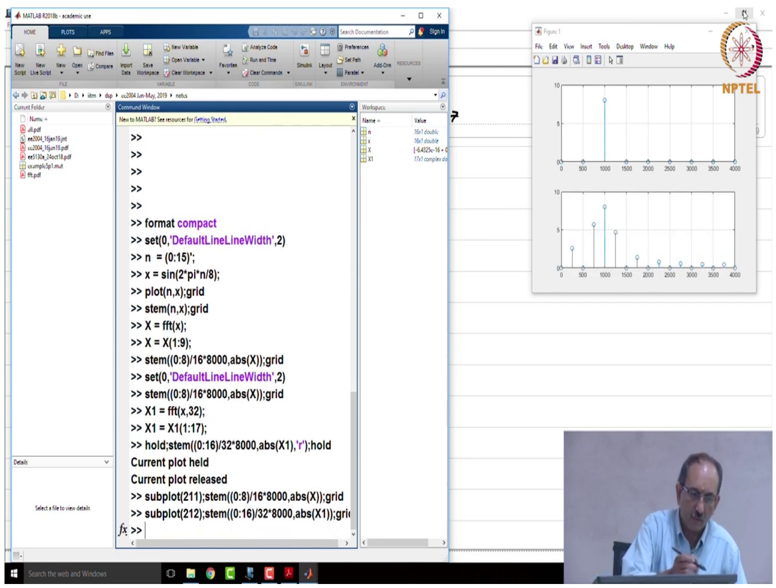
Digital Signal Processing

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

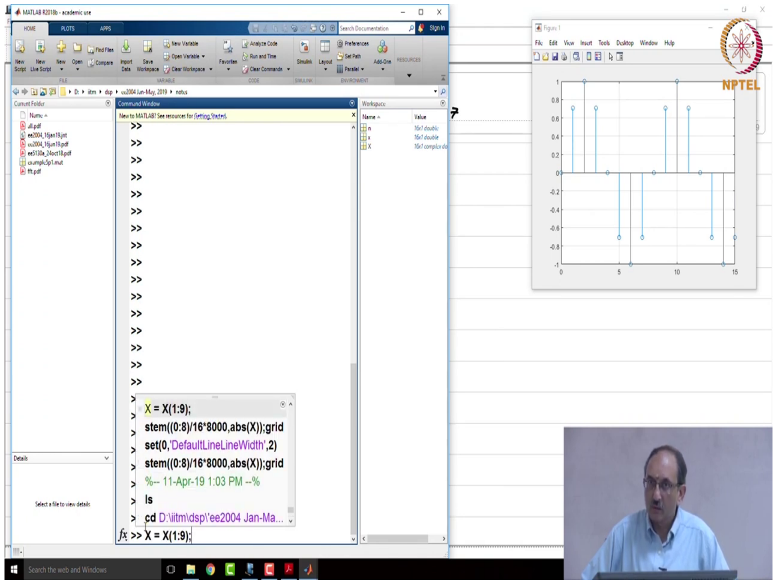
Lecture 77:
The Discrete Fourier Transform (3)
- Introduction to frequency estimation

(Refer Slide Time: 00:26)



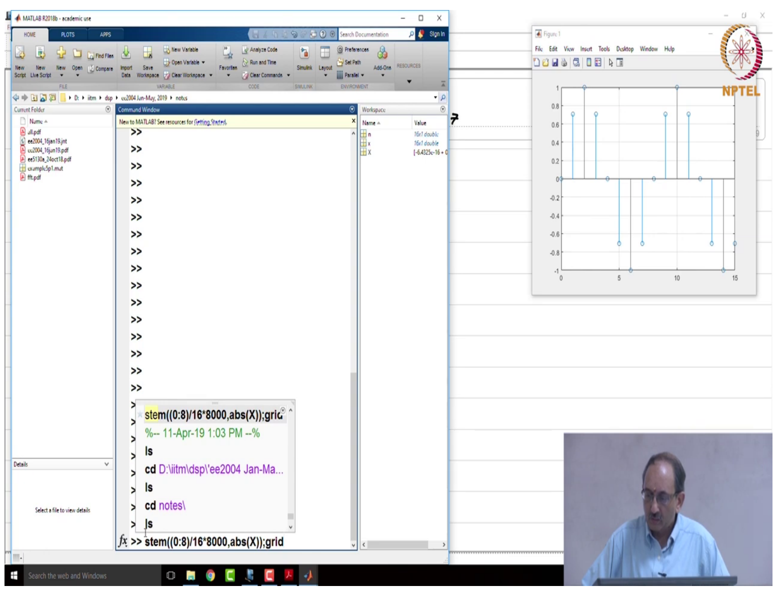
All right, let us get started. We were looking at a matlab example in last class and the signal in question was this. So, this is nothing but 1 kilo Hertz sine wave sampled at 8 and we are looking at 2 cycles here. So, this is the plot of the time domain sequence, you can also show this as a stem plot which is more typical.

(Refer Slide Time: 01:10)



And then, we looked at the spectrum and the DFT in matlab is implemented using the FFT algorithm, we are going to talk about the fft algorithm after this. And because of the symmetry, it is enough if you retain only the samples from 0 to $N/2$. So, we retain samples from 1 to 9 and remember, matlab is 1 based whereas in all our equations, we mean using 0 based index. Therefore, if you want to use matlab's index, you have to convert that from 1 base to 0 based.

(Refer Slide Time: 01:58)



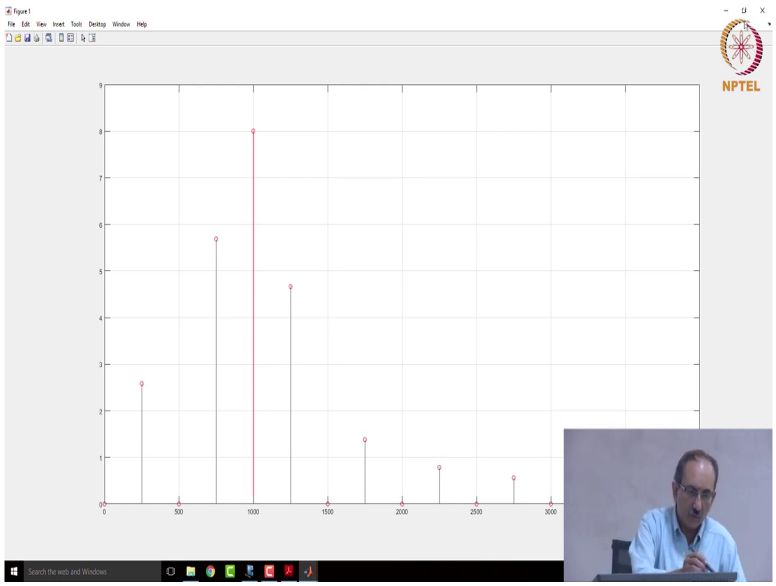
And then, remember, to show the bin index in terms of true frequency, we use the formula $k/(N * F_s)$. So, the index is captured here k/N; $N = 16$, F_s is of course 8 kilo Hertz. And then, we are plotting the absolute value and this is the stem plot. And now, let us look at the same thing but now, we will zero-pad. And remember, if you zero-pad, that is equivalent to taking more samples of the DTFT.

So, now what we will do is, we will take the original signal and then in fft, if you just give it one argument, it will compute the end point DFT of the argument by computing the length of the vector that has been supplied. The vector that was originally given to the fft function was x which is of length 16 and hence, it computed the 16 point DFT of which we plotted the first 9 points; 0 to $N/2$.

Now, we will compute the 32 point DFT. So, now, if we say $ff(x, 32)$, $X1$ will have the 32 point transform rather than the 16 point transform. And $X1$, you need to restrict yourself only from 0 to $N/2$ therefore, you restrict it from 1 to 17 because $N = 32$. And then, you plot this on top of this.

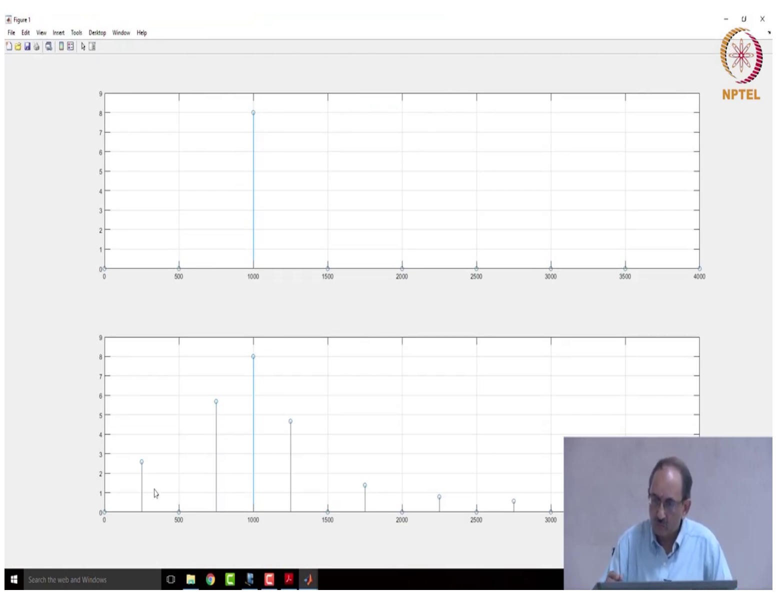
So, we are going to plot from $(0 : 16)/32$ because $N = 32$ times F_s and now you have to plot the absolute value of the new 32 point transform and then we will show it in a different colour.

(Refer Slide Time: 04:59)



And, this is what you see and if you recall the earlier plot, maybe since its over lying and it is exactly overlapping, you are not able to see the difference. Let me plot it as two subplots. So, this is what the original is.

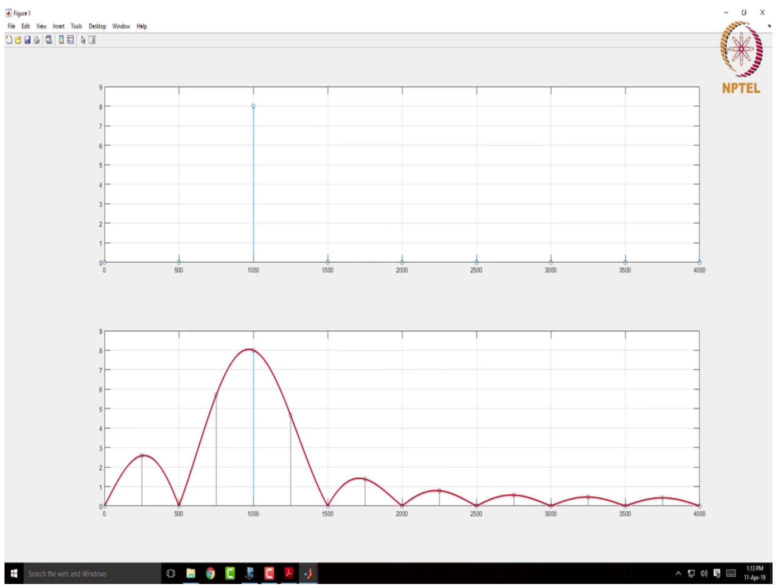
(Refer Slide Time: 05:55)



Now, we are able to compare them. Therefore, here you had 9 points, 0 to $N/2$; 0 to 8. Here, you have 17, 0 to 16 and if you see this point is the same as this and at 500, you have this. So, wherever the original frequency points are there, you have those samples as before. But, in between, you have new samples. So, between 500 and 1000, there was none before whereas, you have this. Similarly between every old sample, you have a new sample value from the 32 point DFT.

I can repeat this with now making it as 64 points and now what will happen is, between 500 and this, you will have a new sample here. So, you will start filling the gaps here with newer and newer sample points. And this actually represents the sampling of the DTFT and to show that, let me plot the DTFT. And I will plot the DTFT and show you the final plot and then we will make a comment on that.

(Refer Slide Time: 07:39)



So, this is the DTFT and clearly you are able to see that these samples are indeed these samples of the DTFT. So, as you zero-pad more and more, you will be sampling the DTFT finer and finer and you will get more closely spaced points. And notice that in the earlier case, rather the first subplot and since was the red curve is the underlying DTFT, you sampled it at this point and got the peak at the frequency of 1000 Hertz and every other sample in the top plot corresponds to the sampling of the zero crossings.

So, that is why in this plot, you had only one non-zero value and every other value of the DFT was 0 and that is because you are sampling the underlying DTFT precisely at the zero crossings. In the next instance, when you zero-padded it with 16 zeros making the total length 32, you have the underlying DTFT now sampled at 32 points. So, that is how you get these in between samples.

Now, one question that should have occurred immediately to you is, remember, I did not show deliberately how I got this and I said this is the DTFT. So, what is the question that immediately comes to your mind?

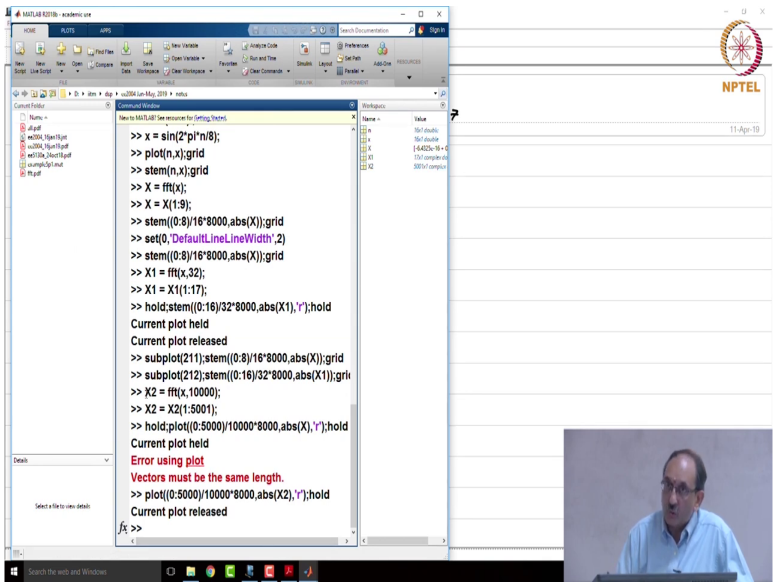
Student: (Refer Time: 09:45) how do we get?

Yeah, how did I get it. Remember, the DTFT cannot be computed on a machine and this surely is being computed on a machine and I am showing this plot. Therefore, this red curve which is the DTFT or so I claim is really the?

Student: DTFT (Refer Time: 10:09).

Very good. DTFT that has been heavily zero-padded and rather than using the stem plot, I am using the regular plot and what the regular plot does is, join consecutive points by straight lines. Therefore, this apparently smooth looking curve. If you really zoom in, it will contain set of straight lines which are connecting points that are closely spaced, that is all. But as you see zero-pad more and more and more, you will approach the DTFT.

(Refer Slide Time: 11:03)



So, this exactly what I did. $X2$ is what I have plotted, $X2$ is $fft(x, 10000)$. So, the total length is 10000; the first 16 samples are the samples of the sequence that I generated, the remaining are zeros. And then, I computed the 10000 point DFT and then I took the first 5001 points, which corresponds to the index in the range 0 to 5000 and then I have plotted it and I have plotted using plot rather than stem.

So, this should, this is what the plot shows and anything else that is kind of making a note of in the DTFT or heavily zero-padded DFT which is basically the practically the DTFT, anything else that kind of seems interesting from this plot.?

Student: (Refer Time: 12:30) 5000 (Refer Time: 12:31)

Say that again.

Student: (Refer Time: 12:39).

Student: Except the (Refer Time: 12:38).

Remember, what is the frequency of the underlying sequence? It is?

Student: 1 kilo Hertz.

1 kilo Hertz and if you had an infinite duration sinusoid of frequency 1 kilo Hertz, you would expect an impulse at 1 kilo Hertz. Now, what is happening here is, you can think of this sequence which we generate at 16 samples off from 0 to 15, you can think of this as being an infinite duration sinusoid

multiplied by a rectangular window and hence in the transform domain, what will happen is if you multiply in the time domain, you will convolve in the frequency domain.

In the frequency domain, you are going to convolve the two transforms. One of the transforms is the transform of the sinusoid, it will have impulses at 1 kilo Hertz and -1 kilo Hertz. The other function is the transform of the rectangular window and the rectangular window's transform is $\dfrac{\sin(N\omega/2)}{\sin(\omega/2)}$. Therefore, what you are really seeing here is, your convolving this rectangular windows transform which is sinc (Dirichlet kernel) convolved by the two impulses, which means it will shift the transform which is originally centered around the origin.

Because, you have the rectangular pulse, it is $\dfrac{\sin(N\omega/2)}{\sin(\omega/2)}$, of course you have the phase term and all. And that is centered around the origin. Its peak occurs at $\omega = 0$. When you convolve with these 2 impulses, each of these impulses will shift the window transform to the locations that they are located at, the impulses are located at $+1$ kilo Hertz and -1 kilo Hertz and hence you will get the transform like this.

The only thing is, we have shown from 0 to $F_s/2$. So, this is the sinc that is now centered around 1 kilo Hertz, you will have an image of this at -1 kilo Hertz. So, this makes sense, because this is the transform of the rectangular windowed sequence. The other thing that worth noting is, two things; one, the impulse has now broadened into a sinc. And if you did not know that the data contained a sinusoid of frequency 1 kilo Hertz, but if we are told that the data contains sinusoids and then if you are asked to find out what the frequency is, you would compute the DTFT and in practice, you would compute the DFT.

And then, you look for peaks in the spectrum because peaks in the spectrum are caused typically by sinusoids; narrow peaks are typically caused by sinusoids. And if you look at this and if you want to estimate the signals frequency, if you just took the 16 point DFT and look for peaks in the spectrum, you would estimate the peak to be at 1 kilo Hertz and then if you zero-padded it enough and then if you look for the peak, is the peak at 1 kilo Hertz or is it shifted?

It is slightly shifted and again this shift makes sense because this is analogous to what we had seen in the context of the resonator. There, 1 pole by itself, there would be no shift in the peak. But, because of the tail of the peak due to the other pole, this pole's peak got shifted and vice versa. Similarly, here you have a sinc that has been shifted to 1 kilo Hertz by this impulse, there is also another sinc shifted to?

Student: Minus.

-1 kilo Hertz, the same thing that was at 0 has been shifted to -1 kilo Hertz. And hence, you can expect the tail of the sinc at -1 kilo Hertz to interfere with the peak of this and that is exactly what is happening here. If you observed more and more cycles of this sinusoid rather than 16 samples in which you observe 2 cycles.

If you generated this sequence for say 64 samples and then zero-padded it enough and obtained this picture. Remember, this is now equivalent to taking the underlying infinite duration sinusoid and applying a larger rectangular window, because now you are having 64 points. So, that is like applying a 64 sample rectangular window, the rectangular window has now a larger duration. Larger in time which means in the frequency domain, it will be narrower.

If it is narrower, the tail interference will be less and hence when you observe the same sinusoid for a larger duration; more cycles. So, this sinc will have a narrower width and the peak will be shifted less. For a given length, if the frequency were lower which means this peak would move to the left and the image's peak will move to the right which means the two sincs are coming closer to each other, you can expect more interference and hence the peak will have more shift.

On the other hand, if the frequency were larger, the peaks will move away and the interference will be less. But again, the furthest two things can be apart if their real valued is $F_s/4$. So, if now the frequency is much closer to say 4000, you will have interference from the other image that has come close to 4000 from the other side. Therefore, the way to look for sinusoids in a signal is to compute the DFT and look for peaks. And suppose, you had another sinusoid that is say at 2 kilo Hertz, but much smaller in amplitude.

Then, this smaller amplitude sinusoid that is farther away, that will get affected by the side lobe of the stronger nearby sinusoid. Therefore, these side lobes will affect nearby sinusoids and the effect on them will be larger, if the interfering sinusoid is stronger and the sinusoid that is interfered with this weaker. And farther they are apart, the easier it will be for you to distinguish them. Closer they are, the interference will be more.

Therefore, these are some of the things that you can draw from this simple example and some of the intuitive arguments that you have made and you can verify this in matlab. You can create sinusoids, you can create a weak sinusoid, add to this and then look at the spectrum and so on. And for a single sinusoid itself, we can move this frequency closer to 0 or closer to 4000 and again observe shift in the peak. So, you can zoom in on this plot and find out where the actual peak is and you will find that it is not at 1 kilo Hertz. So, this is a very very brief introduction to estimation of frequencies of sinusoids given the observed data.