Digital Signal Processing

Prof: S. C. Dutta Roy

Department of Electrical Engineering

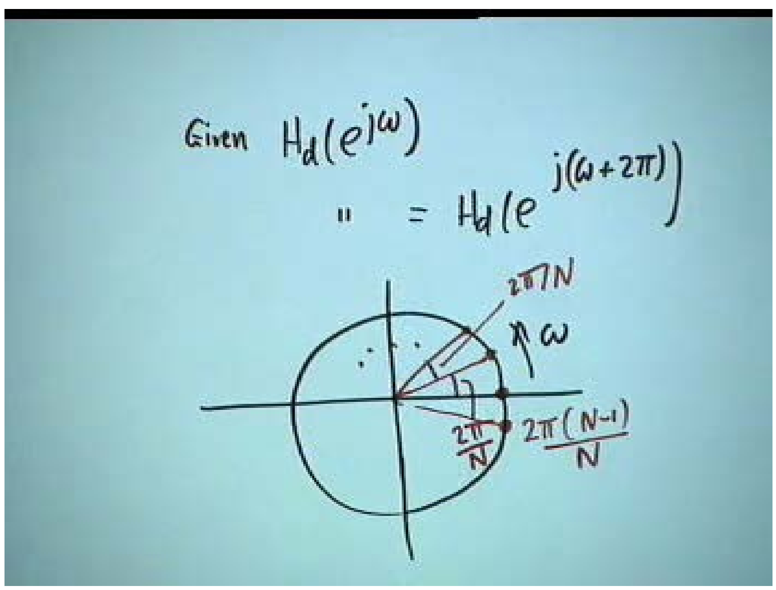
Indian Institute of Technology, Delhi

Lecture – 42

FIR Design by Frequency Sampling

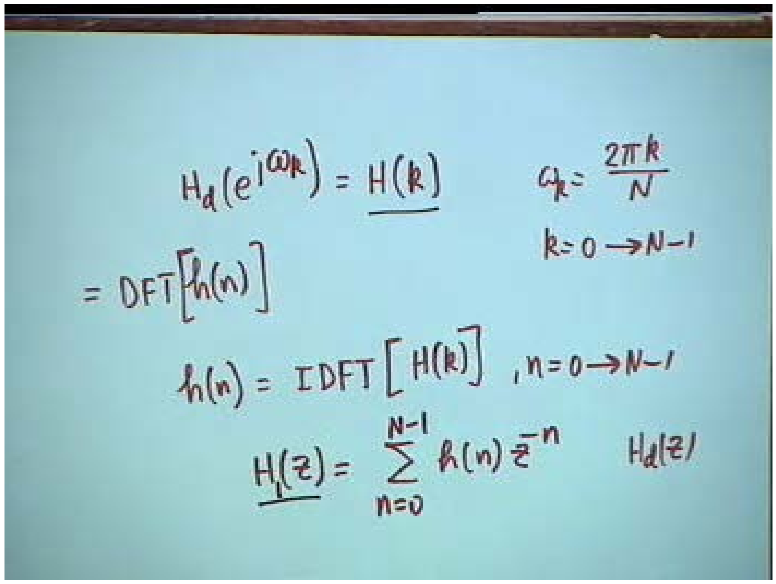
We are going to talk about FIR design by frequency sampling technique. Conceptually it is a very simple technique. Consider the given desired frequency response $H_d(e^{j\omega})$, which, as you know, is periodic, with a period of 2π.

(Refer Slide Time: 01:19 – 03:00 min)



Let $H_d(e^{j\omega})$ be sampled, uniformly, at N number of points within one period 0 to 2π. Therefore our starting sample would be, at ω = 0; the next sample would be at ω = 2π/N. The next sample will be at ω = 4 π/N, and so on; the last sample would be at ω= 2π(N – 1)/N. Let us call these samples $H_d(e^{j\omega_k})$ as H(k), where $\omega_k = 2\pi k/N$; k = 0 → N – 1.

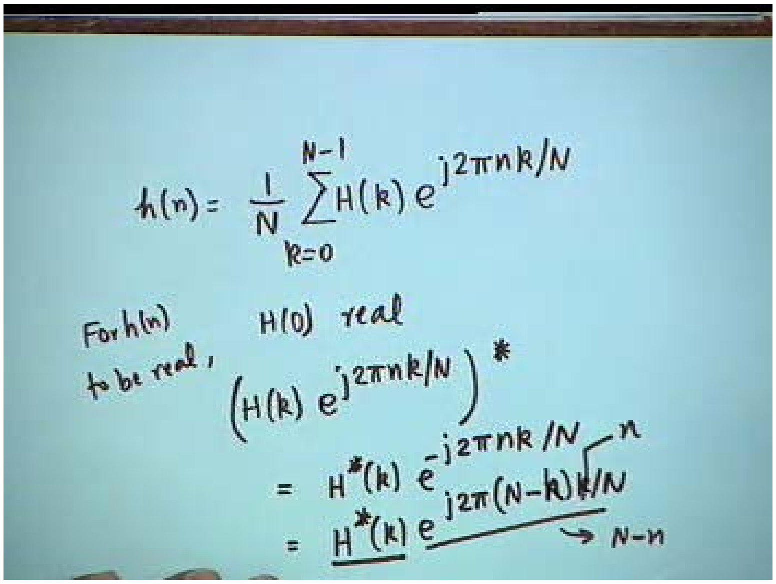
(Refer Slide Time: 03:03 – 06:43 min)



As you know, these samples represent the DFT H(k) of a certain sequence h(n). And if you find h(n) by IDFT of H(k), this shall also be of length N, which forms an FIR filter. Now if you design a filter with these impulse responses, then its transfer function $H_1(z) = \sum_{n=0}^{N-1} h(n)\, z^{-n}$ will be a filter which approximates the desired frequency response $H_d(z)$. They are not identical because $H_d(e^{j\omega})$ was not specified to be of finite impulse response. It is an arbitrary specification. It can be satisfied by FIR and IIR but there is no label on $H_d(e^{j\omega})$. And therefore if you invert $H_d(e^{j\omega})$, you get $h_d(n)$ which in general will be of infinite length. You are getting a finite length through the artifice of DFT. Conceptually it is simple; H(z) should be an approximation to $H_d(z)$, but because you are using only a finite number of impulse response samples, it cannot be exact. But we shall show that the frequency response at the DFT points are identical for both $H_1(z)$ and $H_d(z)$.

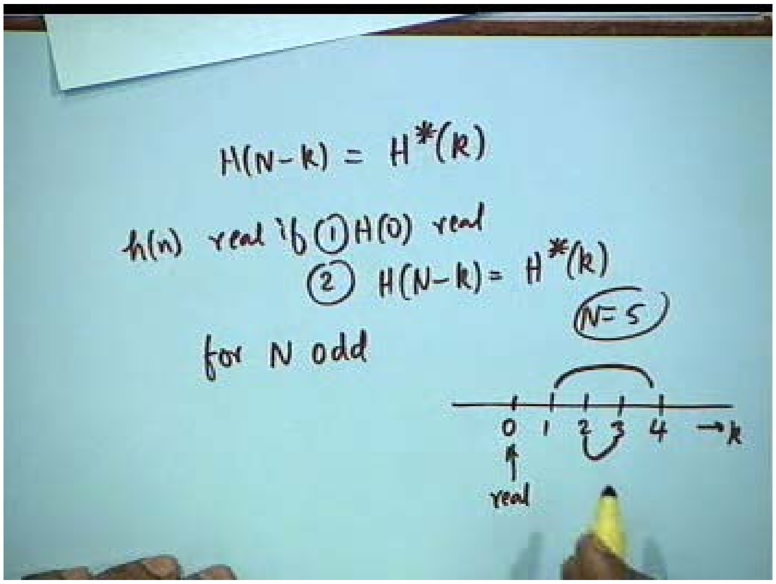
Note that I have used $H_1(z)$ instead of H(z) for the frequency response of h(n). Why? I have already used H(k) for the samples of $H_d(e^{j\omega})$ and using H(z) would have meant replacing k by z; this would have been wrong!

(Refer Slide Time: 06:46 – 10:14 min)



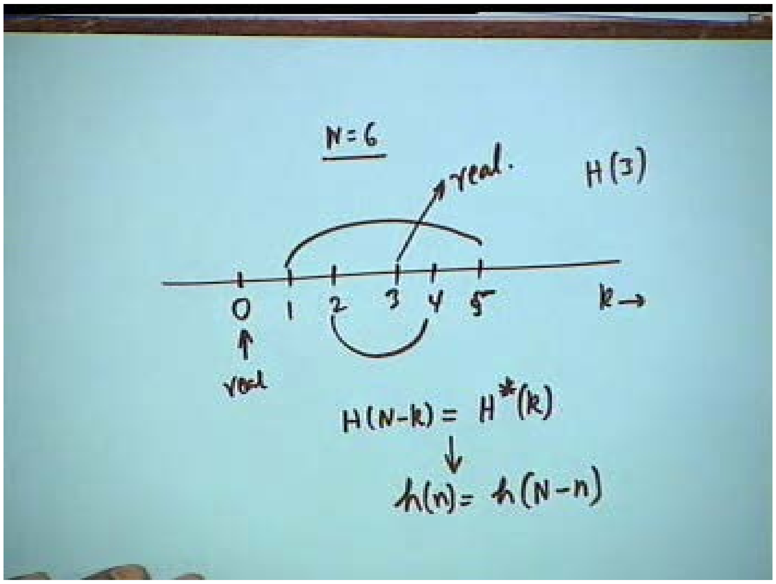
We find h(n) as the inverse DFT of the samples H(k). In other words, $h(n) = (1/N) \sum_{k=0}^{N-1} H(k) e^{+2\pi nk/N}$. It becomes complicated when we handle complex impulse response or complex coefficient digital filter. So in all probability h(n) has to be real; this is not guaranteed here because H(k) is complex, and the summation is not guaranteed to be real. But you notice that H(0) = h(0), which must be a real quantity. So, for h(n) to be real, one condition is that H(0) must be real. The other quantities are complex and if we can find a pair of them which are complex conjugates of each other, then we make sure that the sum would be real. Corresponding to $H(k) e^{j2\pi(nk)/N}$, we must have its conjugate in this summation, i.e. we must have a term, namely $H^*(k) e^{-j2\pi nk/N}$ which I can write as $H^*(k) e^{j2\pi(N-k)n/N}$.

(Refer Slide Time: 10:17 –11:44 min)



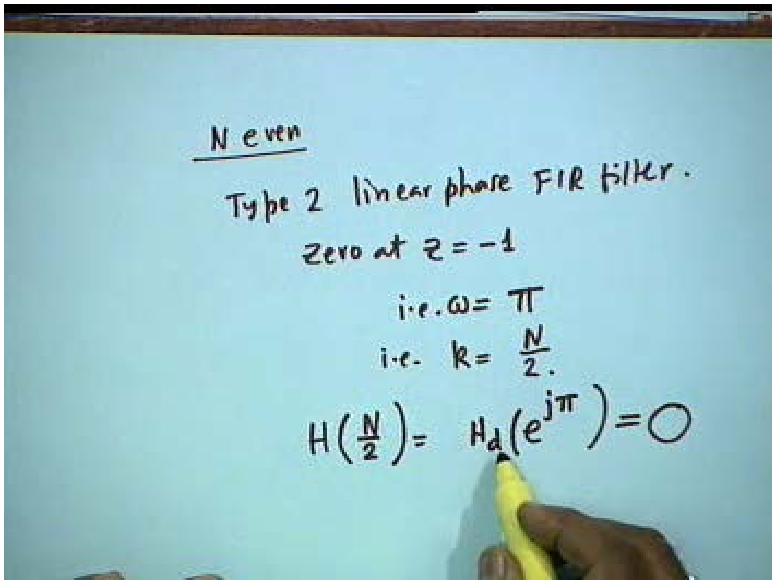
This will be the (N – k)th term provided H(N – k) = H*(k); then h(n) shall be real. Therefore the conditions for h(n) to be real are: H(0) real and H(N – k) = H*(k). Note that for these conditions to be satisfied, N must be odd. For example, if N = 5 then we have k = 0, 1, 2, 3 and 4, and we require H(0) = real, H(1)* = H(5 – 1) = H(4) and H(2)* = H(3).

(Refer Slide Time: 11:46 – 13:46 min)



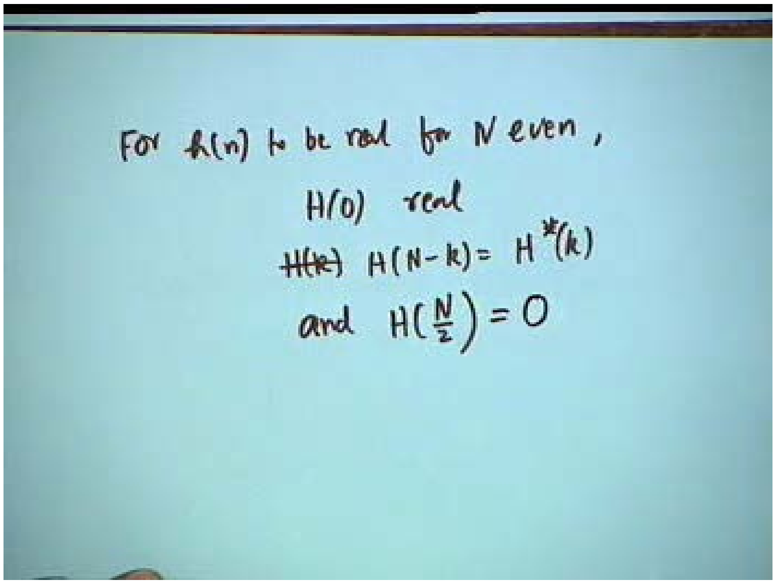
However, if I take N as even, say N = 6, then we require H(0) to be real, H(1)* = H(5) and H(2)* = H(4). What about H(3)? Should it be 0? Not necessarily. It should be real. The condition that H(N – k) = H*(k), i.e. frequency domain complex conjugation symmetry, reflects in the time domain as the symmetry of the impulse response coefficients (and you can very easily prove this), i.e. h(n) = h(N – n). Therefore conjugate symmetry in frequency domain corresponds to linear phase FIR filter which we do require. This corresponds to a linear phase FIR filter. What kind of linear phase FIR filter is this? It is symmetric even length, so it is type 2.

(Refer Slide Time: 13:49 – 15:39 min)



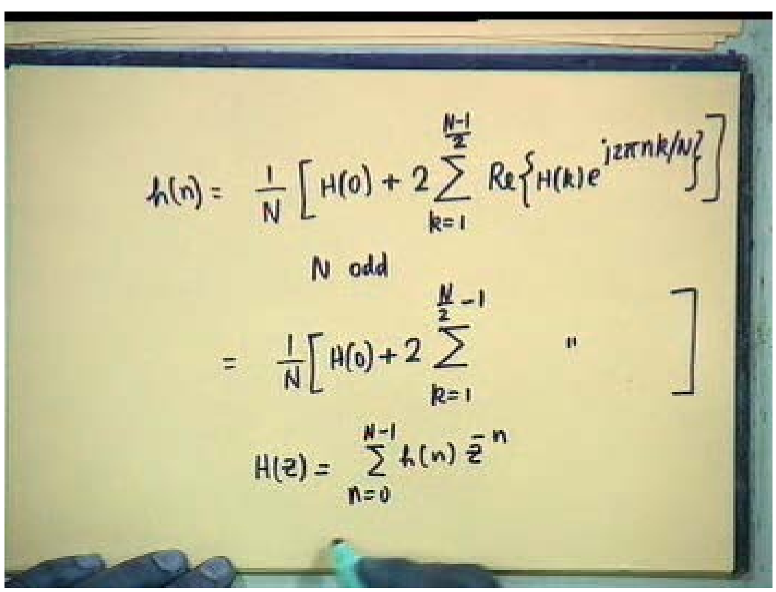
Therefore for N = even, we get a type 2 linear phase FIR filter. And a type 2 linear phase FIR filter has a zero at z = – 1, that is ω = π, and it cannot therefore be used for high pass (and bandstop) filters. In general, for N even, we should have H(N/2) = $H_d(e^{j\pi})$ = 0. Why zero and not any other real quantity? Because H(z) has a zero at z = – 1 i.e. ω = π.

(Refer Slide Time: 15:43 – 16:28 min)



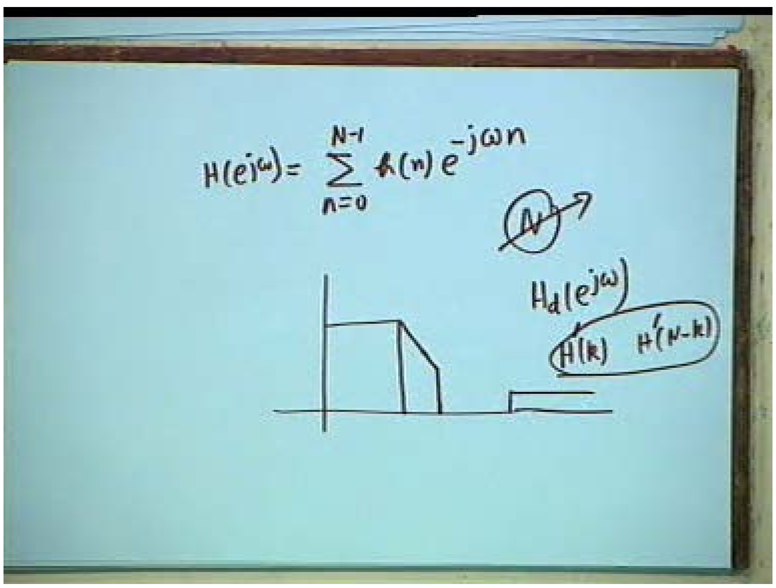
To repeat, the sample in the frequency domain which corresponds to $H_d(e^{j\pi})$ must be 0 and therefore for h(n) to be real, for N even the conditions are H(0) real, an arbitrary value, H(N – k) = H*(k) and H(N/2) = 0. Of course, if it was not linear phase, you can allow H(N/2) to be non-zero, but real.

(Refer Slide Time: 16:39 – 18:44 min)



To find an expression for h(n), note that if there are complex conjugate pairs then obviously the sum of them would be twice the real part of one of them. So h(n) would be h(0) plus twice summation real part of H(k) $e^{j2\pi nk/N}$. The limits of the summation would depend upon whether N is even or odd. If N is odd, the limits should be from k = 1 to (N – 1)/2. There is perfect matching between the corresponding values at k = 1 and N – 1, at k = 2 and N – 2 and so on and there is no middle sample. On the other hand, if N is even, the limit shall be from k = 1 to (N/2) – 1, the (N/2)th term being zero. Once you have obtained h(n), then the next step is to obtain $H(z) = \sum_{n=0}^{N-1} h(n)\, z^{-n}$.

(Refer Slide Time: 18:45 – 24:35 min)



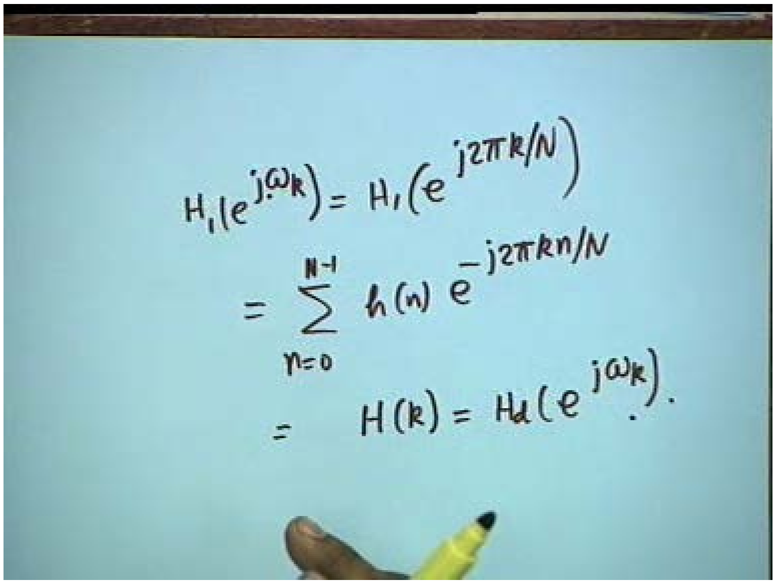
Next, calculate $H(e^{j\omega}) = \sum_{n=0}^{N-1} h(n)\, e^{-j\omega n}$ and then plot it to see whether the tolerances are satisfied or not. As far as the number of points is concerned, it should be governed by the same formula that was used for windowing design of low pass and high pass FIR filters. Formula for band pass and band stop shall be looked into later.

If with the chosen N, the design satisfies the tolerances, then you are lucky. More often than not, the tolerances will not be met. In that case, what are the things that you can do? Remember that the formula for N is only an estimate. You have control over N and you can increase this. As in windowing technique, since N cannot be accurately estimated, generally start from an ideal filter H_d, i.e. we want abrupt transitions from pass band to stop band. There is no guarantee that increasing N will satisfy the specs because of the nuisance of Gibbs Phenomena, which leads to overshoots and undershoots. One of the ways is, not to allow abrupt transitions. In other words, you allow a transition region, make the transition smoother rather than making it abrupt. As an example, you could make the transition linear. You could make it smoother than linear by assuming a Butterworth characteristic. You could also assume it to have a Chebyshev characteristic. In other words, if the obtained $H(e^{j\omega})$ on the basis of estimated N does not satisfy the specifications, then change N or change the nature of the $H_d(e^{j\omega})$. And go on doing this iteratively till you get what you want.

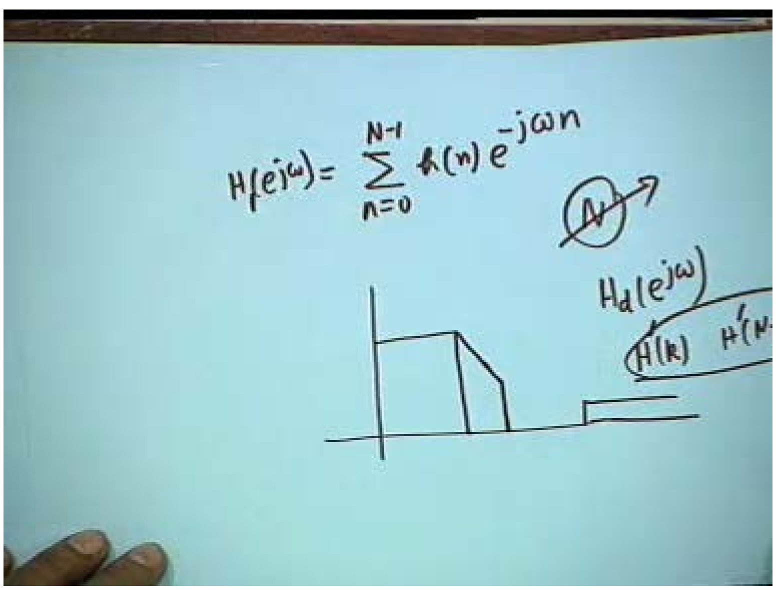
IIR design is much more elegant but FIR design always has an uncertainty factor. The other thing you can do when you are very close to the specification, is to pick up one h(n) and perturb it. That is, change h(k) to some h′(k). There is no deviation from linear phase because you are obtaining a conjugate match in the frequency domain. This arbitrary, but small change may bring the obtained characteristic closer to what you want. So there are three choices: change N, change the shape of the desired characteristic or pick up one or more h(n)'s and perturb it or them to obtain the desired tolerances. One or a combination of them should work.

Frequency sampling design is not as complicated as it appears. There are standard programs available on how to proceed. On the monitor screen, you shall see the obtained frequency response.

(Refer Slide Time: 24:41 –26:28 min)

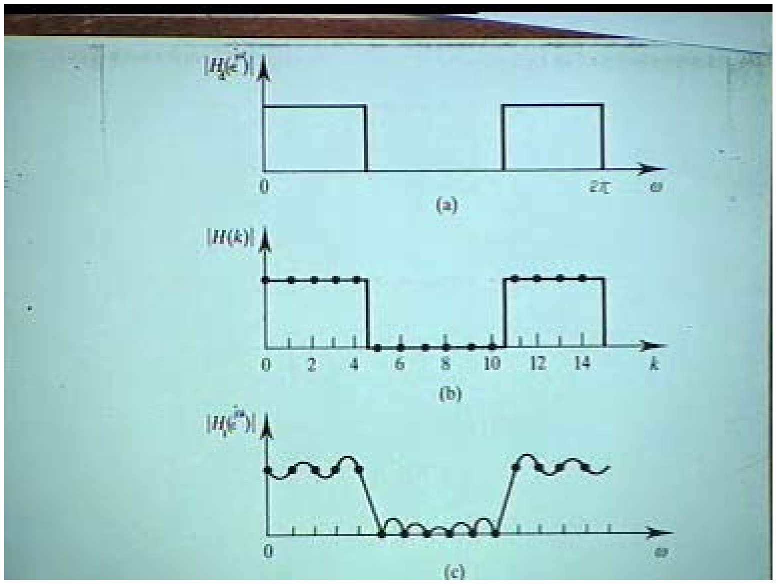


(Refer Slide Time: 24:51– 24:54)



Notice that H(k) is the same as $H_1(e^{j\omega_k}) = \sum_{n=0}^{N-1}$ h(n) $e^{-j2\pi kn/N}$. Therefore the obtained frequency response is exactly equal to the desired frequency response at the sampling points. At intermediate points, $H_1(e^{j\omega})$ is an interpolation, and there shall be deviations. What one has to ensure is that this deviation does not exceed the tolerance limits. A typical plot is like that shown in the next diagram.

(Refer Slide Time: 26:36 – 29:45 min)

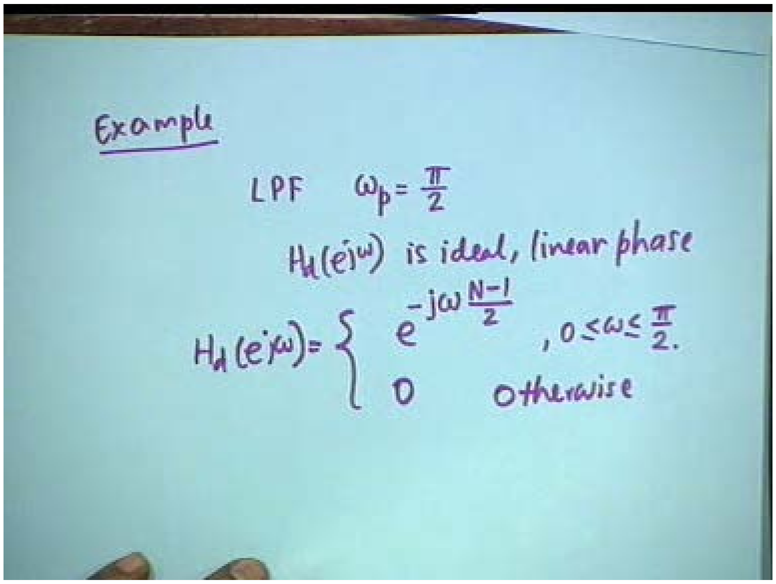


This is a case of design of low pass filter; we have taken the frequency response from 0 to 2π and we have aimed for an ideal magnitude which has a brick wall shape. One samples it at equal intervals. It stops at k = 14; k = 15 corresponding to ω = 2π. Therefore the length N, is 15. Once you obtain $H_1(e^{j\omega})$ from these 15 points, notice that the obtained frequency response exactly matches $H_d(e^{j\omega})$ at these sampling points.

In between, there are overshoots and undershoots. In the stop band also, there are variations and therefore what you have to check is the tolerance. Does it exceed the tolerance limits? Naturally, you can say that if N is increased there is a better possibility of satisfying the tolerance scheme. But again there is no guarantee; you have to test it at every N.

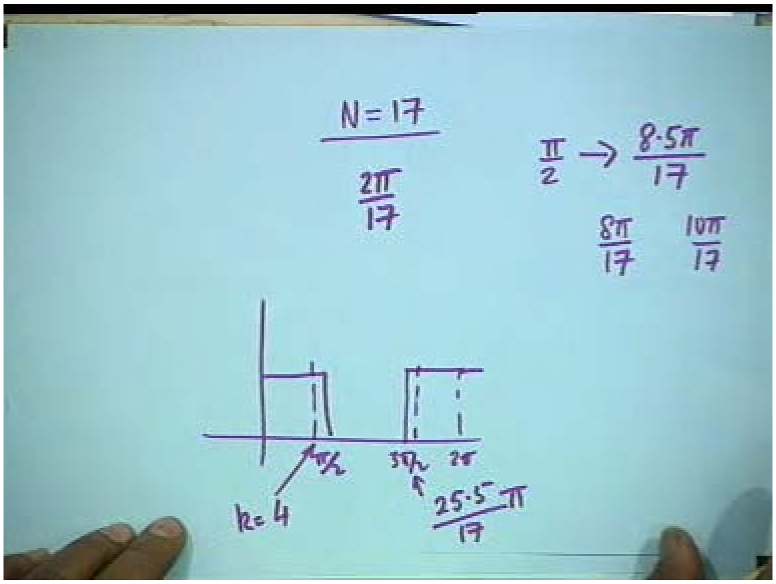
Even if you bring two points closer; the undershoot may be higher or overshoot may be higher. This is an uncertainty and you shall have to plot it. Your digital computer does not calculate a continuum of values. In plotting $H_1(e^{j\omega})$ you must take a large number of points, at least ten times N; this is a rule of thumb. Even if you take ten times N, in between two points, there will be an excursion. Here is an example.

(Refer Slide Time: 29:50 – 31:42 min)



We will aim at a digital low pass filter with $\omega_p = \pi/2$; assume that $H_d(e^{j\omega})$ has an ideal linear phase characteristic, which means that $H_d(e^{j\omega}) = e^{-j\omega\tau}$, where $\tau = (N - 1)/2$, for $0 \le \omega \le \pi/2$, the end of the pass band and is 0 otherwise. Now let us see how the samples are taken. Obviously in these samples you need to take the magnitude and also the phase. That is, H(k) in general shall be complex but because of this ideal assumption the magnitude is 1 in the pass band. It is angle which we shall have to take care. One has to be very careful in this angle because you are taking angle from 0 to 2π. When the angle from 0 to 2π is taken the chances of mistakes are very high. As far as the magnitude is concerned there is no problem.

(Refer Slide Time: 32:15)

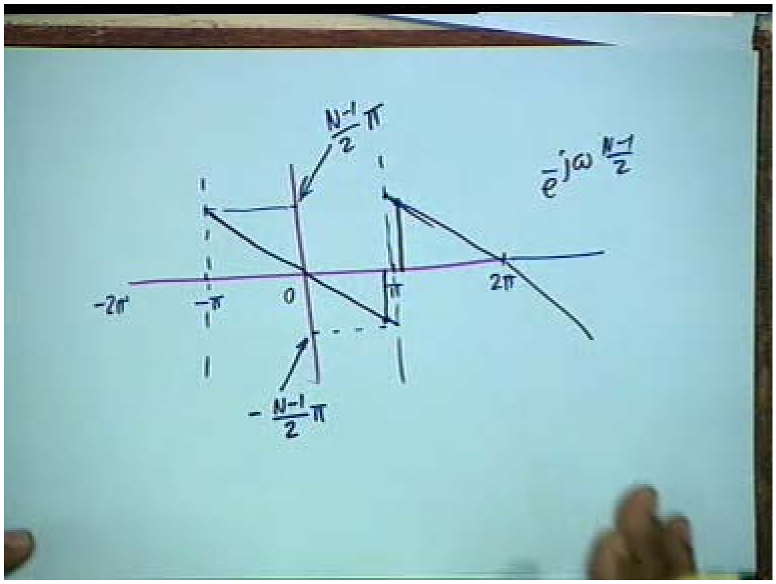


For example, if N = 17, an odd number, then what does π/2 correspond to? The samples are at multiples of 2π/17; so π/2 corresponds to 8.5π/17, where no sample exists. The last sample in the pass band would be at 8π/17. The next sample is at 10π/17, which is in the stop band.

There are two pass bands now because you are going from 0 to 2π. Where will be the other pass band? It would be from 3π/2 to 2π. 3π/2 would correspond to 25.5/17π. The sample at 25π/17 is in the stopband, so the first sample in the second passband is at 26π/17.

The values of k in the first passband range from 0 to 4, and in the second passband it ranges from 13 to 16. Now what about the phase at these points? The magnitude is 1 at all these points, and since you are asking for linear phase it is guaranteed that H(k)* shall be H(N – k). Therefore in the computation, you need not go beyond k = 4. In general, you go up to (N – 1)/2 which equals 8 here but k = 5 to 8 fall in the stopband, where the magnitude is zero. So you only have to go from k = 1 to k = 4. Now face the problem with the phase.

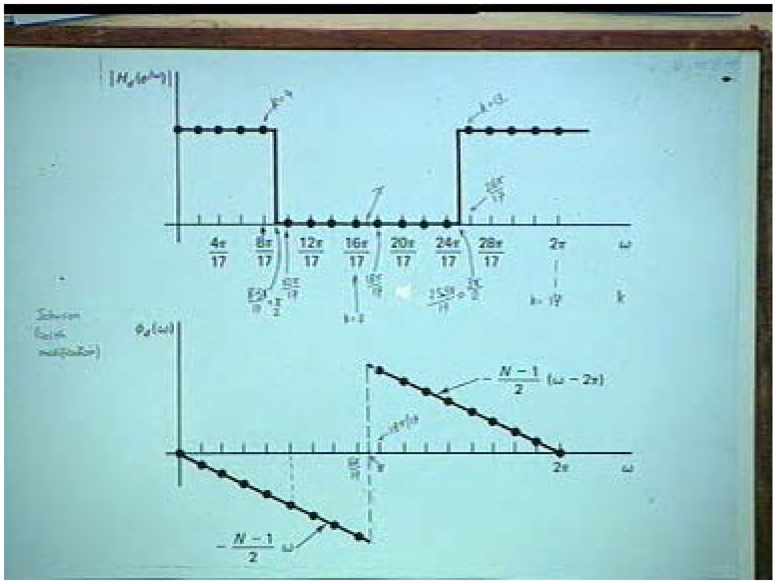
(Refer Slide Time: 35:54– 38:42 min)



As far as the phase is concerned, a little thought will show that it will go from 0 to – ((N – 1)/2)π at ω = π in a linear manner.

The phase is an odd function, and $H_d(e^{j\omega})$ is periodic, which means that its magnitude is periodic and its phase is also periodic. So at ω = π, the phase must jump to [(N – 1)/2]π, and then drop linearly to 0 at ω = 2π. There is a phase jump of (N – 1) times π at ω = π. This is the point that has to be remembered carefully. You had a sample close to π on both sides. Exactly at π, there was no sample. Fortunately, you do not have to go upto π because in our example, $\omega_p = \pi/2$. So you have to go upto ω = π/2 but you require the phases beyond ω = π to verify whether the samples are complex conjugates of each other or not. The diagram for the particular case N = 17 is shown in the next figure.

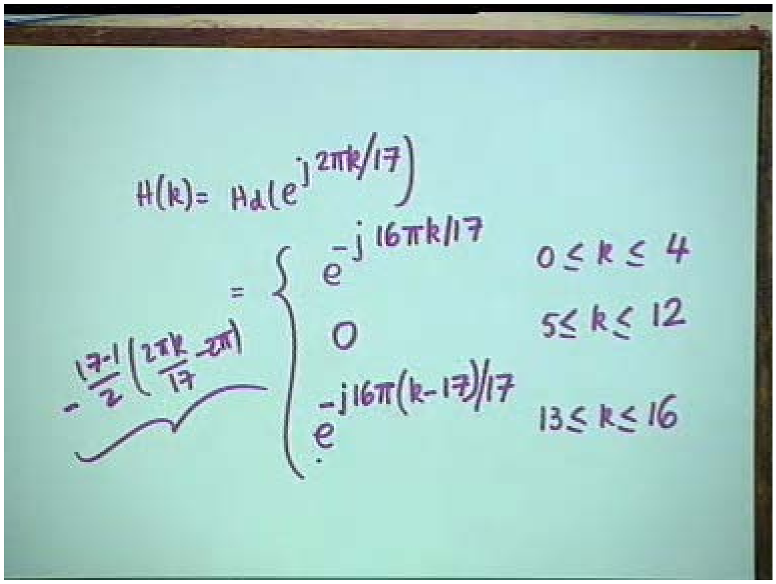
(Refer Slide Time: 38:50–41:27 min)



π/2 corresponds to 8.5π/17 and there is no sample there; the previous sample is at 8π/17 which means that your pass band goes from k = 0 to k = 4. The rest is stop band. The middle point is π where the phase transition occurs. And π corresponds to 17π/17 so it is exactly midway between k = 8 and k = 9. the sample below π is at k = 8 and that above π is at k = 9.

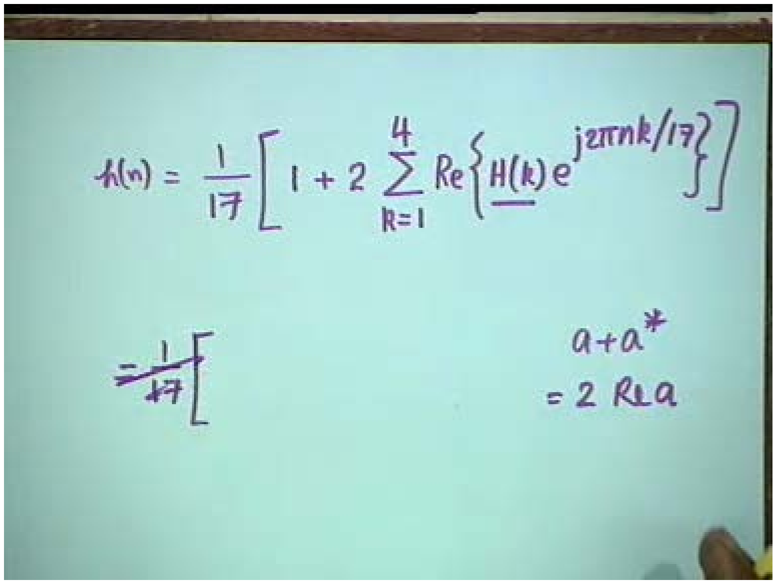
And if the phase is – [(N – 1)/2]ω between 0 and π, then the phase must be – [(N – 1)/2] (2π – ω) between π and 2π so that at ω = π, we get the value [(N – 1)/2]π and the value 0 at ω = 2π. It is a different matter that in the stop band, you do not have to consider phase, but you must construct the phase taking care of discontinuity, particularly for band pass and band stop filters. It is extremely important because for a band pass you will not have two pass bands; you will have four pass bands, and you shall have to take values in between. Therefore you must construct the phase plot carefully.

(Refer Slide Time: 41:31 – 43:38 min)



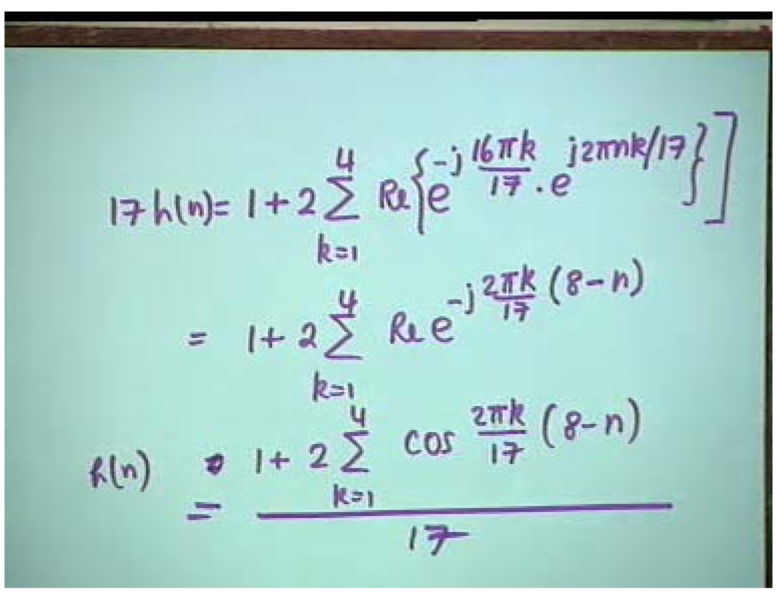
Therefore for H(k) k = 0 to 16, there are three ranges that we have to specify: the first pass band, the stop band and the second pass band. H(k) would be $e^{-j16\pi k/17}$ between 0 and 4. Then it is 0 from 5 to 12, and between 13 and 16, the value is $e^{-j16\pi(k-17)/17}$. You can verify that H(k) and H(N – k) are complex conjugates of each other; otherwise, you will not get real h(n). Also it is automatically guaranteed that we get a linear phase solution.

(Refer Slide Time: 43:45 – 45:42 min)



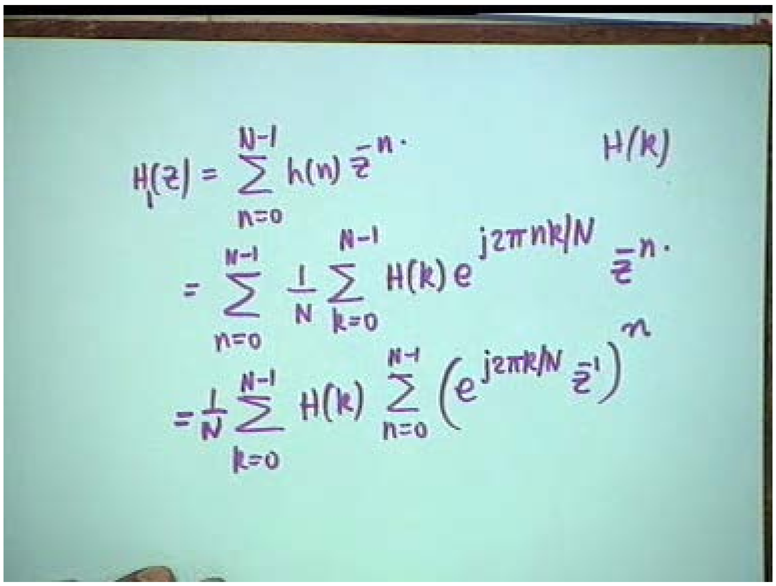
Once you have obtained the samples, then how do you obtain h(n)? h(n) is 1 for n = 0. For n $\neq$ 0, you shall have to go up to 8, but the magnitudes are 0 beyond n = 4. This means that you have to go only up to 4. The H(k) term is combined with H(N – k). The sum is twice the real part of either term.

(Refer Slide Time: 45:43 –47:26 min)



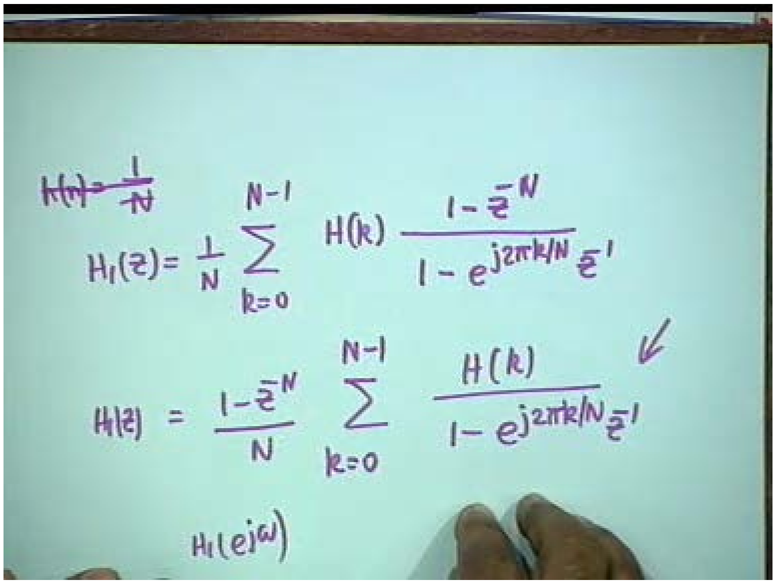
The result can be put in the form 17 h(n) = 1 + 2 $\sum^{4}_{k=1}$ real part of ($e^{-j16\pi k/17} \times e^{j2\pi nk/17}$) = 1 + 2 $\sum^{4}_{k=1}$ cos(2πk/17) (8 – n). Now you can calculate h(n). How many points? You have to calculate n = 0 to 16. You must calculate 17 points and then do the frequency response. Do not worry about the points at which you took the samples and do not compute them because they will be exactly identical; this is guaranteed by design so you compute in between.

(Refer Slide Time: 48:29 – 50:35 min)



You can simplify the expression for the transfer function $H_1(z)$ as follows. You can of course calculate h(n) and then find summation h(n) z^{-n}, n = 0 to N – 1. There is a simpler alternative where you do not have to calculate h(n). You can calculate $H_1(z)$ directly from H(k). $H_1(z) = \sum_{n=0}^{N-1}$ h(n) z^{-n}; replace h(n) by the inverse DFT of H(k). Then $H_1(z) = \sum_{n=0}^{N-1}$ (1/N) $\sum_{k=0}^{N-1}$ $H(k)e^{j2\pi nk/N}$ z^{-n}. Now the summations n and k are interchangeable because the variables are different; so bring this summation k = 0 to N – 1 at the beginning and then write the rest, which is (1/N) summation (n = 0 to N – 1) $(e^{j2\pi k/N}\, z^{-1})^n$.

(Refer Slide Time: 50:38 – 56:36 min)



This is a geometric series and the result is $H_1(z) = 1/N\sum_{k=0}^{(N-1)} H(k). (1 - z^{-N})/[1 - (e^{j2\pi k/N} \times z^{-1})]$. $1 - z^N$ is independent of k therefore you can bring this out and get $H_1(z) = [(1 - z^{-N})/N] \sum H(k)/(1 - e^{j2\pi k/N} z^{-1})$, $k = 0$ to $N - 1$. In order to calculate $H_1(^{ej\omega})$, you can simply use this formula directly.

We shall simplify this formula further. That is, the frequency response $H_1(e^{j\omega})$ can be simplified further so that we do not have to handle complex numbers. This is linear phase, it is guaranteed to be so; so you can take the linear phase term $e^{-j\omega(N-1)/2}$ out and the rest would be a real quantity which will be its pseudo magnitude. And it is the pseudo magnitude that you have to compute. So the computations would be further simplified.

The expression that we have got here is a parallel decomposition of the FIR filter, each block being recursive. Now where are the poles of each term? They are on the unit circle. You cannot realize a filter with a pole on the unit circle because it will be unstable and therefore it does not give you a simpler realization of the filter. But you know it is linear phase and therefore half of number of delays shall be required. The direct method or cascade method can be used. It is not a parallel recursive realization of $H_1(z)$; earlier people who worked on it decided to bring the poles slightly inside the unit circle and then realize it.

We have taken this formula only for computation purposes. We shall simplify this formula next time and show that $H_1(e^{j\omega})$ computation simply means computation of a real quantity which can be either positive or negative.