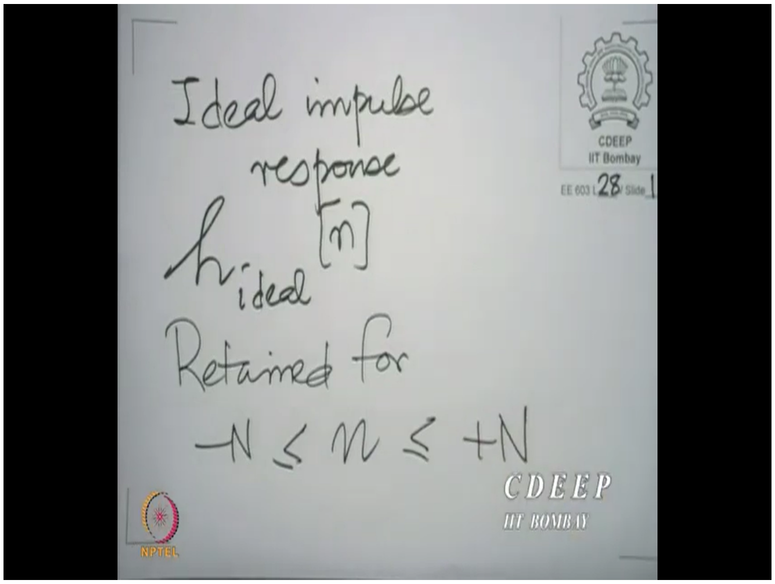
Digital Signal Processing and its Applications
Professor Vikram M. Gadre
Department of Electrical Engineering
Indian Institute of Technology Bombay
Lecture 28 A
Introduction to Other Window Shapes and Exploring the FIR filter

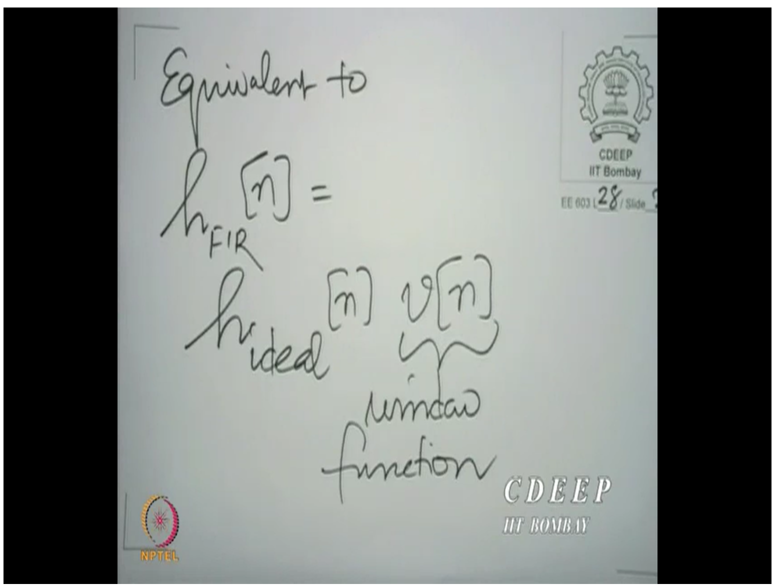
So, warm welcome to the 28th lecture on the subject of digital signal processing and its applications. We have commenced the discussion of FIR filter design in the previous lecture. And we had looked at the simplest approach to FIR filter design, taking a cue from how we represented irrational numbers in finite precision.

The easy way to do it is simply to truncate up to a certain level of accuracy. So, for example, we agreed that we would simply allow a certain number of impulse response samples to be retained. And naturally the choice would be, if you were given how many samples you could retain, then you would choose, in some sense, the most significant of the samples, the samples that made the most difference.

And intuitively the samples with maximum magnitude are expected to make the most difference. And typically, in most of the responses that we noticed, the most important samples lie around N equal to 0. And therefore, we chose a certain number of samples placed around N equal to 0, depending on what length we are permitted to make the impulse response of the finite impulse response filter. Now, we were trying to understand the consequence of this on the frequency response.

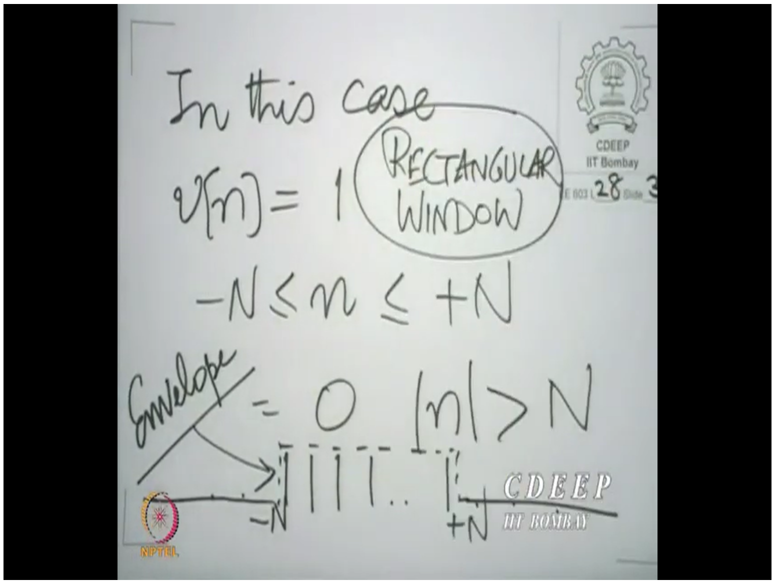
(Refer Slide Time: 01:56)





And we had noted that if we considered an ideal impulse response, $h_{ideal}[n]$, and retained this for $-N \leq n \leq N$, this is then equivalent to $h_{FIR}[n]$ or the impulse response of the FIR filter, given by $h_{ideal}[n] * v[n]$, where v[n] is a window function. And of course, in the specific case where we are retaining some samples and throwing out the rest, this window function is essentially just a constant in that interval.

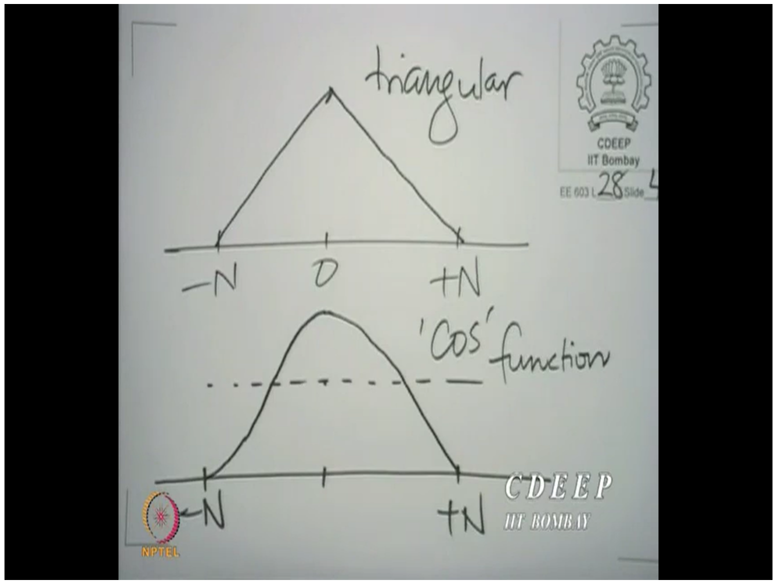
(Refer Slide Time: 03:12)



So, in this case, v[n] is 1 for $-N \leq n \leq N$ and 0 else, for $|n| > N$. Now, if we sketch the sequence, which is easier to do. This is the point minus N, this is point plus N, it is 0 else. And if we put around this sequence, what we call an envelope, like this, an envelope, which when sampled at the integers results in the window.

Then this envelope is rectangular. It is essentially a rectangular pulse. And therefore, we call this a rectangular window. So far, we have not seen any other window, but we can of course, conceive of other windows by putting other shapes in the same place.

(Refer Slide Time: 04:45)



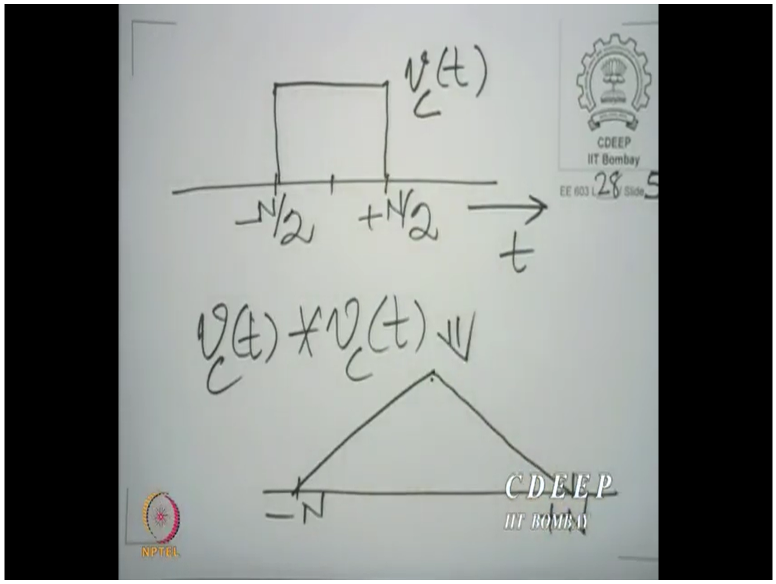
So, for example, between minus N and plus N, we could put a triangle, if we desire or we would put a cycle of a cosinusoidal, a cos wave, you know, you could put the cos wave displaced by some amplitude. And therefore, you could have the cosine function go this way.

And of course, we could conceive of many other windows like this. These are all examples of shapes that one can put between minus N and plus N. And then sample, each such shape, when it is put between minus N and plus N, is then sampled at every integer. So Vn could be essentially samples of different shapes, put between minus N and plus N taken at the integers.

Naturally, as capital N grows, the number of samples increases. So, in the previous lecture, although we had introduced the idea of a window, we had not seen why we should write a general sequence Vn, now we know. We could have shapes other than rectangular. So, what we had seen last time was the rectangular window. We could have other shapes.

In fact, if you look at the triangular shape, it is very easy to see that the triangular shape can be obtained by conforming the rectangular shape with itself.

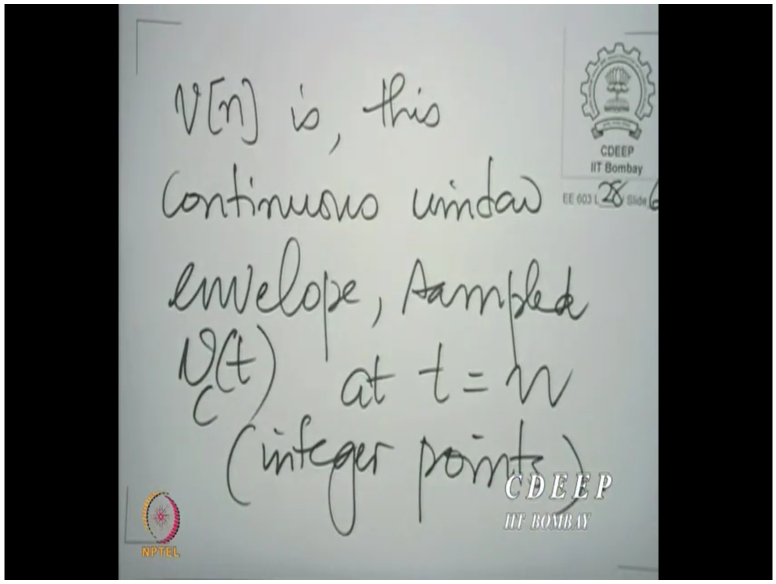
(Refer Slide Time: 06:53)



So, if you have a rectangular shape between minus N by 2 and plus by 2, now, since this is an envelope, since it is a continuous function, let us say of the variable t, let us call it $V_C(t)\ *\ V_C(t)$, would give us the triangular window. So, we can see there are relationships between the windows as well.

And that will also help us obtain the discrete time fourier transform in case we decided to obtain it for the triangular window appropriately sampled. Anyway, whatever it will be, we will agree that you know, all these windows, once your sampled that envelope at each of the integers, you get a sequence. That sequence is called v[n]. Let us write down what v[n] is.

(Refer Slide Time: 08:13)

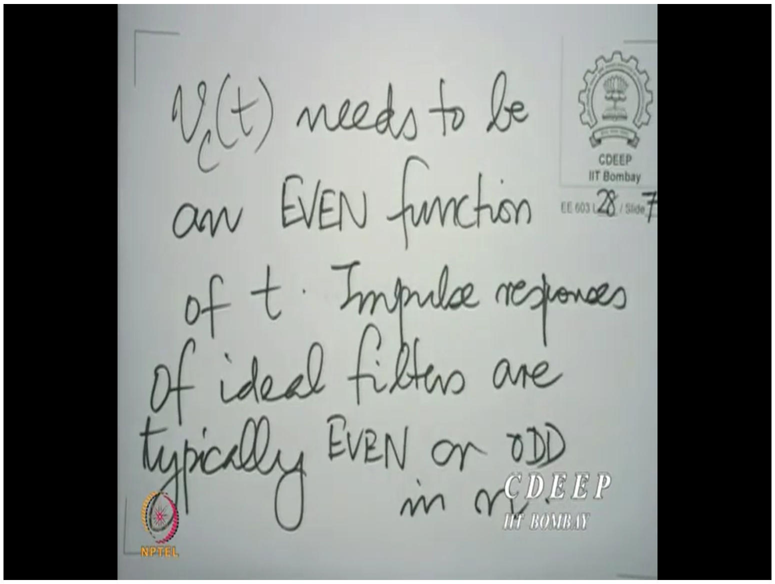


v[n] is this continuous window function, window envelope, if you like to call it that, sampled. So, of course, we have called the envelop $V_C(t)$ sampled at t = N, the integer is N. And of course, v[n] has a discrete time Fourier transform. Yes, please, there is a question. So, the question is if the window be other than rectangular, would it distort the original impulse response?

Well, indeed, it would change the impulse response samples at the points at which they are retained, but that change might, as we will see shortly, be for the better. So, it would of course change the samples, but our ultimate objective is not so much to retain the samples as to get a frequency response which is as close to the ideal as possible.

And therefore, what we need to see is what effect does the choice of window have on the frequency response, not so much the raw samples. Anyway, so you see one thing is very clear.

(Refer Slide Time: 09:40)



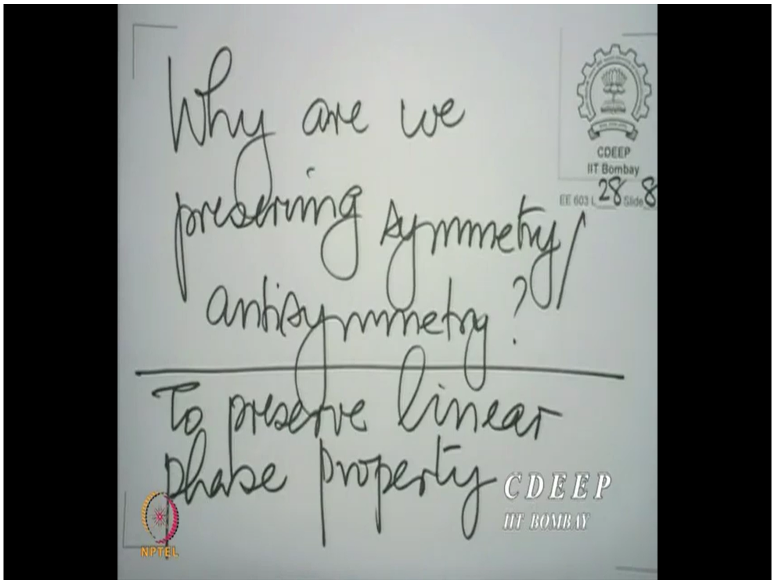
Vc(t) needs to be an even function of t. This is because typically impulse responses of ideal filters are typically even or odds in n. So, if we take all the four piecewise constant ideal filters, they are all even in n. We shall see some other non-piecewise, constant responses, which are odd in n, but either even or odd symmetry is, if the impulse response is even, we call it symmetric around n equal to 0. And if the impulse response is odd, we call it anti-symmetric.

So, typically we find that impulsive responses are symmetric or anti-symmetric about the center. And if we wish to preserve the symmetry, anti-symmetry, as it is, so symmetry should be preserved the symmetry and anti-symmetry as anti-symmetry. Then the window function needs to be even, that is not too difficult to see.

Because of the symmetric, of course, multiplying by Vn will preserve the symmetry, but if it is anti-symmetric, it will still preserve the anti-symmetry because corresponding points on the positive and negative side of n will be multiplied by the same number. So, anyway, this preservation, now why is this preservation of symmetry important?

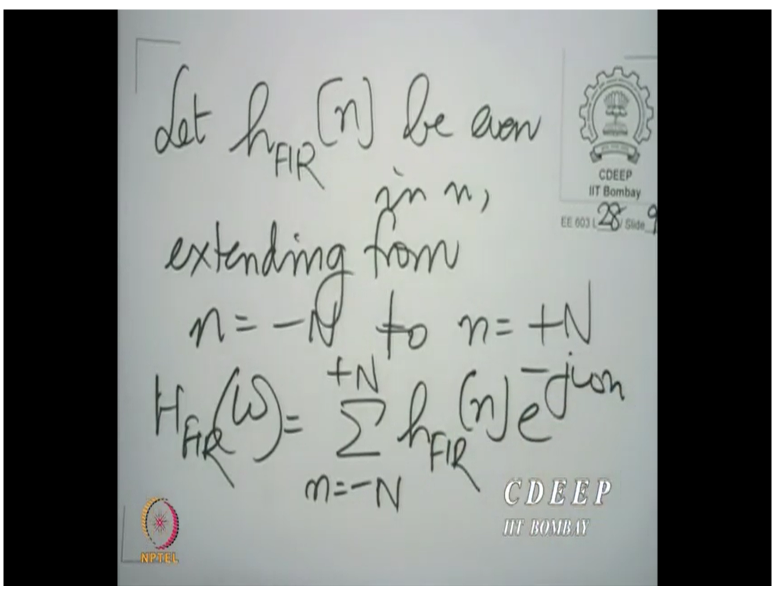
We shall see it has something very strongly to do with the phase response. In fact, we can see that right away, and this also tells us why we are in the first place looking at FIR filters.

(Refer Slide Time: 11:50)



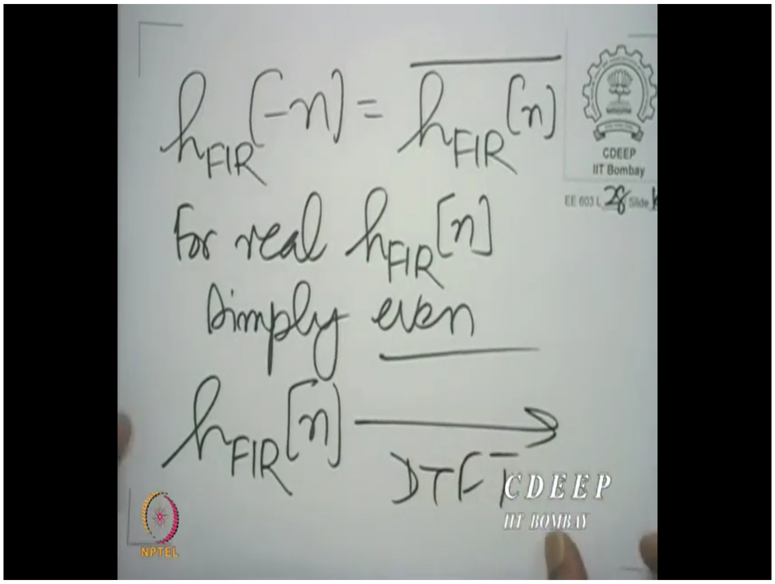
Why are we looking at, why are we preserving symmetry or anti-symmetry? And the answer is, to preserve what is called the linear phase or pseudo-linear phase property. Let us assume an FIR filter has a symmetric response, that means its response is even in n.

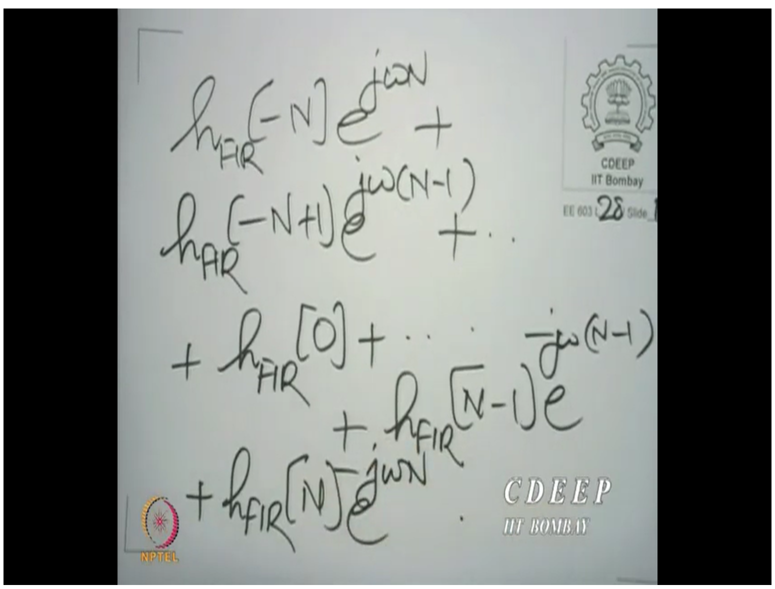
(Refer Slide Time: 12:32)



Let $h_{FIR}[n]$ be even in N extending from n equal to minus N to plus N. The frequency response $h_{FIR}[\omega]$ or $e^{j\omega}$, as you please to call it, would be $\sum_{n=-N}^{+N} h_{FIR}[n]e^{-j\omega n}$. Now, of course, we are assuming evenness with the underlying assumption of the impulse response being real as it is for all the ideal filters. But if it is not real, then we will assume that this conjugate symmetry.

(Refer Slide Time: 13:36)



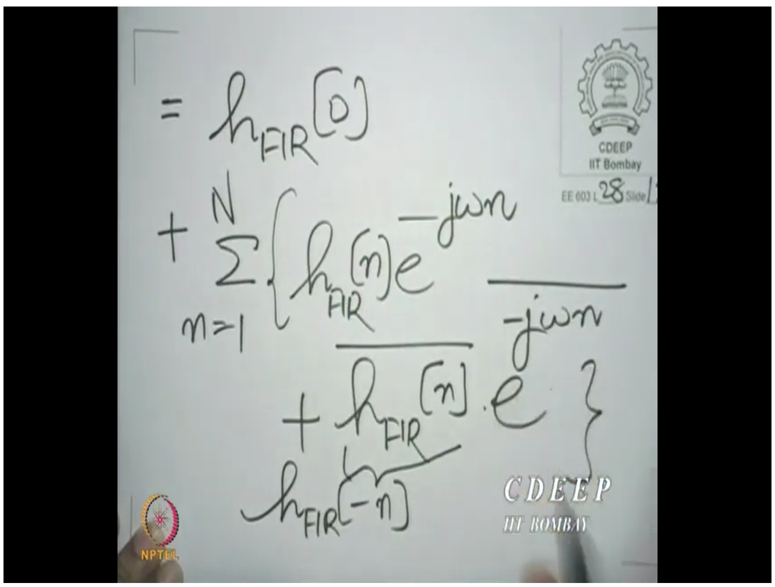


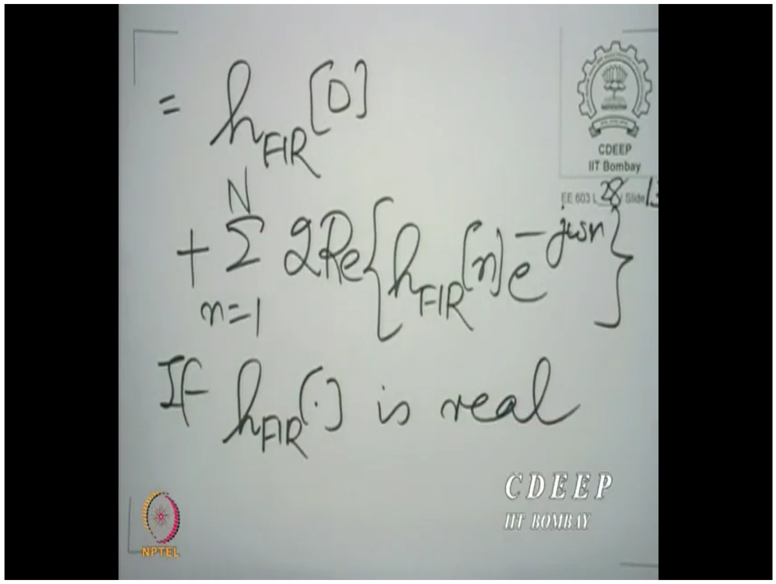
So, we will say, in general that $h_{FIR}[-\ n]$ is $\overline{h_{FIR}[n]}$. For real $h_{FIR}[n]$, it is simply even. Now, let us look at t frequency response from the point of view of its amplitude and phase part. So, $h_{FIR}[\omega]$, $h_{FIR}[n]$ has a DTFT given by h_{FIR} minus, let me expand it now, $e^{j\omega N}$ plus

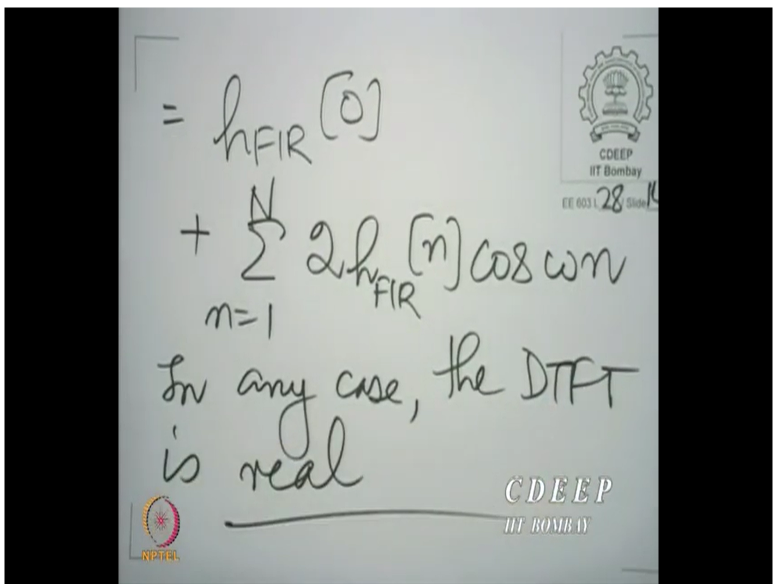
$h_{FIR}[-N+1]e^{j\omega(N-1)}$, and so on, $h_{FIR}[0]$ plus and then you have, $h_{FIR}[N-1]e^{-j\omega(N-1)}$ plus finally $h_{FIR}[N]e^{-j\omega N}$.

Now, expanded showing some typical terms. And what we do is to combine corresponding terms now on the negative and positive side. For the moment, let us assume that this is real to make matters simple. Or actually, we do not really need to, let us take in general that they are conjugate symmetric. So, in that case, each of the corresponding negative samples is going to be the complex conjugate of the positive sample.

(Refer Slide Time: 15:31)







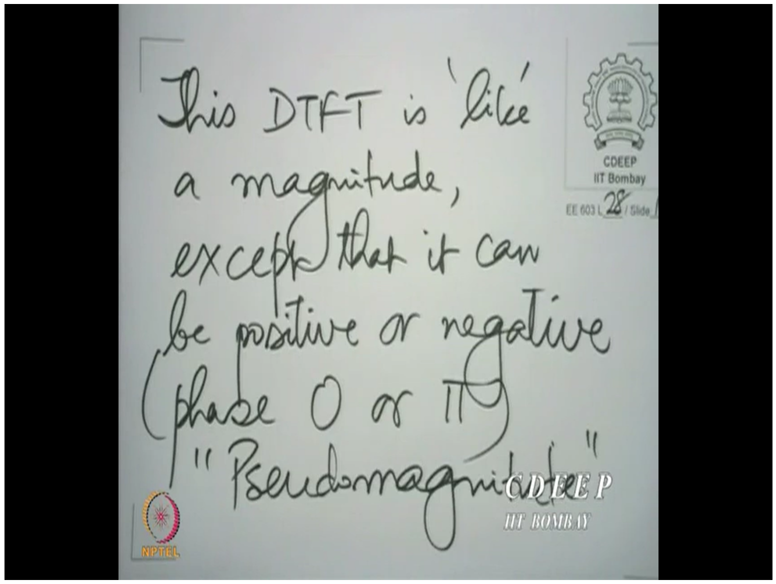
And therefore, we are going to have a combination of N terms of the following nature. This can be combined as $h_{FIR}[0] + \sum_{n=1}^{N} (h_{FIR}[n]e^{-j\omega n} + \overline{h_{FIR}[n]e^{-j\omega n}})$ because associated with minus, you see, $h_{FIR}[-n]$ is $\overline{h_{FIR}[n]}$.

And the corresponding exponential term associated with it is $e^{j\omega n}$, that can be written as $\overline{e^{-j\omega n}}$. We have capital N and such terms here. Now, clearly, this is the complex conjugate of this. So, this term in its entirety is the complex conjugate of this. And therefore, their sum is 2 times the real part of this.

That is easy to see, and that can be written as $h_{FIR}[0] + \sum_{n=1}^{N} 2Re\{h_{FIR}[n]e^{-j\omega n}\}$. And of course, if $h_{FIR}[n]$, then we have a very simple expression. We have this becomes $h_{FIR}[0] + \sum_{n=1}^{N} 2h_{FIR}[n]cos(\omega n)$.

what is noteworthy anyway, whether h FIR n is real or not, is that this impulse, this frequency response is real. You know, the either case, in any case, the DTFT is real. This is noteworthy. Now, when the DTFT is real, its phase can either be 0 or π. So, this quantity, this expression that we have here, or this DTFT itself, if you would like to call it that is also called the pseudo magnitude.

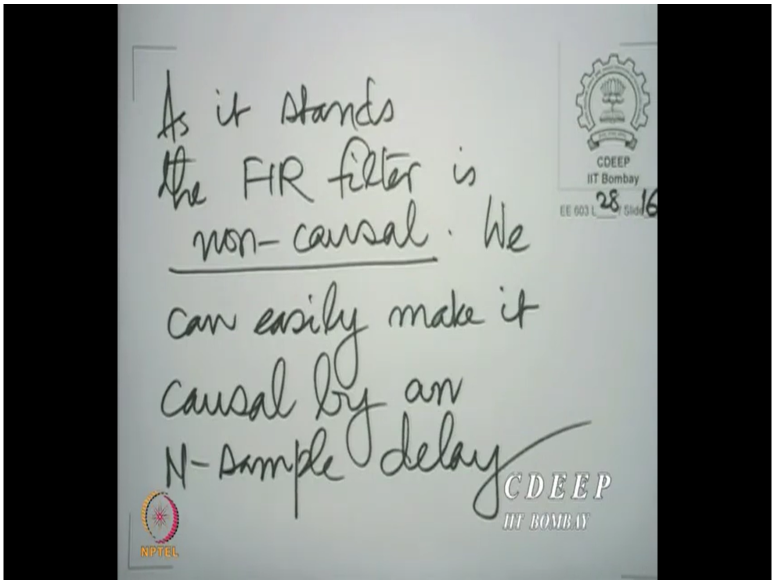
(Refer Slide Time: 18:38)



This DTFT is like a magnitude, except that it can be positive or negative. The phase is either 0 or π but nothing else. And therefore, it is called a pseudo magnitude. If it happens to be positive all over, then it becomes a magnitude. The magnitude of the DTFT is itself.

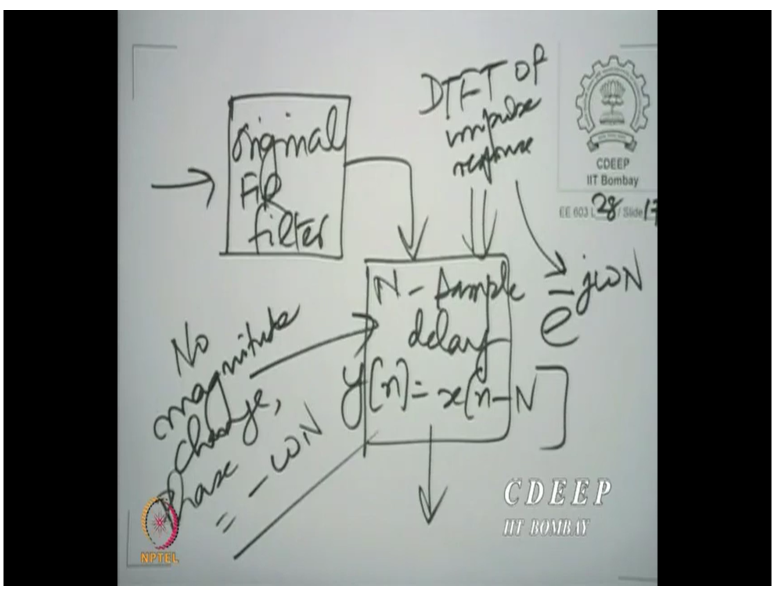
Now, it should be noted that when you thus truncate, so when you have taken the ideal impulse response and multiply it by a window function to retain samples between minus N and plus N, thus the corresponding FIR filter is non-causal. Then we have a problem or we think we have a problem there.

(Refer Slide Time: 19:58)



You know, if as it stands, the FIR filter is non-causal. But it is very easy to make it causal. All that you need to do is to delay the output by N samples. We can easily make it causal by an N sample delay. And N sample delay is very easy to understand in the frequency domain. All that N sample delay does is to impose an additional linear phase.

(Refer Slide Time: 21:02)



The other way of looking at it is, you are saying, you have the original FIR filter, you subject it to an N sample delay. That means $y[n] = x[n - N]$. And this introduces mode change in magnitude. This is a linear shift in variance system, very clearly and it is very easy to find its frequency response. The phase response simply becomes $-\ \omega n$. In fact, you can write down its DTFT.

The frequency response, DTFT of impulse response or frequency response is the $e^{-j\omega n}$. So, it has a magnitude of 1 and a phase of $-\ \omega n$ and therefore all that this does is to add a linear phase. Now, of course, this is not very difficult to understand intuitively. If your delay the output, obviously the nature of the output is unchanged, it is just going to come later.

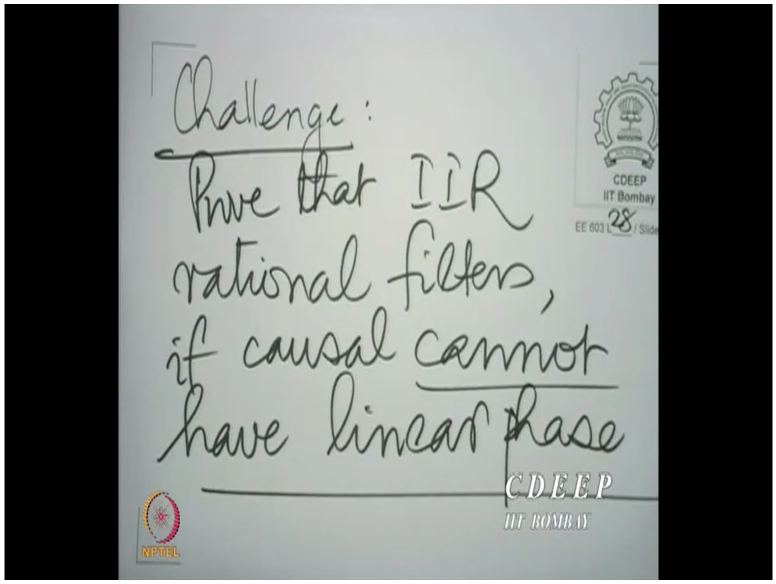
And that coming later is reflected in the phase. So, now you also see what happens when you have a phase response. The role played by a phase response is essentially shifting in time. If all components are shifted by the same amount in time, then the phase response is transparent. All that you will see is that your output comes later, but if some components come earlier than others, then you have a problem.

And that is what you mean by non-linear phase response. If the phase response is non- linear, then some frequency components come before the others. And then we have a problem then the fundamental nature of the signal changes. So, obviously the desirable situation is to have as linear as possible, a phased response. And in fact, even after thus making the FIR filter causal, the phase response of the FIR filter is almost exactly linear.

The only little bit of non-linearity is caused by the fact that the original DTFT or the original frequency response could be either positive or negative. So, it is a pseudo magnitude multiplies by a linear phase. And that pseudo magnitude is a magnitude, then you have exact linear phase. If the pseudo magnitude becomes negative at places, then you have a problem that the phase becomes 0 plus a linear phase at some places.

And pi plus a linear phase at other places, that is much better than having non-linear phase. This is clearly the reason why, FIR filters have a great advantage over IIR filters. Unfortunately, it is impossible to design infinite impulse response filters with linear phase. In fact, I put this as a challenge before you.

(Refer Slide Time: 24:25)



Prove that IIR rational filters if causal, cannot have linear phase. So, the precludes a good phase response and we saw that in the Butterworth and the Chebyshev filters. There was a trade-off. When you went from the Butterworth to Chebyshev, you would not do any worse in terms of resource as far as meeting magnitude specifications were concerned.

But when it comes to meeting phase response, which we have not specified at all, Chebyshev filter cannot do better than the Butterworth filter. So, although the Chebyshev filter does not do worse than the Butterworth filter or might even do better in terms of magnitude specifications. That means it conserves on resources to meet magnitude specifications. It unfortunately degrades the phase response. And in fact, as you go from Chebyshev to Elliptic, that is even more true, though we have not talked about elliptic filter design.

So, this whole problem manifests itself in IIR filters and in FIR filters, we obliterate it entirely, we do not have a problem of phase response at all. We can get exactly linear phase. And of course, I told you right at the beginning, when we started the engineering part of this course, nothing comes for free in engineering systems.

What we will do is, in terms of, I mean what you gain in terms of phase response here, you lose them in terms of resource. Ultimately, a lot of engineering systems come down to the simple maxim, invest more gain more. I can scarcely think they are very, I mean, it requires a lot of ingenuity, it requires a lot of deep understanding to violate this principle to some extent, of course, we do to an extent and places.

But by and large, it is always a game of investing more and gaining more or investing less and then also losing something. Of course, the game also pertains to what you can lose, but you know, sometimes you do not mind losing something and then you can gain something else. And that is also a gain then, what are you willing to lose?

Anyway, here, if you wish to gain in terms of phase response, you have to lose in terms of resource. So, what typically happens is that to meet the same magnitude specifications, you would have to keep a much larger length for the FIR filter. And in the design exercise that has been administered to all students in the schools you are of course, expected to observe that.

So, you would be designing an IIR filter with given specifications and you would also be designing an FIR filter with the same magnitude specifications. And you would compare the requirement of resource when you design it with the IIR approach and when you design it with the finite impulse response approach to meet the same magnitude specifications. But anyway, we now need to complete our discussion on how you can meet a set of magnitude specifications.