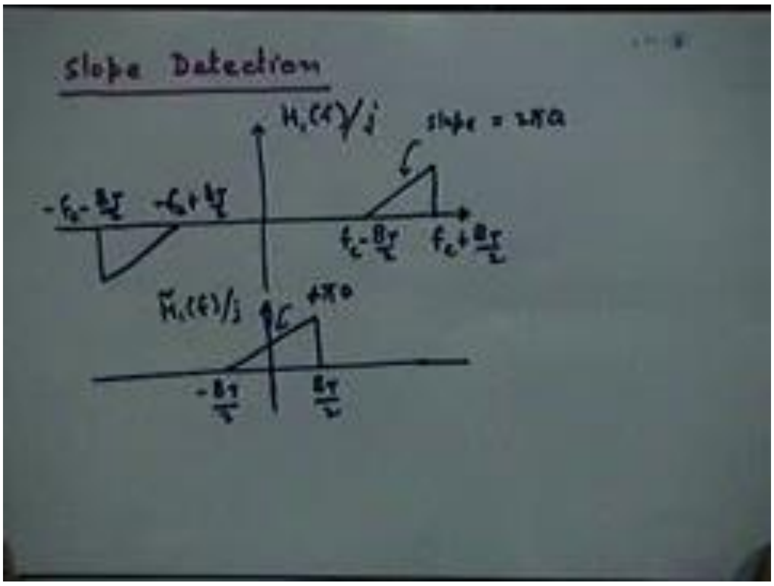
Communication Engineering
Prof. Surendra Prasad
Department of Electrical Engineering
Indian Institute of Technology, Delhi

Lecture - 21
Demodulation of Angle Modulated Signals (Contd.)

We were looking at the Demodulation of Angle Modulated Signals, we had discussed several ways of doing this and we would now trying to develop a mathematical model for the balance discriminators that we talked about earlier. If you remember a balanced discriminator is essentially two slope detectors, which are used together to form an overall discriminator. Such that, the output is proportional to the input frequency deviation that is the purpose of the balanced discriminator.

The second purpose of the balanced discriminator is and that is the reason, why it is called balanced discriminator is that the output of this balanced discriminator should be equal to 0, that the input signal has a frequency equal to the carrier frequency. So, that the 0 or d c input response of the FM signal is maintained.

(Refer Slide Time: 02:22)

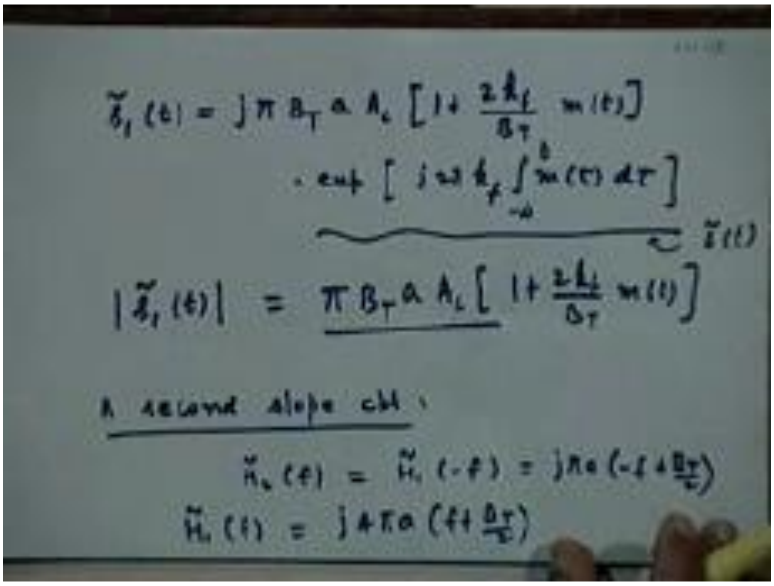


So, if you recollect we were first looking at this idealized model of the slope detector, which is not that of the balanced discriminator, but that of an unbalanced discriminator. The idealization was, if we assume that we had a slope circuit, this frequency response must proportional to the frequency deviation from f c. but, it was going only on one side, well it was going on both sides, where you it was having a d c response equal to a non 0 d c response corresponding to f c. So, the f c is somewhere here and as your input frequency is varied around f c, the response of the circuit could vary linearly in this fashion.

That is the slope detector, this of course; this circuit is followed up by an envelope detector, whose output is proportional to the modulus of this. And therefore, the output is proportional to the frequency deviation of the input signal as we want, but with that difficulty that had the character frequency the response is not 0, let us have finite value.

We are arise rise the circuit is mathematically by using a complex envelope representation of this filter, this circuit, which is like this and we went through this equations describing the input and output.

(Refer Slide Time: 03:52).



And finally, reach if we recollect this equation for the output complex envelope, in terms of input which was taken to be this. This is the complex envelope of the input, this is the complex envelope of the output and as you can see, this is s tilde t here, the complex envelope of the input, as you can see the modulus of this, which it presence the envelope of the output. We have formed is given by pi B T and basically this term, the multiplying term, because the modulation of this will be unit always.

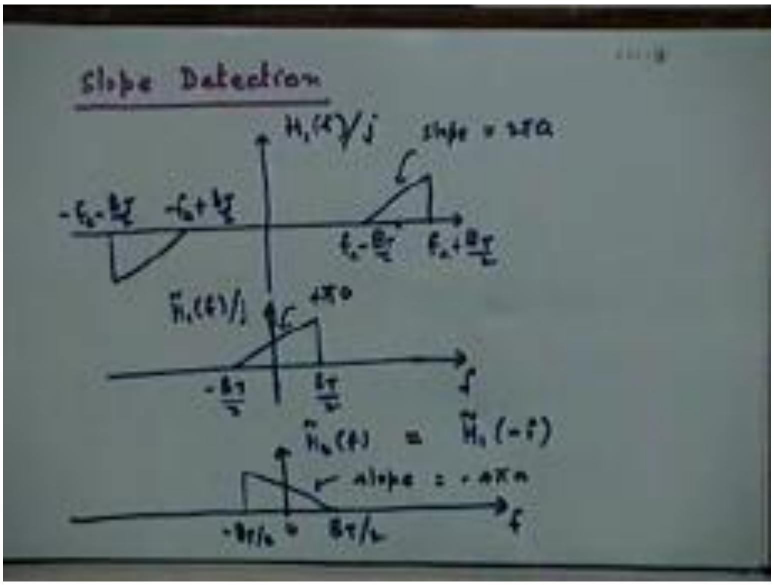
So, that expression for the output in terms of the input signal and as you can see this equation shows that the output envelope of this circuit is dependent linearly on the input message signal m t. As we want with the problem, that we have already discussed, there is a DC component which we did not want. Basically, if the same thing that we physically argued earlier that this slope circuit converts the FM signal or the angle modulated signal into an amplitude modulated signal, this represents the amplitude modulation.

And this represents the carrier, the only thing is in this case the carrier is not one of constant frequency, but the carrier itself is an FM modulated carrier. However, the way the envelope detector works, we have seen how envelope detector works that is a diode followed by an RC circuit. It hardly matters, whether the carrier has the constant frequency or whether the carrier phase keeps on varying.

The operation of the detector is essentially based on detecting the peaks of the carrier, we tries to track the peaks and it would not matter, whether the at times the cycles come close together or cycles are slightly far away from each other. The operation of the envelope detector will be such and then it will produce an output, which is proportional to the instantaneous envelope which is precisely this. So, the modulus of this is equal to this, which is an envelope detector output.

Now, what can we do to take care of the fact that there is a DC component present here, which we did not want and as we know the answer is to have a second slope circuit. With somewhat different characteristics than the slope circuit that we using here shows to balance this out that is what we need to do. So, what can we do to balance it out, so let us look at this slope circuit again that we have considered.

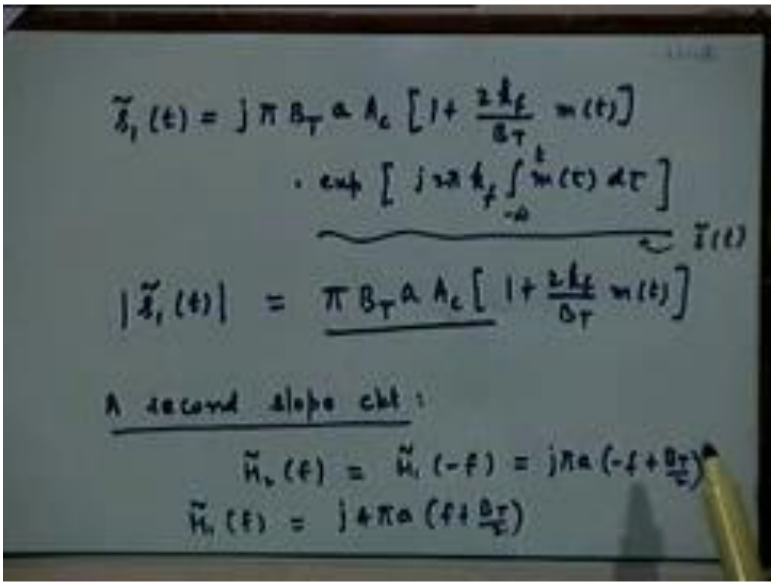
(Refer Slide Time: 06:47)



This one a slope circuit we took now at this stage, so what should we do to balance this out, all we need to do is following, we need to take a rear image of this response, this frequency response around f equal to 0. So, if we consider an H 2 f tilde, which is like this, which is equal to H 1 tilde minus f. Basically, I have taken inverted image of this, around f equal to 0.

If you do that, that is the slope instead of being plus 4 pi a is slope varies minus 4 phi a and this is the exact relationship between the earlier circuit and this circuit, then we will have the very interesting situation. Let us see, what will happen, so we will proceed with an analysis with this kind of a slope circuit in mind.

(Refer Slide Time: 08:16)



So, we are considering a second slope circuit or slope filter, whose complex frequency response of the corresponding low pass filter, you see the actual filter is a bank pass filter this is the complex envelope representation of that, this is the low pass representation of that more precisely. We are looking at a low pass representation and the corresponding complex bank pass representation would be something similar, except that it will be like this, may be if I think draw it.

But, let me first write down the equation for a low pass case, so H to tilde f, which is H 1 tilde minus f and if you remember the equation for H 1 tilde f, what was it, H 1 tilde f was j 4 pi a, f plus B T by 2, that was the expression for H 1 tilde f. So, what will be the expression for H 2 tilde f, which is the same thing as this except that f is replace by minus f.

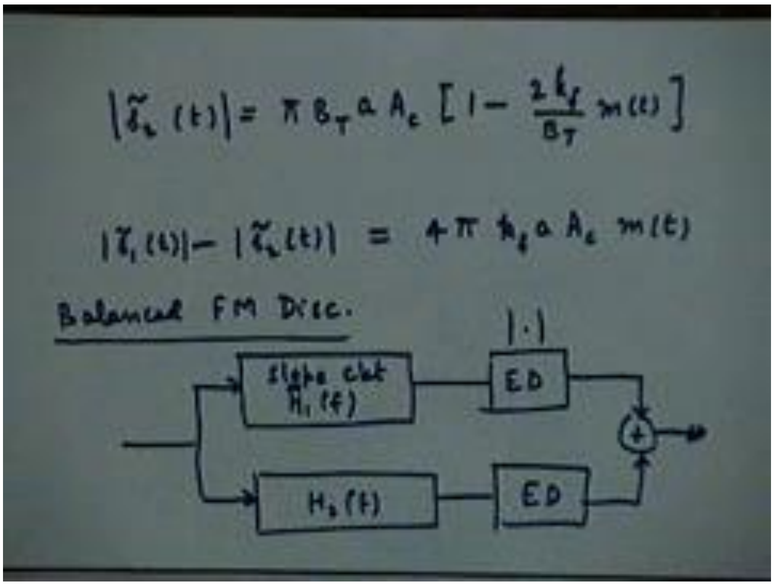
So, this becomes j 4 pi a minus f plus B T by 2 and as you can see, this matches with what we have drawn here, that f equal to B T by 2, the output will be 0, that is response is 0 and f equal to minus B T by 2 corresponds this value which is this 1. So, this second slope circuit that we are going to have. Now, if you compare this with some of the equations that we had looked at in this connection earlier at it, which point will the things change. If you remember, basically what we have seen is since H 1 tilde f is like this and I think, let me see which is right equation, this is a 1.

This is the equation that we use to compute the output envelope in terms of the input envelope and the frequency response of the circuit. So, everything remains a same expect that f 1 tilde f will get replaced with H 2 tilde f. So, what will happen instead of j 2 pi a f plus B T by 2, you will have j 2 pi a minus f plus B T by 2. This f will get replaced with minus f and if we go to the time domain, which was the next step.

The next step was to go to the time domain, so the time domain equation here would become instead d s tilde by d t will have minus d s tilde by d t, because this is j 2 pi a f s tilde f. The corresponding time domain operation is differentiation operation with the negative sign, how about this will remain is the positive sign and everything else is same.

So, you will get the same expression as before with a minor difference with the coefficient corresponding to this term, will turn out be negative that is about the only difference and which was that term. If you look at the final expression out of these two terms, which is the one, which is coming from the derivative, which is the second term, the first term is the constant term, coming from here. So, the first remains as before the second term also remains as the as before, except that this plus sign gets replaced with the negative sign.

(Refer Slide Time: 12:28)



So, I do not have to repeat the analysis that we did earlier, we can directly write the result. The result would be that s 2 tilde t would be equal to pi times B sub T into A sub C into 1 minus 2 k sub f upon b t into m t. This is the envelope of the response; I want you writing the envelope, everything else is same. So, the second slope circuit will produce an output like this, so what do I need to do to get to balance out the d c, all have to do is sub track the 2 outputs.

So, if I look at s 1 tilde t envelope of that minus the envelope of s 2 tilde t, you will get an output it is equal to 4 pi k f a A C into m t, B T will cancel out is that correct and the overall balanced FM discriminator. Becomes then, you need to have a slope circuit with frequency response H 1 tilde f or H 1 f, I think here I should not write H 1 tilde f simply H 1 f.

That is the bank pass filter, will be the second slope circuit H 2 f will follow them up with envelope detectors basically modulus operation and subtract the 2 outputs. So, that is the mathematical theory, behind the balanced discriminator. So, as we can see the balanced discriminator, that we discussed earlier, essentially tries to realize this.

Remember, we had two tune circuit is, one which was approximating one slope circuit and the other which was approximating the other slope circuit that is all that we have. And that was precisely, what series mathematical theory, now precisely explains, why the difference in the envelopes of the two will be your required demodulated output. So, I hope things now for in place and you have fairly good understanding of how those staggered tuned balanced FM discriminator work.

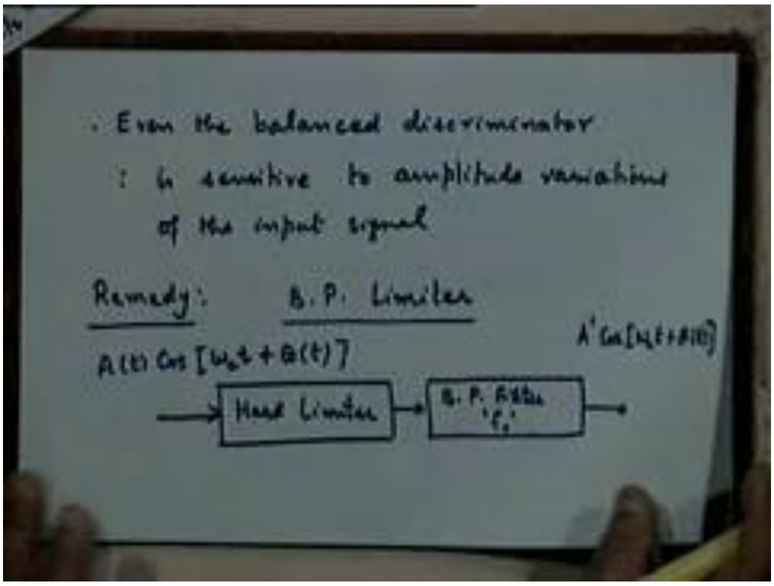
This is the basic theory and as I mentioned earlier and I like to repeat it today, this is the most basic form of a discriminator, the other improved versions of the balanced discriminator about which I request you to read yourself, the self study exercise. So, I will I hope that you will do that, if you have not all done. So, that is about FM discrimination or demodulation of an angle modulated signals.

Now, let me come back to a point, which I mentioned sometime at the beginning of this discussion on FM discrimination or beginning of FM demodulation. You see, in all this theory also, you would have realize that, a there is a diff derivative involved at various stages, there is a differentiation involved, which we have used. Some form of the other, differentiation comes with the picture; d s tilde by d t and because of that, your output envelope will have a dip.

In fact, it not only because of the derivative, but also because of this term, so both these terms, you can see that, if the input signals, it has envelope variations. They will be either passed on as such to be output or were still they will be extenuated, because of the derivative operation. So, sometimes even though the envelope variations themselves will be small, the rate of variation derivatives sometimes may be quite large and then they can get extenuated.

So, your output in that case will be proportional not only to m t as it should turn out to be, because this entire query was based on the assumption that input signal had a constant amplitude. If the input signal does not have a constant amp amplitude has an envelope of it is own, time varying envelope of it is own, not a constant envelope, then those variations will that passed on to the output, which we do not want. So, the balanced discriminator that we have discussed so far is sensitive to amplitude variations in the input signal, so that is something that you must appreciate.

(Refer Slide Time: 18:19)



So, even the balanced discriminator, whether it is balanced or unbalanced, it does not matter. The kind of discriminators that we have discussed or sensitive or this is sensitive to amplitude variations, which we do not want. And we also discuss the remedy for this problem what was the remedy, the bank pass limiter and basically I give you the intuitive field for how does bank pass filters work and what is it do.

Let us, develop a slightly better understanding of that to basically polish of our discussion on FM demodulation, because this is an important a very almost an essential component of the FM demodulator. Before, you pass the FM signal for demodulation into the FM demodulator or into that balanced discriminator; you must pass it through a bank pass limiter.

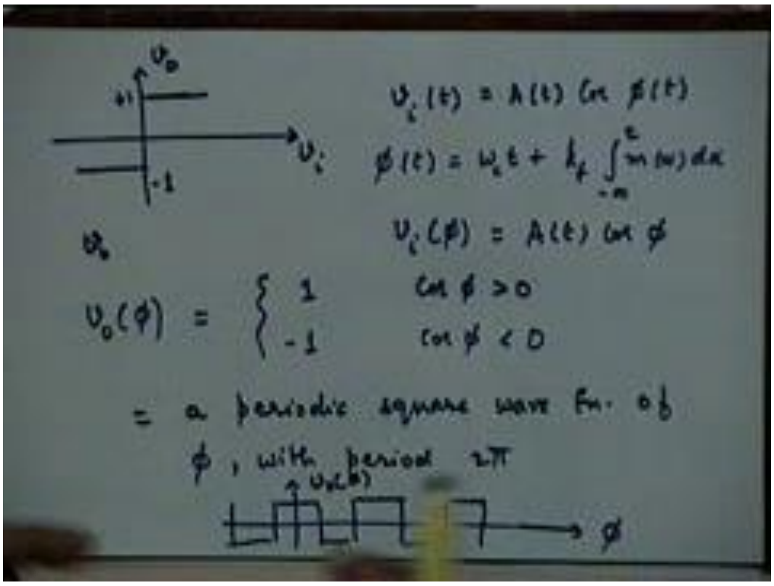
The purpose is to recollect to converting the input signal into the constant amplitude signal, input signal, which may have some envelope variations. Either, because of some distortion introduced by the system or because of noise or whatever reason, you like to kill, if they are in any amplitude variations present, we like to kill those amplitude variations. At the same time, we like to get that FM signal back intact and that is what a bank pass limiter does for us.

If you remember the block diagram of a bank pass limiter is something like this, you have a hard limiter, remember what is a hard limiter, followed by a bank pass filter, tune to a the frequency f c, the carrier frequency f c. This system was a whole is known as bank pass limiter. So, the input to this is, we are saying is a signal, which not only contains your angle modulation that is cosine omega c t plus theta t.

But, the instantaneous amplitude which should have been a constant value, may have some temporal variation, may have some variation with respect to time, which we want to kill. Because, if you do not do that d s tilde by d t and s tilde t that are present in the relationship that we discussed, will produce at the output some component, which is not proportional only to m t, but also additional variations because of these things.

And what we therefore, hope the bank pass limiter to do for us is the produce an output, which is some a prime, cosine, omega c t plus theta t. We should keep our FM signal intact, but amplitude variation should get, this is what you want, I think precisely what this circuit does, what this system does. Let us see, how what is your query behind this, how does it work to produce this, precise this, fairly the simple theory, but let us go through it, because it is important to have a very clear field for these things.

(Refer Slide Time: 22:33)



Let us start with the hard limiter, hard limiter is a device as I mentioned earlier is some something like a highly saturated amplifier or you could say, it is a clipping circuit which clips at 0.

So, if you about to plot the output versus input characteristics, output voltage versus input voltage on an instantaneous basis, then while amplify it is a straight line passing through the origin. For a hard limiter, it is a highly non linear system, this is what the characteristics are, if the input is even slightly positive, the output becomes 1, the input is even slightly negative, the output becomes minus 1 that is the definition of the hard limiter.

As against hard limiting, we can have a soft limiter, which goes something like that, it also limit is ultimately that does not limit is the moment the signal becomes either positive or negative, it carries out the operation most smooth. So, in our case, what we are saying is our input v i t is some a t into cosine of let us call it phi of t, the phi of t is omega c t plus k f, let us say it is an FM signal, this is the input to the system.

So, we can think of the output, now as a function of phi, the input is a function of phi, so think, if element is the input not as a function of time, but as a function of this variable phi, the phi is this complete function of time, phi of t depends on time. So, let us say, there is a parametric relationship between phi of t and t, so you can think of v i as a function of phi, it is simply something a t cosine of phi, let us try to understand this function that we get.

Now, let us particularly, let us try to understand, what is v not as a function of phi, what can we say, what we are saying is, this is going to be equal to 1, when v i phi is equal to is greater than 0 is that correct, v not phi is going to be equal to 1, if v i phi is greater than 0. And when will this be greater than 0, we are assuming of course, a of t is always positive, because it is the envelope.

So, essentially, whether vi phi is positive or negative will depend on whether cos phi is positive or negative, so it is equal to 1, if cos phi is greater than 0 and it is equal to minus 1, when cos of phi is less than 0. Now, if you forget the temporal relationship, for a moment forget about the time dependencies that we are talking about. So, what can we say as a function of phi, what is the nature of the function v not phi as a function of phi.

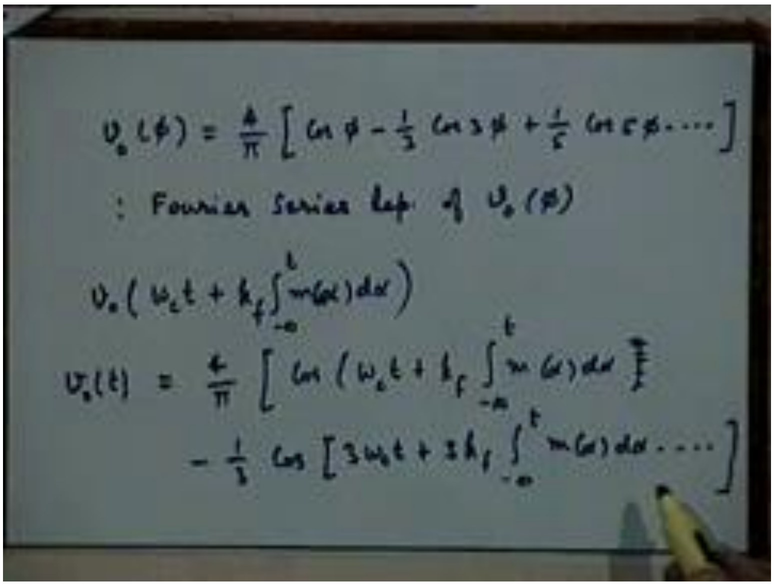
It is periodic, because cos of phi as a function of phi is periodic and this as a dependence on the nature of cos of phi, cos of phi is equal to is greater than 0. For half the time, as we know as phi, goes from 0 to 2 phi and it is equal to less than 0. For the other half of the time, I should use the word time for half of the cycle and for the full cycle as phi goes from 0 to 2 phi. So, in as much as cos of phi is a periodic function of phi, v not phi is also a periodic function in phi.

And clearly therefore, in what kind of a periodicity is this, suppose you about to plot this periodicity as a function of phi, it is a periodic square wave. So, v not phi is a periodic square wave function of phi with period 2 phi, this is clear to everyone. If, I plot v not phi as a function of phi is essentially a square wave, mind you I am not plotting this as a function of time, I am plotting v not as the function of phi, v not caring, how phi depends on time that is a separate issue.

As far as phi is concerned, it is a constant, we do not even have to worry about it, always saying in fact, we are not worrying about, what is a t, we are worrying about what is v not phi, v not phi will not have any dependence on a of t. We can have a dependence only on cos phi and cos phi is greater than 0, it is plus 1, a of t is just disappears from the picture in our consideration, it is absolutely of no consequence.

So, this is an important point that this thing as a function of phi is this square wave function with a period of 2 phi and therefore, I can represent this function in terms on a Fourier series, not as a function time as a function of the phase phi. And I know, what kind of a Fourier series as square wave function has, it do it contains only our arguments of the argument and the precise representation is as follows.

(Refer Slide Time: 29:23)

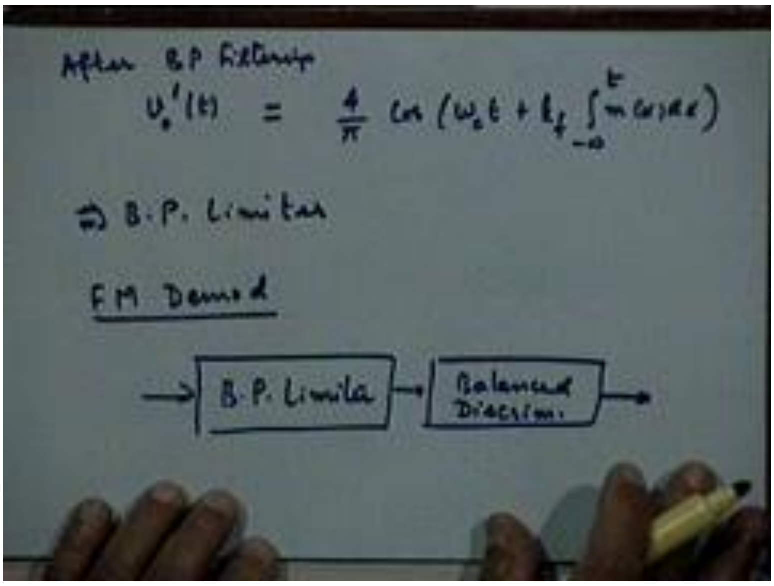


So, v not phi, which is the square wave function as this Fourier series representation, 4 by pi into cosine of pi minus 1 by 3 cosine of 3 phi plus 1 by 5 cosine of 5 phi and so on, this is the Fourier series representation of v not phi. So, now let us substitute for phi, you know what is phi, phi is this function of time. So, we can write therefore, that v not as a function of what phi is which is omega c t plus k f integral m alpha, d alpha. So, this is substituting for phi v not as a function of this.

So, I am thinking the time dependency back, now if I substituting for phi is equal to 4 by pi and you will have all these terms cosine of omega c t plus k f integral m alpha d alpha, you should not close this bracket here minus 1 by 3 of cosine of 3 omega c t etc and so on. You can think of this as a function of time now, so v not as the function of time, because this is not the function of time. So, I can instead of saying this is a function of pi, I am writing this as a function of time, because the time dependence is very explicit in this relation, this is what we have.

So, this is a series representation for the output signal at the output of the hard limit, what are we following it out with, we are following it off with the bank pass filter tune to f c. And as we can see, all other components of the Fourier series representation of v not t, they are around 3 omega c phi omega c as so on and so forth. So, if the bandwidth is sufficiently small to only theta for the bandwidth of the FM signal and reject the higher frequency component, you are simply remove all this, you are filter them out.

(Refer Slide Time: 32:41)



And therefore, your v not t, after bank pass filtering, let me call it v not prime t could be only the first term of this Fourier series, which is 4 by pi cosine of omega c t plus k sub f integral m alpha d alpha. So, as you can see independent of what is a value of a of t, the output has a fixed amplitude equal to 4 by pi and your FM signal, the FM variations remain intact as before and that is how the bank pass limiter works.

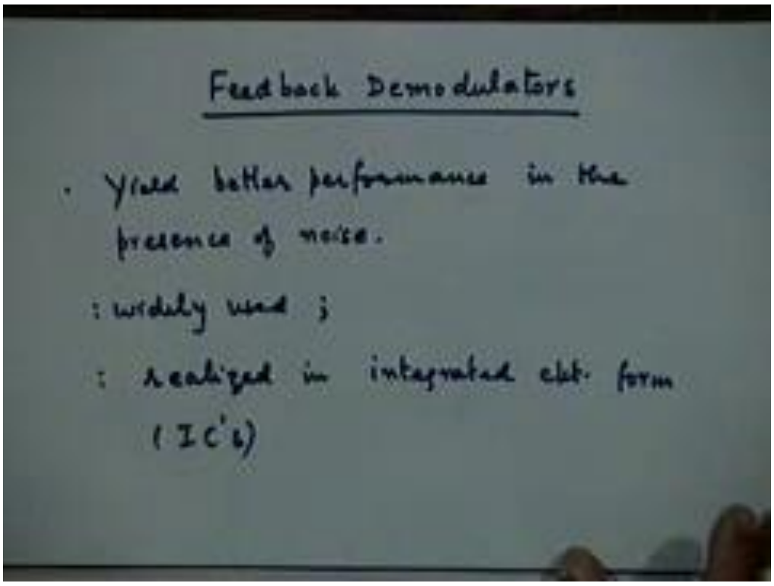
So, I hope now with this discussion, it is clear the role, how the bank pass limiter works and it is it should also be clear that an overall FM demodulator must contain both these devices. It must first contain bank pass limiter, followed by preferably balanced discriminator and together, they form your FM demodulator. In fact, many of the advantages of FM signal that we will talk about later.

Particularly, it is increase tolerance to noise comes from the fact that the demodulator is preceded by the bank pass limiter, without this, those advantages will not be available to us. And of course, is it is a required and necessary component for a proper FM demodulation, particularly in the presence of channel imperfections. If the channel is perfect, we do not need this, if the channel causes no amplitude distortion, no amp envelope variations of the input signal, we do not need a bank pass limiter.

But, invariably can willing will impose some distortions will include some noise and therefore, they will be some amplitude variations and therefore, we need to have a bank pass limiter, any questions at this point. So, can we assume that the principles of FM discrimination and bank pass limiter are sufficiently clear now? We are not through FM demodulation as such because, what we are discuss, so far is one class of FM demodulators, which are basically open loop systems.

They are systems in which we are trying to device circuit is such that the output is dependent on the input frequency deviation with respect to the center frequency. They are also close loop FM demodulators and I will which also called feedback demodulators.

(Refer Slide Time: 36:19)



So, let us spend some time on those things, because what we are going to discuss in this collection is going to be important from several points of you. This feedback demodulator for initial motivation for the discussion is to show the very important classes of FM demodulators or angle demodulators. However, they also have other applications for example, applications to which we referred to earlier, when we are discussing synchronous amplitude demodulation.

We talked about the need for a local carrier, which is synchronized with the incoming carrier, particularly when we wanted to demodulate d s p as a series. So, how do you achieve that synchronization; that is again achieved through the same feedback devices, which we are going to use for FM demodulation. So, although our initial motivation is to show, how this so called feedback demodulators can be used to carry out demodulation of FM signals will also discuss some other applications of these separately or as we go along.

Basically, the motivation for using these kinds of demodulators over the normal discriminators that we discussed so far is that they yield in two performances, better performance in the presence of noise. It is improper to discuss it fully here, because we have not yet considered the effect of noise on FM, an AM for that matter. But, when we do that and let me just that discussion little bit that we give you a preview of the discussion, when you do that will find a very important characteristic.

The important characteristic that we will study away is that, while the FM in general is better in the presence of noise, which gives superior performance in the presence of noise and that is why it is the preferred modulation to use for FM broadcast in music broadcasting, it also has the disadvantage. The disadvantage other than the bandwidth disadvantage which of course it has, because it uses much larger than it.

The other disadvantage it has is, it displays a phenomenon which is called Threshold effect, the Threshold effect refers to the fact that the output is of very noise, if the output is good, because the input signal to noise that is why reasonably large. So, there is noise present, but not to an extent that something bad happens and this is the problem. In FM, if the input signal as I show deteriorates beyond a certain point, becomes very bad. Let us say, then the output of the FM demodulator, I mean simply face to resemble the input at all the FM demodulator, simply breaks now.

So, there is a certain Threshold SNR, which you must maintain at the input of an FM receiver, if your signal is not strong enough, then the output will be of very bad qualities or in a good quality. So, this is the Threshold effect which an FM receiver exhibit is the results for which will discuss later, but it does that. So, as you know, usually when an input SNR goes down the output SNR also will go down.

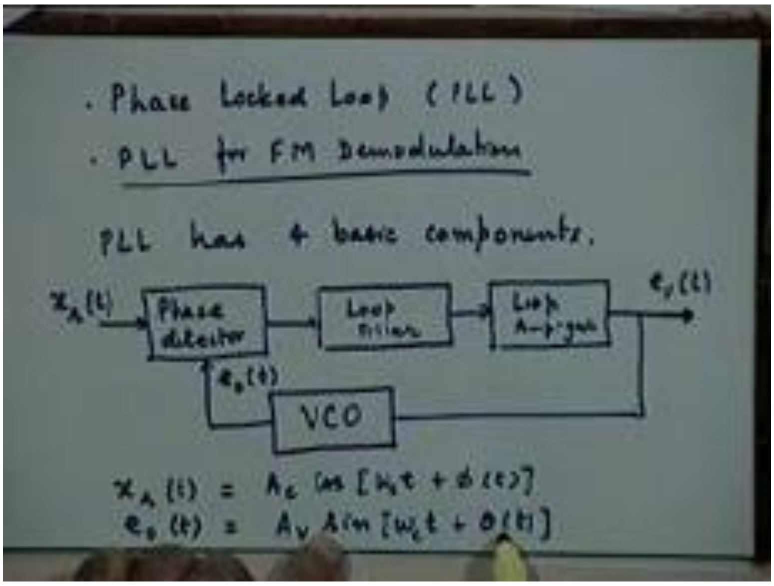
There is nothing wrong with that; that will always happen in any system, there is more noise at the input, there will be more noise at the output, because typically we are working as a linear behavior that will always happen. But, usually this relationship will be a graceful relationship, a proportional relationship that is, if the input SNR is goes by certain amount, the output SNR is also goes by the same amount, whatever the actual values.

But, if this relationship becomes worst than the graceful dependent, then that phenomenon we called the Threshold effect and that is what happens in FM. It is very nice up to a certain point, but it is very badly on that point. So, it is either very good or very bad nothing in between, something like that has to why that happens, we will see that later.

And one of the advantages of going for feedback demodulators is that it can extend the Threshold, it can lower the Threshold for us. So, that is that is one of the reasons. So, it is widely used, so therefore, feedback demodulators are widely used. And second advantage of these feedback demodulators is that they can be easily realized in integrated circuit form, so in the form of IC's.

The reason for that is typically in these kinds of feedback demodulators, we do not use any inductors, your FM the balanced discriminator that we discussed so far is based on the use of tune circuit is staggered tune circuit. Is the two slope circuit, that we have talked about, they will be typically relies by using a linear portion of the FM tune of a resonant circuit, parallel LC circuit. Whereas, there will be no LC component in this realization, so that is the advantage of this feedback demodulators.

(Refer Slide Time: 42:25)



There are two basic kinds of feedback demodulators and we will discuss first the most important of these, the more important of these two is called the phase locked loop or PLL. So, we are going to discuss the use of PLL for FM demodulation, but mind you PLL is a very, very important component and it has many applications. FM demodulation is just one of them, PLL has four basic components will discuss as to why we call it by this name little later.

We will be first describe, what it is, it has four basic components and there is a feedback involved amongst this components. The four components are there is a phase detector which operates from 2 inputs, let us say they are two sinusoidal inputs here, the output of the phase detector would be proportional to the instantaneous phase difference between these 2 inputs that is why we call it a phase detector. The output here would depend on the instantaneous phase difference between these 2 inputs; this is followed by a filter which is called a loop filter.

We also have some amplification in the loop, so we call it a loop amplifier, sometimes this is club together with the filter and say the filter itself has some gain. So, it has some associated gain and then it closes a loop like this and the feedback path closing a loop. We have a VCO remember, what is a VCO a voltage control oscillator, which we have discussed in the context of FM signal generation.

So, a VCO is present in the feedback path of this circuit and the input to the VCO at this point, will argue can serve as if you feed electron signal as the input, this input signal can be taken as the output. Let us call this output as e sub v t input to the VCO signal; we will see that under certain condition e sub v t will be proportional to the message signal m t, where the input signal happens to be an FM signal.

Now, how does it work, you can discuss in sufficient more detail as to how it works and what are it is characteristics, but let me just give brief idea today, as a input you have an FM signal. Let us say input signal is x sub r t, which is A sub C cosine omega ct plus phi of t, it contains the FM modulation. It contains the message signal, in terms of it is dependence on, in terms of the effect that phi of t will depend on m of t, it will be proportional to integral of m t any may FM signal.

Your VCO here in this loop will produce some other signal of similar kind; we do not what it will be exactly. So let us say this VCO signal, let us call it e sub 0 t or e sub o t, that will be some a sub v, sine of omega c t plus theta t. And the phase detector; who tries to produce an output as close to phase term is sensitive to the instantaneous phase difference between these two.

So, it will be sensitive to phi of t minus theta of t and the feedback loop and the feedback loop, as any feedback loop will try to do is will try to derive this error to 0. This way is design, such this whole is thing is design in such a manner, that it will try to derive this error to 0. If suppose that happens, we have to discuss whether that happens or not, suppose that happens, basically what we are saying is this theta of t will try to follow phi of t. And, if theta of t phi is it follows phi of t, what is a relationship between the message and the phase.

Messages are derivative of the phase, is it not, it messages so; that means d theta by d t will follow the message, that means the input to the VCO. Because, after all the instantaneous frequency of the VCO will depend on the input voltage; that means this input voltage will try to follow the message signal, that is the basic idea. So, if we can ensure that this loop is said works in a manner that the instantaneous phase of the VCO output tries to follow the instantaneous phase in the input signal.

Then, the input to the VCO signal will try to follow the message signal and therefore, you would have got at this point the message signal, that you that you are looking for.

Have you understood, what I am trying to say, if not tell me, I will repeat again, any questions on this your questions, means what you have not followed or you have followed. You are keeping absolute mum on this, we will stop here, think about it and we start at this point again next time.