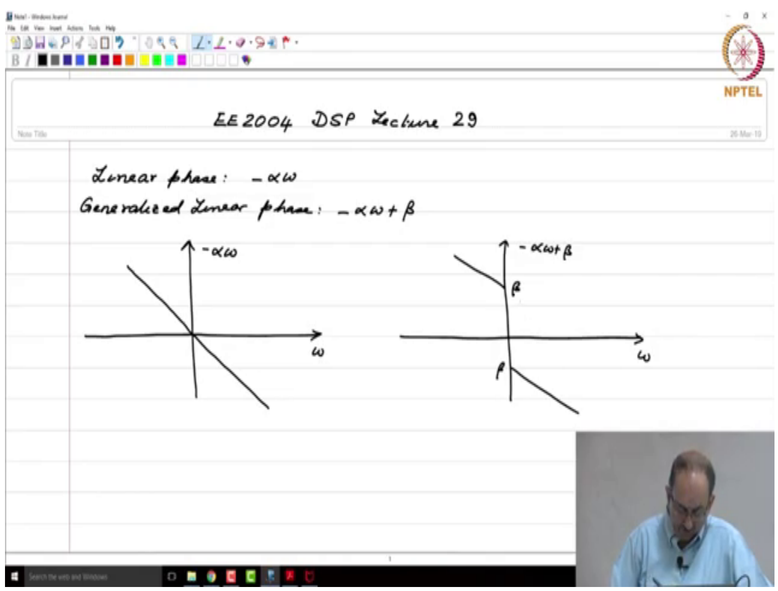
Digital Signal Processing
Prof. C.S. Ramalingam
Department Electrical Engineering
Indian Institute of Technology, Madras

Lecture 62:
Linear Phase (2)
-FIR filters with anti-symmetry
-How the eye and ear react to changes in phase

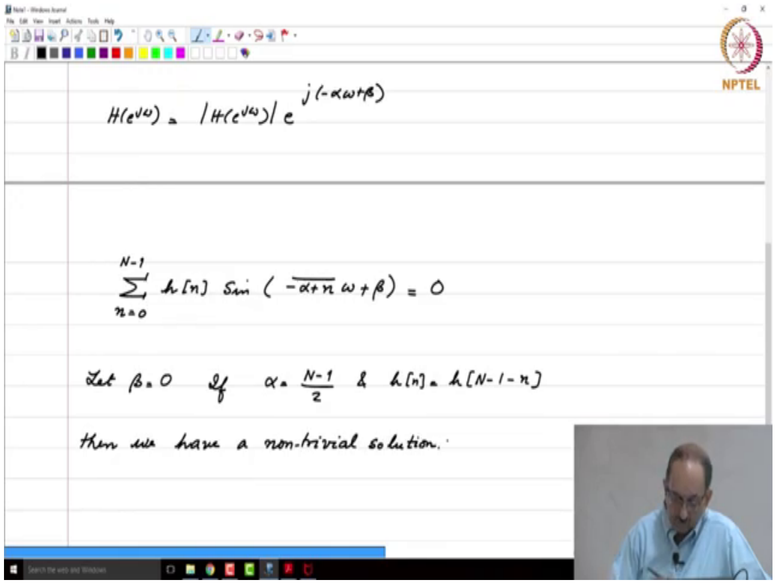
So, we are looking at Linear Phase filters. We saw that linear phases needed, if we want to preserve wave shape.

(Refer Slide Time: 00:33)



And, we saw two types of linear phase response. One was the regular linear phase, which was $-\alpha\omega$. Then we saw generalized linear phase which was $-\alpha\omega + \beta$. And if β where 0, generalized linear phase defaults to linear phase and the picture associated with these two types is this. So, this is $-\alpha\omega+\beta$. And then we saw that, by imposing this linear phase on the frequency response, i.e., we forced the frequency response to have this value and then this in turn gave rise to this condition; i.e., $\sum_{n=0}^{N-1} h[n] \sin\left(\overline{-\alpha+n}\,\omega\right)$.

(Refer Slide Time: 01:39)

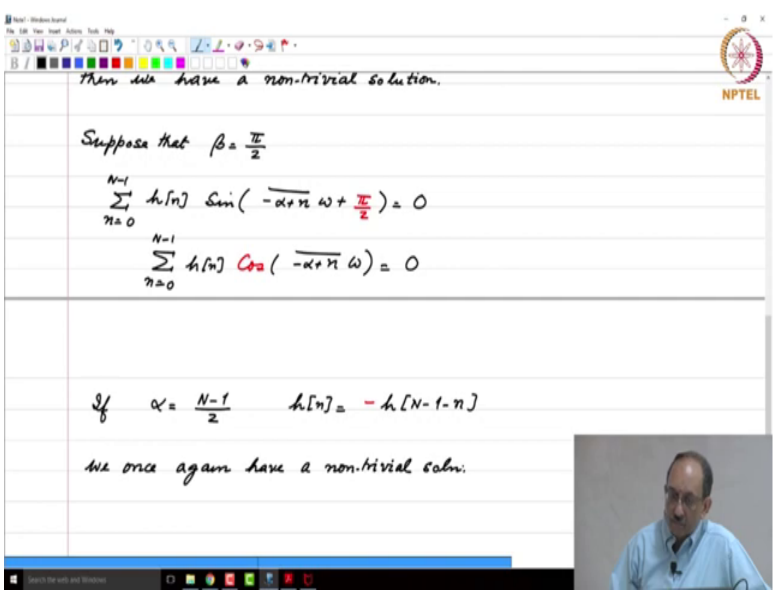


This was the condition that had to be satisfied, if we assume the phase is generalized linear phase of the form $-\alpha\omega + \beta$ and of course, this has the trivial solution $h[n] = 0$. That is mathematically correct, but practically useless.

To simplify this problem, rather than having an arbitrary n over which we have to sum. We restricted the filter to be FIR, causal FIR, so therefore, this became 0 to $N - 1$. And in trying to get a non-trivial solution, we further assumed $\beta = 0$ and then we saw that if $\alpha = (N - 1)/2$ and $h[n] = h[N - 1 - n]$, then we have a non-trivial solution. And then we saw simple examples where the impulse response was of length 4 and length 5 and we saw how pair wise cancellation happened there.

And we also made the remark that; this condition is true not only for $\beta = 0$, but also for $\beta = \pi$. Now, we left last class with the question, what about other values of β?

(Refer Slide Time: 04:11)



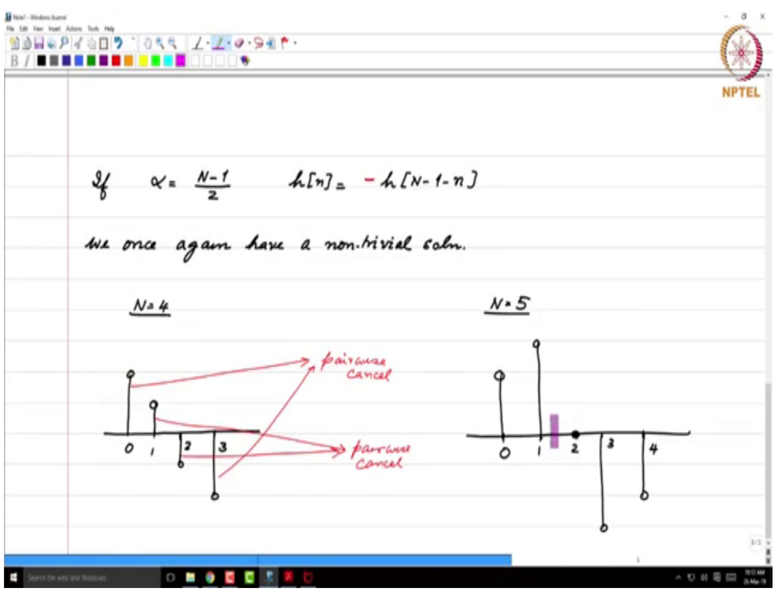
Now, suppose that $\beta = \pi/2$ therefore, we are trying to look for a non-trivial solution to this equation; $\sum_{n=0}^{N-1} h[n] \sin\left(\overline{-\alpha + n}\, \omega + \frac{\pi}{2}\right)$. So, this is the same as $\sum_{n=0}^{N-1} h[n](\;)$. So, $\sin(\theta + \pi/2) = \cos(\theta)$ therefore, this now becomes $\sum_{n=0}^{N-1} h[n] \cos\left(\overline{-\alpha + n}\, \omega\right)$. So, this is the equation for which we need to find the non-trivial solution.

Since, we have made considerable progress for the earlier case, we can invoke similar ideas here. Again, pair wise cancellations is what will work. And in the earlier case because $\sin(\theta)$ was an odd function so that $\sin(-\theta) = -\sin(\theta)$, we saw that you could get pair wise cancellation if $h[n] = h[N-1-n]$ where the impulse response had even symmetry.

Therefore, the argument changing sin, sin being an odd function, cancellation took place. Whereas, now we have an even function and therefore, it is easy to see that, if α is once again $(N-1)/2$, but now we have $h[n]$ not being equal to $h[N-1-n]$, but you need this to be anti-symmetric. You need to have a minus sign here because cos being an even function, $\cos(-\theta) = \cos(\theta)$.

So, that would not effect a change in the sign of the two terms and hence the change in the sign of the two terms has to be effected by change in the sign of the impulse response. So, if this condition is satisfied, we once again have a non-trivial solution.

(Refer Slide Time: 07:17)

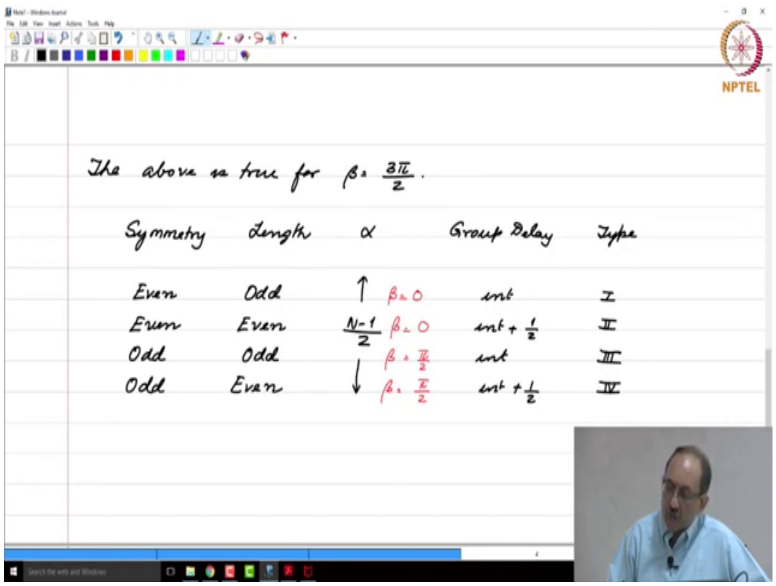


Again just to quickly get a feel for this. Let us look at two simple examples. Therefore, if you had say $N = 4$, you need odd symmetry. Therefore, this goes from $0, 1, 2$ to 3. You need odd symmetry and hence, if $h[0]$ were this, $h[3]$ has to be equal in magnitude, but opposite in sign and if $h[1]$ were like this, $h[2]$ will be like this. And as before, this and this pair wise cancel, similarly this and this pairwise cancel and once again that equation is satisfied.

Now, on the other hand, if you have now $N = 5$. So, we have $0, 1, 2, 3, 4$; 0 to $N-1$. So, if this were $h[1]$, $h[4]$ has to necessarily be like this. And if this were $h[1]$, $h[3]$ has to necessarily be like this. And remember, now we have an odd symmetric sequence and if you have an odd symmetric sequence about a certain point at that particular point, the sequence value has to necessarily be 0. Therefore, this term will always be 0.

Whereas, in the earlier case, independent of what this value was because you had this multiplying the sine function, that became 0. Whereas, now this term is going to multiply cosine and cos(0) is 1 therefore, for the product to be 0, this has to be 0. Again, everything is consistent. This is odd symmetric, this is the point about which you have symmetry and hence everything is consistent and satisfied.

(Refer Slide Time: 10:09)



And, just to complete the statement, regarding what other values of β for which this is valid, the above is true for β equal to, what other value of β for which this is true? $3\pi/2$. And now, we will just make this statement, we will not prove it. There is no other value of β for which you will get a non-trivial solution.

So, the only values of β for which you will have non-trivial solutions are, $\beta = 0$ and $\beta = \pi/2$ of course, $\beta = 0$ is also the same as $\beta = \pi$. Then $\pi/2$ is the same as $3\pi/2$, therefore, in effect we are only considering $\beta = 0$ and $\pi/2$. No other value of β exists for which you have a non-trivial solution for this set of equation. We will just make the statement and leave it at that.

Now, let us further get some more insight into this. Let us first summarize, what we have seen so far. Based on these simple examples, we have considered, we are able to see that we have considered the lengths of the filter to be either odd or even and the symmetry can also be either odd or even. When $\beta = 0$, you need even symmetry and for that even symmetry case, we had considered even length and odd length. Similarly, when $\beta = \pi/2$, again the symmetry was not even, but odd. You need symmetry, but in this case, the symmetry turns out to be odd.

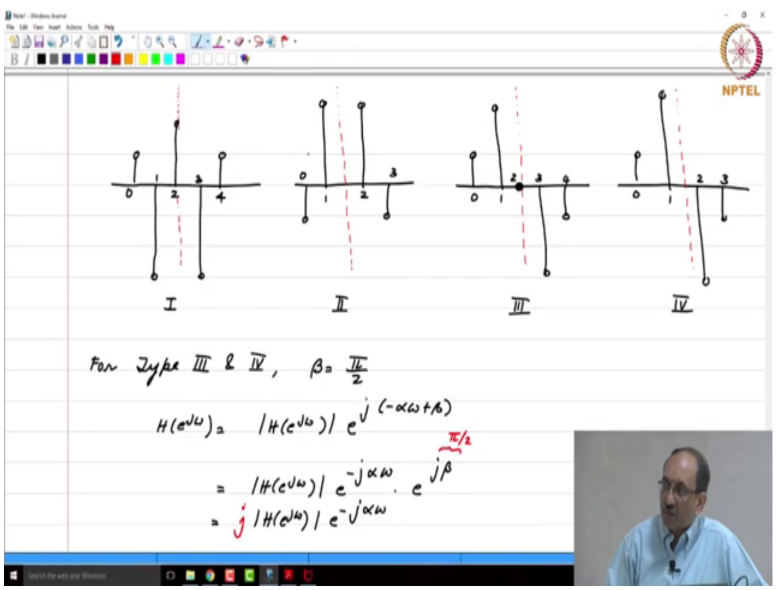
And once again, we consider two different cases; length even and length odd. Now let us summarize these two. Therefore, symmetry length, the value of α and then group delay. So, symmetry is even, length can either be odd or length can be even. The other case was symmetry can be odd length can be odd, symmetry is once again odd, but the length is now even. In all these cases, the group delay is $(N-1)/2$.

The first two cases correspond to $\beta = 0$, the second two cases correspond to $\beta = \pi/2$. And remember, the group delay also is α because, the phase response has the form $-\alpha\omega + \beta$ in general and negative derivative with respect to ω of the linear phase gives you α which is $(N-1)/2$. The important point here is now, if length were odd that is N were odd, then $(N-1)/2$ is an integer, therefore, group delays an integer.

On the other hand, if the length were even, $(N-1)/2 + (1/2)$. Same case for the odd symmetric filters as well. So, the important point to note here is, if you now have a signal and then pass it through a linear phase filter, if the length were even, the output will suffer a group delay that is integer number of samples. On the other hand, if the lengths were even, then the group delays integer plus half. Therefore, the output will have half sample delay. So, this is very important.

We will mention a little more about this half sample delay as we go along. These filters have been named along these lines. So, this is called type I filter, Type II, Type III, Type IV. So, these four filters have been named as types I, II, III and IV. Types I and II have even symmetry. Types III and IV have odd symmetry and types I and III have odd length, types II and IV have even length.

(Refer Slide Time: 16:17)



So, this is Type I, because the symmetry is even and the length is odd. And point of symmetry falls on a sample value. So, this sequence is symmetric about this particular sample and it falls on a sample value and it is easy to see the kind of point we are going to make for the other cases. So, this is Type II. So, Type II is again even symmetry and in this case, the point of symmetry is midway between samples. So, this is Type II.

Similarly, we can quickly complete the picture for Types III and IV. So, this is Type III and this is the point of symmetry. And Type IV is again odd symmetric. And the point of symmetry similar to Type II is an imaginary point in that, it falls midway between samples. And again things are consistent, as they should be. For Type III and IV, we obtained them by letting $\beta = \pi/2$ and recall that the frequency response can be written as $|H(e^{j\omega})|e^{(-\alpha\omega+\beta)}$.

And for Types III and IV, this is $e^{(-\alpha\omega)}e^{j\beta}$ but, β is really $\pi/2$. And therefore, this is really $|H(e^{j\omega})|e^{(-\alpha\omega)}e^{j\beta}$ and $e^{j\pi/2}$ is nothing but j. Now, why does this make sense? It makes sense because, the sequences in this particular case, real and odd therefore, the transform must be; one thing you can definitely say for sure about the transformer is, it must be imaginary.

And sure enough the $e^{j\pi/2}$ gives you the factor of j that is needed. So, everything is consistent. I do not know, whether this has struck you, this phase response is linear. And the linear, the slope is $-\alpha$ and the group delay is $+\alpha$. And $\alpha = (N-1)/2$. So, one thing that immediately jumps out from this fact is that, this phase response is dependent only on? phase response is dependent only on what?

Only on length, right. Phase response is linear, it is $-\alpha\omega + \beta$ and $\alpha = (N-1)/2$. And the phase response is purely dependent only on the length N and is independent of and its independent of?

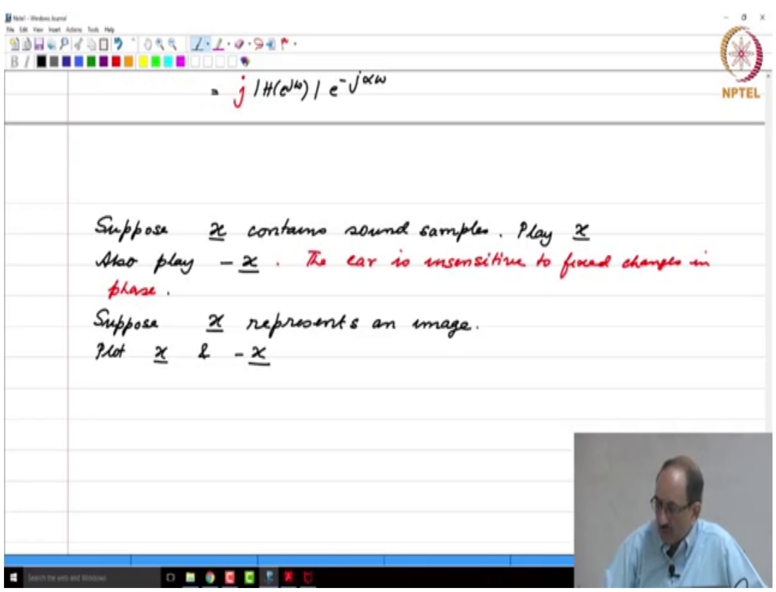
Student: (Refer Time: 21:54).

Very good, it is independent of $h[n]$. So, you can have random coefficients, as long as the length is same and you have symmetry, you will get exactly the same phase response. So, the only change that the change in $h[n]$ we will bring about is the change in the magnitude response. Therefore, $h[n]$ controls only the magnitude response. If this is symmetric either being symmetric or anti symmetric, the phase response is always $-\alpha\omega$ or $-\alpha\omega + \beta$ and the group delay is $(N-1)/2$.

Therefore, independent of $h[n]$, you have linear phase with that particular slope. Very easy exercise to spray in MATLAB. Just generate two or three filters with random coefficients, keep the lengths same and then form the entire filter with completing the given random sequence by either symmetry or anti symmetry.

And then, look at the phase response, they will be identical. And some general remarks can be made about the phase response with respect to how it affects us. If you have a signal, say in MATLAB if you store it as a vector.

(Refer Slide Time: 23:42)



So, suppose, x contains sound samples, you have a sentence or a piece of music or whatever and then you can play this in MATLAB. We have seen this command soundsc which scales the vector and then, if you give it the correct frequency sampling frequency, the audio port will play this waveform. So, you can play this vector, you can also play $-x$. You can multiply this vector by -1 and then play the file.

And remember, when you multiply by -1, you are effecting a phase change of 180 degrees.

And now, if you close your eyes and then somebody else plays these two files at random, x and $-x$. You will not be able to tell the difference. If you are able to tell the difference, you need to see a doctor. So, the ear is completely insensitive to fixed changes in phase. So, this is one aspect why is it so? We do not know, we are like that. Now have an image, you can plot the image in MATLAB.

So, x can now represent an image. Suppose, now x. So, let me just complete this. The ear is insensitive to fixed changes in phase. Again as an illustration, if you go to this *https://pages.jh.edu/*, which I had refer to when, I was talking about convolution. That has one applet called listen to the Fourier series.

So, what they have done is, they have taken a periodic square wave form which you can play and then they have broken it down into its Fourier series components. And then what they have done is, for each Fourier series component, they have randomized the phase. And then you can play the signal by reconstructing it after randomizing each phase of each harmonic. Again, if you close your eyes and listen, you will not be able to hear any difference. So, this is another illustration of the fact that the ear is insensitive to phase.

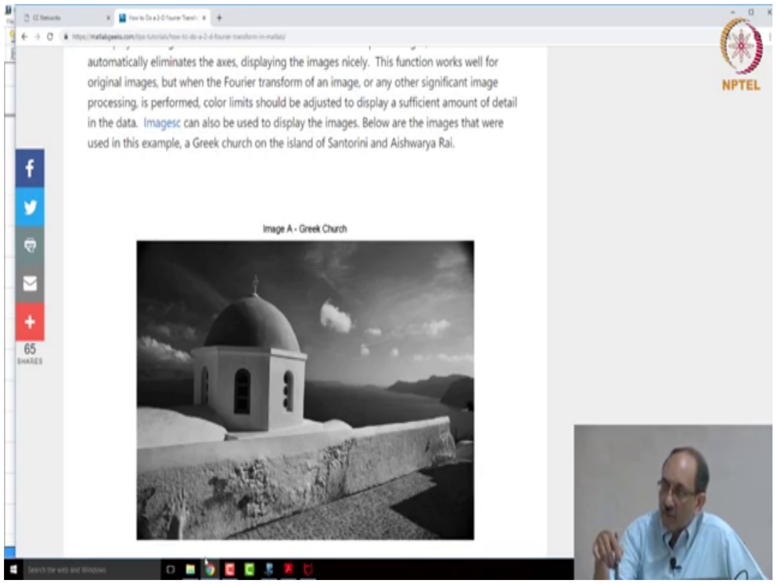
Now, suppose x represents an image. In this earlier case, if this were a sound sample, typically this will be a vector, 1D vector. Whereas, now if x represents an image, it will be a matrix, will be a 2D variable. Now you can plot x and $-x$. So, what do you think will happen?

Student: (Refer Time: 27:41).

Very good, so what you will see is, when you plot $-x$, you will see the negative image. So, eye is supremely sensitive to phase. The ear is completely insensitive, the eye is supremely sensitive to phase. Now this can be very dramatically also illustrated by this example. Even before we do this, this, listen to the Fourier series example; if you randomized the phase and if you close your eyes and listen, you will not be able to tell whether the original Fourier series expansion was played or the reconstructed waveform after randomizing the phase was played.

In that case, if you look at the waveform, the square wave that you start off with, when you randomized the waveform and reconstruct the signal, that signal will no longer be the square wave. It will have some shape. Depending on the phase, shape will be different. So, there you are looking at the waveform, with your eyes. So, clearly the change in phase is effecting the shape of the waveform and you are able to see that. But if you play the same thing on an audio device, you will not be able to hear the difference. So, this is how our systems are.

(Refer Slide Time: 29:11)



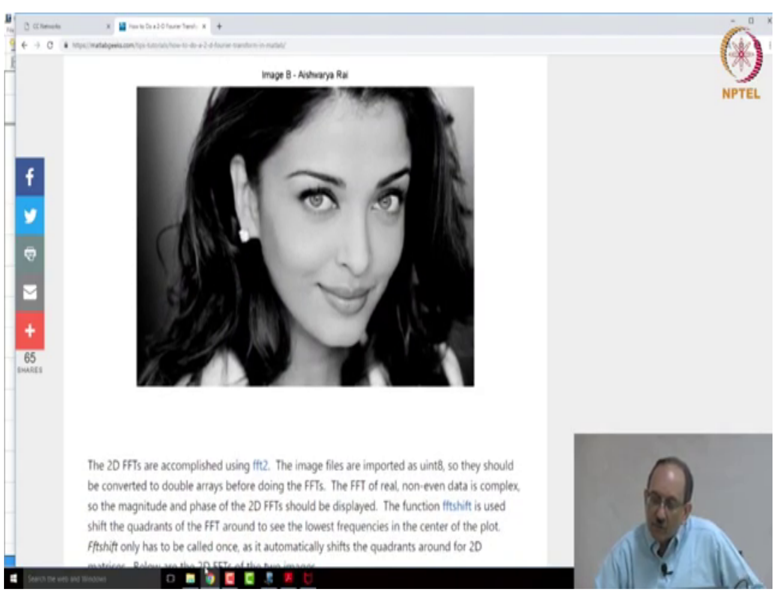
Now, let us look at this. So, I will just show you a couple of images. So, somebody has taken the trouble to show this. So, this is a Greek church. So, what we are going to do is, we are going to take the 2D Fourier transform of this, plot the magnitude and plot the phase. And then, what we will do is, we will take the magnitude of image A and then take the phase from image B and then combine them and then reconstruct the image and then we will see what happens.

So, you have two different images, we will take the magnitude from one image, phase from the other. So, you will form magnitude times $e^{j(\text{angle})}$, $e^{j(\text{angle})}$ comes from the second image and we take the inverse Fourier transform. You will get the image back and we will see what you will get, yes.

Student: Second image (Refer Time: 30:13) possible.

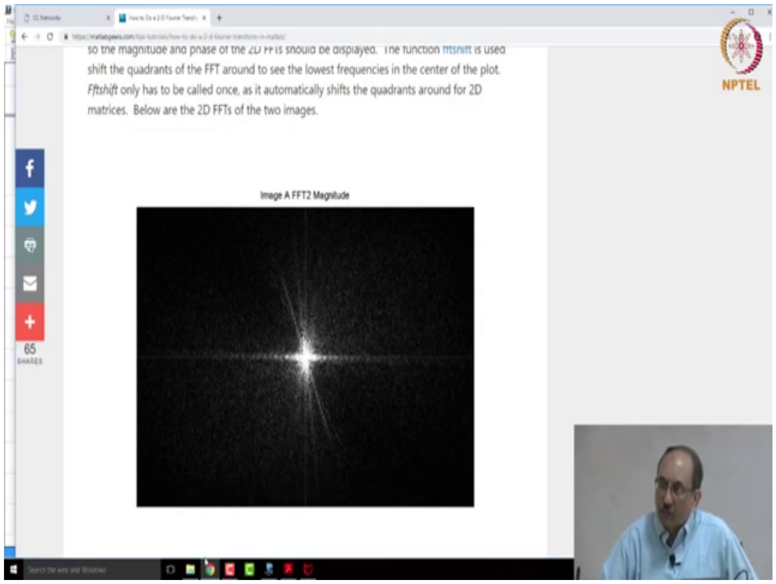
No second image is independent. So, this is Greek Church, and this is the image of a phase.

(Refer Slide Time: 30:24)



Somebody apparently called Aishwarya Rai.

(Refer Slide Time: 30:32)

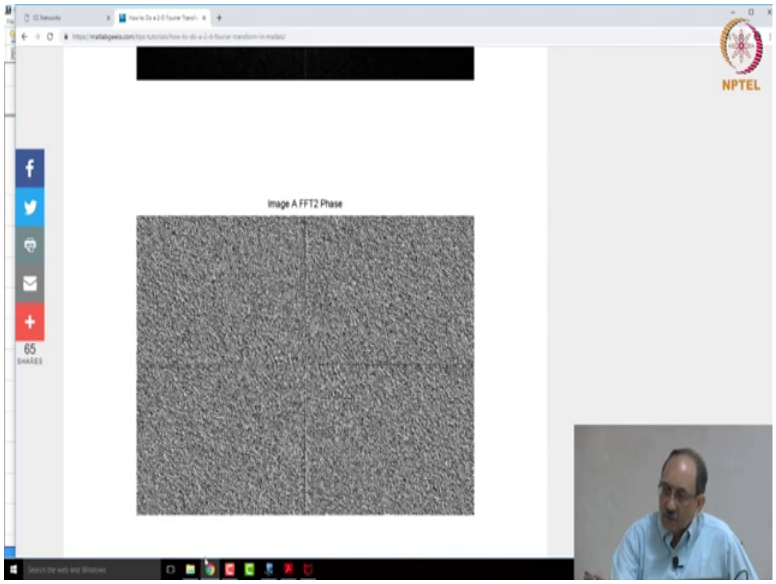


So, this is image A magnitude. So, remember, when you take a time domain signal, plot the 1D Fourier transform which is the DTFT, then that will give you the frequency distribution of the signal. That is, depending on whether the signal has low or high frequency or a mix, you will see those components showing up in those points in the frequency domain. So, if the signal were mainly low pass, then you will have components mainly around ω at 0.

If the signal were mainly high pass, you will have components, mainly around $\omega = \pi$. Now this is 2D Fourier transform. And again the x and y axis are indeed frequencies, but this is what is called spatial frequency. So, same thing will happen here, if you had a surface that had a sinusoidal variation and then if you took its 2D Fourier transform, you will see align component, only at that particular spatial frequency.

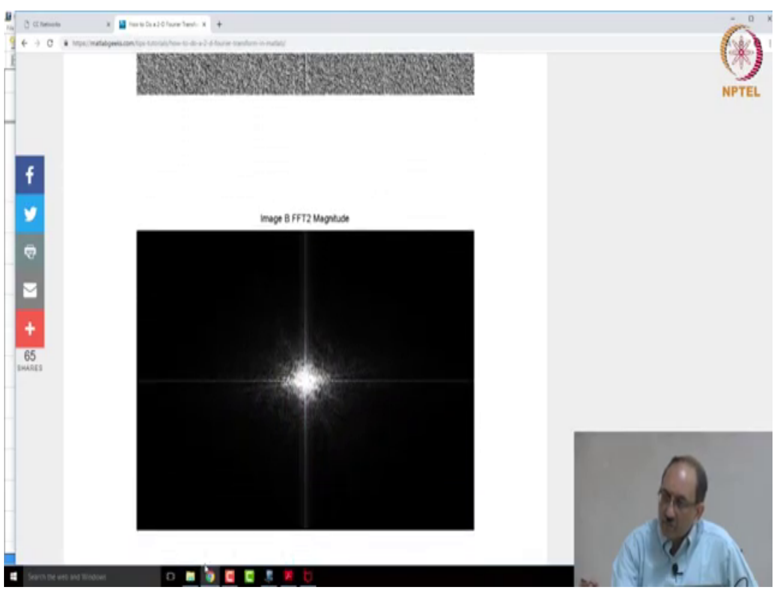
So, this corresponds to the x's spatial frequency, this corresponds to the y's spatial frequency. This point always will be $(0, 0)$. So, this the first quadrant, second, third and fourth. And you see this is the magnitude frequency response and mainly this is around $\omega = 0$. Therefore, the dominant components in this image have low frequency components.

(Refer Slide Time: 32:11)



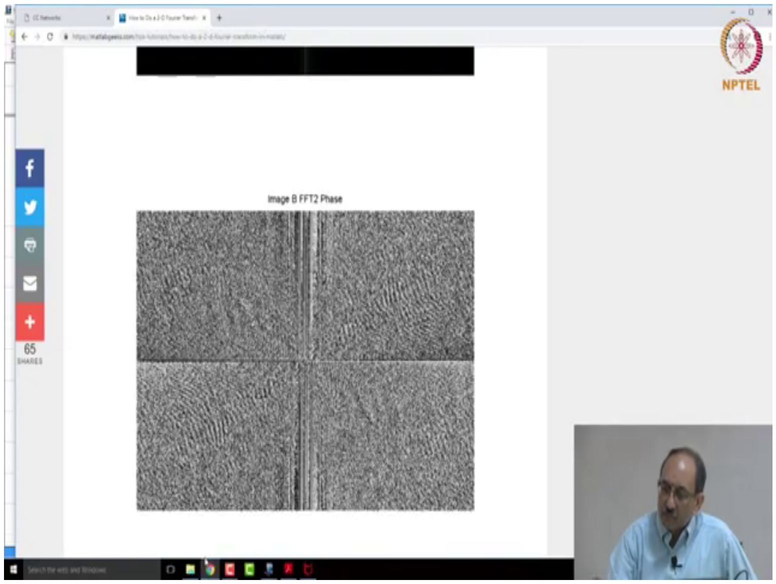
So, this is the magnitude response for the first image. And here he has only shown the phase response for image A. So, this is the phase response. Again, this looks like garbage, because this is probably the principal phase between $-\pi$ and π. Even if this were the unwrapped phase, I do not think, you will expect to see any particular pattern. So, this is image A phase.

(Refer Slide Time: 32:38)



And this is image B. So, this is phase as in p h a s e. So, this is image 2. Image 2 is the phase as in f a c e and this is the magnitude response. Again this is largely low frequency.

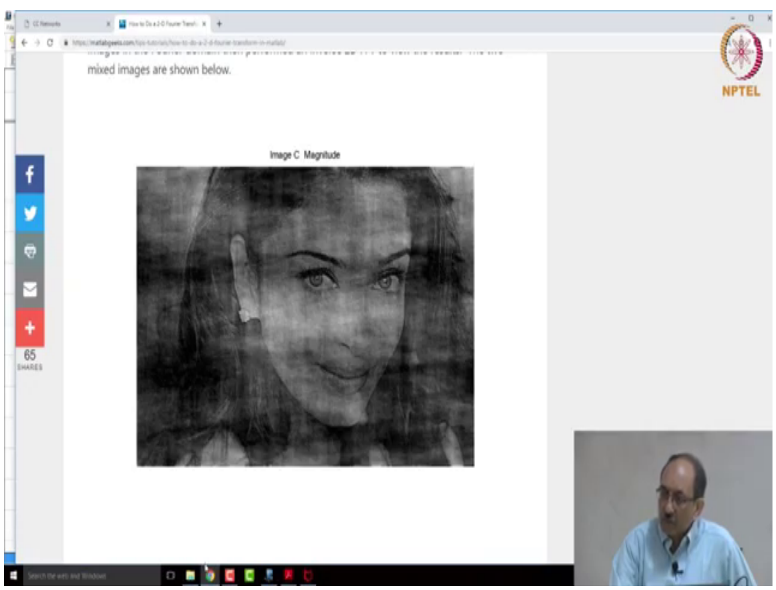
(Refer Slide Time: 32:55)



And this is the phase of the second image, again nothing much jumps out. Now what we will do is, we will take magnitude of one times $e^{j(\text{angle})}$ of the other image and then compute the inverse Fourier transform.

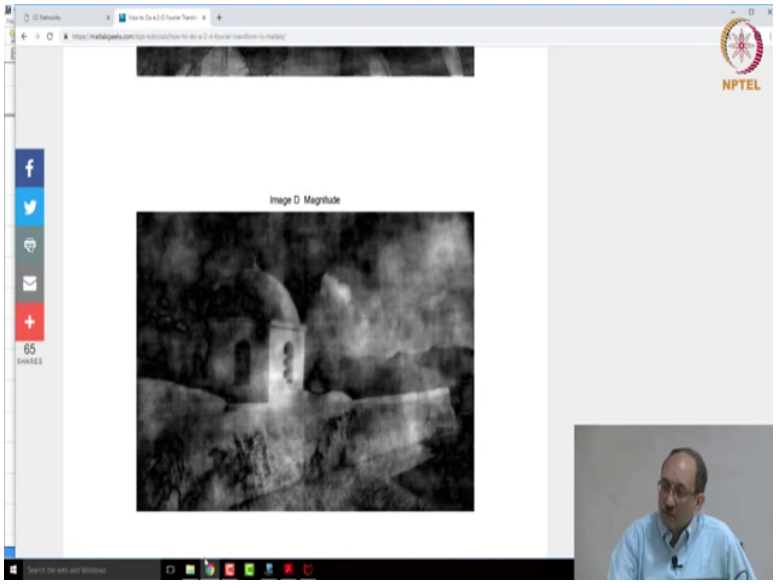
So, clearly it would not be a clean nice image, you can expect it to be distorted. But the question is, are you able to recognize it as is it similar to one or the other or it is a just random. Can be any one of these three possibilities.

(Refer Slide Time: 33:35)



So, this is the image C magnitude. And clearly you are able to recognize this as the phase as the second image, but notice that this has been obtained by taking the magnitude of the first image which is the church and the phase component is coming from the second image and it has been inverse Fourier transform and this what you see. And, clearly this is very close to the image from which the phase was taken from, p h a s e.

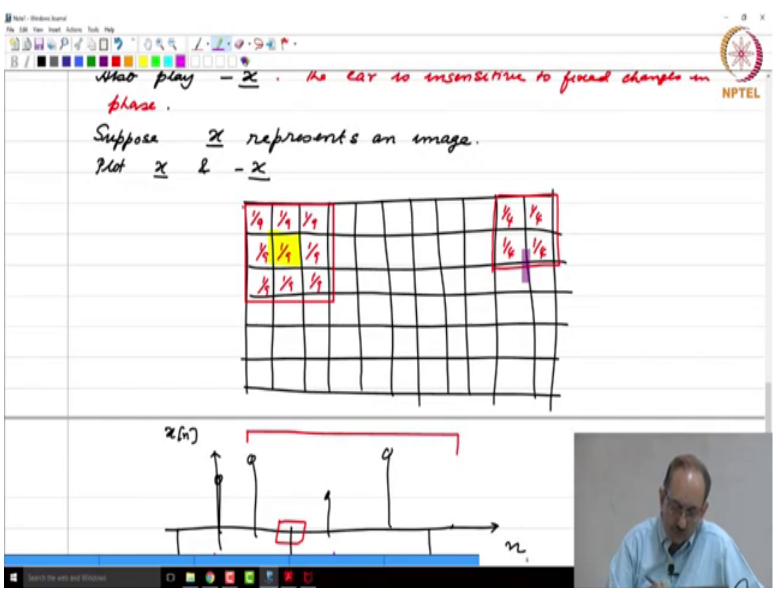
(Refer Slide Time: 34:12)



And this is the church image, but this has been obtained by taking the magnitude of the second image which is the face, f a c e and the phase p h a s e from the first image. And, clearly this looks much closer to the church. Therefore, this kind of tells you, how the eye is so sensitive to face. Whereas the ear is so completely insensitive, the eye is exact opposite, it is supremely sensitive to phase.

And this is one of the reasons why in image processing, you will always work with linear phase filters. Because you want to preserve wave shape and non-linear phase, non-linear phase and p h a s e, that will cause significant distortions which your eye can perceive. And hence, in image processing, you will always work with linear phase filters and filters in image processing are 2D filters.

(Refer Slide Time: 35:34)



And now, we can quickly get a feel for what this is. So, now, what I am plotting is, I am plotting pixels here. So, I have an image with pixel values. So, if this were a 256×256 image, then you will have in MATLAB, 256×256 matrix with entries. So, these are the pixel values, these are the pixel values. Now, suppose, I want to do low pass filtering and simple low pass filtering is just averaging.

In time domain, when you want to average, you will just take say 5 samples, add them and divide by 5. So, this is 5 point averaging and so on. Suppose, you want to do low pass filtering by averaging the pixels, then what you will do is, you will create a 2D low pass filter. Suppose you want to average pixels, what you will do is, you can do this. You can consider this 2D low pass filter and then, here this is a 3 by 3 filter.

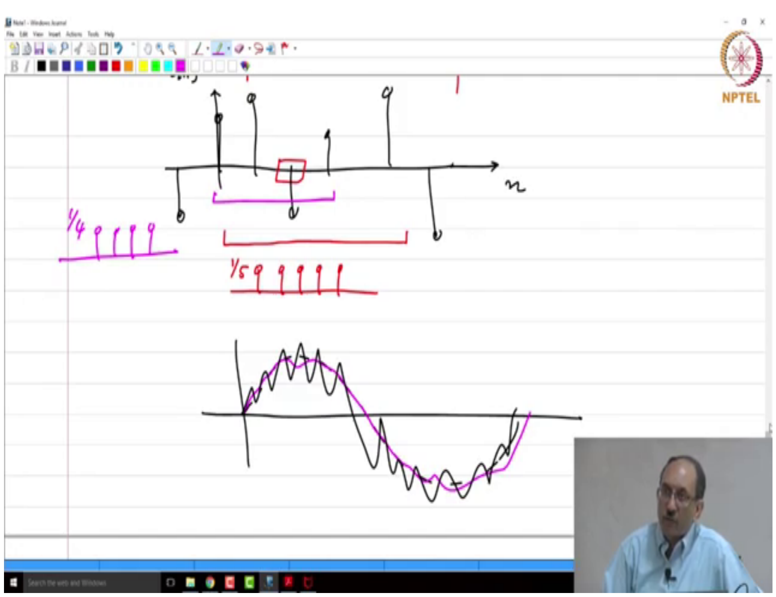
Suppose, the weight is all 1/9, right. Now, you want low pass filters. So, what you will do is. So, this is the filter, remember what do you do in, when you want to filter, you keep one signal fixed, take the other, flip it around and then slide it across. In the 1D case, you have to slide to the left and to the right to cover all possible shifts. In the 2D case, you have to shift left and right, up and down. So, now when you do this, you will take this, what is called a 3×3 mask.

So, you will take this up multiply this with this image and then basically what you do when you convolve the whatever filter is, you multiply sample by sample and then you add. So, when you multiply sample by sample and add, effectively what you are doing is, you are averaging all the 9 samples, because the weight is 1/9.

And now, what you will do is, you will move this mask one sample to the left and then you will move this like this and so on. And, then you will also move it down. So, when this mask covers all possible positions, you are done with low pass filtering. Now first of all, note that, there is symmetry in this filter. Therefore, this is linear phase and you want linear phase because, if you have non-linear phase, you will introduce wave shape distortion for which they are used supremely sensitive.

So, this is one point. The other point to note is that, when you average this, all these 9 pixels you will get one number. You have to now assign this to one place in the result. You want to get the filtered image back. So, when you average these 9 pixels, you get one pixel value. Where would you place that pixel value to? What position will you place it intuitively? You will clearly place it here, this is where you will place. Now, if you go back and look at this, what you have done is, we have actually placed it at the centre of the filter and this is the analog of placing the output at $(N-1)/2$.

(Refer Slide Time: 40:17)



Now, how do we see that? Suppose you had a signal like this. So, this is now the 1D counterpart, I am now talking about. So, this is time and now, I am going to average this. Suppose I take 5 samples and average. So, what I will do is, I will take these 5 samples and average. And then, so this means I will, I am actually convolving it with the sequence of all 1's with amplitude 1/5.

This is the filter with which I am convolving. So, when I flip this and slide it across point by point, for whatever position when I multiply this, I am multiplying each sample of the filter with the corresponding samples of the sequence and adding them up. Which means I am effectively averaging 5 samples for this particular example.

Now, I will slide the filter across. So, in the next position, I will do this. Again, note that, when I do the averaging, again I get one number, right. If I take 5 samples and I average, I will get only one number. Where would I place that resulting number to, what position would I place it? Intuitively, where do you think would be the place to assign the result the one number that you get?

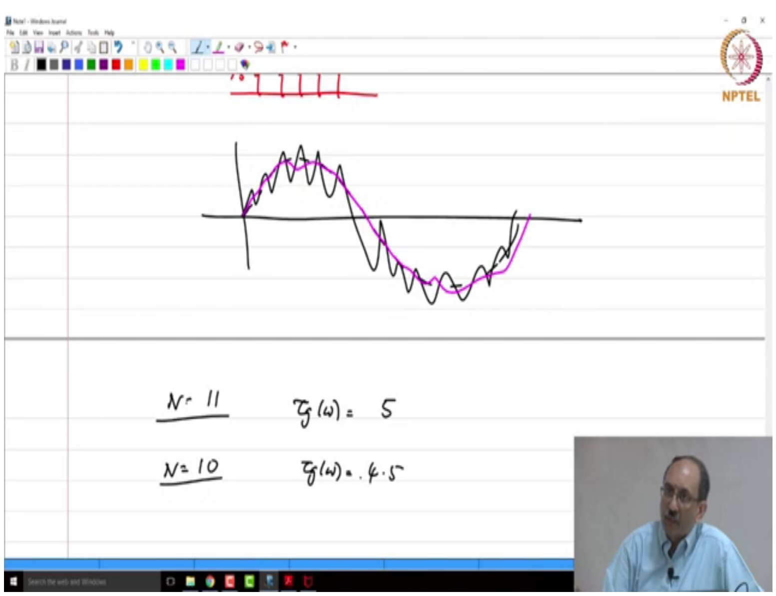
Student: (Refer Time: 41:48)

At the centre point, correct. So, you will place the average at this location, and this is precisely $(N-1)/2$. Now suppose, instead of taking a 5 sample averaging filter, suppose I took 4 samples, suppose I did this. Now, if the filter in this case would be all 1s, but now I have only 4 samples and the weight will be 1/4. So, that will correspond to averaging 4 samples.

Suppose, I take 4 samples and average, again I will get one number. Where do I place the result? The closest point to the centre that I can place is either here or here. And this would have effectively introduced what is called as the half sample delay. And the way to see this is, another way of seeing this is, suppose you have sinusoid again I am drawing continuous curves, what I really mean as samples.

Then suppose, I have a noisy sinusoid like this, I take a clean signal, add noise and then I pass it through a low pass filter. If I pass it through an linear phase FIR low pass filter, say Type I. Then, this signal will get cleaned up. Because, whatever noise that falls in the stop band of the filter will get eliminated, only in-band noise will be present. So, you can expect the output to be somewhat smoothed version of this.

(Refer Slide Time: 43:40)



Maybe, it will be something like this. But, the output will suffer a delay. And the group delay suffered by the filtered output will be $(N-1)/2$. We have seen, that is what the filter does. Linear phase FIR filter will give you a group delay of $(N-1)/2$. Suppose the filter length is 11. Suppose, if $N = 11$. Then, the group delay is $(N-1)/2$, so this will be 5.

Therefore, if you take the filtered output which is a cleaned up version of the original noisy signal and if you delay it back by 5 samples, you can time align the noisy input and the cleaned up signal. On the other hand, suppose if you take $N = 10$, in this case, this is 4.5. If you now take the cleaned up signal and then try to align it with the input, no integer shift will make you align with the original signal. Similarly here, when you are dealing with images, you do not want to introduce half sample delays, because the filtered image and the original image, you will not be able to line them up.

And instead of a 3×3 mask, suppose you had taken a 2×2 mask. So, all the weights will now be $1/4$. Again you can average, you will get one number, but where will you place the final result. If you place the final result here, here, here or here, you will introduce half sample delay in the image. You will not be able to line up the original image and the cleaned up image. Therefore, two things are important for image processing; one, all these 2D filters have to be exactly linear phase. Otherwise, they will cause significant distortion.

Second, you do not want to introduce half sample delay in these applications. Which means the filters that you choose have to necessarily be an N by N mask, where N is always odd. So, this is extremely important to realize. Certain applications, just do not tolerate half sample delays. And image crossing is an excellent example where, half sample will cannot be tolerated.

It is also another example where you strictly need linear phase. Whereas, in audio, because the air is completely insensitive, you can have a non-linear phase filter and still be fine, the ear will not be able to see or rather hear the difference. So, these ideas are extremely important and things like half sample delay and linear phase are critical in applications such as the ones I have mentioned.

So, whenever you are choosing an FIR filter, you also need to be sensitive to whether half sample delay can be allowed or not if they cannot be allowed, then necessarily if it were a 1D filter, it has to be of length odd. And we will further talk about, whether linear phase is possible for the IIR filter case. So, all that we have done so far is, we have assumed FIR and then shown the results. That $h[n] = h[N-1-n]$ or $h[n] = -h[N-1-n]$, either symmetry or anti-symmetry. And this is all FIR. So, we will further talk about these concepts as we go along.