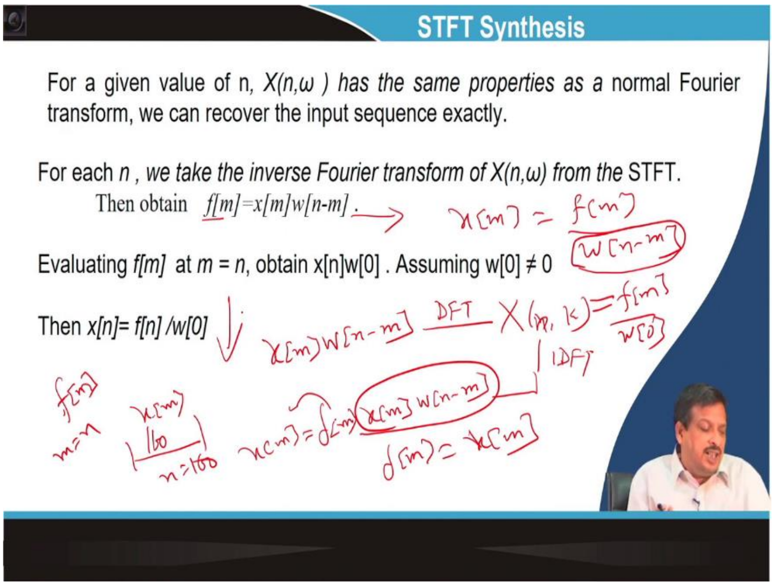
Signal Processing Techniques and Its Applications
Dr. Shyamal Kumar Das Mandal
Advanced Technology Development Centre
Indian Institute of Technology, Kharagpur

Lecture - 29
STFT Synthesis

So, now, in the last class, we talk about STFT Analysis. Now, we talk about STFT Synthesis.

(Refer Slide Time: 00:32)

(Refer Slide Time: 00:33)

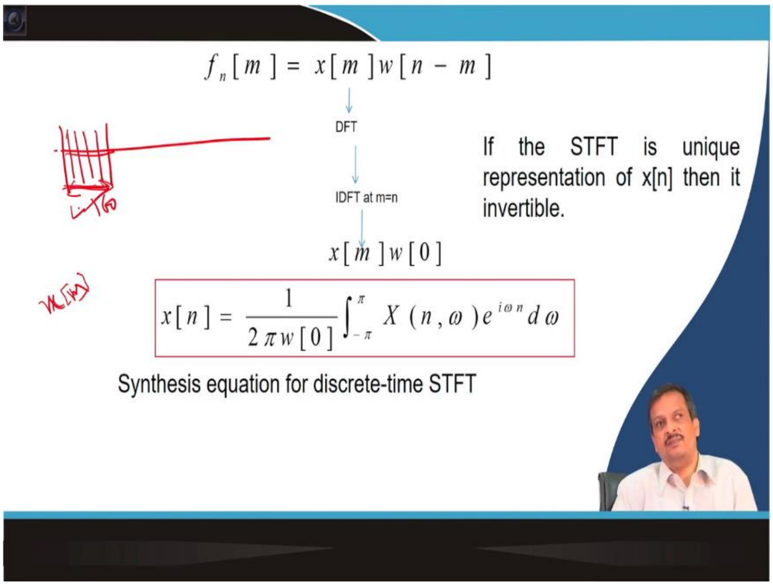


So, what I have done is I have a signal x[m], and I have multiplied by the w[n-m] ok. Then I did that DFT discrete Fourier transform, and I got X(n, k). What do I have to do if I want to get the signal back? I have to take the IDFT of this one. So, that is what I will get back: I will get back x[m] multiplied by w[n-m].

But, my intention is to get back x[m], not that x[m] multiplied by w[n-m]. So, if this is my f[m], this is my f[m]. So, f[m] when it will be equal to x[m]. So, I can say the f[m], if I want to get back the x[m] is here, then x[m] is equal to f[m] divided by w[n-m]. So, if I say that w, how do I eliminate w[n-m]? If I say that ok, f[m] is f[m] is great because every m equals n. What is the n? n is the index of the window.

So, if it is a 160-sample window, n is equal to 160. So, if I evaluate the window for every m equal to m, that means, if my window is shifted by one sample, then what we will get at m equal to m is nothing but a f[m] divided by w[0]. So, if the w[0] is not equal to 0, then I can say that f[m] equals x[m] for every m equal to m. So, at m equal to n, this will be w[0], which is ok. Now, if the w[0] is not 1 and w[0] is a constant, then I can get the x m back with no problem. So, that is my limitation in STFT synthesis. So, what do I have to do?

(Refer Slide Time: 03:17)

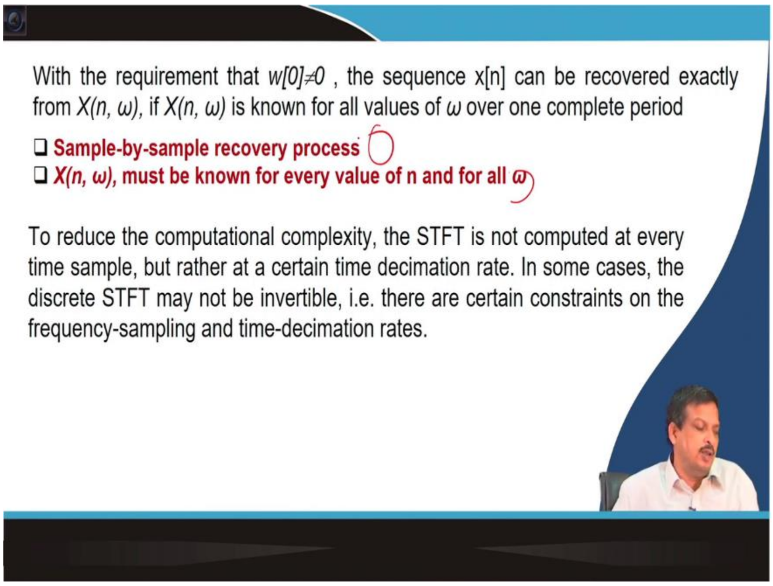


When I suppose I have a signal, I have to analyze it. Let us say 160 samples have been analyzed, but I have to know it for every sample. So, I can only shift the window by one sample only. So, for every sample, I have to compute inverse DFT, and only then do I get back the signal.

So, that is called sample-by-sample recovery. But, it is very difficult; suppose I have an 8-kilo hertz signal; I take the window size of 160, and then for 160 times, I have to compute the IDFT, which is very time-consuming and is not usable. So, what should I do?

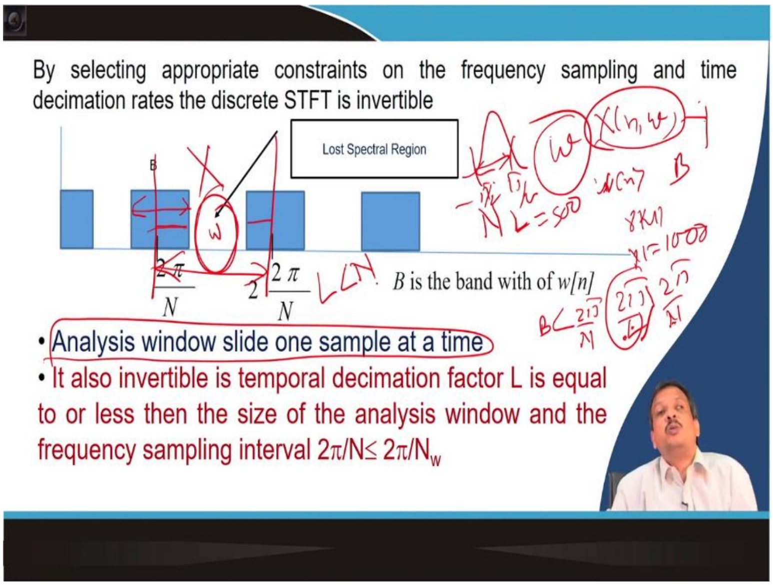
So, what should be the shift of the window? So that, although it has been shifted, I can also get back my x[m]. So, I have to find out the algorithm STFT synthesis what the allowable shift is possible. So that I can get the signal back. So, that is called STFT synthesis constant ok.

(Refer Slide Time: 04:51)



So, let us say what I have said is that I have to know all ωs sample by sample, which is not possible. So, I have to find out how much is shifting. Though I have shifted, I also got the signal back.

(Refer Slide Time: 05:08)



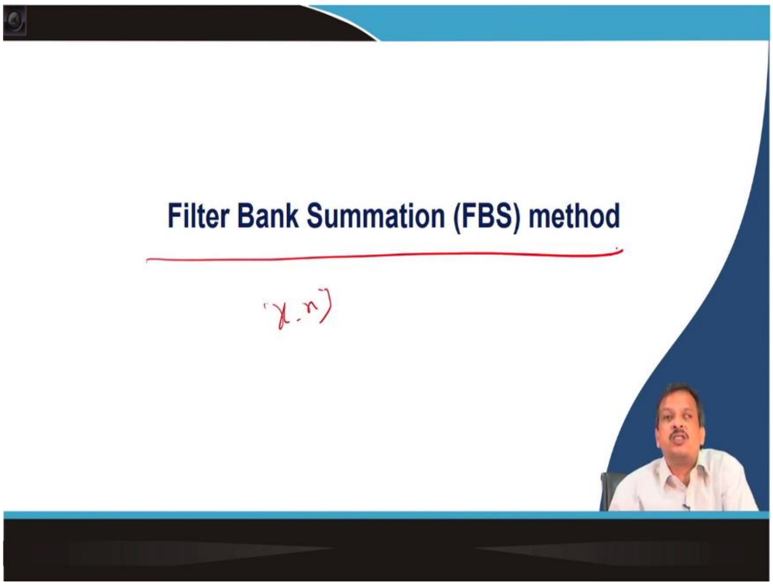
What is that, and what is that limitation? So, this is what I have already explained, one sample by one sample, one sample by sample, but the other also. So, I have to know all ω X[n] ω has to know for all ω. Now, suppose I have chosen a window w[n] whose bandwidth is B. So, this is the bandwidth of the window; when I put the window there, and I have chosen the DF length of the DFT, it is N.

So, the separation between this one and this one is 2π/N. Now, the bandwidth is less than; so, this is the bandwidth, and this is the B by 2 B by 2. So, it is nothing but a B. So, if B is less than 2π/N, this thing will happen. So, this portion of the ω I do not know after the filtering because the signal is limited by the filter. Is it clear? For example, suppose I said I have a signal 8-kilo hertz sampling frequency and N equal to 1000.

Now, say I have chosen an L window length equal to 500 I a window length equal to L 500. Then what is the main lobe; this is nothing but a, this is π by L, this is minus π by L. So, it is 2π/L is the bandwidth; here it is 2π/N. Now I said, so, what I am doing, I am passing the signal 2π/N, and I restricted the bandwidth of 2π/L. So, 2π/L is larger, L is smaller than N. So, this will be larger than this one. So, there will be an overlap.

On the other hand, I can say I have found out that some of the portions of the signal are not. Some of the portion of the frequency is not represented at the output of the X at the X(n, ω). So, I have not known all X[n] for all ω. Now, if this is happen if the L is less than capital N. What is happening in time domain? I choose a signal 1000, but I only multiply by 1 as 500 and for the rest of the 500, I make 0. So, when I choose the next window, I am not choosing the signal. So, what should I do?

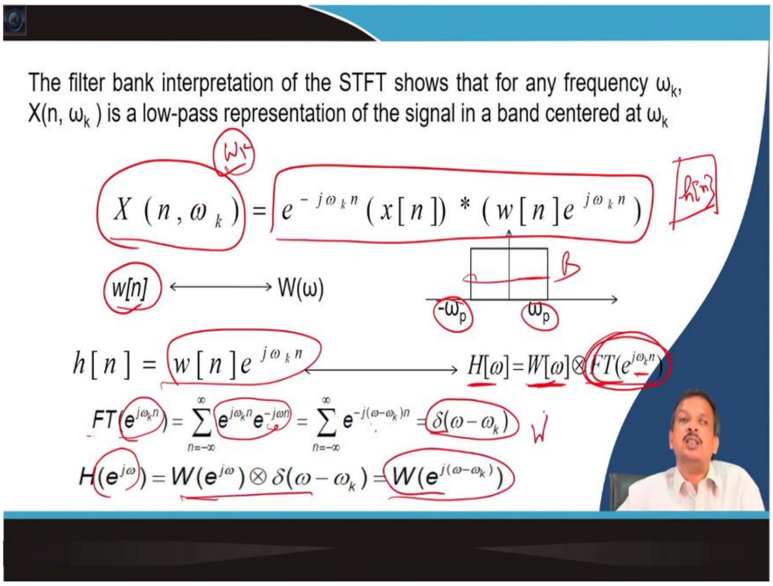
(Refer Slide Time: 08:31)



So, there are two methods for STFT synthesis; one is called the filter bank summation method, and one is called the overlap-add method. Since I will not discuss details here of filter bank summation methods that are already available in another YouTube video when I talk about speech signal processing, you can search on YouTube and find out where you can get the details of filter bank summation methods.

So, I just touched it here because this is a signal processing course. So, I am not emphasising a particular application, filter bank summation method or OLA method. You can read it if you are interested, and then you can go to the YouTube video and read it again. So, filter banks summation methods, such as the DFT or STFT, can be viewed as a summation of a bandpass filter. So, I use that principle to get back the signal x[n], and then I call it the filter bank summation method.

(Refer Slide Time: 09:39)



So, what is mathematics, or what is the analogy behind it? So, X(n, ωk); ωk is a particular frequency when I take the ω0, ω1, ω2, ωk. So, ωk is a particular frequency, and this is my STFT analysis ok, and I said the frequency response of ω n window function is this one.

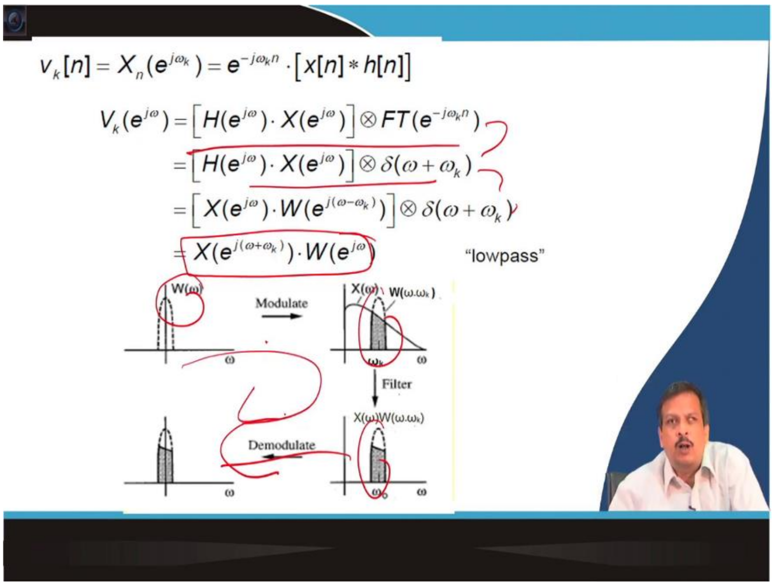
So, it is nothing but a low pass filter whose bandwidth is this is ωp, this is minus ωp, this is the bandwidth B. Then the cut-off frequency is p, upper cut-off frequency and lower cut-off frequency and bandwidth is B ok or not. So, now, if I say what is, let us say a w[n] as an h as a system. The window is nothing but a system which is called a filter, whose frequency impulse response is h[n].

So, I can say H[ω] is nothing but a W[ω], and Fourier transform of $e^{j\omega n}$. This is my response to h[n]. So, what is the Fourier transform of $e^{j\omega kn}$ for a particular frequency? So, it will be again multiplied by $e^{-j\omega n}$. So, I get ω minus ωk. So, it is nothing but a δ function ω minus ωk.

Now,

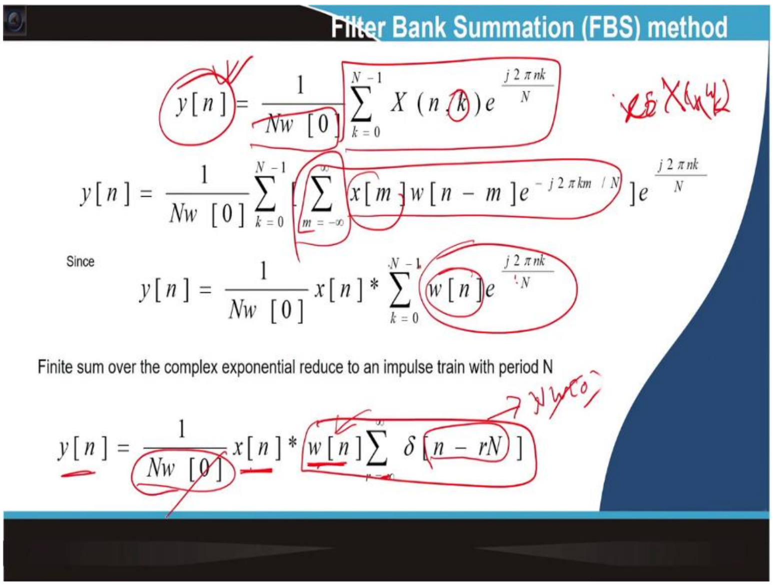
$$H(e^{j\omega}) = W(e^{j\omega}) \otimes \delta(\omega - \omega_k) = W(e^{j(\omega-\omega_k)})$$

(Refer Slide Time: 11:45)



Let us say Vk[n] is my recovery signal. So, V k $e^{j\omega}$ is nothing but this one, and then I just follow that step, and I get this one. So, this is W[ω] modulated to here, then filter, then demodulated, and I get back this one low pass filter. So, this is the process I also mentioned during the filtering view of STFT.

(Refer Slide Time: 12:28)

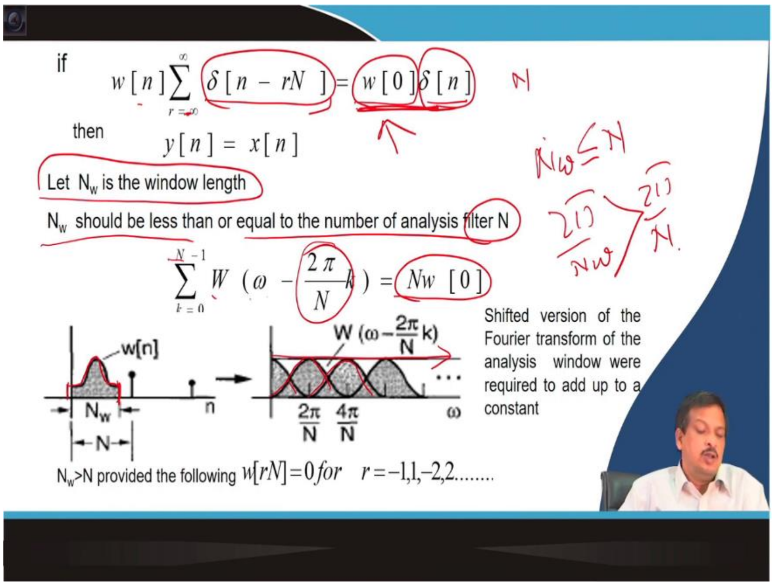


Now, I have come to the recovery process here. So, what I want is y[n]; I want to recover y[n] from x[n] from X[n, ω] or ωk for a particular frequency. So, if it is discrete DFT, then it is discrete frequency k. So, y said 1 by N into ω 0. So, in that case, it will be 1 by ω 0. This one is the recovery of y[n].

Inverse Fourier transform of X(n, k). Now, I said X[n]ow what is X(n, k) I put the value ok; k equal to 0 to N minus 1. So, this I put. So, what I get is that I already know it is, so, if you see, this is nothing but a convolution between the signal and this portion, m equal to minus infinity to infinity is a convolution function. So, I can say y[n] is a 1 by Nω[0] and convolution of x[n] and w[n].

Because if you see the k equal to 0 to N minus 1, this does not depend on k. So, I take it outside, and this is nothing but a δ[n - rN]; r equal to minus infinity to infinity. Now, if it is that, then I can say y[n] is equal to x[n] when this portion becomes N into w[0]; because N w[0], N w[0] will be cancelled, and y[n] will be equal to x[n].

(Refer Slide Time: 14:38)



So, I can say

$$\sum_{k=-\infty}^{\infty} w[n]\delta[n-rN]$$

Then only I get y[n] is equal to x[n]. So, this is the requirement. So, how do I make this portion is w[0] into δ[n]. Let N w be the window length. So, if N w is the window length, then this is the frequency response of the window ok.

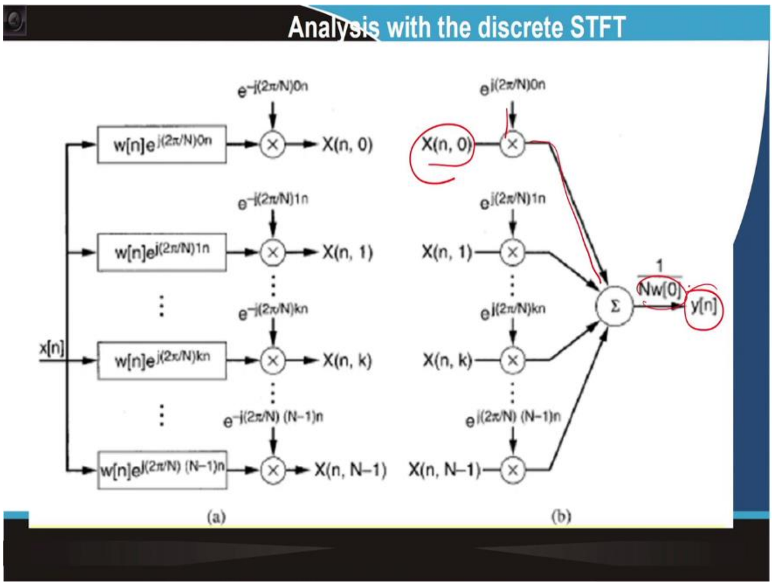
And N w should be less than or equal to the number of analysis filters; then I can say

$$\sum_{k=0}^{N_w-1} W\left(\omega - \frac{2\pi k}{N}\right) = N_w W[0]$$

When it is possible, I can say there ω if there is a low frequency that is left over. So, N w should be less than or equal to the number of analysis windows. So, I can say that I can say what the structure of the W ω. So, all the filters will be overlapped in such a way that when I add them up, they become one.

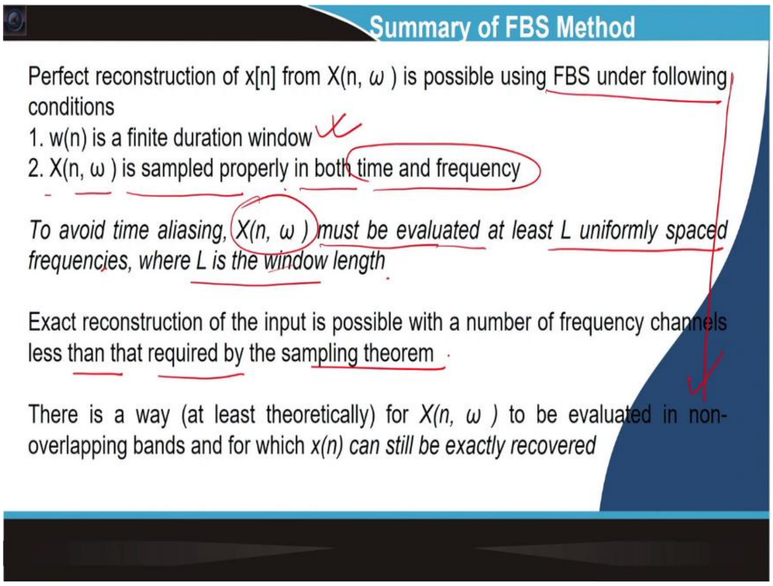
So, that is the filter band summation method. So, N w should be less than or equal to a number of N w should be less than equal to N because what is required 2π/N w must be greater than 2π/N then only can I know all the frequencies.

(Refer Slide Time: 17:01)



So, that is the filter bank summation method. So, what we will do is get an X[n] 0, then we demodulate and then again I divide by Nω[0], and I get y[n]. Ok, these are syntheses.

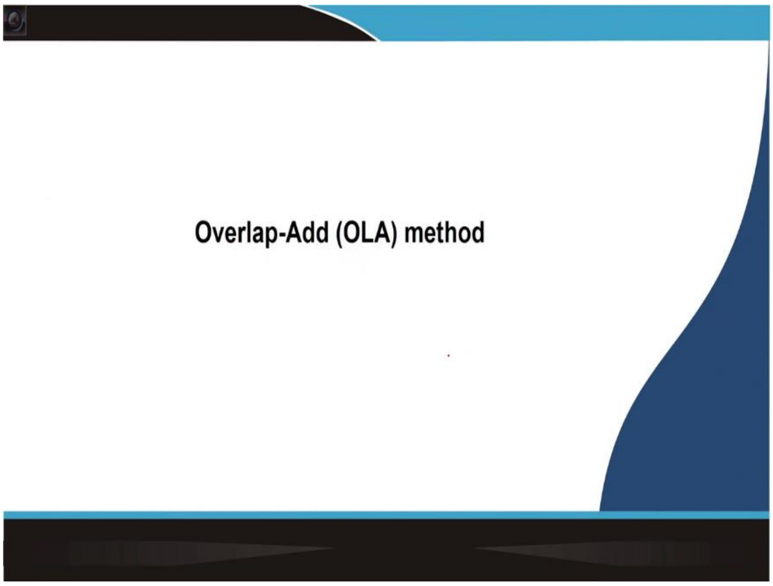
(Refer Slide Time: 17:15)



So, this is a summary of FBS methods. You can read the perfect reconstruction of x[n] from X[n] ω is possible using the FBS method. So, ω n is a finite duration window, and X(n,ω) is sampled properly in both time and frequency. To avoid time aliasing, in the n w X[n] w[n] ω must be evaluated at least L uniformly spaced frequency, where L is the length of the window.

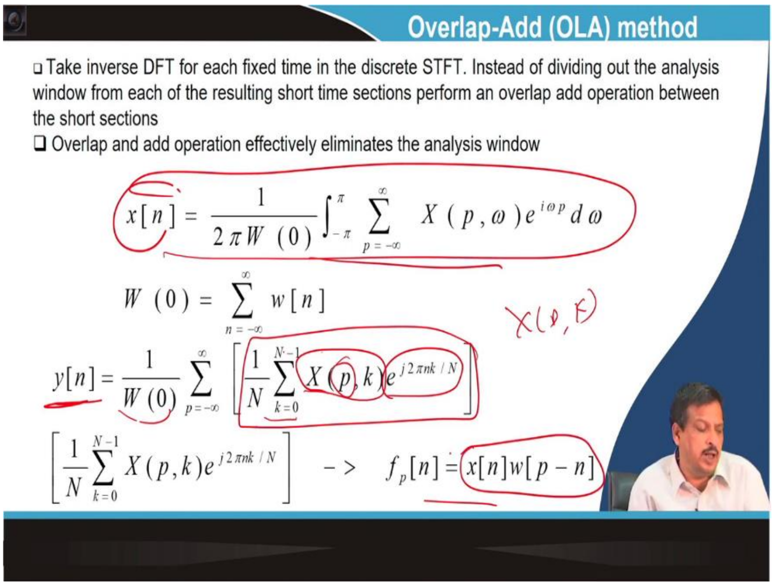
Exact reconstruction of the input is possible with the number of frequency channels less than that required by the sampling theorem. So, that is the summary of FBS methods.

(Refer Slide Time: 18:15)



Then there is a call Overlap-Add methods.

(Refer Slide Time: 18:18)



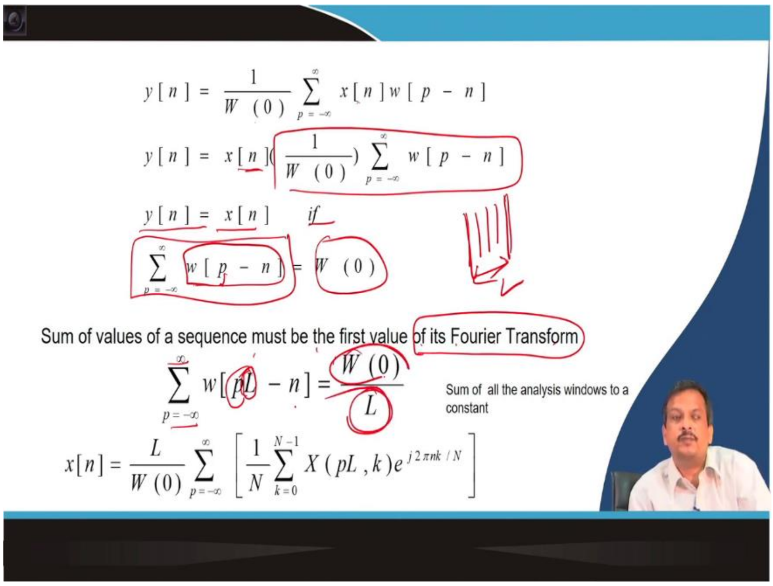
I am not going to detail every method because it is already available on YouTube in another lecture on speech processing.

So, again, I can say my purpose is the same: I have to find out x[n]. So, if I take this as the DFT DTFT view, let us say I want to, and I will start from here. So,

$$y[n] = \frac{1}{W(0)} \sum_{p=-\infty}^{\infty} \left[\frac{1}{N} \sum_{k=0}^{N-1} X(p,k) e^{j\frac{2\pi nk}{N}} \right]$$

So, this is nothing but this one x[n] multiply w and p minus n.

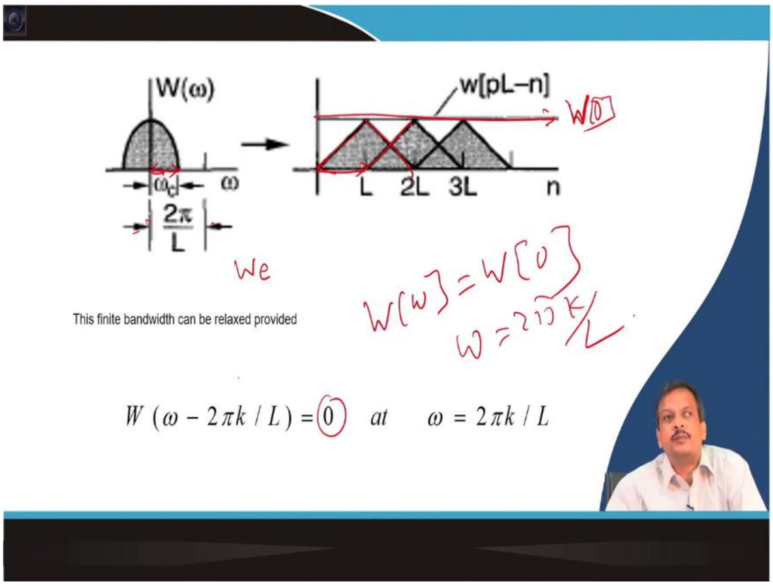
(Refer Slide Time: 19:16)



So, if you see that y[n] is equal to 1 by w[0], this is nothing but an I said x[n] is outside. So, this is nothing but a w p minus n and y[n] equal to x[n] if this one is equal to w[0]. What is the meaning? The sum of the value of a sequence must be the first value of its Fourier transform. The sum of the value of a sequence w p minus n must equal its first energy value. The first value of this is the DC component of the Fourier transform.

So, p is equal to minus infinity to infinity p into L, and L is the dissipation factor if that. So this is for every sample, so I put L instead of every sample. So, I also have to take this sample instead of every sample, and this sample's dissipation is L, which must be w[0] by L. So, what is there? The sum of all analysis windows must be constant.

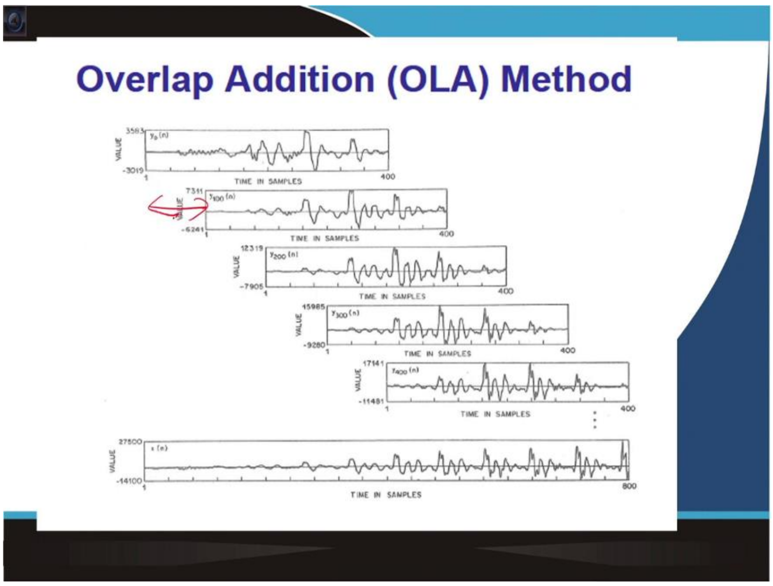
(Refer Slide Time: 20:39)



So, what is the pictorial view? So, if this is my ω c. So, I can say the ω c, and this is my ω L; so, here, if you see that ω 0 by L. So, I can say this is my window impulse response, and this amount of L shifting is possible. Where if I take the sum, I get w[0] first value of the window frequency transform, which is ok.

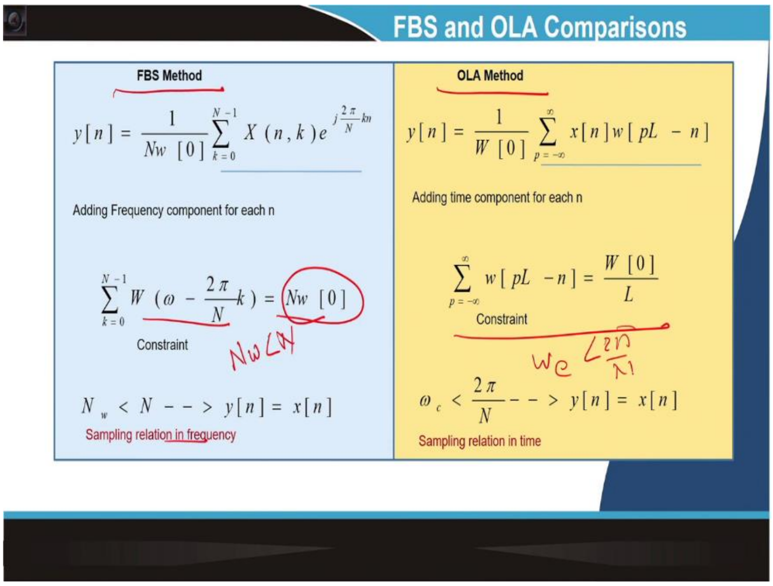
So, where only k is equal to 0, W[ω] is equal to W[ω] will be W[ω] 0, but k is equal to everywhere it will be 0. So, I can say ω must be 2π by k by L ok.

(Refer Slide Time: 21:48)



So, this is the number of shifts that is possible. So, how much shift is possible that depends on the bandwidth of the window function also. So, there is a summary comparison. I think the summary slide is here or not.

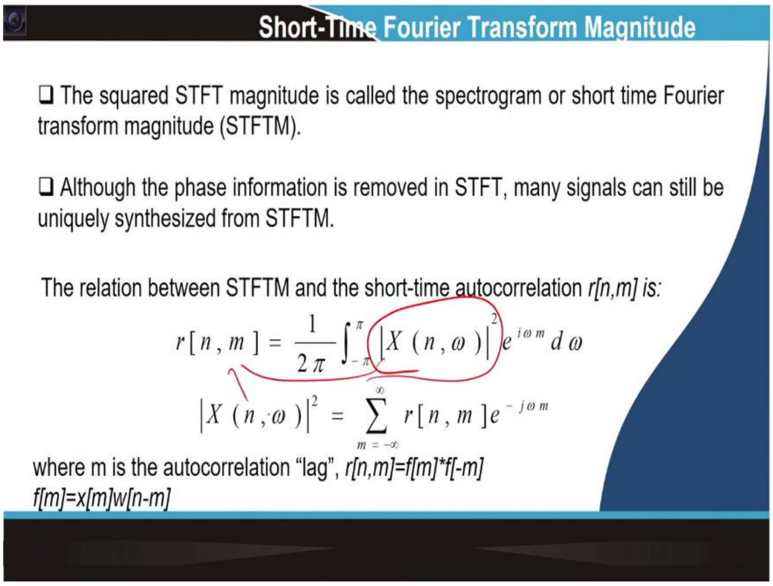
(Refer Slide Time: 22:11)



This is the summary. So, this is the FBS method; this is the OLA method. FBS method constraint is that this portion should be equal to this one, which says that the sampling relation in frequency N w must be less than N. Here is the constraint is this one that means ω c off frequency of the off frequency of the window must be less than 2π/N.

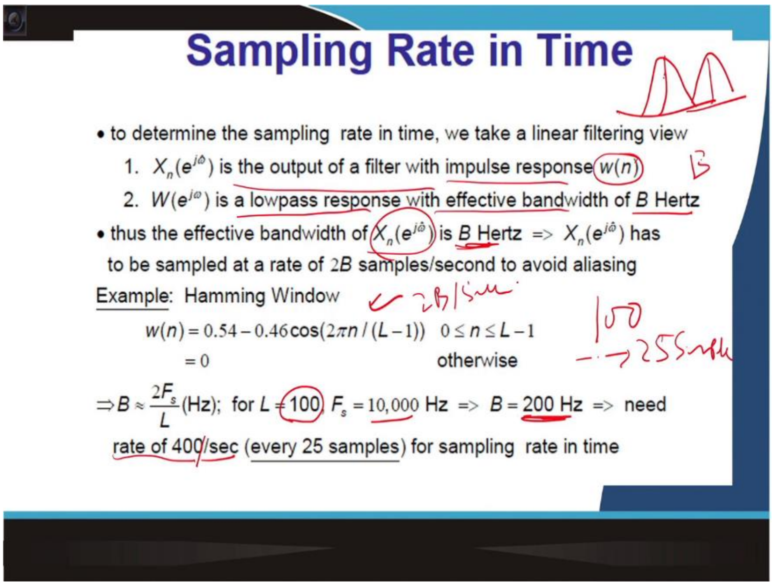
So, that is the time-frequency trade-off between the shift of L, and how much shift is possible you can get from that things.

(Refer Slide Time: 23:00)



I think I have a slide that has an example; I am not going into details on this amplitude magnitude. So, it is also possible that only from the spectral magnitude of the analysis STFT analysis can the signal recovery be proved.

(Refer Slide Time: 23:18)

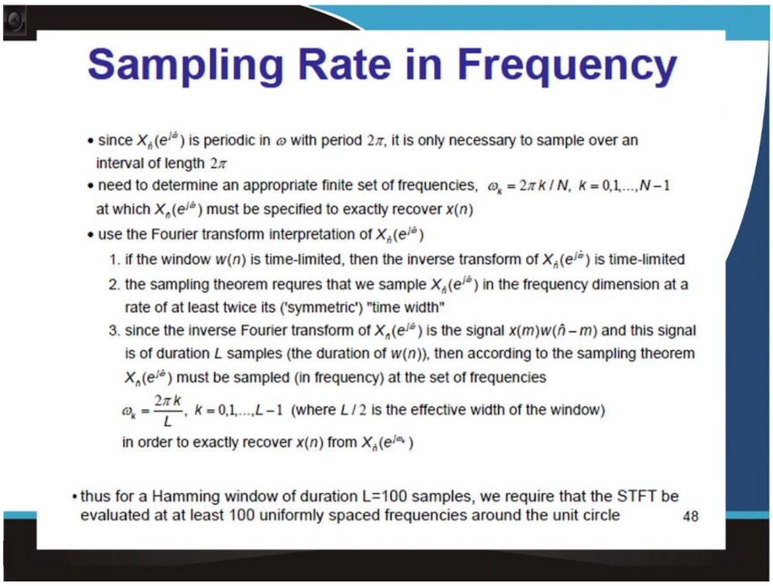


Here, I can say that there is an example. So, X $e^{j\omega}$ is the output of the filter with impulse response w[n]. And w[n] is the low pass response with the effective bandwidth B hertz. So, I have a window whose effective bandwidth is B hertz. Does the effective bandwidth of e to the power B hertz; because the bandwidth is the filter, we will define at the bandwidth of the frequency analysis, ok?

This has to be sampled at 2 B samples per second. For example, let us say I have taken a hamming window. So, bandwidth is twice F s by L twice F s by L. So, for L equal to 100 and F s equal to 10 kilohertz, B is equal to 200 hertz; that means it requires 400 needs the rate sampling rate time sampling rate is 400 sample per second ok.

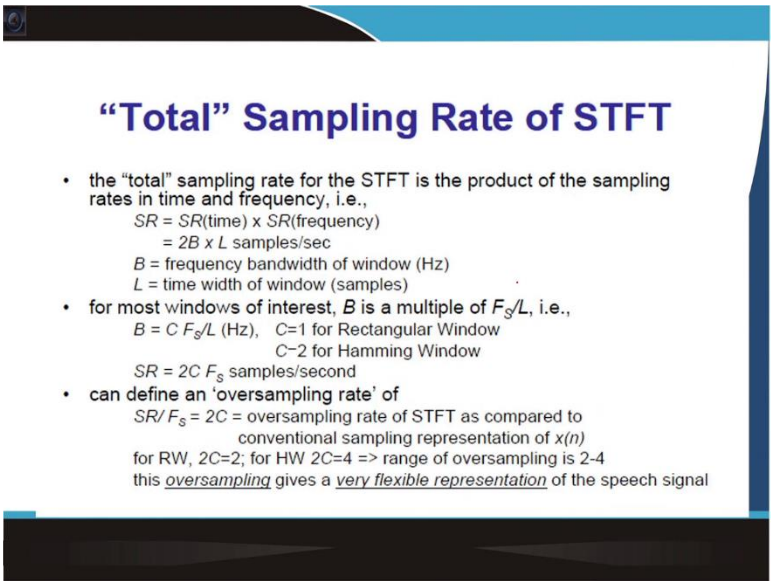
So, that means here I can shift the window. If the length of the window is 100, I can shift the window at 25 samples. Twice F s by L bandwidth is 200 hertz; so, I require 2 B sample per second ok rate in time.

(Refer Slide Time: 25:02)



So, I will supply this material to you.

(Refer Slide Time: 25:07)



If you are interested in the details, read it. You can read it for the details of the STFT analysis, which will be available on YouTube. You can also read it from this slide. So, because I am not loading the course with the particular analysis because of STFT again, it can take another week to explain the details of STFT analysis.

So, thank you.