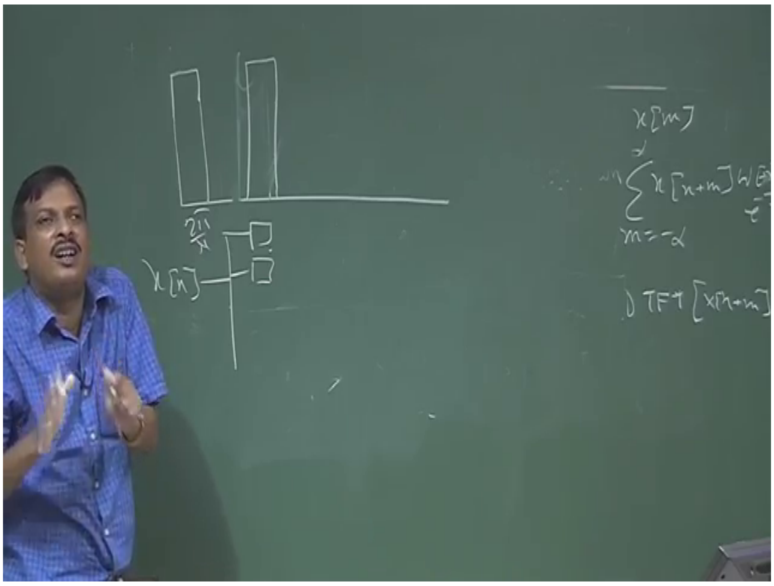
Digital Speech Processing
Prof. S. K. Das Mandal
Centre for Educational Technology
Indian Institute of Technology, Kharagpur

Lecture - 27
Short - Time Fourier Transform Analysis

So next class is that, we will let us go for the filtering view of STFT; so what I said that if I in DFT view that n represent the number of band we have analysis. So, can I not think that I have a signal?

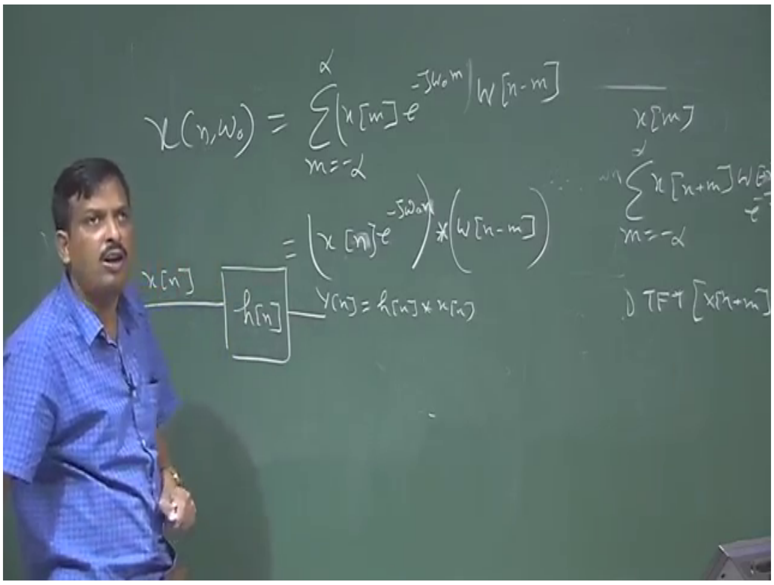
(Refer Slide Time: 00:37)



I want to analyze the signal for a particular band. So, if I say this is a band pass filter. So, I analyze the signal for these bands, then analyze the signal for these band. So, I can say this is nothing, but a if I have a signal pass the signal who were different band pass filter and each band width is nothing, but a 2 pi by N.

So, this is at a nut cell the filtering few of the STFT. So, in mathematically how do you will be represent? So, I can say let us I want to find out the output frequency response for a particular band or a particular signal particular frequency omega 0.

(Refer Slide Time: 01:27)



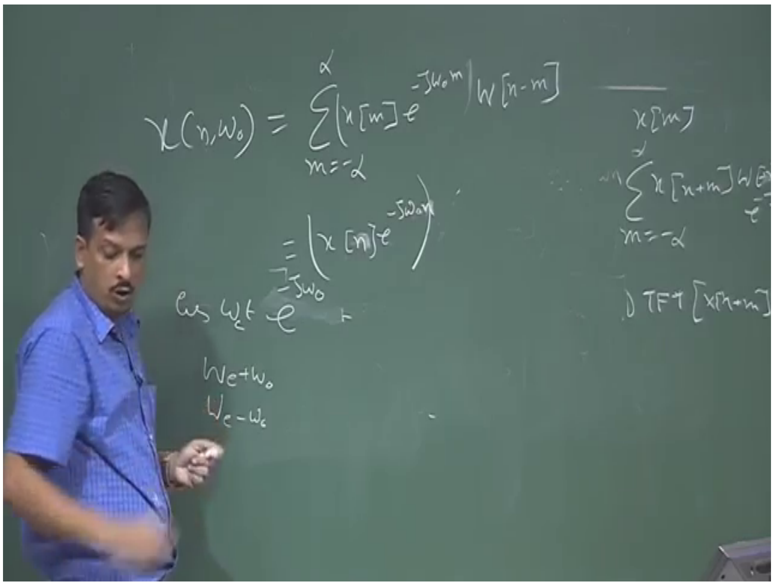
So, I have a signal x n, I want to find out whether this x n contained omega 0 or not whether this x n contained omega 0 or not forget about the windowing, windowing part I am coming later on. So, what I will do you know that to get that output, what I the I then a convolution in time domain I do a convolution for a particular I generate a particular frequency signal e to the power J omega 0 and convolved with x n that is my DFT equation ok.

Now, if it window then the same I have not taken the whole signal at a time. So, I cut the signal using a window. So, I can say x of n omega 0 for a particular frequency, it is nothing, but a m equal to minus infinity to infinity x of m, e to the power minus J omega 0 m, this signal multiply by the window this is my window or not. So, I can say I find out the convolution or the particular component with the signal and then the convolution output pass through a multiply with the window or I can say if this is this is my things if I remove this bracket then I can say it is nothing, but a convolution of x of m, e to the power minus J omega 0 m convolved with my window function this is or not o, if it is n th time instant.

So, I fix the time instant then I can say it is the n th portion of the signal, n th portion of the signal is find out the frequency component omega 0 with the n th portion of the signal and convolved with the window function. It is same when the time instant is fixed now if it is that; what is the meaning of that, what is the convolution. If I have a system h n it is the system response system impulse response, if I have a signal at x n if I pass the x n with h n what is the output? Y n is nothing, but a convolution of h n with x of n input signal.

So, if by x output frequency response is nothing, but a convolution of this term with window function. So, what is this term? This is nothing called a frequency shifting or modulation it is nothing, but a modulation if I multiply lets cos.

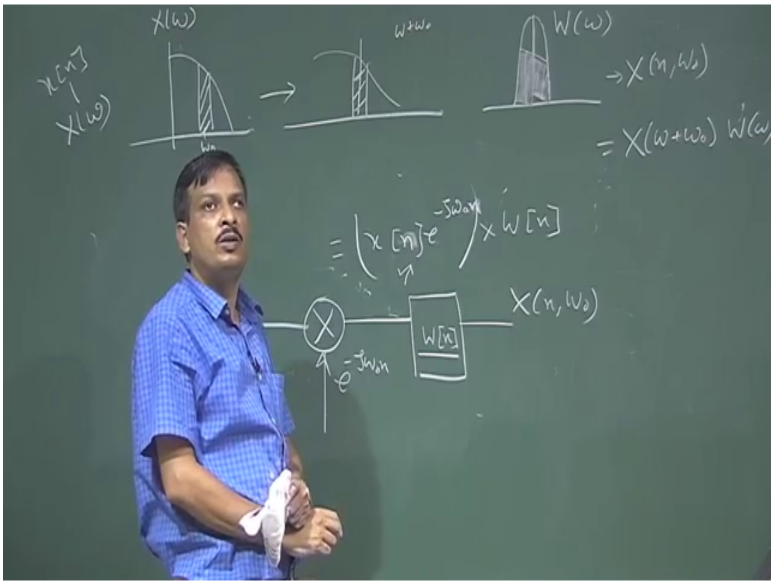
(Refer Slide Time: 04:55)



Omega t multiply by a another cos signal cos omega 0 t what is that? Lets it co carrier signal and multiply with a base band signal. So, output is cos omega c plus omega 0, if I cos omega 0 and with cos omega c minus omega 0. If it is was j instead of cos if I multiply with a shifting so; that means, omeomega c will be shifted to omega 0.

So, if I say that only then I can say this this convolution if I want to draw the sys the block diagram of this convolution function.

(Refer Slide Time: 05:41)



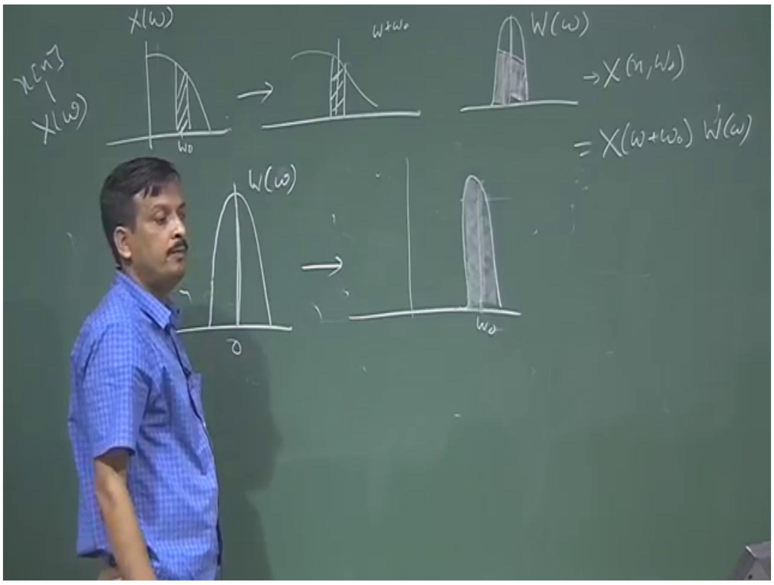
So, instead of n I should write n I should not write minus m because time instant I have taken n n is the fixed time instant. So, instead of minus m I write n. So, if I want to draw the signal diagram of system diagram of this thing. So, I can say this is my x n is coming, it has to be modulated by e to the power minus J omega 0 n at particular frequency, this is the modulation, modulation mean multiplication then it has to be pass through a system whose filter in impulse response is w n and I get the output which is x of fixed time for a particular frequency response ok.

So, if I draw the whole frequency response, suppose the base band signal I delete this one base band signal x m has a frequency response of this, this is the frequency response of x m. So, x m 2 or I can say x l. So, let us this is x n and has a frequency response of x omega. So, this is x capital x omega it has to be modulated by J omega 0 n. So, once I modulated then I can say let us this is my j omega 0, this frequency band instead of omega 0 single frequency lets omega 0 is a band. So, this is my omega 0 ok.

If this is my omega 0, then what will happen each has to be mod once I do the modulation. So, this origin this origin has to be shifted. So, this will come here I get this is nothing, but a omega plus omega 0, modulation is done shifted 2 omega 0 origin omega origin has to be shifted to omega 0 once this is called modulation. Once this modulation done lets I have a filter which impulse response is w n, but frequency response is like this, this is the frequency response of the filter. So, it is nothing, but a capital w omega ok.

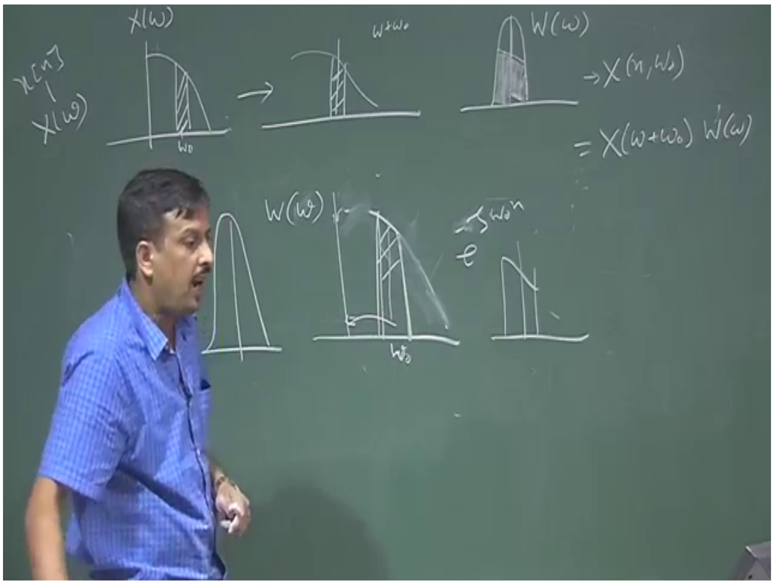
Once this is pass through this time domain convolution frequency domain multiplication. So, this will be multiply here. So, this portion is my output I get as a x of n omega 0 which is nothing, but a x of omega plus omega 0 multiply by w omega or not. Now I can shifted this notation instead of writing this way I can write this way also what is the way? X n omega 0 is nothing, but a e to the power minus J omega 0 n, x of n convolved with the. So, what I am doing instead of convolving the frequency response of the signal lets convolved the frequency response of the window.

(Refer Slide Time: 09:43)



So, instead of doing this what I will do I write this is my frequency response of window w omega oh sorry this is w omega. Let us modulate this thing to here. So, I modulate this thing 2 omega 0. So, instead of omega in 0 frequency I can say let us at omega 0 it is come omega 0 I shifted that things to omega 0 and I pass this thing to a filter which only filter out this portion or what I will do explaining the. So, what I want to do is that instead of writing w I can say I can get the w omega ok.

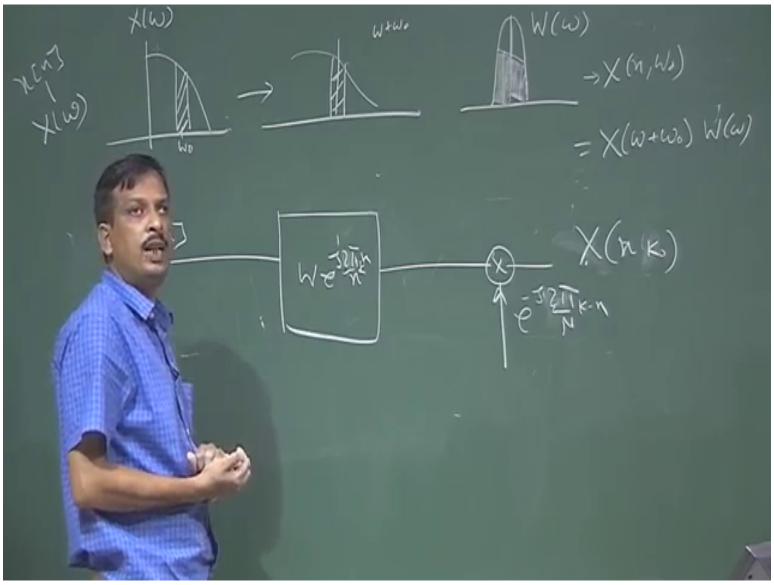
(Refer Slide Time: 10:47)



So, once I get the w omega that is the frequency response of the window, pass the base band signal with the frequency response of the window. So, if the base band signal is frequency response is this, this one then these window will be come here I pass the this signal is passes through this. So, this portion will be there (Refer Time: 11:19) portion will be not there. So, this portion will be not there only this portion will be there which is omega 0 now if it is omega 0 I have to shifted this 2 here. So, what I will what I will do? I will multiply this with the e to the power j minus j omega 0 n. So, what I will do I will come this will come to the here.

So, at output filter this will be shifted or demodulated to come to origin. So, this can be one mean then the diagram will change not should be the diagram the first pass the signal this.

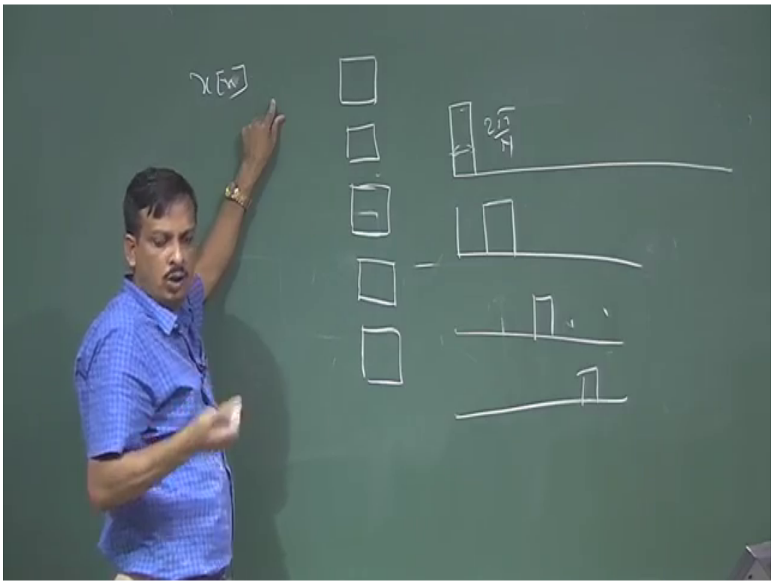
(Refer Slide Time: 12:07)



 is my x n pass the signal with a window frequency response w e to the power j omega 0 n, then demodulated shift the frequency with e to the power minus j omega 0, that also possible then also I get x, x n omega 0. Now if I generalize this equation instead of omega 0 I can write instead of omega 0, I can write k once I write k then this omega 0 is become 2 pi by n k, this omega 0 is become 2 pi by n k into n.

So I can write analysis equation in filtering view that the same that I have a signal.

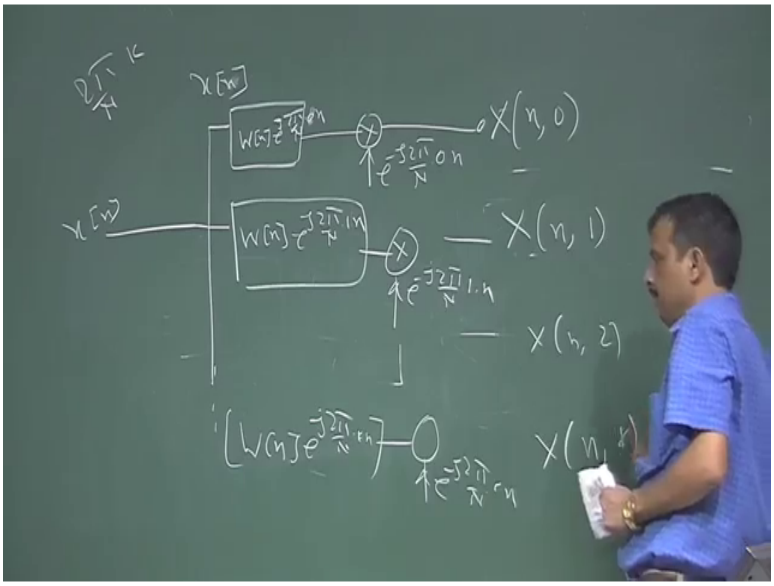
(Refer Slide Time: 13:23)



X n lets x n x n is my signal I have frame the signal first x n is a fixed length signal I am thinking this signal consist of several frequency component and I design a several band pass filter which is a fixed frequency band and each band pass filter is shifted to a particular frequency. So, what I will doing? If this is my frequency scale I design a band pass filter one this one, second one will be here third one will be here fourth one will be here how do I do that each bandwidth is 2 pi by N.

So, this is nothing, but a shifted to shifted frequency. So, I can say let us I write down the synthesis equation, synthesis diagram lets this is my x n.

(Refer Slide Time: 14:22)

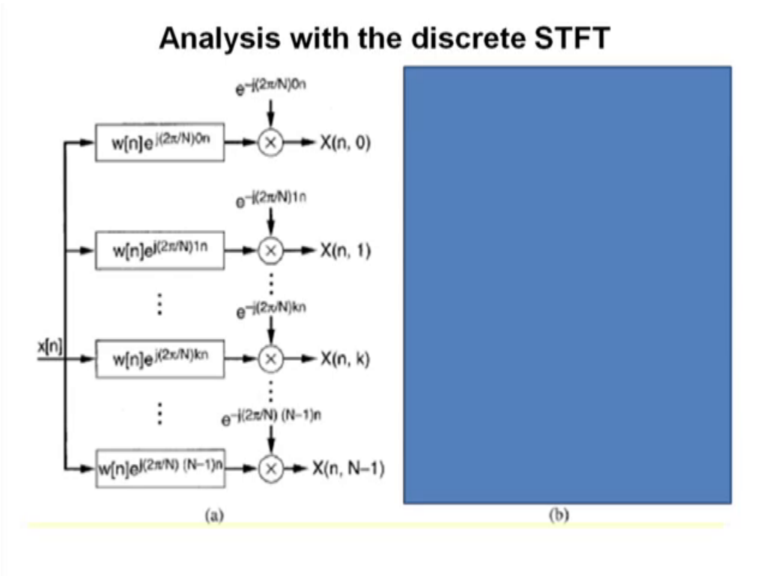


So this is my x n. So, this is 0 th component when k equal to 0. So, 2 pi by n into k. So, k equal to 0 I can write what is this omega 0 the sorry w o n into e to the power minus e to the power J 2 pi by n into k, k equal to 0 into n when pass demodulate it is with e to the power minus j 2 pi by n k equal to 0 into n then I can write again w n, e to the power j 2 pi by n into 1 n again demodulate it e to the power minus j 2 pi by n k equal to 1 n. So, dot dot dot dot I can write w n e to the power j 2 pi by n into k n, multiply by e to the power minus j 2 pi by n into k.

So, each of the output this is nothing, but a x if it is n n th instant with 0 frequency, this is output of x n th instant with plus k equal to 1, then x n 2 then it is nothing, but a x of n. So, this called synthesis path. So, analysis with discrete f to f. So, I can say this portion sorry this portion is called analysis of STFT with filtering view. So, I can say I apply the x n signal, I can find out the 0 th component first component second component third component and k th component this way. So, this is x 0 1this is forget about this portion do not see this portion.

Let us try to hide this portion I just hide this portion, lets design insert say you hide this portion this portion I will discuss later.

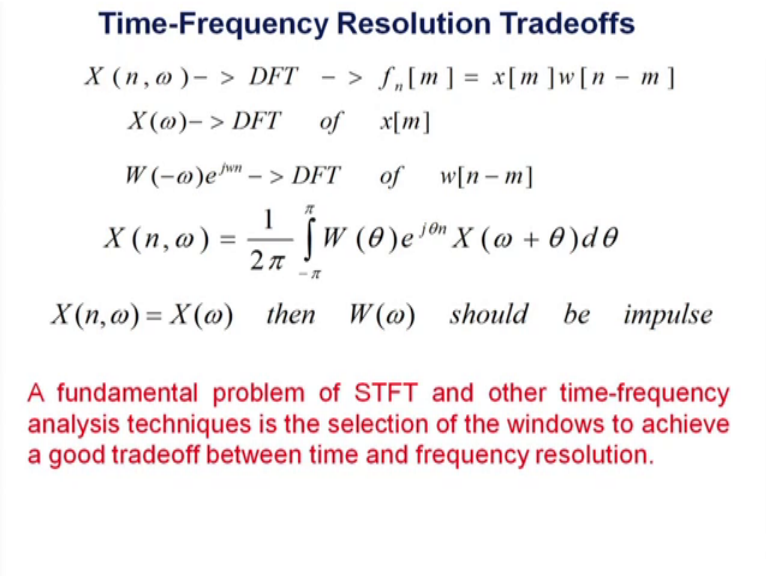
(Refer Slide Time: 16:56)



So, now, this portion is the analysis portion. Now in synthesis portion what is my requirement from their I have to apply some process by which I can some up them and apply the inverse transform, I should get back the x n there is a synthesis portion. So, if I remove it if you see when I remove it that x n, those inputs are going there and some kind of re inverse transform is taken care and that inverse transform is written here I will discuss the details in our transform and then I get the x n that if omega 0, w 0 is not zero then we exactly recover the signal that we will discuss which is call synthesis.

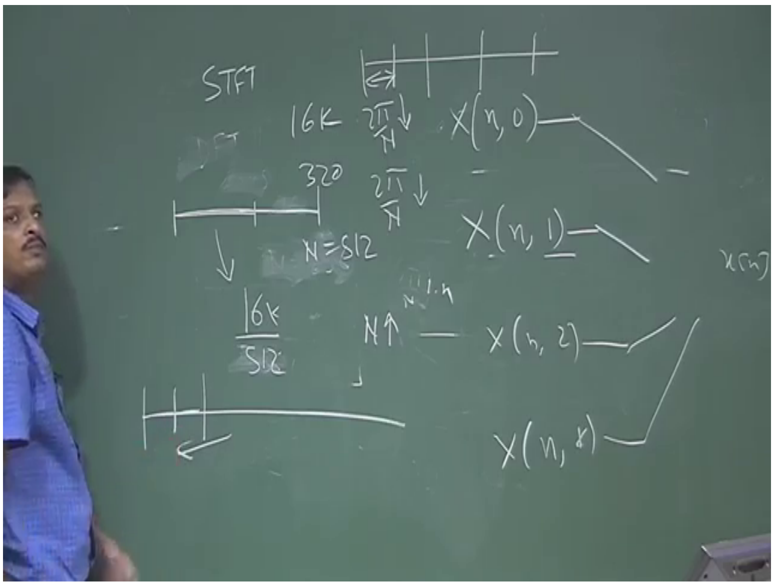
So, l STFT analysis this is the STFT analysis part synthesis part, I will discuss in the next class . So, there is a some property if you find the any book there is a some property of the STFT analysis.

(Refer Slide Time: 18:09)



I have not reading the slides you can read the slides whether. Now come to the another things which is called time frequency trade off do not think about the mathematics there is a mathematical part is there in the slides I am not I have just discuss I have describe what is time frequency trade off.

(Refer Slide Time: 18:35)



if you see in any DFT or STFT analysis or STFT is TFT, there is a 2 words one is call time I have taken a small segment of the piece signal and I have apply a frequency transform number of band that is k.

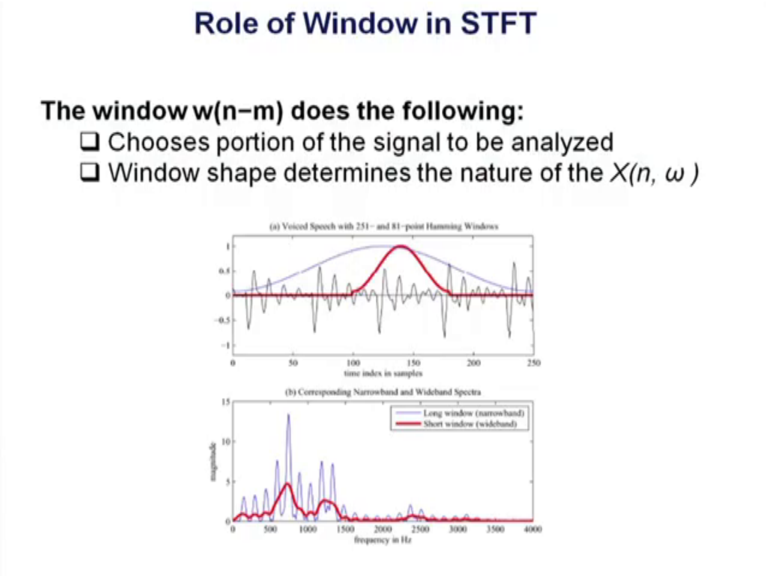
So, let us I take a signal of 10 millisecond which content 160 sample . So, numb n length of the DFT is the nearest 2 to the power something unless I have to put more 0 padding and the amplitude will be vanish. So, what I will do nearest 2 the power something. So, if it nearest then I can say it is 256 lets 2 to the power 256 n equal to 256. So, I take the n equal to 256. Now what is the resolution? If the signal is 16 kilohertz sample n equal to 256 then I can 16 k divided by 256 is my resolution.

Now, if I want to increase the frequency resolution. So, instead of 10 millisecond if I take 20 millisecond then it is a 320 sample, then I can say n is equal to 512; if it is 512 this is the resolution. So, now, resolution come 512. So, resolution is increases what is the resolution? Resolution means this is the frequency scale if I divide this long window. So, that resolution is less if I want to increase the resolution. So, division has to be very small size. So, 2 pi by n value if it is decrease then resolution is increase.

So, if this will be decrease 2 pi by m will be decrease if n is increase. So, once I increase the n length of the DFT, then I require a large of signal at a time. So, if I take a large number of signal this is the time axis. So, instead of short segment if I take a long segment then what I loosing? I loosing the time resolution. So, if I increase the frequency resolution I decrease the time resolution. So, that is call time frequency trade up. If I take a larger amount of signal at a time my frequency resolution will be very high, but I loose the within that time the variation of the signal I loose. So, the time resolution will be decreases.

So, if I increase the frequency resolution time resolution decreases, if I increase the time resolution frequency resolution will be decreases. So, that is the time frequency time up in STFT analysis.

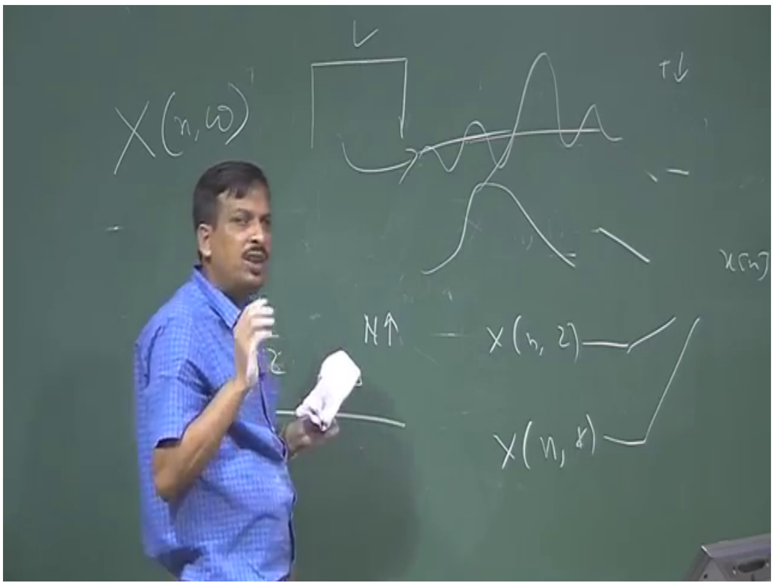
(Refer Slide Time: 21:36)



Then there is a another topic is call effect of window. Different window has an different effect what I said what about the frequency response I get which is x n k lets if I this is my output of frequency analysis n omega this is nothing, but a convolution. So, what I am doing, a signal is multiply by a window. So, time domain multiplication frequency domain convolution; that means, this is nothing, but a convolution of frequency response of the window and the signal frequency response.

So, what is meaning suppose if I have a window of rectangular window.

(Refer Slide Time: 22:27)



You know the frequency response of the what is the frequency transform the rectangular window I nothing, but a sink function. So, there is a peak main lope and there is a side lope. So, that will be and the main lope with depends on the length of the window. So, that will be convolved with the frequency response of the original signal. So, the effect of the window will be always there, once I multiply by the window. So, the frequency response is not the exact frequency response is signal. So, it is a frequency response of the main lope multiply by the side lope.

So, what I want I want to increase the main lope should be flat ideally if the main lope is flat then the side lope effect of side lope it is negligible, then I can say the frequency response whatever I get it is exactly the frequency response of the window. So, if you see there is a different kinds of window is there.

(Refer Slide Time: 23:46)

Window Function for FIR Filter Design

Name of Window	Window function
Bartlett(triangular)	$1 - \frac{2\left\|n - \frac{M-1}{2}\right\|}{M-1}$
Blackman	$0.42 - 0.5\cos\frac{2\pi n}{M-1} + 0.08\cos\frac{4\pi n}{M-1}$
Hamming	$0.54 - 0.46\cos\frac{2\pi n}{M-1}$
Hanning	$\frac{1}{2}\left(1 - \cos\frac{2\pi n}{M-1}\right)$
Kaiser	$\frac{I_0\left[\alpha\sqrt{(M-1)^2 - \left(n - \frac{M-1}{2}\right)^2}\right]}{I_0\left[\alpha\left(\frac{M-1}{2}\right)\right]}$

If I if you remember that I have given in the review of DSP that what are with the Blackman window, Hamming window, Hanning window Kaiser Window. And different window has different kinds of frequency response. So, my intention is that main lope with has to be increased and side lope attenuation must be very high side lope should be very slow disturbing then I can say the analysis of frequency analysis of the signal is almost (Refer Time: 24:15).

So, that is my effect. So, I want to choose such a window whose main lope with is very high. If you see in the spectrogram when I discussing about the spectrogram analysis in this software there is a if I just open a sample English file, English voice file and if you see that is setting spectral analysis setting if you see there is a window, hamming window triangular window. So, depending on the window function that frequency response will be different, because the effect of window frequency response will be transfer 2 there. So, if it is hamming window something else, if it is Hanning window if it is hamming then it is different.

So, depend on the window the frequency response is different even if I do the spectral analysis analyze spectral by scan see that there is a different window selection, Hamming window, Hanning window, then Blackman window, then Kaiser window or Gaussian and this is the FFT size. So, this is the choice of n if I this choice this is the frequency resolution is very low, if I go their frequency resolution is very high. So, depending on the choice is there window size and window and windows.

So, window size I window size will come from the length of the your FFT if n is equal to if the window length is 20 millisecond; that means, 320 sample is there then my n is fixed to almost goes to 320s 256 you know 512, because it is implemented in FFT. So, length of the window has tied with the frequency time resolution, and type of the window I will choose what kind of details analysis I one maximum cases in speech and Hamming and Hanning window is used.

Next class I will do discuss about the STFT synthesis, ok.

Thank you.