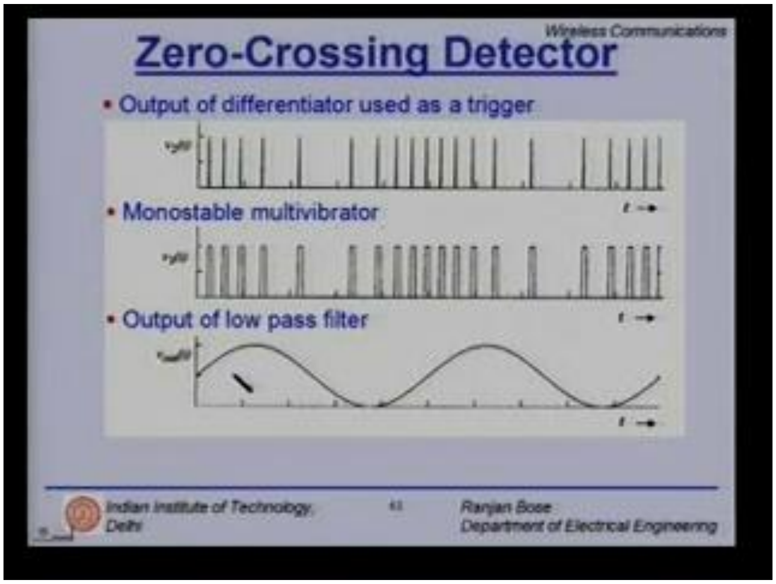
What do we do next? We differentiate it. It's just a simple differentiation gives you the following and then we do an envelope detection on this. The moment you do envelope detection, you get v_{out} as V_1 2 pi f_ct + d by dt theta t. if you do an envelope detection, you get v_1 2 pi f_c + 2 pi k_fm(t). You have recovered back your message. So it's a very simple and elegant method. You differentiate and then follow it up with envelop detector. It's a very simple method. It is called the slope detection.

(Refer Slide Time: 00:24:31 min)



Now the second method is the zero crossing detector. Here again we have a limiter followed by differentiator, then a monostable vibrator and a low pass filter. What does it do? It physically counts the number of zero crossings. Now if the frequency is more, there will be more number of zero crossings in unit time. It is as simple as that. So what is done is if you have an input FM signal as depicted here, clearly the frequency is changing depending upon the amplitude which modulated it, correspondingly, based on the zero crossing. After the limiter, you obtain a pulse train. Clearly if my signal has been of a lower frequency here, I will have broader pulses. For a higher frequency, I will have narrower pulses. Now I must differentiate this to find out where the pulses begin.

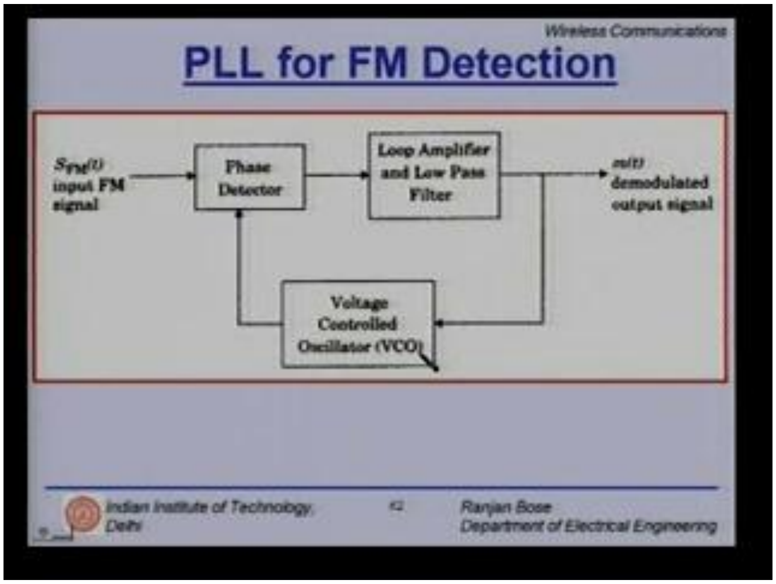
(Refer Slide Time: 00:25:46 min)



So after you differentiate it, you exactly get the location of the start points. In effect, you found out where are the zero crossings. Wherever the frequency was high, I have more closely spaced impulses whereas, wherever the frequency was low, I have far apart pulses. Pass it through a monostable vibrator and kind of average it out by passing through a low pass filter. You obtain your modulating signal. Wherever there was a low amplitude you had sparse density of impulses. Wherever there was a high amplitude you had closely spaced impulses. It is a very simple method by just counting the zero crossings. Hence it is called the zero crossing detector.

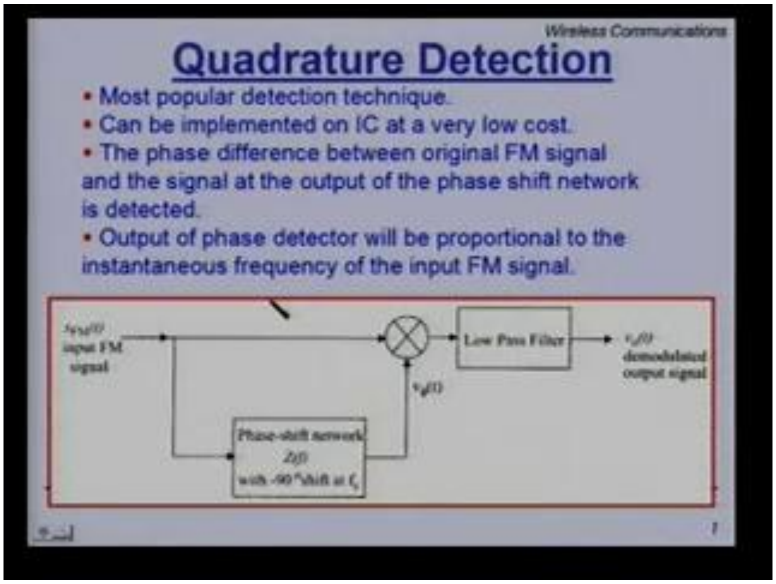
Conversation between student and professor: The question being asked is, is there is smart technique to directly go from V_2 to V_{out}? Is there a circuit or an algorithm which takes this pulse train or impulse train and converts it into this one (Refer Slide Time: 27:14)? Yes. If you possibly pass it through a low pass filter right at this point (Refer Slide Time: 27:23), you will get something like this but maybe a little jittery. So through a monostable vibrator, you have basically put in some more energy in those impulses and then when you average it out, you get a smooth function. So in principle you can go from V_2 directly to V_{out} also.

(Refer Slide Time: 00:27:42 min)



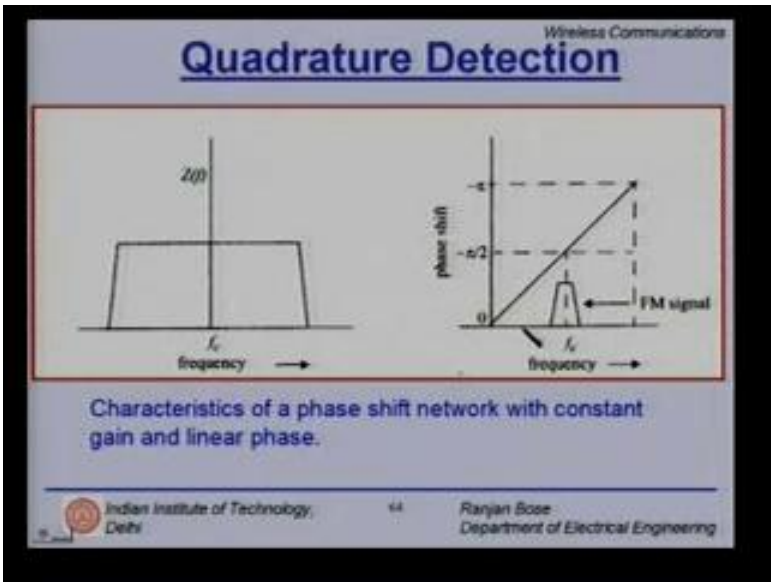
Now the third and a popular technique is to use the phase locked loop or PLL for FM detection. The philosophy is very simple. You have an internal VCO or voltage controlled oscillator. You generate your own frequency. You have a phase detector which checks the error in the phase with respect to the input modulated FM signal and your own self-generated oscillator and your job is to minimize the error. If you have been able minimize the error, you are able to track the phase and if you have been able track the phase, the input to the voltage control oscillator is actually the voltage of m(t). It is the demodulated signal. It is a very simple yet elegant method of demodulating FM.

(Refer Slide Time: 00:28:46 min)



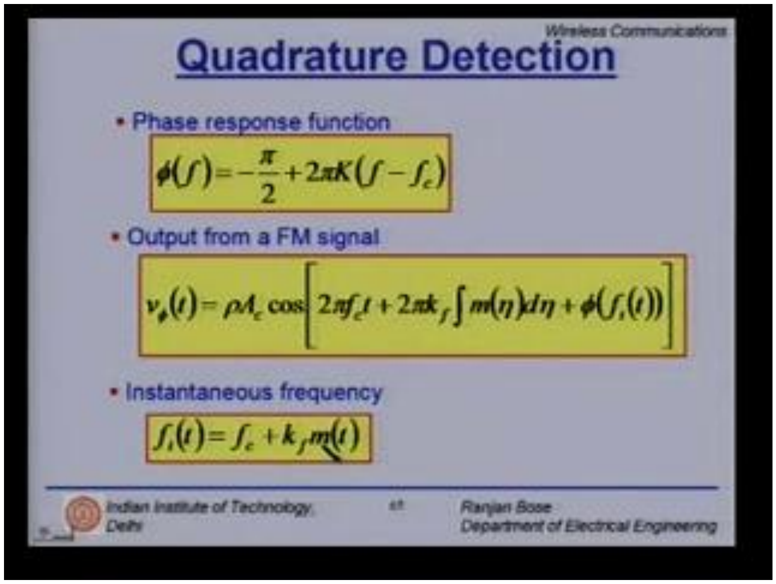
So let us now look at the fourth method which is called the quadrature detection. It is one of the most popular detection techniques simply because it can be implemented on an IC at a very low cost. The phase difference between the original FM signal and the signal at the output of the phase shift network is detected. Please look at the diagram below. So we have an input FM signal going to a multiplier. Here it is passed through a phase shift network with a – 90 degree shift at FC. In the following slide, we will look at how the phase characteristic of this phase shift network looks like. But then you multiply it, pass it through a low pass filter and you automatically get the demodulated signal output. Please note that the output of the phase detector will be proportional to the instantaneous frequency of the input FM signal, thereby demodulating the signal.

(Refer Slide Time: 00:30:10 min)



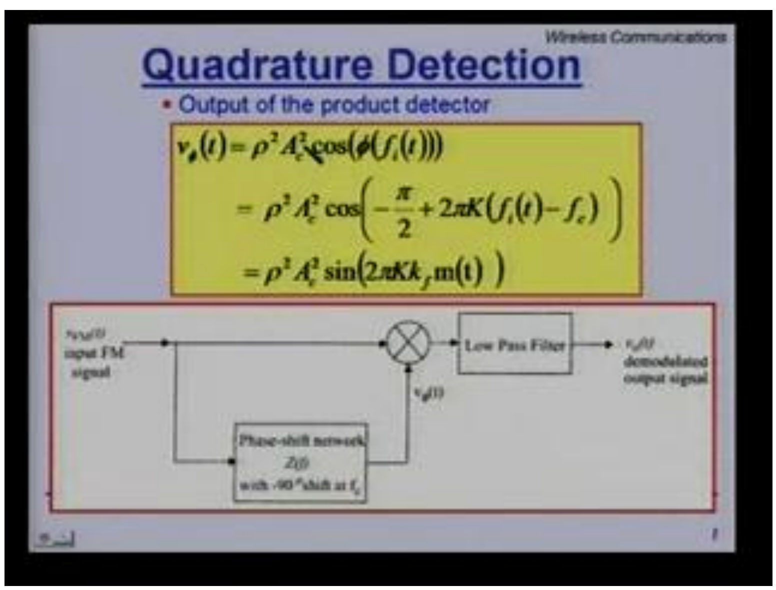
This slide shows the characteristic of a phase shift network with constant gain and linear phase. In this diagram, if you look on the y axis is the phase shift and on the x axis is the frequency. Here is the FM signal and this is the linear phase shift.

(Refer Slide Time: 00:30:37 min)



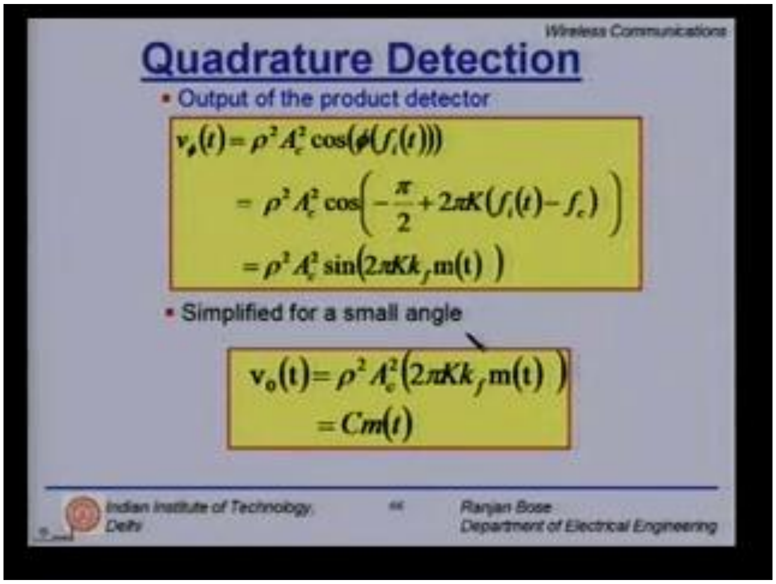
So the phase response function for a quadrature detection method is given by phi as a function of frequency is - pi by 2 + 2 pi K_f - f_c. This is the error in the frequency. The output from an FM signal is given by the following. The instantaneous frequency f_i is $f_c + k_f m(t)$.

(Refer Slide Time: 00:31:13 min)



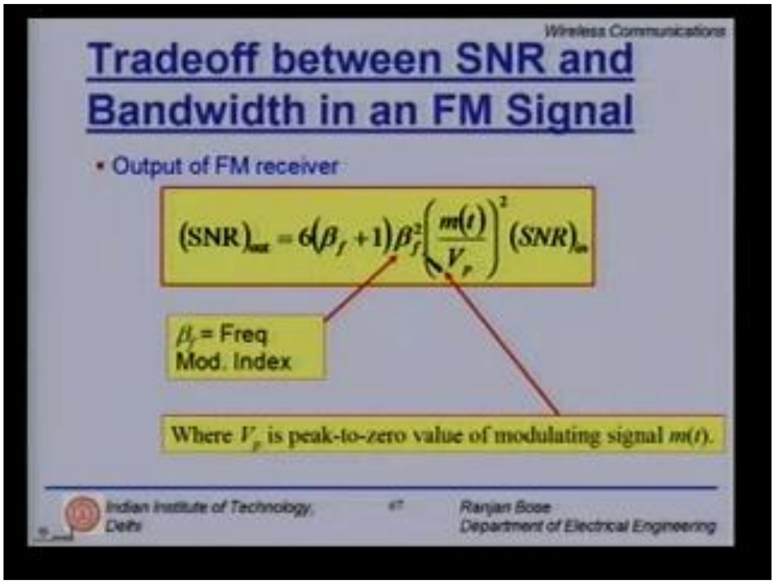
Now what happens at the product detector? At the product detector, if you remember, this is the diagram of the product detector. So the v phi t is rho square A_c squared cosine phi $f_i(t)$. It is simply given by this and if you work it out, you can write it as rho squared Ac squared sin 2 pi k k_f m(t).

(Refer Slide Time: 00:31:43 min)



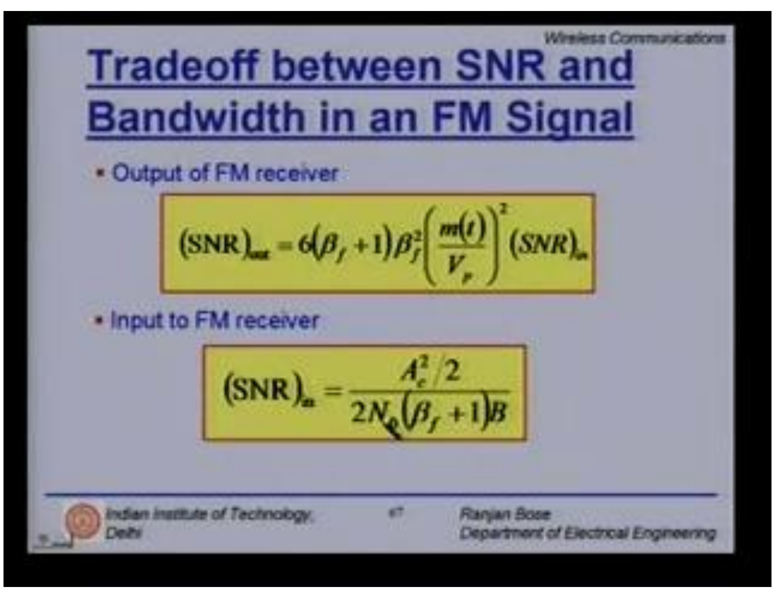
When you simplify it for a small angle, then the V $_{out}$ can be written like this. Well, there are lot of terms here but at the end, rho squared, Ac squared, 2 pi, K, k $_f$ are all constants. You are left with a constant times m (t). So this quadrature detection technique is a very simple method to obtain m (t). As you can see, it is fairly easy to implement in hardware. There is just a phase shift network, a multiplier and a low pass filter and hence its popularity. So the bottom line is if you do the basic math, you come up with V $_o$ (t) as some constant times m (t) and once we have got this we have recovered our modulating signal. Our job is done.

(Refer Slide Time: 00:32:46 min)



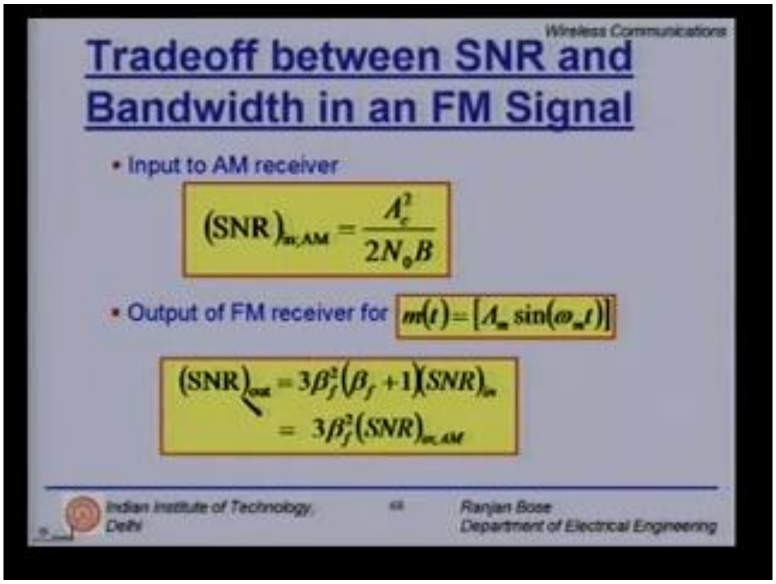
Now let us spend some time talking about the trade- offs between the signal to noise ratio 'SNR' and the bandwidth in an FM signal. so if you work out the output of an FM receiver, it is given by $(SNR)_{out}$ as 6(beta $_f$ + 1) where beta $_f$ is the frequency modulation index and V $_p$ is the peak to zero value of the modulating signal m(t) times the signal to noise ratio at the input. So post processing, this is your SNR. Please note there is a beta $_f$ and a beta squared here. So somewhere it is proportional to beta $_f$ cubed. Remember beta $_f$ is the modulating frequency.

(Refer Slide Time: 00:33:57 min)



The input to the FM receiver $(SNR)_{in}$ is given by Ac squared by 2. This is nothing but the power of the carrier and here it is the noise perspective density multiplied with the bandwidth. so this is the $(SNR)_{in}$. So using these two, we can obtain the $(SNR)_{out}$.

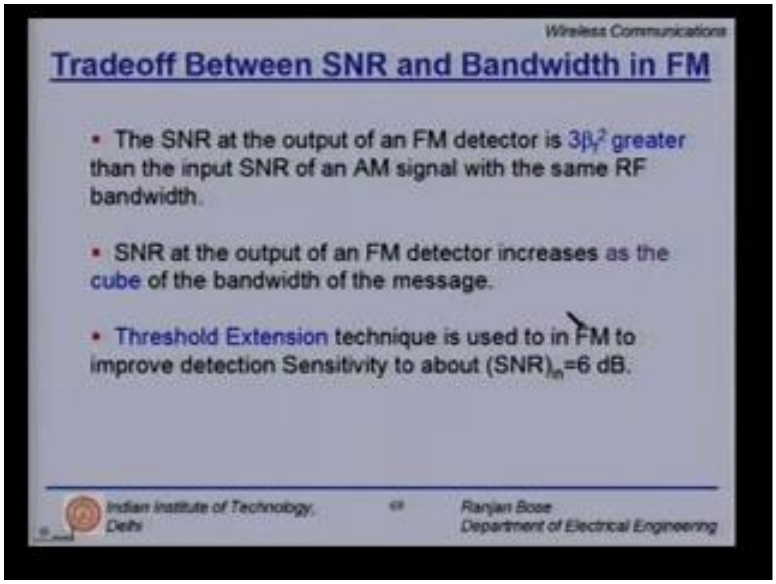
(Refer Slide Time: 00:34:31 min)



Now if you want to compare, the input to an AM receiver is nothing but Ac squared by 2. The 2 should go up. And N_0 times B. B is the bandwidth. The output of the SN-FM receiver for a sinusoidal m (t) is simplified as $(SNR)_{out}$ by plugging into the previous value, you get 3 beta $_f$ squared $(SNR)_{IN, AM}$. so as you can see that, with respect to the SNR $_{IN\ AM}$ given by this formula, I have a flexibility, a design possibility to improve my $(SNR)_{out}$ simply by playing and increasing the beta $_{f.}$ so the question being asked is: why are we considering this SNR and this SNR (Refer Slide Time: 35:48)? So there are two things. One is the SNR at the input and then post processing, there is an SNR at the output. This is for both AM and FM. However, what we are trying to do in this slide is to see what is the relative advantage of FM over AM. Now we need some common platform to compare.

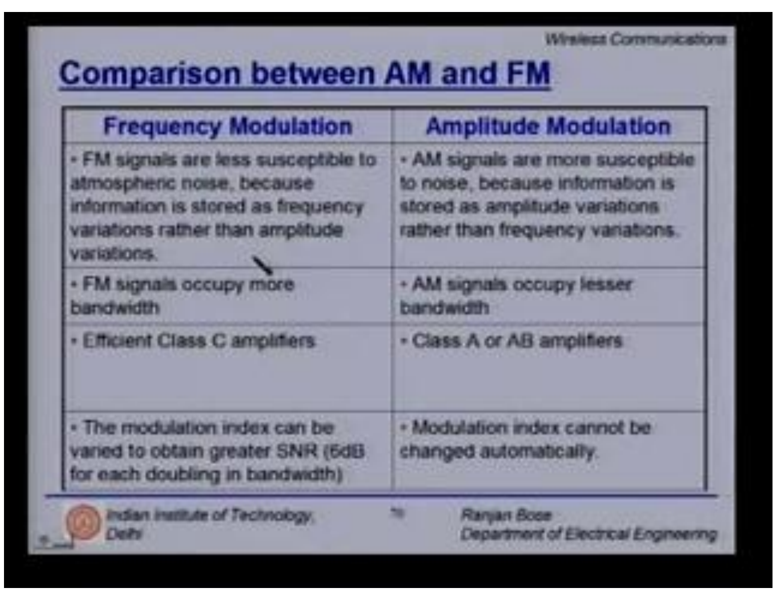
So, what we find out from very basic calculation that the $(SNR)_{IN}$ at AM is given by this formula. We saw last time that this SNR is signal to noise ratio which is the signal power divided by the noise power. Signal power is Ac squared by 2 and noise power is N_0 beta. We are assuming additive wide Gaussian noise. Then if you write it in terms of your previous equation, this SNR out for this one (Refer Slide Time: 36:51), you obtain SNR $_{out}$ for FM as 3 beta f squared $(SNR)_{IN}$ for AM. So I have a possibility to tweak and improve my SNR performance by just changing beta $_f$. so I can very easily compare and tradeoff. Please remember beta $_f$ comes at the cost of bandwidth. So I can play with the bandwidth and hence improve beta $_f$ and thereby improving my SNR. There is no such luxury for the case of AM.

(Refer Slide Time: 00:37:36 min)



So the SNR at the output of an FM detector is 3 beta $_f$ squared greater than the input SNR of an AM signal with the same RF bandwidth. SNR at the output of an FM detector increases as the cube of the bandwidth of the message. The threshold extension technique is used in FM to improve the detection sensitivity to about (SNR) $_{IN}$ of 6 dB. These are relative advantages of FM clearly over AM.

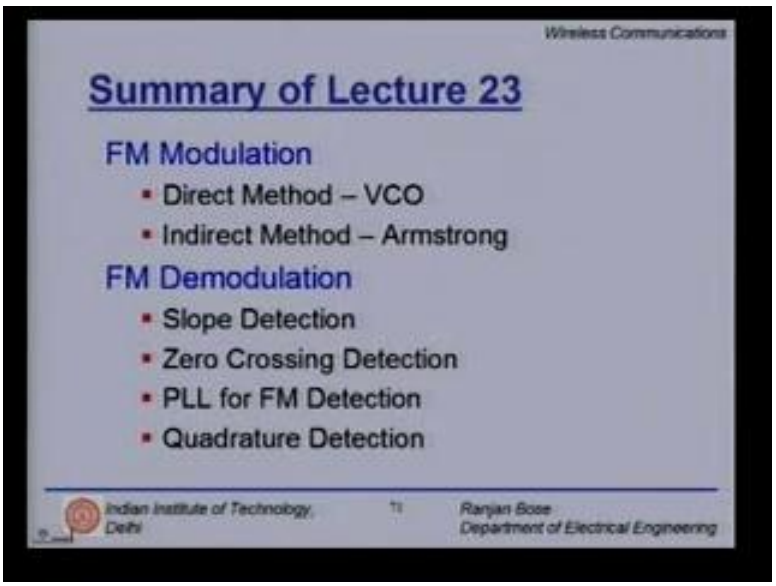
(Refer Slide Time: 00:38:17 min)



A brief slide on the relative comparison. On the left we have the attributes of frequency modulation and on the right is amplitude modulation. So FM signals are less susceptible to atmospheric noise because information is stored as frequency variation rather than amplitude variation. But noise as you know, is additive wide Gaussian noise hits the amplitude directly. AM, on the other hand is most susceptible to noise because of the amplitude encoding. FM signals clearly occupy more bandwidth.

AM signals are more bandwidth efficient. We saw last time that for frequency modulation techniques we can use efficient class C amplifiers at the transmitter which is cheaper. Whereas for class A and B amplifiers must be used for amplitude modulation. The modulation index can be varied to obtain a greater SNR. 6 dB for each doubling in bandwidth. For AM, modulation index cannot be changed automatically. So we have a flexibility to improve our output SNR for FM.

(Refer Slide Time: 00:39:40 min)



So let us summarize today’s lecture. We started off with the different frequency modulation techniques. Specifically we talked about the direct method using the VCO or the voltage controlled oscillator. Then we talked about the indirect method, the one proposed by Armstrong. We looked at four popular FM demodulation techniques. We started off with the slope detection followed by the zero crossing method, the phased lock loop for FM detection and finally the quadrature detection. We will conclude our lecture here and in the next lecture, we will talk about digital modulation techniques.