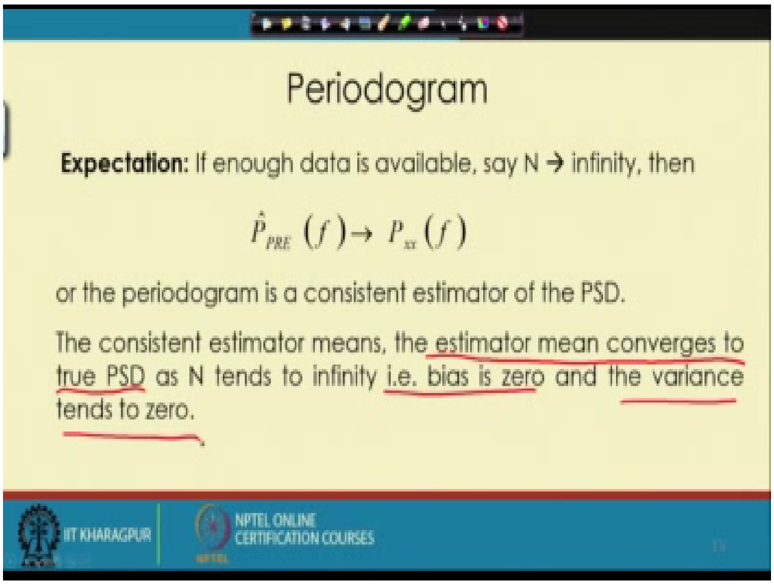
Biomedical Signal Processing

Prof. Sudipta Mukhopadhyay
Department of Electrical and Electronics Communication Engineering
Indian Institute of Technology, Kharagpur

Lecture – 36
Frequency Domain Characterization (Contd.)

So, now what we see periodogram provides a very good estimate, we can replace the classical definition of PSD with periodogram and if we have sufficient amount of data, it will go towards the ideality; that is it will give us the exact energy at a particular frequency of our choice and when we are having finite amount of data which is the case in that case, we will get actually a band pass version of the that PSD the actual PSD; that means, the energy around the that particular frequency given by a band pass filter output which is quite reasonable. So, now, we try to find out more properties of periodogram.

(Refer Slide Time: 01:24)

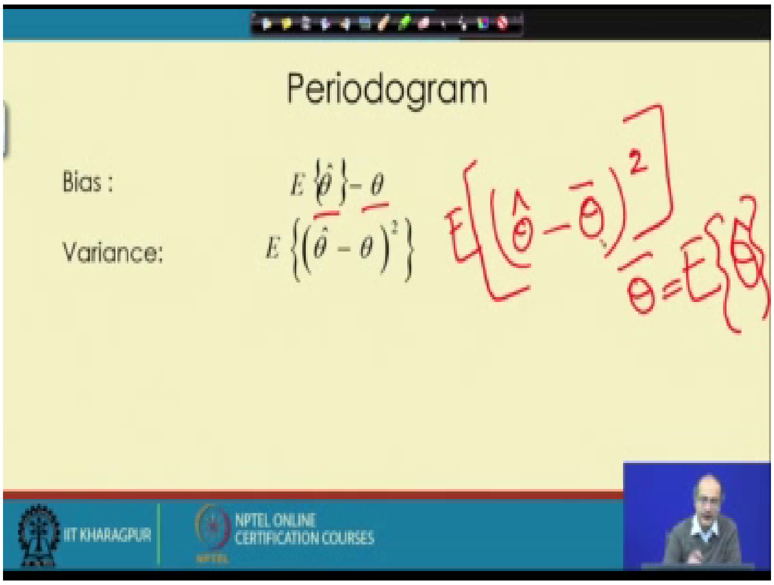


So, if we have enough amount of data say N tends to infinity, then we expect the program is going towards the that ideal PSD that P x x f and to create the difference two things we have done that we have used the suffix here for periodogram PRE and because we are computing it, we are getting just an estimate out of the given data. So, we have used this hat ok. So, these two things, we have done and then we can get it that is happening towards actually merging towards the ideal value of periodogram which seems to be almost there after the previous result we can say that in it other words that it is a consistent estimator of PSD ok.

Now to be consistent two things actually we need to satisfy the estimator mean should converge to the true PSD, the estimator mean converges to the true PSD that is one as n tends to infinity; that means, the bias should be 0 and the variance also should tend to 0 because it as an estimator, it can have some small error which can be given in terms of bias and variance, but if it has to be a consistent estimator then both of them should go towards 0 as N tends to infinity.

So, now let us look at that what are the definitions of bias and variance to take it forward.

(Refer Slide Time: 03:50)

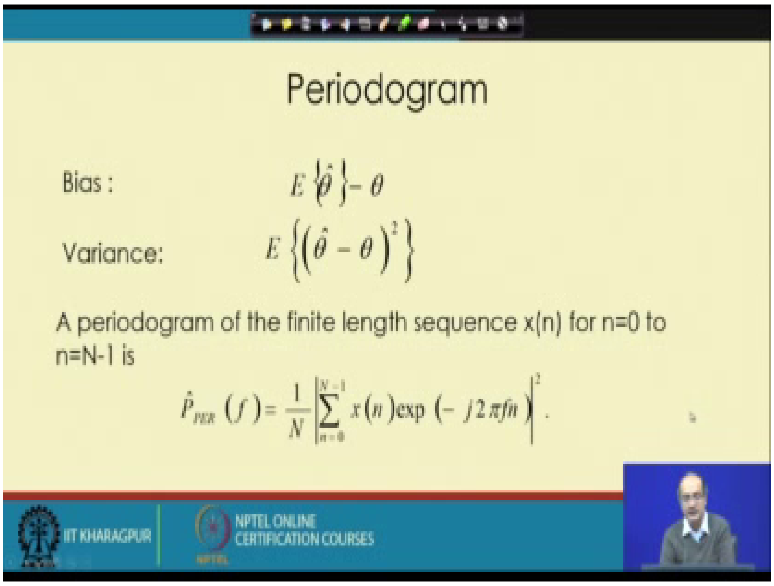


The bias means a say first of all let us take that here is the variable theta if we want to estimate and if the estimated value is theta hat ok. So, in that case the expected value of theta hat is providing the mean. So, difference of expected value of theta hat minus the true value, it gives us the bias that mean the mean of the estimates how far it is from the true value this is called the bias ok.

So, if there is any difference all the time we can say this is the bias and if someone has the idea that how much would be the bias you can have a corrections, also for example, for a particular range if we know your estimates are always higher by say some certain value. So, you can apply a correction all the time we see that when the in the market when the vegetable sellers they sell something and they are balances are not perfect they use small weights to balance it.

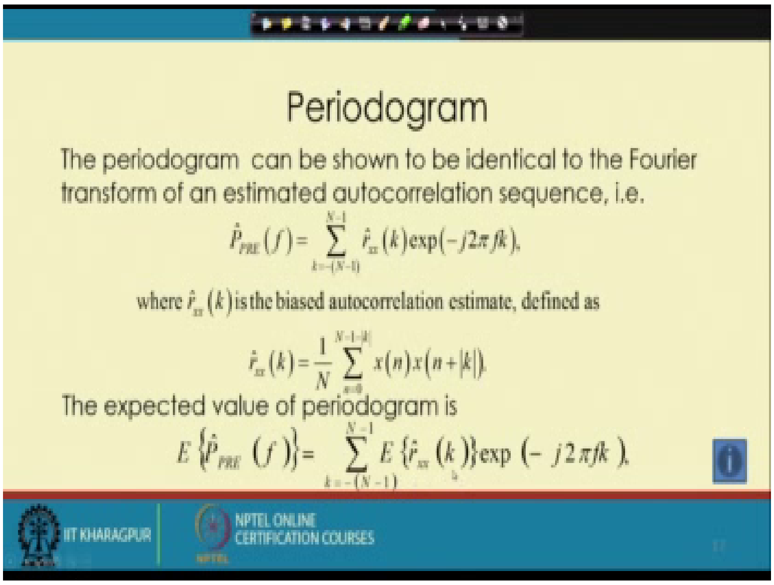
So, that we can call bias that every time they do the measurement they know that results will tend to one side they are trying to compensate it. So, that is a simple example of bias that can be corrected the next thing is variance the variance we already know, it gives us a spread of that estimated result from the true value ok, here we have taken actually the difference from the true value assuming that that it is going to give us the that the bias is going to 0 otherwise you can actually write it the difference between theta hat minus theta bar square and expectation of that where theta bar is expectation of the estimates the mean of the estimates here we have assumed that it is 0 mean. So, we have directly written it as theta otherwise we should have written it in this way ok.

(Refer Slide Time: 07:06)



So, we get the expression of bias and variance and now let us look at that our working definition of the period of gram using the sequence x n for n varying from 0 to capital N minus 1 ok, from this definition. Now we will try to find out that how the bias and the variable variances they come up ok.

(Refer Slide Time: 07:44)

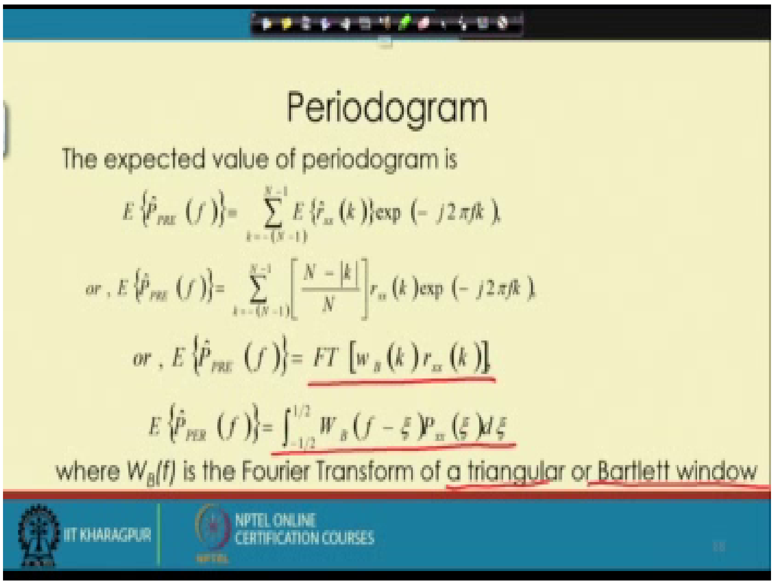


So, first we will look at the bias the periodogram, we can actually get as a identical to the Fourier transform of the estimated autocorrelation sequence please note this term its not the ideal autocorrelation sequence, but estimated autocorrelation sequence we have already shown these things earlier. We have shown that previous derivations starting from the value that we have that starting from the original definition of PSD. And from there that we have derived we got that for that case it is going towards that the PSD with ideal autocorrelation sequence, but here the things are little different we do not have that n tends to infinity here in this case expectation operator is also dropped.

So, we are getting actually we have to be contended only with the estimated autocorrelation sequence and that estimated sequence is given here that what would be the autocorrelation estimate here. However, what we find this is a biased estimate why it is biased because we see as the lag increases; that means, the value of k increases the sum within this summation term it decreases, it does not remain to be same whereas, this term of scaling that remains to be constant 1 by N ok. So, we do not get actually the real value.

So, if we now take the expected value of the that periodogram which is required to calculate the bias; then we get that expectation operator will work on this estimated autocorrelation sequence. So, now, let us move further that what we get out of it.

(Refer Slide Time: 10:53)



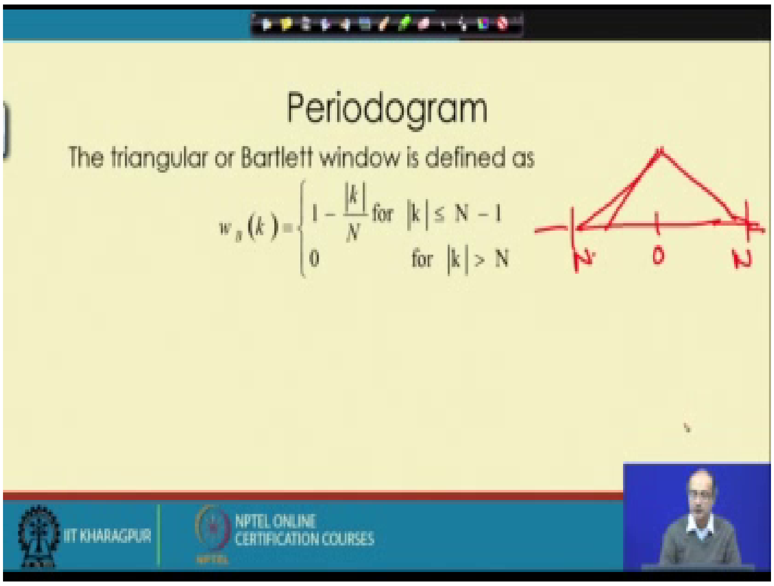
As we take the expectation for that we can replace this one with the real value. In fact, what all we are doing we are making use of this expression of that autocorrelation estimate and if we take the expectation of it we are just counting that here within this summation this term will give us the real autocorrelation function, if you take the expectation now how many times that will be repeated that will get by the this term ok.

So, taking them together we get this expression which would be maximum at the zeroth lag and it will get upper down as it is going increasing value of k ok. So, what we get here we get a new function like this which is present along with our the original expected value that is the autocorrelation and the that Fourier kernel which should give us the estimate of the frequency along with that we have a new function.

And that function if we tell that is W B; it is a windowing function or weighting function applied on the autocorrelation estimate. Then we can tell that here we are computing the Fourier transform of the product of the window or weighting function multiplied by the autocorrelation of the signal. So, that much we can get from here where W B, it has a particular form and with the help of that we can actually write in the frequency domain from this term that this term; from the Fourier transform when we have the product of the two terms they become convolution in the frequency domain.

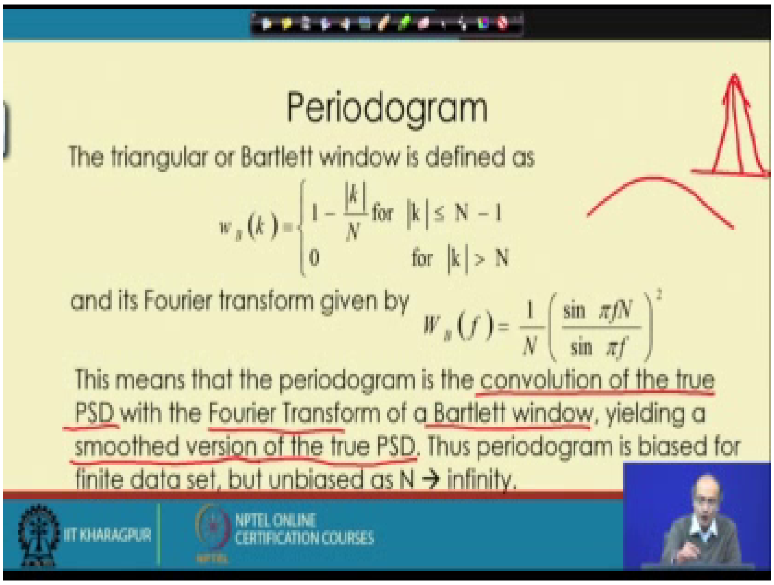
So, that gives rise to this the next line where we get that it is convolution between the function W B and the PSD that is a Fourier transform of the autocorrelation function that convolution is occurring in the frequency domain where this W B is the Fourier transform of a triangular function which is also known as Bartlett window. So, we get something that what we were looking for we are getting not exactly that; that it is not giving us the PSD, but PSD the original PSD which is convolved with a window function. So, if actually we can say a bit corrupted form of the original thing, but somewhat close to that it depends on how this function would be ok.

(Refer Slide Time: 15:30)



So, this triangular or Bartlett window, we can look more, carefully, it is a window which is having the highest value at 0 and it is when it is becoming actually, linearly, it is decreasing in both the sides its symmetric and beyond the value N, it will become 0 in both the side ok. So, it is a triangular window and the frequency transform of it.

(Refer Slide Time: 16:19)

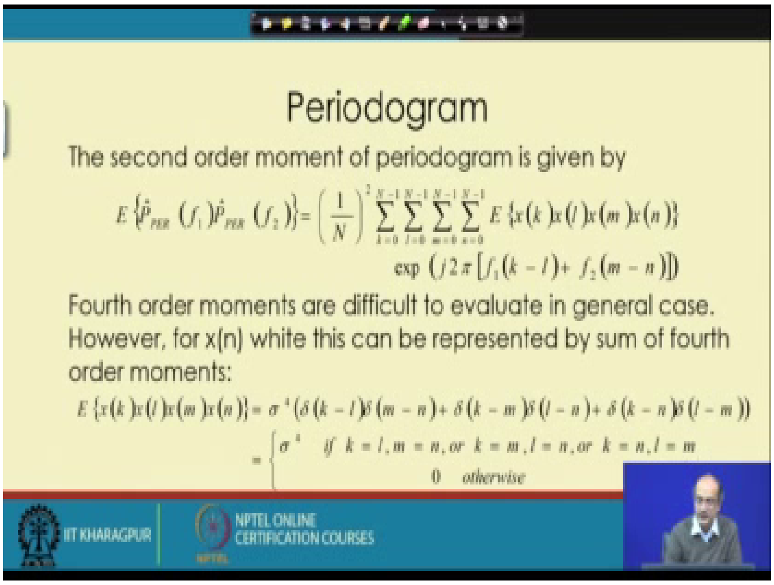


It comes in the form of a sin function which would be centered at 0. So, this means our periodogram, it is a convolution of the true PSD with the Fourier transform of Bartlett window and what convolution gives us is gives us a smeared view in a positive way to look at it is it is giving us a smooth version ok, if you are very picky then you will tell that the details of the spectrum is lost if you look it in a positive way then you will tell the roughness of the spectrum is removed. So, we are getting a smooth version of the true PSD.

Now, if the PSD is a continuous one say it is like this then smoothing will not probably make much change, but if you think if it is a monotone signal. So, it is like this single frequency is present then after the this convolution, it will give us value like this the output would be like this ok. So, in sense what we get out of it that periodogram is a biased estimate for finite set of data, but unbiased as N tends to infinity as we have seen previously that periodogram is giving us the windowed result or convolved with a Bartlett window in the PSD, but benefit is that if N is large and it tends to infinity these triangular function; it will change and ultimately that the frequency transform of this Bartlett window it will convert stored the impulse response.

So, we should get a unbiased estimate for infinite amount of data. So, starting from that point that we told that if we have infinite amount of data whether periodogram will give us that consistent estimate or our result will converge to the true value the answer. So, far is no if that data is finite, but it is yes if it is infinite amount of data is present. So, so far we can tell about the bias; that means, bias will go to 0. So, that say again a very good positive result. So, with that in mind we go forward to see that what happens to the variance.

(Refer Slide Time: 20:18)

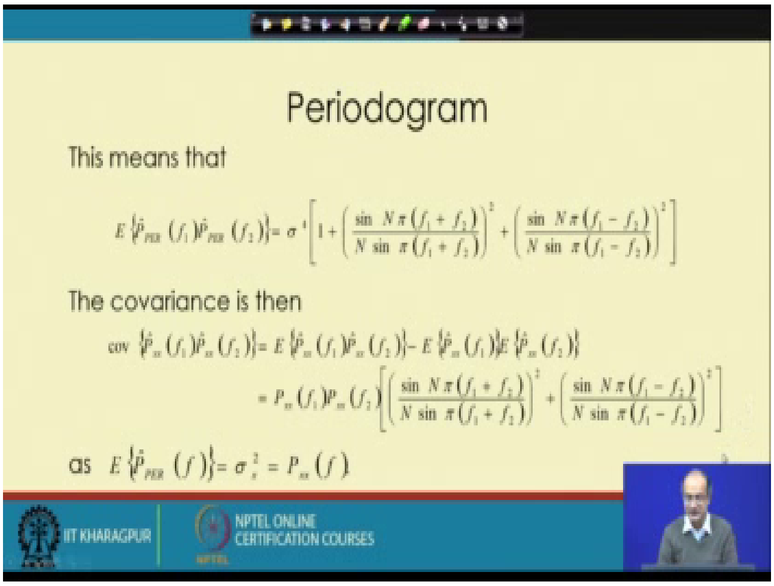


So, for that we look at the second order moment of the periodogram that second order moment or covariance at two frequencies f 1 and f 2, if we take it is given by a bit complicated expression given here because we saw each of these term, it involves actually two sample terms.

So, here we have four terms of x and thereby having four indices of summation and in the same way that we have four indices in the exponential term also ok. So, it becomes a little more complicated and we can unable to actually evaluate, it in a general form, we cannot take it forward in that way. So, we take a easier route we take a very specific example where we can deal with it if we take x n is white if x n is a white noise, then this expression becomes much more simple and tractable in that case; that we get this most complicated term that is expect expectation of these four terms.

It gives us actually product of some delta functions which can be given in a simpler way that this is equal to sigma to the power 4 when k equal to l and m equal to n or k equal to m and l equal to n or k equal to n and l equal to m all the other places it would be 0. So, we could get somewhat tractable value of this and using that. Now we will see that how we get the second order moment.

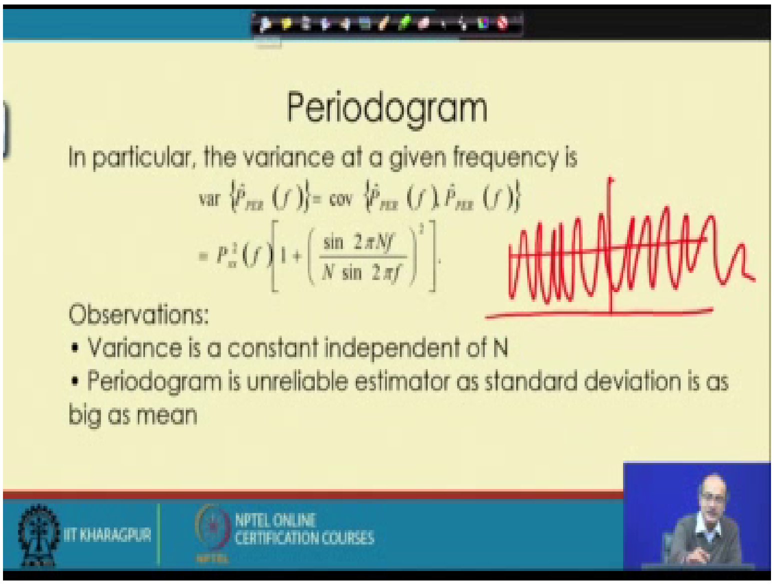
(Refer Slide Time: 23:12)



Now, using that we get this expression for the white noise ok; so, we get the two terms at two different frequencies, we can say the sin functions we could get and now using these we can write the covariance what we are looking at covariance is nothing, but the term what we have taken the expectation of the two terms minus expected product of the expectation of each of these terms.

Now, this one we already know that it would be sigma to the power 4 we get rid of this term. So, we are left with this expression now this is somewhat simple, at least, we could get some form.

(Refer Slide Time: 24:47)



But it is not so, simple that we could be very contented about it. So, let us look at now the case of variance what happens for the variance we have f 1 equal to f 2 equal to f. So, in that case out of the two terms in terms; one will vanish where we had the term f 1 minus f 2 and the other term that f 1 plus f 2 will give rise to this expression ok. So, the first observation is variance is a constant independent of N. Now this is; however, is not a very good news we were expecting as N tends to infinity that variance should come to 0, but because of the presence of these term we cannot say that N is actually affecting these variance and it is going to 0 ok.

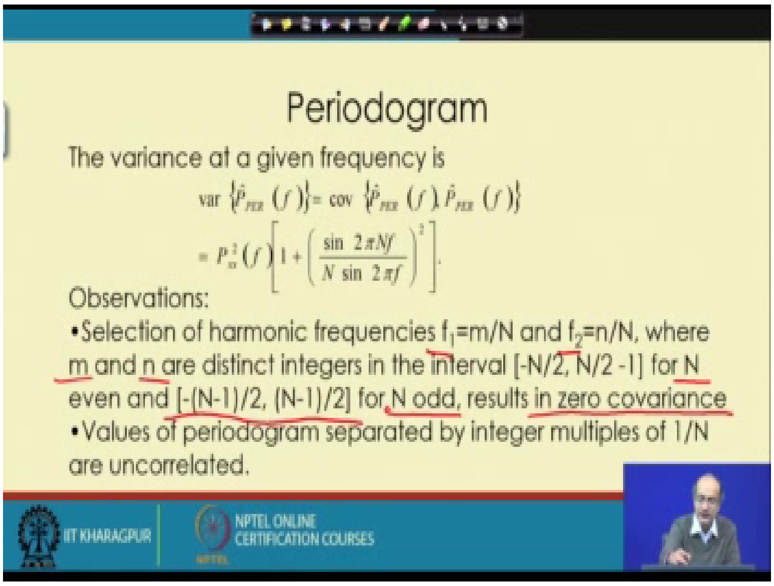
So, what we get? Periodogram is a unreliable estimator as the standard deviation is as big as mean why it is so? We know that the first term if you look at this is the actually the mean P x x f is the true value. So, what essentially we are telling that whatever the mean value is they are variances as much as that.

So, that immediately gives us that view that whatever we are getting out of periodogram, it has lots of fluctuations and the values we get that is very much unreliable. And at this moment you may doubt my words that what sir is suggesting how that can be true; what we are teaching in a class all of us, we are very confident we take the Fourier transform take the absolute value and we comment on the that the frequency content of the signal; I think all of you have done that at some point of time.

Please take some simulation environment whatever you are comfortable with it may be MAT LAB it may be psi lab, it may be any other tool take a random sequence rather to be more precise it would be a pseudo random sequence you take you take data from 0 to N minus 1 take little large value. So, that you get a good estimate, compute the Fourier transform of that white noise and you take then the absolute value of it and plot it in front of you.

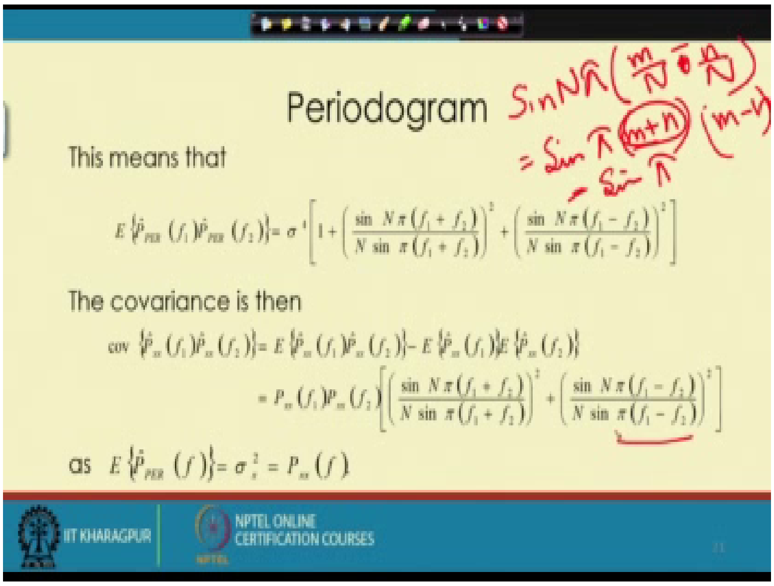
Look at the plot what you are expected for a white noise in the frequency domain it should give us all the frequencies at the equal magnitude, but you will find plot looks like this, it not at all a constant value that is because you have used periodogram and periodogram is not a reliable estimate after doing this explained, I hope that you would be convinced about what we are teaching in this class that it is not reliable because the variance is as big as the mean.

(Refer Slide Time: 29:52)



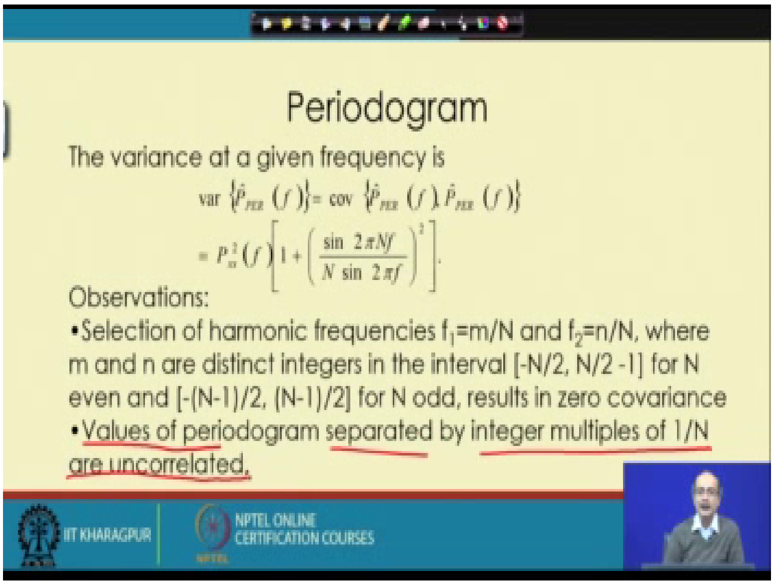
Now variance at a given frequency, we know the expression. Now we look at that why such fluctuations are happening let us take some frequencies f 1 and f 2 which are harmonics given by m by n and small n by capital N ok. So, you get that these are the actually frequencies; what you typically get when you take a d f t that the d f t values you get in this way. So, were m and n are distinct integers in the interval minus N by 2 to N by 2 minus 1 for N is even and minus N minus N plus 1 by 2 to N minus 1 by 2 for odd value and it results in zero covariance, how do we get that for that we have to go a little backward we to look at the expression of the covariance once more in the covariance 1.

(Refer Slide Time: 31:34)



We got the two terms here that now here if we replace f 1 is the first term that numerator we are taking sin of N pi f 1 is say m by capital N; the other one is small n by capital N. So, we can write sin of pi into m plus n; m and n both are integer. So, this is equal to sin pi which is 0, same is true for this one there you will get the expression m minus n would replace actually m plus n this term would be replaced by m minus m because we will have here minus. So, again when m and n; they are integer and different, it will be same as sin pi. So, it is a very easy to derivation we could get that the values of the covariance would be 0.

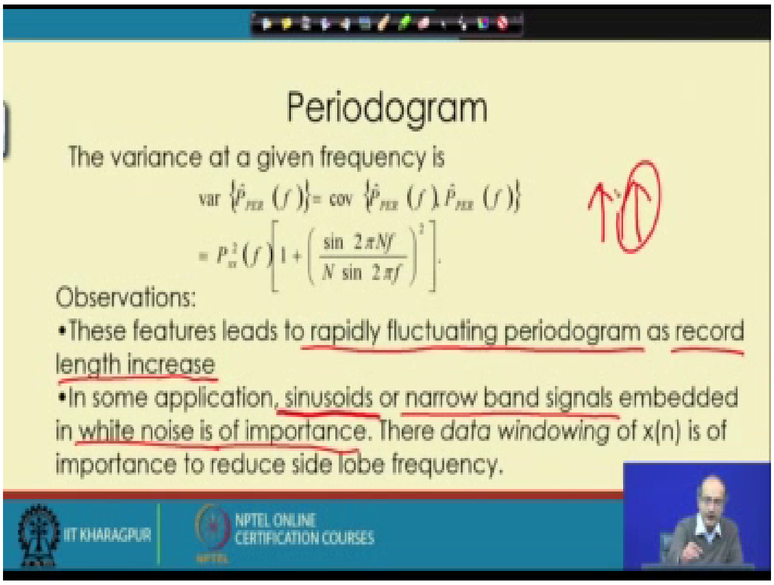
(Refer Slide Time: 33:00)



So, from there, we have to accept that fact that two frequencies, if we take harmonic frequencies the covariance would be 0. So, the values of the periodogram separated by an integer multiple of 1 by N are uncorrelated if the covariance is 0; that means, they are uncorrelated.

Now you get that why we would expect to get random noise kind of pattern instead of a constant frequency or a constant value of the signal energy for PSD when we take the Fourier transform of white noise and look at the absolute value of the Fourier transform because the values are uncorrelated and they have high variance. So, we are getting instead of constant value somewhat uncorrelated values which looks like random noise and truly they are expected to be so, because we have taken periodogram as an estimate estimator.

(Refer Slide Time: 34:35)



So, here few more small features we will look at. So, what we are expected to get that we will get rapidly fluctuating periodogram ok. So, instead of getting better and better estimate as we expected with the signal length we expected the bias is going to 0. So, we would get more and more close to the true value instead of that what is happening? We get periodogram becomes more and more rough it is becoming more with fluctuations ok. So, that is the benefit we are getting with lengths of data or increase in data and in some cases it can be really difficult specially when we are looking for sinusoids or narrow band signals embedded in white noise.

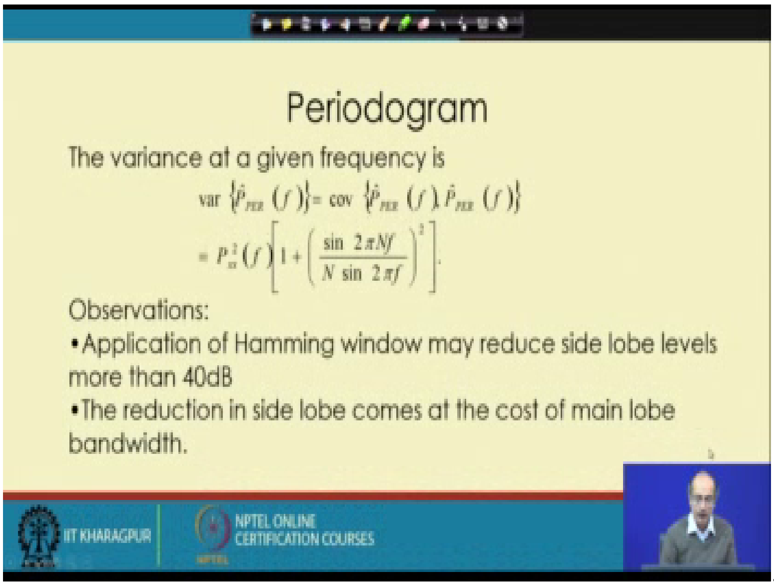
One such application comes immediately in our mind that is the rudder application where we inject a frequency and look for that the returning waves. And those sinusoids they could change the wave frequency because of the Doppler shift, but the change would be very small. So, that the wave that is getting reflected from a stationary barrier, it would have the same frequency as the that the transmitted wave and for a moving target it will give a little different frequency.

But so, two sinusoids; two actually frequencies, we are supposed to get or two impulses, we are supposed to get which would be close by and white noise is there which have are having fluctuations. So, what can happen these two may get merged because of these fluctuations. So, you would not get that there is a moving target ok.

So, in this case that the data windowing becomes very important to reduce the side lobe ok; so, that is one of the thing we realize that if we have to use periodogram, we need to use some kind of windowing without that we cannot go ahead and this windowing is just to take care of this rapidly fluctuating periodogram.

So, earlier we are trying to get rid of that wanted to go for ideal case by increasing the data yes by increasing the data we have more information, but because of these fluctuations we need to go for some windowing to reduce the side lobe frequencies. So, that those peaks does not get actually merged ok.

(Refer Slide Time: 38:00)



So, we can look at different kind of windows and having window is one of them which gives good reduction in the side lobe, it gives up to 40 dB or more, but it comes at the cost of the main lobe bandwidth ok, we leave it here we will come back after a break.

Thank you.