Digital Signal Processing

Prof. S. C. Dutta Roy

Department of Electrical Engineering

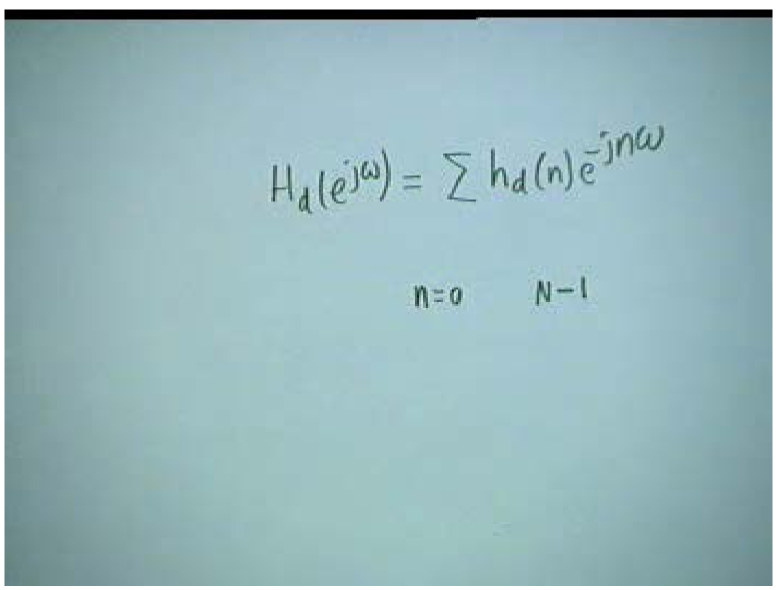
Indian Institute of Technology, Delhi

Lecture – 40

FIR Design by Windowing

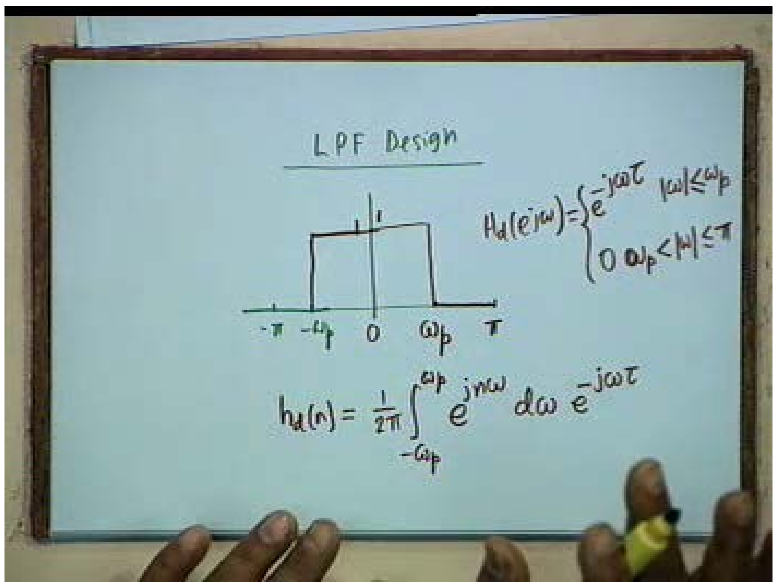
This is the 40th lecture and our topic for today is FIR Design by Windowing which we have already started and we shall also introduce the frequency sampling techniques, if time permits. In the previous lecture we talked about the basic concept of windowing, that is the given desired frequency response is expanded in Fourier series which is nothing but the Fourier Transform of the desired impulse response.

(Refer Slide Time: 01:28 – 03:51 min)



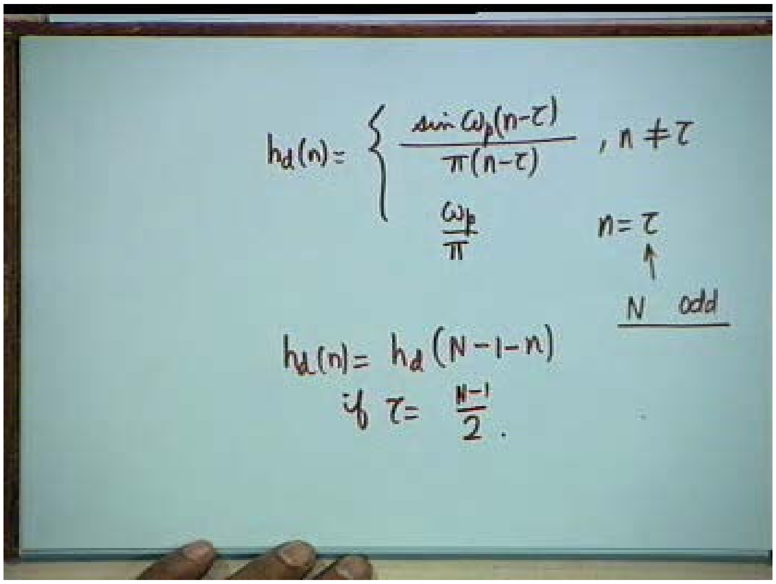
So you find $h_d(n)$ which is infinite in length. You have to truncate this from some point (n = 0) to (n = N – 1). The length is N, which is to be estimated by empirical formula. You are not going below (n = 0), because then it becomes non-causal or unrealizable. Therefore you start at (n = 0). If you truncate abruptly at N – 1, we say that this is the application of a rectangular window. And any abrupt termination leads to Gibbs phenomenon which we illustrated last time. The attempt is to reduce Gibbs phenomenon to reduce the overshoots and undershoots in the frequency response by tapering the window in a smoother manner. We have talked about a number of windows: the Modified Rectangular, the Hann window, the Hamming window, the Generalized Hamming window, the Blackman window, and the Bartlett window, which is triangular in shape. And we said that the optimum windows are the Kaiser and the Dolph Chebyshev.

(Refer Slide Time: 03:54 - 06:57 min)



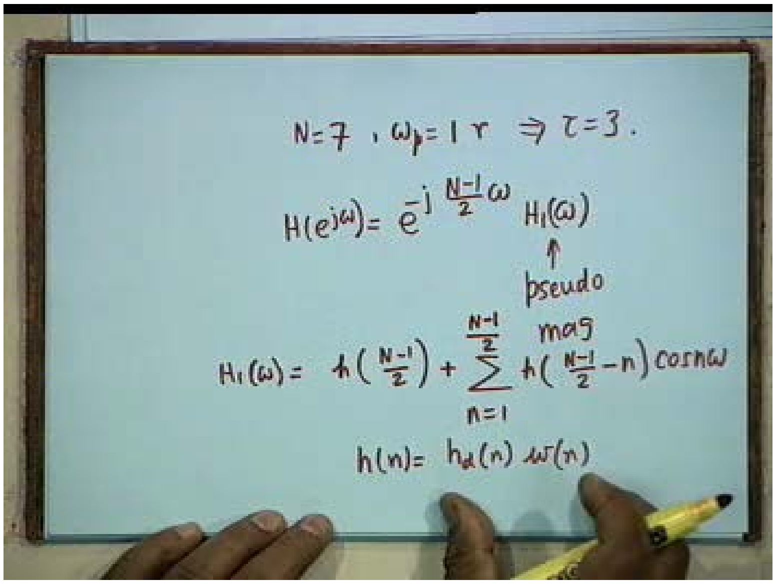
Then we started an example of low pass filter (LPF) design and since we have no control over the tolerances we aim for the ideal response. That is, we assume that our magnitude response is 1 from 0 to some frequency ω_p and then from ω_p to π, it is exactly equal to 0. Now you have to find $h_d(n)$ by the inverse Fourier Transform formula over one period and the easiest thing to do is to take the limits from $-\pi$ to $+\pi$. You can take from 0 to 2π also and the result shall be the same. Although our range of vision is 0 to π, in order to obtain $h_d(n)$ we have to extend from $-\pi$ to $+\pi$ and then $h_d(n)$ becomes $[1/(2\pi)] \int_{-\omega_p}^{\omega_p} e^{jn\omega}\, d\omega$. Now this is true if we only want a zero phase response. We want a linear phase and therefore we take $H_d(e^{j\omega})$ not as 1 but as $e^{-j\omega\tau}$ for $|\omega| \le \omega_p$ and 0 otherwise, i.e. for $\omega_p < |\omega| \le \pi$. After all, FIR design is resorted to only if linear phase is a constraint. So we pre-assume that we are going to approximate a linear phase.

(Refer Slide Time: 07:01 - 09:16 min)



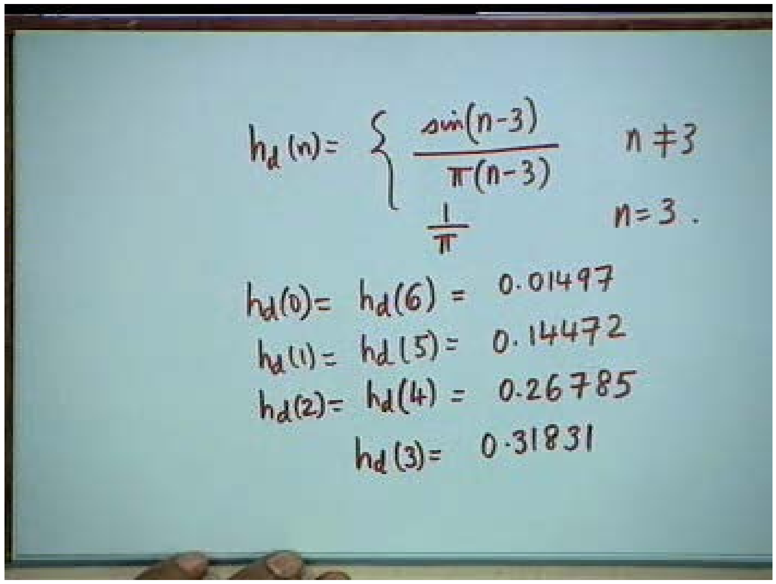
Therefore if we carry out the integration, the result, as we had shown last time, is $h_d(n) = \sin(\omega_p(n - \tau))/(\pi(n - \tau))$, $n \neq \tau$ and it is ω_p/π if $n = \tau$. This case will arise only when τ is an integer because N is an integer. In other words this case will arise when N is odd and this is usually the case. We assume N to be odd because the delay is an integer. The half sample delay creates its own problem. Therefore, we assume that N is odd. Last time, we started with linear phase, obtained $h_d(n)$, and verified that $h_d(n) = h_d(N - 1 - n)$ requires $\tau = (N - 1)/2$.

(Refer Slide Time: 09:23 - 12:13 min)



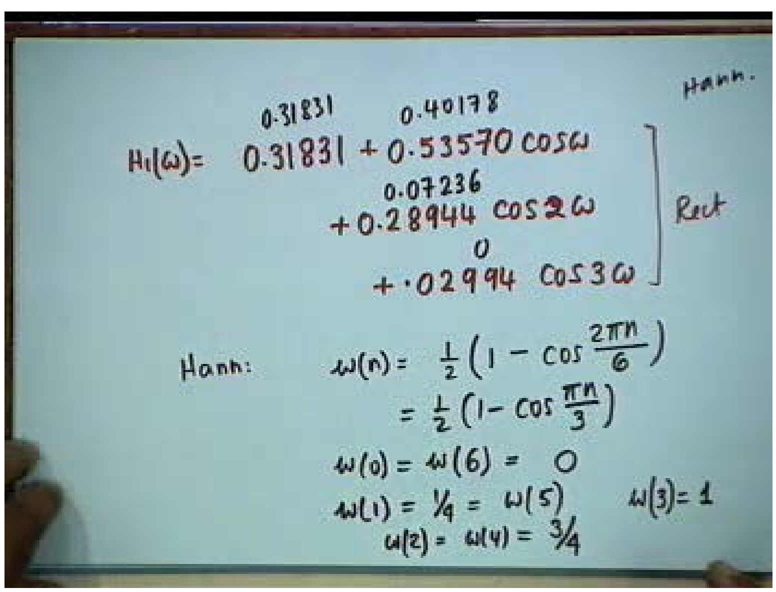
The example we shall carry throughout is that of N = 7 and ω_p for the low pass filter is 1 radian; it is not 1 radian per second. N = 7 therefore leads to τ = 3. In order to obtain the frequency response curve we require the expression for this linear phase filter of type one, that is, symmetrical impulse response and odd length. If you remember for type one, the frequency response is given by $e^{-j\omega(N-1)/2}H_1(\omega)$, where $H_1(\omega)$ is the pseudo magnitude; it is not a magnitude because it can change its sign. $H_1(\omega)$ is given by h((N – 1)/2) plus other terms which match each other to produce a cosine series and this is $\sum_{n=1}^{(N-1)/2} h((N-1)/2 - n) \cos n\omega$. You require this formula to estimate what you have achieved by the windowing technique. h(n) is to be obtained from $h_d(n)$ and the chosen window function w(n). Once you obtain these, you substitute here and then you get the frequency response.

(Refer Slide Time: 12:16 - 14:05 min)



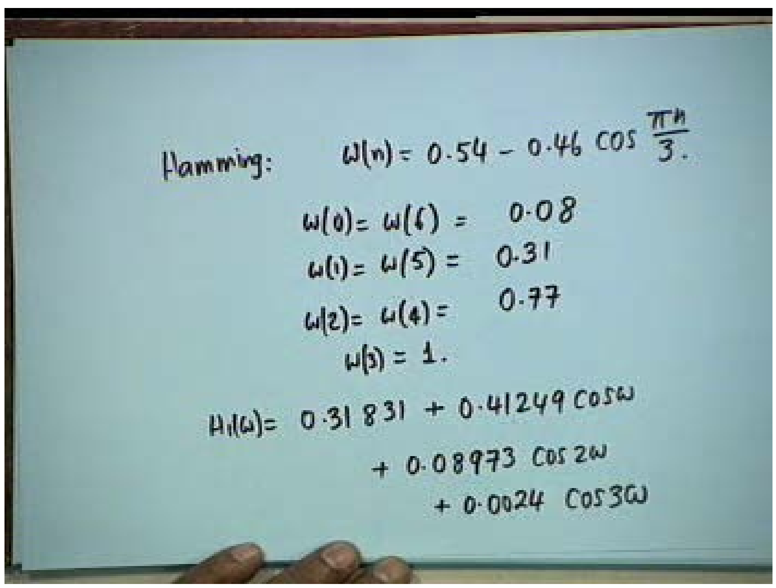
In this particular case, $h_d(n) = \sin(n - 3)/[\pi(n - 3)]$, $n \neq 3$, and $= 1/\pi$ for $n = 3$ and this gives rise to the following: $h_d(0) = h_d(6) = 0.01497$; $h_d(1) = h_d(5) = 0.14472$; $h_d(2) = h_d(4) = 0.26785$; and $h_d(3) = 1/\pi$, that is 0.31831. Now, if we use rectangular window then $w(n) = 1$ and therefore these themselves determine the frequency response.

(Refer Slide Time: 14:07 - 19:09 min)



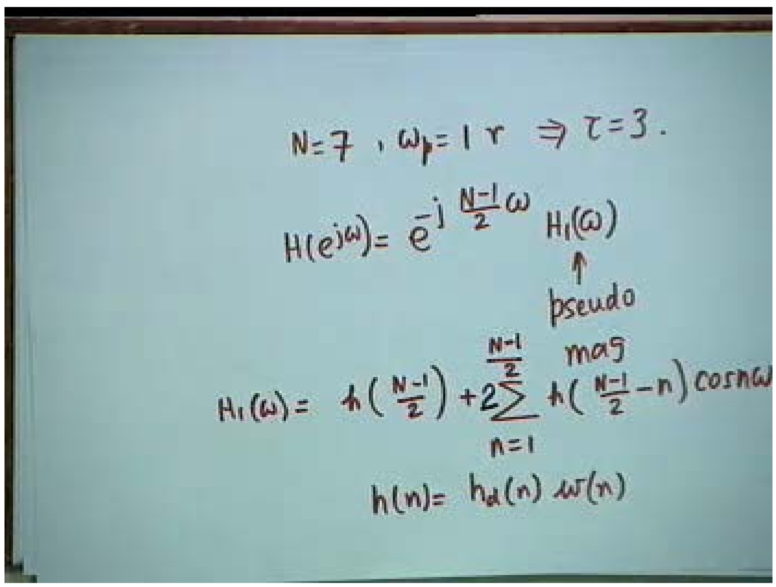
So $H_1(\omega)$ the pseudo magnitude becomes equal to 0.331831 + 0.53570 cos ω + 0.28944 cos 2ω + 0.02994 cos 3ω. This is the result for rectangular window. Now, if you apply any other window, for example, the Hann window, then w(n) is (1/2)(1 – cos(2πn)/(N – 1)). And therefore w(0) = w(6) = 0. Hann actually uses two lengths less because the terminal values are 0. w(1) comes as (1/2) (1 – cosπ/3) = 1/4 and that is w(5) also. Half the number have to be calculated because of symmetry. Also, w(2) = w(4) = (1/2)(1 – cos 2π/3) and that calculates out as ¾, and finally w(3) = 1. It is not required to calculate the middle point that is normalized with respect to 1. In Hann window therefore this term shall remain the same. The coefficients of Hann in $H_1(\omega)$ are written above those in the expression for rectangular window (refer slide).

(Refer Slide Time: 19:11-22:58)



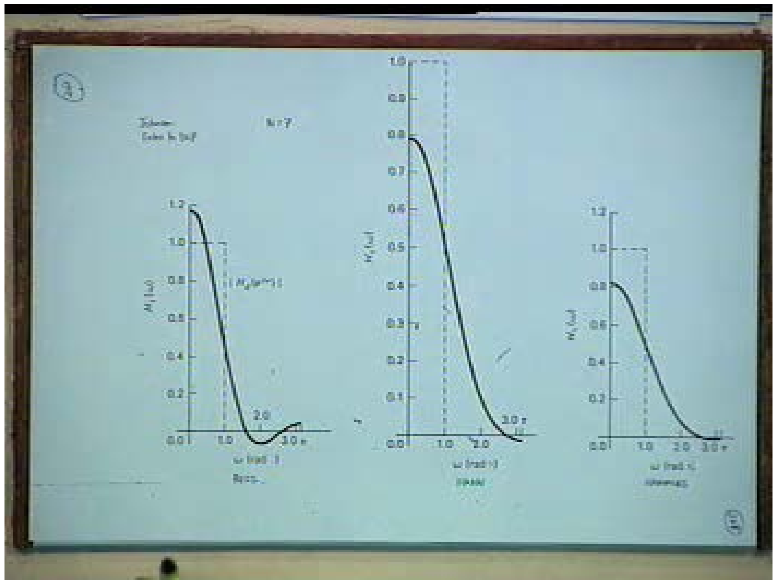
In Hamming, the window values are w(n) = 0.54 – 0.46 cos πn/3 because the length is the same and we know these values without calculating anything. Here w(0) = w(6) = 0.08; w(1) = w(5) = 0.31; w(2) = w(4) = 0.77; and w(3) = 1. Now I must tell you why I suddenly changed to two places of decimals. There is so much of uncertainty because N may not be correct and we have no control over tolerances; therefore ultimately what you get is a rough design. If it satisfies the specs, you consider yourself lucky, but generally you have to iterate; you have to go to higher orders. The only control that you have is the order and the window. So, start with Hann, then go to Hamming. If Hann suffices you are very happy because the hardware requirement is reduced by two. That is why Hann is not discarded from the beginning, it is used. Hann, Hamming and perhaps one more window generally suffice, unless you want to go to the optimum. If you want to go the optimum, then all these calculations are not needed. They are simple calculations for the other windows and can be used if the requirement is not very stiff and if the tolerance can be relaxed. In this particular example, using Hamming, the frequency response becomes $H_1(\omega) = 0.31831 + 0.41249 \cos \omega + 0.08973 \cos 2\omega + 0.0024 \cos 3\omega$.

(Refer Slide Time: 22:33 - 22:42 min)



And the three pseudo magnitudes have been plotted here side by side.

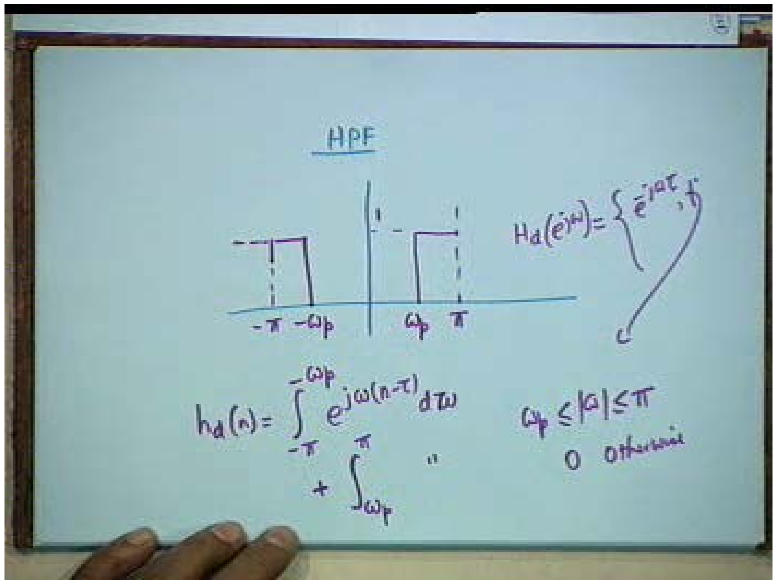
(Refer Slide Time: 23:04 - 25:06 min)



If you notice carefully, these figures cannot be compared straightaway because they are not normalized. The maximum amplitude has not been normalized to 1; that should be done before comparison. They should be plotted on the same graph so that you can compare them. But one thing that you notice is the zero crossing which does not change with normalization. The zero crossing is between 1 and 2 in the rectangular window, but it is between 2 and 3, closer to 3 than 2 in the Hann as well as Hamming which is a reflection of the widening of the main lobe.

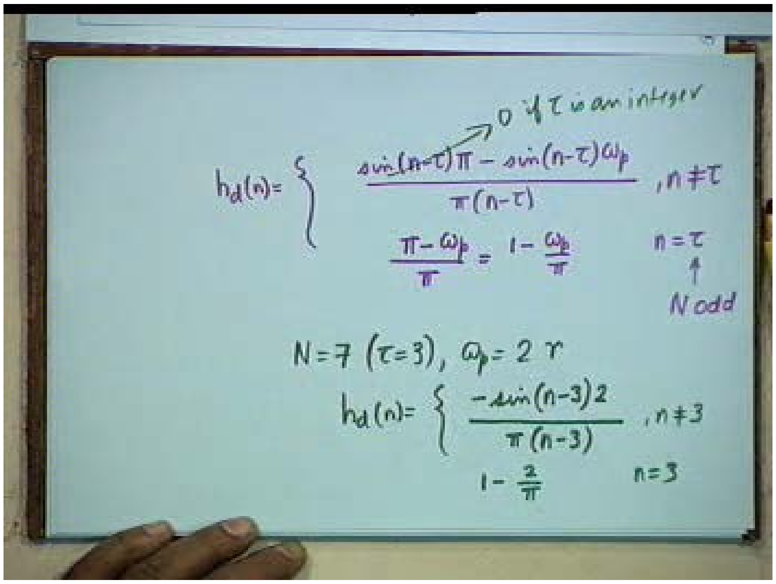
Now, what is the effect on the low pass filtering? The cutoff becomes slower; rectangular window gives you a sharper cutoff than the other two. The side lobe height is reduced considerably in the Hann and Hamming. In Hamming, for example, the side lobe is hardly discernable. In Hann, the scales are different. This is not the type of plot you should make, you must make the scales the same; normalize the maximum amplitude to 1 then compare and weigh against the tolerance scheme that is specified and then decide which shall be used. If Hann suffices then you need not go further. Let us look at a high pass filter (HPF) design by windowing. For HPF you must note the specs carefully.

(Refer Slide Time: 25:25 - 27:19 min)



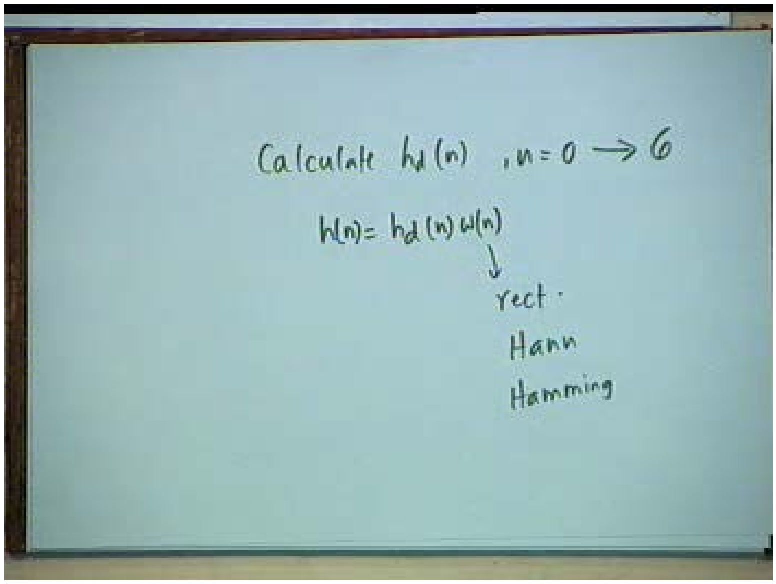
For HPF, we want unity amplitude between ω_p and π. I am not bringing the plot down at $\omega = \pi$, because if I extend, it goes up to $\pi + \omega_p$. I have the mirror image between $-\omega_p$ and $-\pi$. The magnitude is 1, but my transfer function $H_d(e^{j\omega})$ is linear phase, that is $H_d(e^{j\omega}) = e^{-j\omega\tau}$ for ω between ω_p and π and between $-\omega_p$ and $-\pi$. So my integral now shall have two parts $h_d(n) = \int_{-\pi}^{-\omega p} e^{j\omega(n-\tau)}\, d\omega + \int_{\omega p}^{\pi} e^{j\omega(n-\tau)}\, d\omega$ The integrand is the same but we have to put the limits carefully.

(Refer Slide Time: 27:23 – 29:46 min)



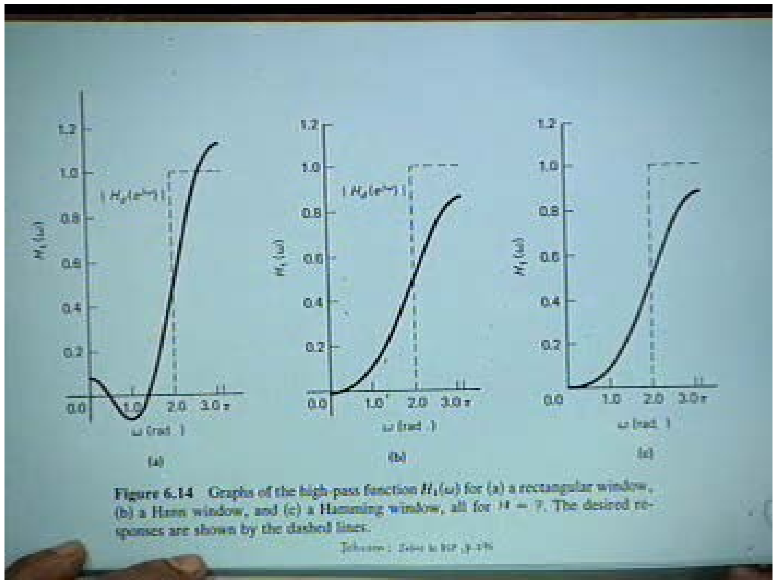
If you do that, then the result comes as $h_d(n) = [\sin(n - \tau)\pi - \sin(n - \tau)\omega_p]/[\pi(n - \tau)]$ provided $n \neq \tau$. If $n = \tau$ then it is simply $(\pi - \omega_p)/\pi = 1 - (\omega_p/\pi)$ for $n = \tau$ which indicates that N is odd. You also notice that if N is odd, then the term $\sin(n - \tau)\pi$ here is identically equal to 0 but this is not the case if N is even. Recall that there is another reason why an odd length is preferred. What is it? I leave it to you to recall. Suppose N = 7 that is $\tau = 3$ and $\omega_p = 2$ radians; then $h_d(n) = -\sin[(n - 3) \times 2]/[\pi(n - 3)]$ for $n \neq 3$, and it is $1 - (2/\pi)$ for $n = 3$. So you can calculate h_d and can apply the various windows.

(Refer Slide Time: 29:57 - 30:59 min)



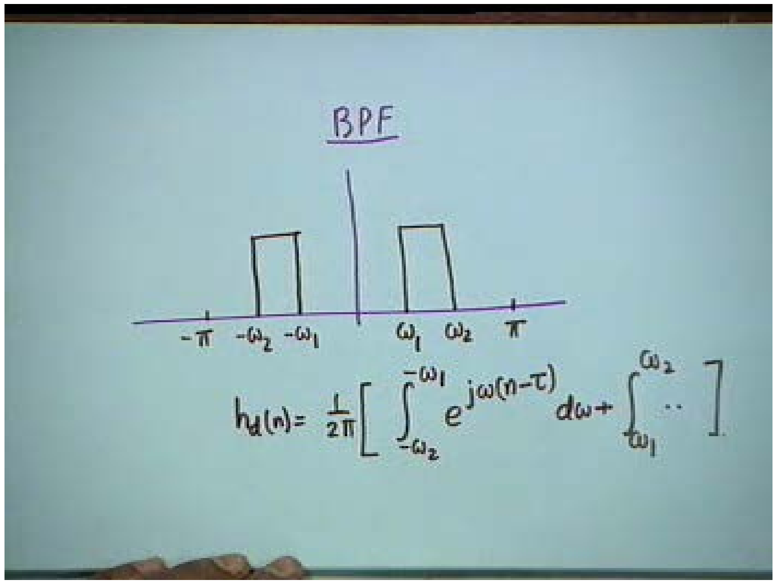
The steps would be to calculate $h_d(n)$, $n = 0 \rightarrow 6$ in which obviously you have to calculate only for n = 0, which is $1 - (2/\pi)$, and for n = 1 and 2. And then make $h(n) = h_d(n)\ w(n)$ where w(n) is rectangular, or Hann or Hamming window, which are the simplest windows. You have other choices. Even in Hamming, you do not have to assume the first term as 0.54; you could assume any α, so there is lot of flexibility. You can calculate the pseudo magnitudes in the three cases. The results are plotted here in the next slide.

(Refer Slide Time: 31:14 - 33:12 min)



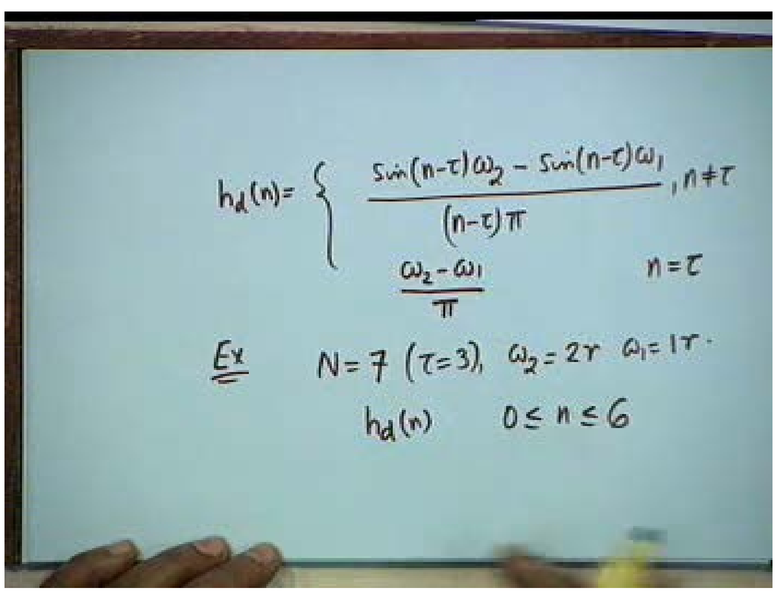
Notice carefully what happens to the frequency response. Once again these three figures should not be compared as they are, because the maximum response has not been normalized to 1. Maximum response in rectangular case is about 1.16. The maximum response in the other two plots does not go to 1. Notice that Hann and Hamming are close to each other because the windows have been tapered. But the Hann window will always give you a slightly inferior response because it uses a smaller length. The undershoot in Hamming is almost unnoticeable. The slope of the main part would be positive here because it is a high pass filter. The slope deteriorates with Hann and Hamming. There is hardly a ripple in the upper part. In all probability, rectangular window will not work, therefore you have to use either Hann or Hamming window. Next we take the case of a band pass filter.

(Refer Slide Time: 33:19 -34:49 min)



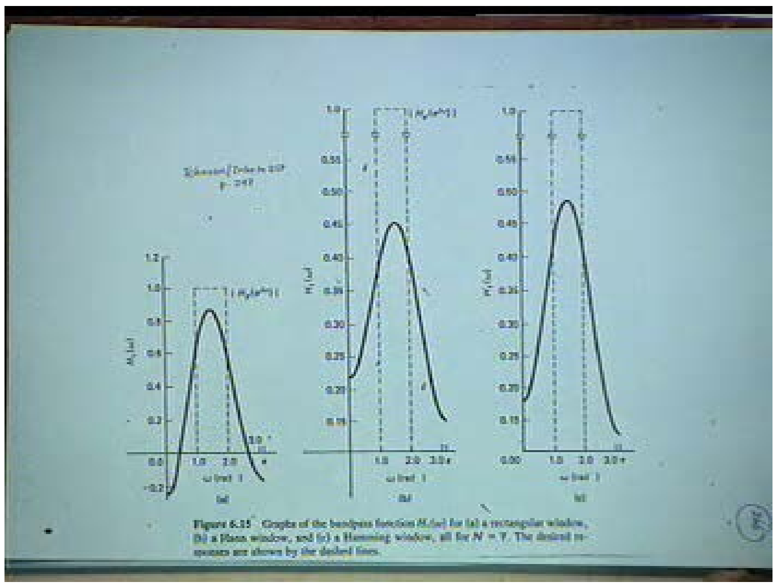
For a band pass filter, you want a rectangle between ω_1 and ω_2 and another rectangle between $-\omega_2$ and $-\omega_1$ (refer slide). Also, ω_2 is less than π (Refer Slide). This is my band pass filter and therefore $h_d(n) = [1/(2\pi)]\ [\int_{-\omega 2}^{-\omega 1} e^{j\omega(n-\tau)}\, d\omega + \int_{\omega 1}^{\omega 2} e^{j\omega(n-\tau)}{}_{d\omega}$. If you calculate this, it comes out as $h_d(n) = [\sin(n-\tau)\,\omega_2 - \sin(n-\tau)\,\omega_1]/[(n-\tau)\,\pi]$; $n \neq \tau$ and when $n = \tau$ it is simply $(\omega_2 - \omega_1)/\pi$.

(Refer Slide Time: 34:52 - 36:29 min)



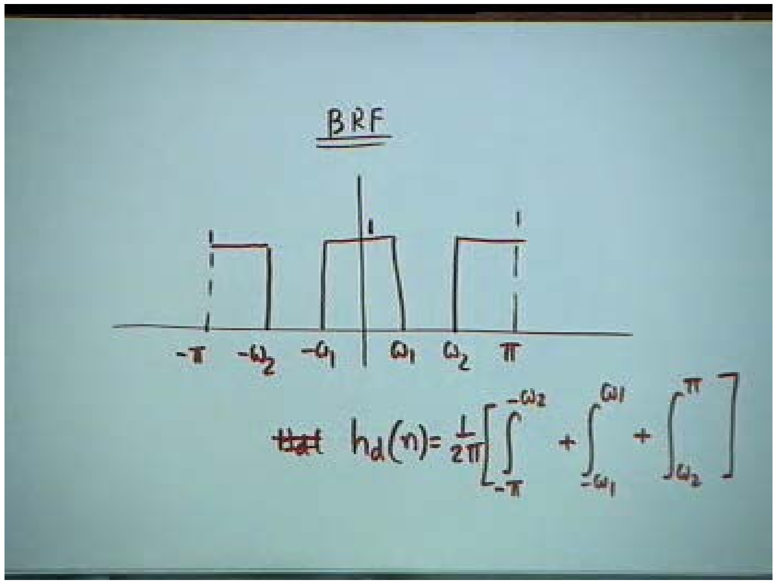
To take an example, suppose N = 7; it is a very small length and is only used for illustration. N = 7 means $\tau = 3$; let $\omega_2 = 2$ radians and $\omega_1 = 1$ radian. Then you calculate $h_d(n)$ $0 \le n \le 6$ in which you have to calculate only three of them and the rest would be symmetrical. Then apply the Hann window and the Hamming window. The pseudo–magnitude is not plotted on the same graph; it is plotted side by side for Hann window and the Hamming window.

(Refer Slide Time: 36:44 – 37:21 min)



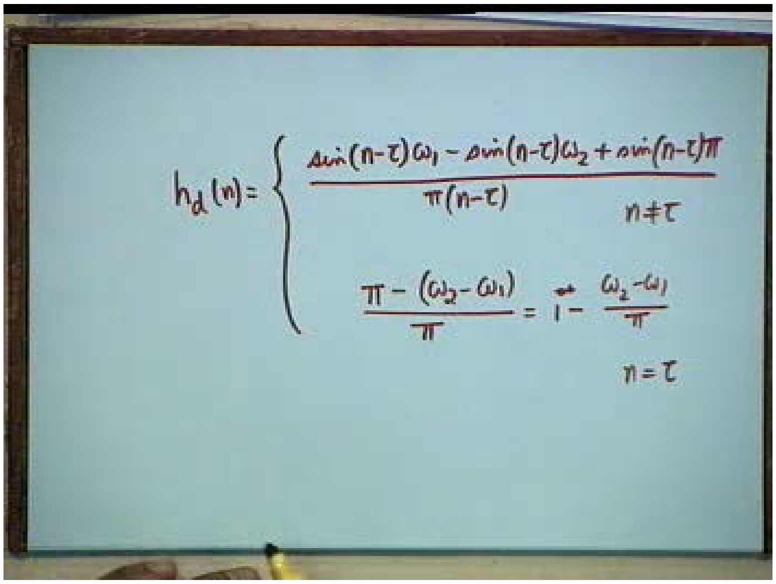
Invariably, the pseudo magnitude does go negative with rectangular window but with Hann and Hamming, it does not. The slope of rise and slope of fall are sharpest in rectangular case. What you pay here for reducing side lobes is a widening of bandwidth. As I said N = 7 is a very small number and it cannot achieve any useful purpose. Finally we look at band reject filter.

(Refer Slide Time: 37:25 - 39:08 min)



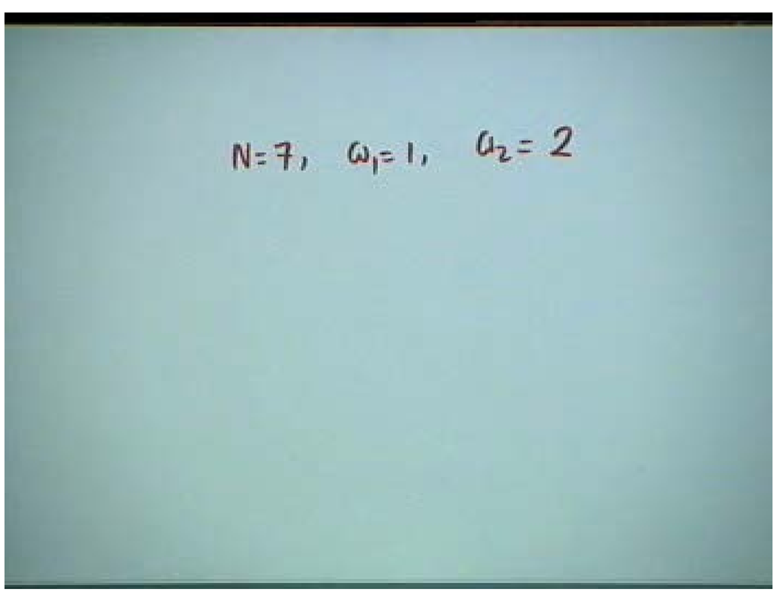
Now you shall have a pass band between $-\pi$ and $-\omega_2$, $-\omega_1$ to $+\omega_1$ and ω_2 and π (refer slide); the stop band is between ω_1 and ω_2. There is a pass band before that and a pass band after that. This is the band reject filter and the passband magnitude is 1. But you assume that the transfer function is $e^{-j\omega\tau}$ so that it has linear phase. So $h_d(n)$ shall now consist of three integrals; the integrand is the same as we have seen earlier. The integrals are from $-\pi$ to $-\omega_2$, then $-\omega_1$ to $+\omega_1$, and finally ω_2 to π. The integrand is $e^{j\omega(n-\tau)}\,d\omega$.

(Refer Slide Time: 39:13 - 40:57 min)



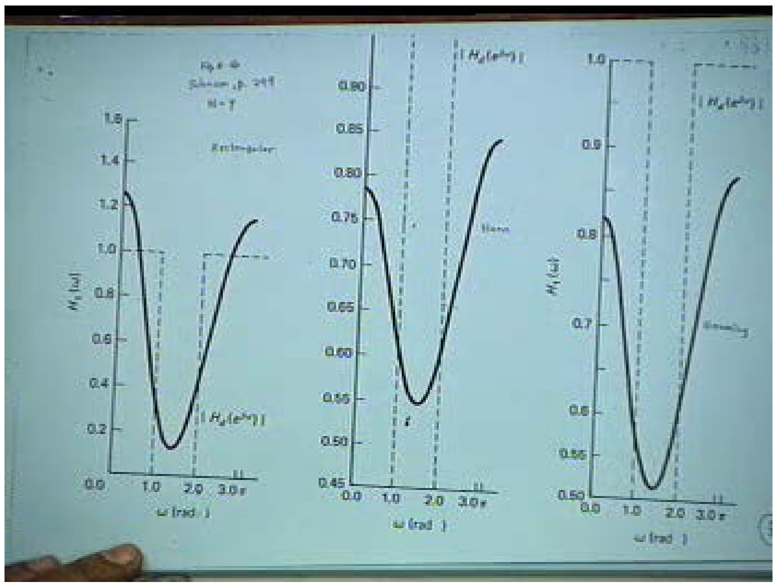
The result is that $h_d(n)$ becomes equal to [sin(n —τ) ω_1 – sin(n – τ) ω_2 + sin(n – τ)π] (This last term appeared in high pass filter also and it will disappear if τ is an integer) divided by π (n – τ), for n ≠ τ; and for n = τ it is [π – ($\omega_2 - \omega_1$)]/π = 1 – (($\omega_2 - \omega_1$)/π). $\omega_2 - \omega_1$ is the stop bandwidth in this particular case because we are aiming at an ideal response.

(Refer Slide Time: 41:01 - 41:16 min)



Once again we take an example of the same length; N = 7; let ω_1 = 1 radian, ω_2 = 2 radians and you calculate h(n) and find out the pseudo magnitude for the three windows: rectangular, Hamming and Hann. This is what the result looks like.

(Refer Slide Time: 41:30 - 44:00 min)

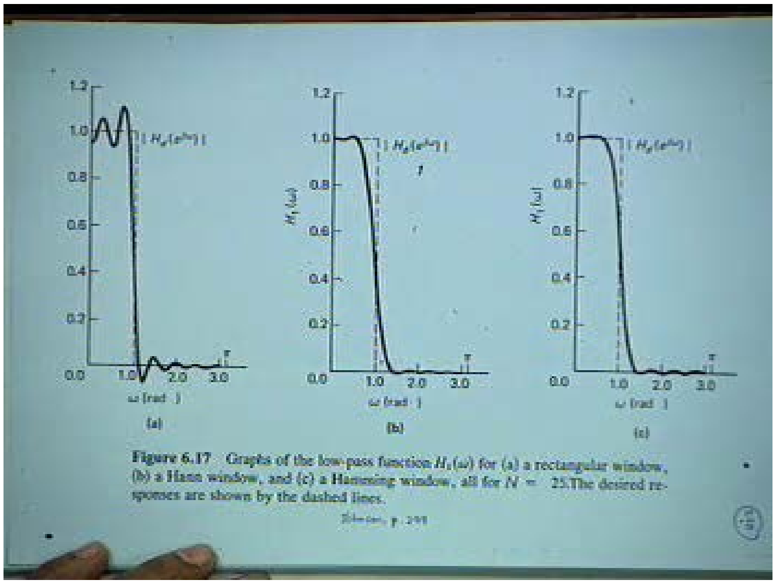


Once again if you want to compare you have to plot them on the same figure. You have to normalize the maximum magnitude to 1. The maximum magnitude for the rectangular window occurs at $\omega = 0$; for the other two, it occurs at $\omega = \pi$. In the normalization also, you shall to be careful. In the rectangular window case, for example, the maximum is not unity. The pseudo magnitude is not symmetrical around the rejection frequency. In the IIR case, there was geometric symmetry.

What made it symmetrical in the IIR case? Why was it that in IIR, we always got symmetry of a special kind, i.e. geometrical symmetry? It occurs because of the property of the analog filter from which we derived that filter; geometric symmetry was continued in the digital filter also. But here we have no control; there is no symmetry, and there was no parameter to control the symmetry.

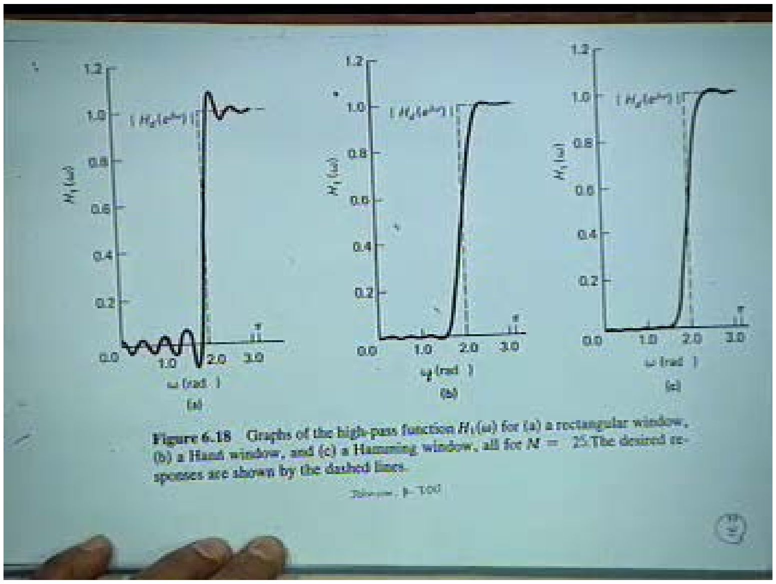
You also notice the tolerances; normalization is needed to compare how close the dip goes to zero and that will determine the stop band tolerance and the stop bandwidth. So a comparison of the three plots requires further work; just plotting them individually is not enough. Now, if I increase the length, you would see what happens. I will now project a number of figures one by one and show what the effect of length is. The next figure shows the case of length equal to 25.

(Refer Slide Time: 44:27 - 46:16 min)



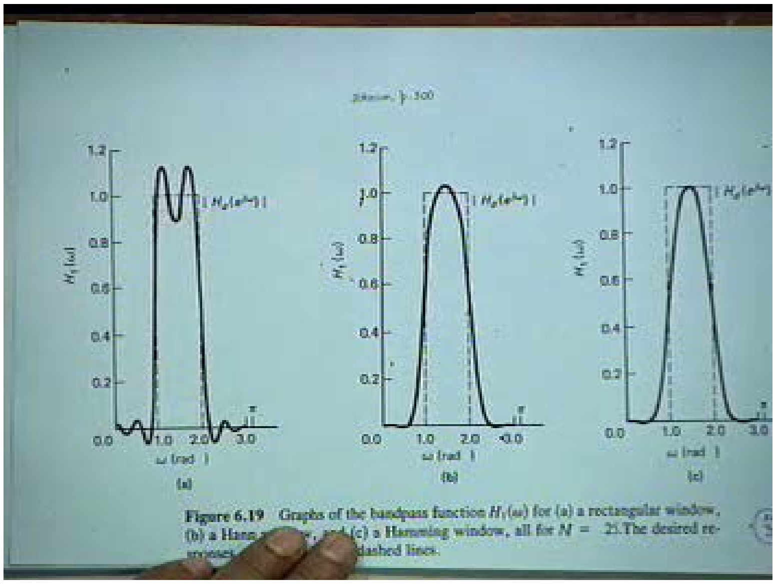
Here normalization has been done. That is, the maximum amplitude has been made equal to 1. There are ripples in the pass band and there are ripples in the stop band, so it is more like an elliptic filter. The amplitude of the ripples now shall go down as N increases because the transition is not abrupt, it is smoother than in the rectangular window. In a rectangular window, however much you increase the length, the overshoot and the undershoot would be 17.8%. In contrast, you see that in the Hamming window, pass band ripple is hardly discernable. Obviously this design is the best of the three, but if Hann satisfies the specs, then it is advisable to accept that. The next figure shows what happens to a high pass filter and the results are very similar. There is hardly anything to distinguish between Hann and Hamming.

(Refer Slide Time: 46:31 - 46:42 min)



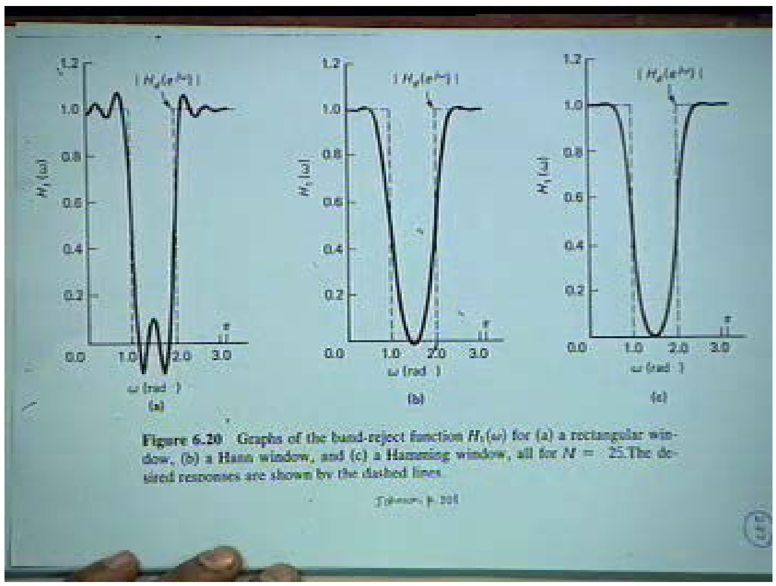
The rectangular one shows the same kind of ripple in high pass also. The next slide shows the band pass case with length N = 25.

(Refer Slide Time: 46:58 - 47:47 min)

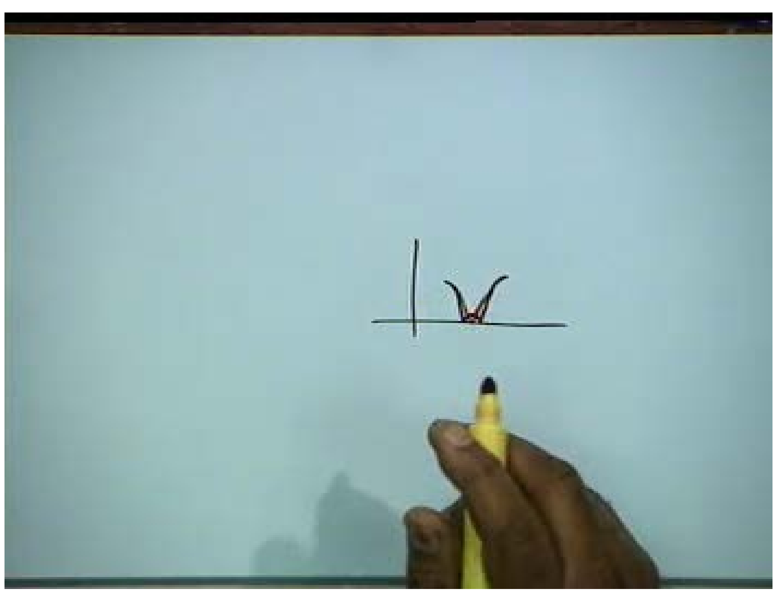


Here the magnitudes are not normalized. When N increases, Hamming normalizes the maximum magnitude to 1. But for Hann there is no control. Hann also goes towards 1 but this is two length less. Once again, you can observe the sharpest cutoff in the rectangular window. Finally consider the band reject filter whose response for N = 25 is shown in the next slide.

(Refer Slide Time: 47:57 - 49:57 min)

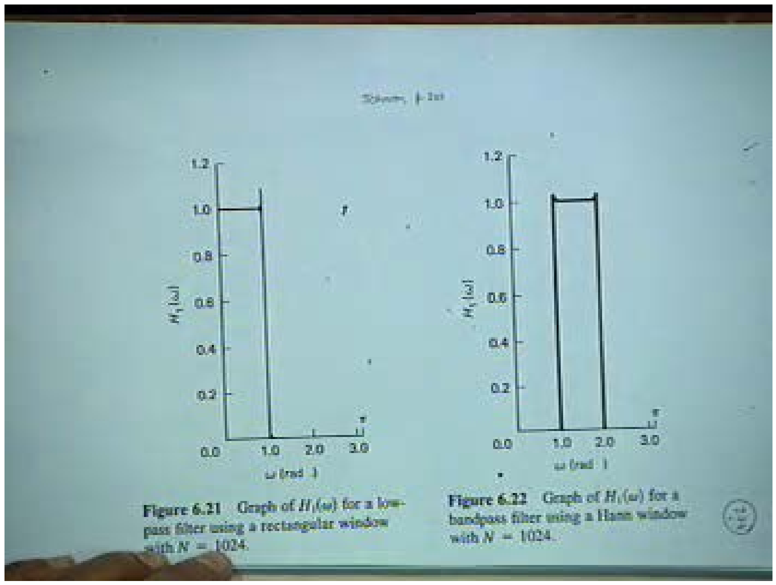


(Refer Slide Time: 48:44 - 49:21 min)



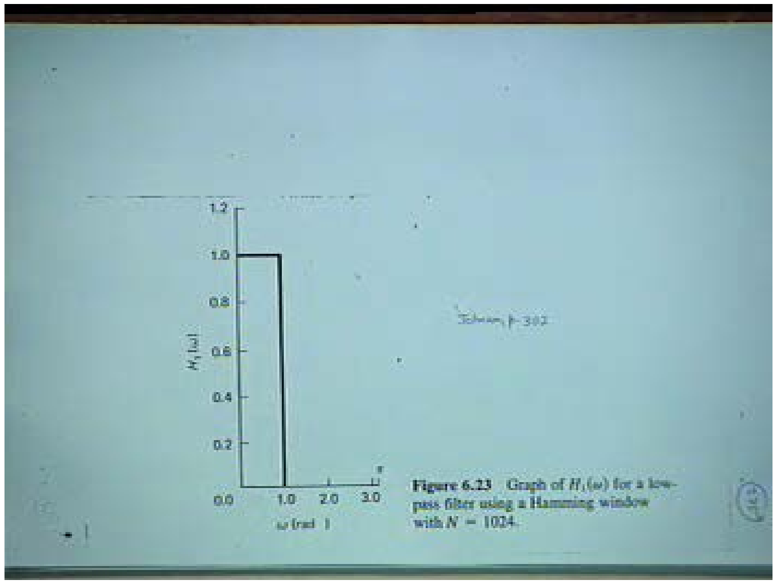
Finally, we show a figure where the length has been increased to 1024.

(Refer Slide Time: 50:12 - 51:19 min)



The first plot is that of a LPF using a rectangular window; there are no ripples, but an extended rod still remains at the transition point and it is approximately 17.8%; this is what Gibbs phenomenon is. The next plot is that of a band pass filter using the Hann window. Finally we show a low pass filter with Hamming window of the same length 1024.

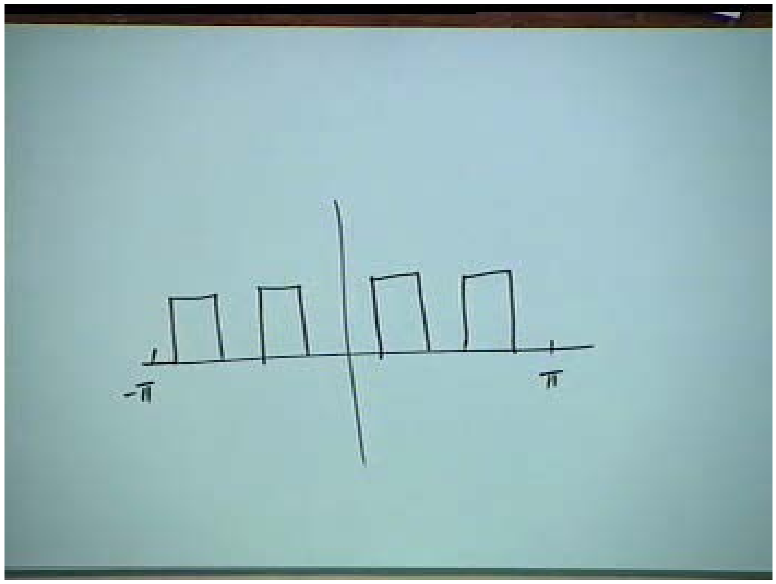
(Refer Slide Time: 51:30 - 52:45 min)



This is almost ideal in FIR design. If you start with sufficiently large N and if your estimate is 800 for example then you start with 1000 and see how it goes. If it over satisfies, then you come down and if it under satisfies, then you go up. What are the disadvantages of large length? One is hardware and other is speed of processing and therefore you need to bring FFT hardware into it and then compute the convolution. You want to design a filter and it has an input, it has an output, it has an impulse response of length 1023. So you make FFT of input and FFT of the impulse response, multiply the two, and take the IFFT; this will speed up the processing.

One important point is that, in IIR design we only took cases of low pass, high pass, band pass and band stop. The transformations were either from analog-to-digital or from digital-to-digital. But we assumed a limited number of pass bands and limited number of stop bands. In a band pass filter for example we assumed one pass band and two stop bands. In a band stop filter we assumed one stop band and two pass bands. If you go to multi-pass band filters, two pass bands and the corresponding number of stop bands between 0 and π then you shall have to design your own transformation. The degree has to be increased. The transformation from analog low pass to analog band pass is second order. If you want two band passes in the range then you will require a fourth order and you shall have to design your own transformation.

(Refer Slide Time: 54:11 - 54:42 min)



On the other hand, in the FIR design, this is not a problem at all; you can have multiple pass bands. For example, you have something like that shown in the slide, you just integrate the same integrand in four different regions. Obtain the corresponding formula and go ahead. So FIR filter design is not as bad as it poses to be, but anything that works finally is good enough for an engineer. We have already seen what happens with an increase in the length to 1023; the filter that we get is almost ideal. We can never get such a response with an IIR filter. IIR filter will have some tolerance in the pass band and some tolerance in the stop band. But with FIR you can almost approach the ideal provided you use the correct window. The order can be reduced by using one of those optimum windows, namely the Kaiser and the Dolph Chebyshev.