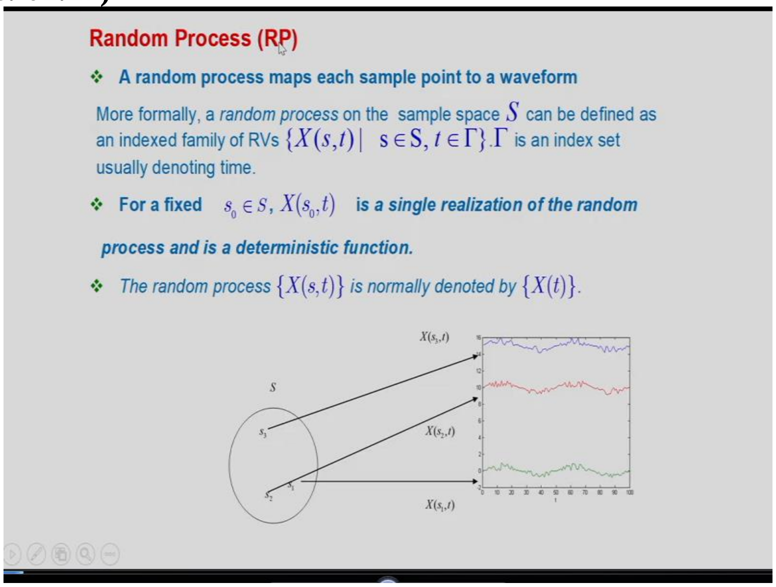
Statistical Signal Processing
Prof. Prabin Kumar Bora
Department of Electronics and Electrical Engineering
Indian Institute of Technology, Guwahati

Lecture-4
Random Processing

Hello students, so far we have discussed the basic concepts of probability and random variables. In the last lecture, we discussed the linear algebra of random variables today will give an outline of random processes in fact in statistical signal processing we model the signal as a random process. And therefore, a statistical signal processing is the processing of random processes will start with the definition of random process. Random process that is RP will abbreviate it as RP.

(Refer Slide Time: 01:27)

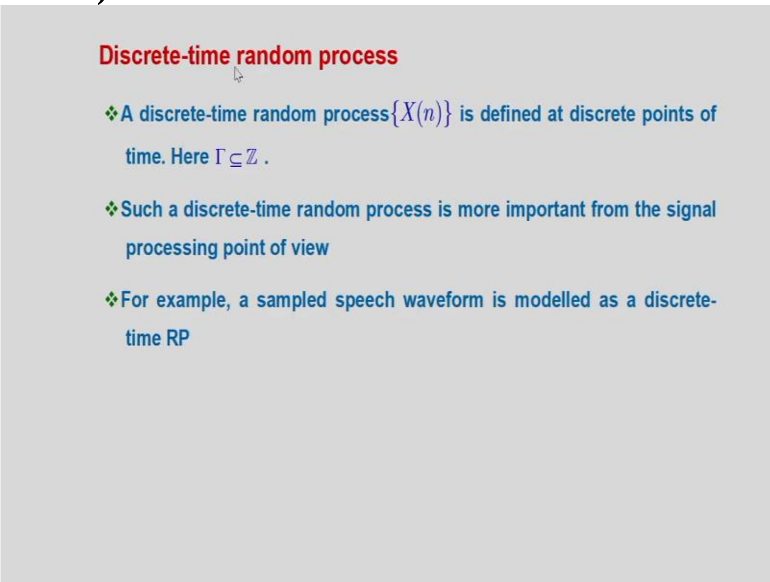


A random process maps each sample point to a waveform unlike a random variable which maps a sample point to a point on a real line. Here it will be mapped to a waveform like this suppose the sample point is mapped to this waveform this is the time axis. So there waveform varying with time and one sample point will be mapped to one of the waveform more formally in a random process on the sample space as can be defined as an index family of random variables.

So a random process comprises of X st such that s belong to the sample space and t belongs to some index set tau. So tau is an index set and tau usually denotes time. So essentially it is a functions of both s and t, so it varies with s as well as it varies with t if we fix the sample point for peak sample point s 0, X s 0t realization of the random process and it is a deterministic function because we are considering only one s 0 and our practical realizations whatever signal we observed that is modeled as a single realization of a random process.

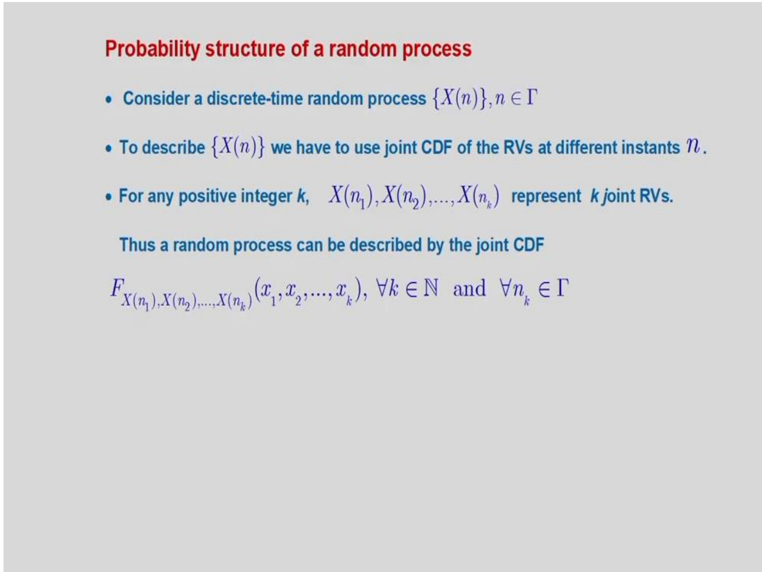
This is an important point any observed signal is considered as the realization of a random process like in the case of random variable we omit s therefore X st is normally denoted by X t here is the illustration of random process this is the sample space where showing three sample points and each sample point is mapped to a sample function like this wave form like this.

(Refer Slide Time: 03:54)



Our concern is discrete time random process. A discrete time random process Xn is defined at discrete points of time. So, here the index set tau is a subset of z, z is the set of integer. So on in integer points the random variables are defined such a discrete time random processes is more important for signal processing point of view. Therefore we will be discussing only discrete time random process as an example a sampled speech waveform is modeled as a discrete time random process. So, sample speech waveform that we can represent as a random process.

(Refer Slide Time: 04:45)

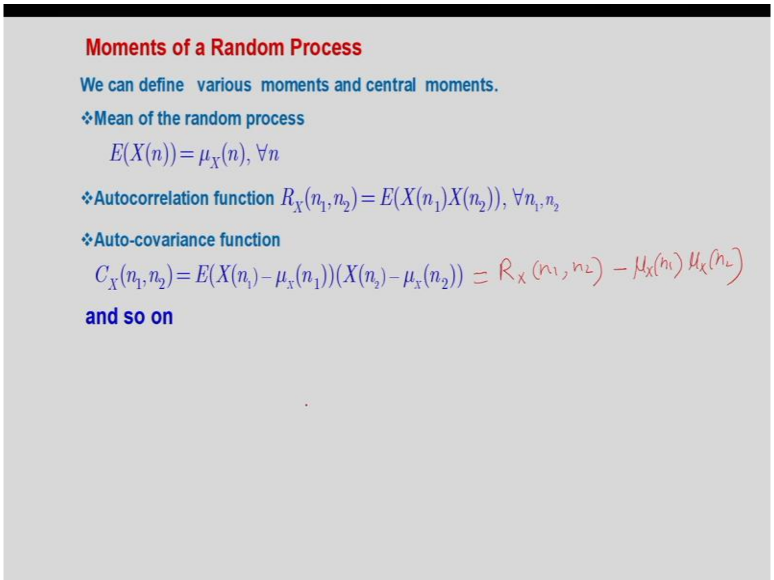


Probability structure of a random process, we know that random variable is described by CDF similarly random vector it is described by joint CDF and similarly we define also probability density function probability mass functions etcetera. How to have such characterization for a random process to describe Xn we have to use joint CDF because it is an index family of random variables.

So, we have to consider joint CDF only and these joint CDF we have to considered at different instants of n. Suppose you consider any positive integer k then Xn 1 Xn 2 up to Xn k represent k joint random variables because we are considering the random variable at instant n 1 n 2 and so on up to n k. So, this is there are k joint random variables. Therefore they can be characterized by joint CDF.

This is the joined CDF of the random variable Xn 1 Xn 2 up to Xn k at point x 1 x 2 up to x k. Now this CDF must be defined for all k belonging to n because this index set is an infinity. So all possible values of k we have to consider, similarly this n 1 and n 2 etcetera. We can place also anywhere on the time axis. So this for all n k belonging to tau we have to considered. Therefore all these considerations make the joint CDF characterization of a random process difficult.

(Refer Slide Time: 06:48)

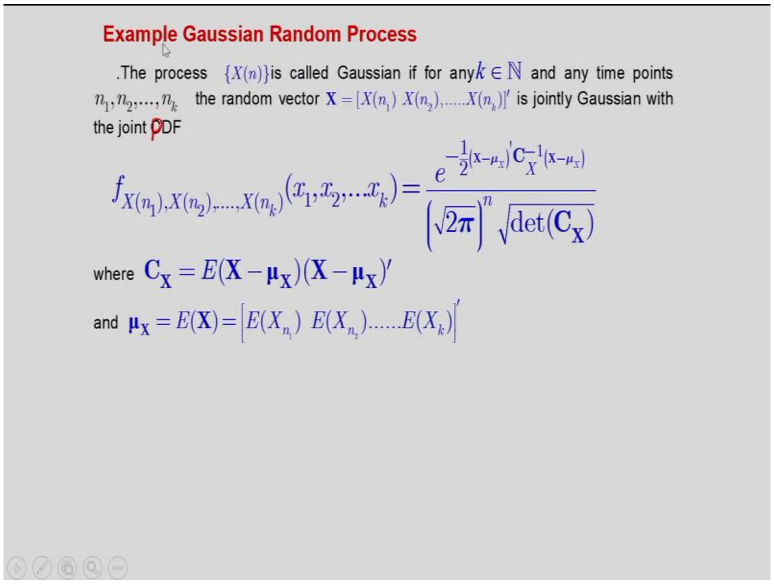


So, we have to have some simpler description of a random process. We can define various moments and central moments of a random process. For example you can defined mean of the random process E of Xn. Now this we have to consider at all possible instant of time. So E of Xn is equal to it will be again a function of n for all instant of time n. Similarly other moments also we can consider for example autocorrelation function this is denoted by R Xn 1 n 2.What is this E of Xn 1 into Xn 2.

This is the joint expectation of Xn 1and Xn 2. This parameter R X n 1 n 2 also we have to consider for all possible values of n 1 n 2 and autocorrelation function it gives the similarity of 2 random variables at different instant of time Xn 1 into Xn 2 and then they were taking the average bellow, we can also characterize a random process by auto covariance function C X of n 1 n 2 = E of Xn 1 - mu x n 1 into Xn 2 - mu x n 2 so you subtract the means from the corresponding random variable and then you find joint expectation.

And this we can show that this quantity is equal to R X of n 1 n 2 - mu x n 1 into mu x n 2 and similarly we can define the different moments. For example higher harder moments for harder moments etcetera.

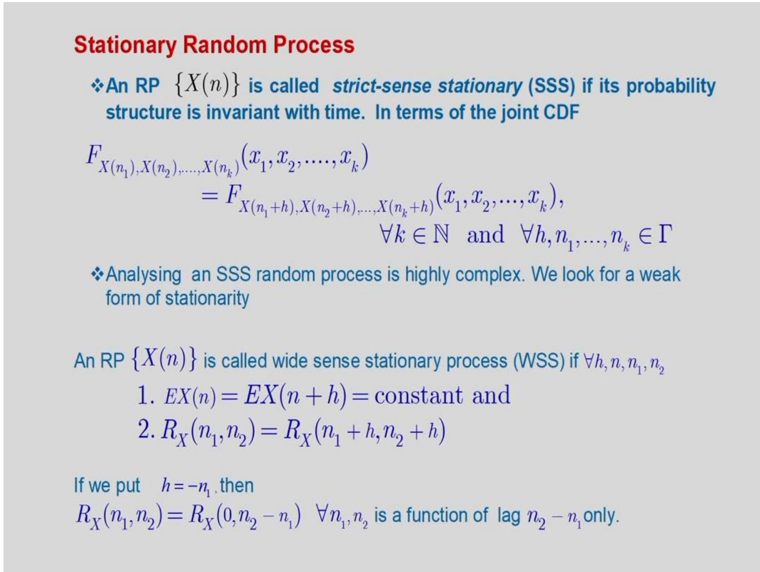
(Refer Slide Time: 08:57)



Will give an example Gaussians random process the process Xn is called Gaussian if for any k belonging to N and any time points n 1 n 2 up to n k the random vector given by this column vector with elements Xn 1 Xn 2 up to Xn k is jointly Gaussian with joint PDF this is the PDF we have already discussed about this PDF. So PDF is the Gaussian PDF and given by e to the power - half X - mu X transpose into C X inverse of the covariance matrix into X - mu X.

And divided by this normalization factor root over 2 pi to the power n square root of determinant of C X as we know C X is the covariance matrix that is given by E of X - mu X into X – mu X transpose. Similarly mu X vector it is the mean vector and it is the vector comprising of mean of is component random variables. So that this jointly Gaussian random processes important for us and it has comparatively their description because we have a for any N or any k we can have this joint PDF.

(Refer Slide Time: 10:37)



Because the probabilities structure of the random process is very complex. We have to have some simple assumptions one such assumptions is stationary random process. A random process Xn is called strict sense stationary that is abbreviated SSS. If its probability structure is invariant with time, so whatever probability description is at this point of time suppose the same probability description will hold for another instant of time may be after 1 month in terms of joint CDF.

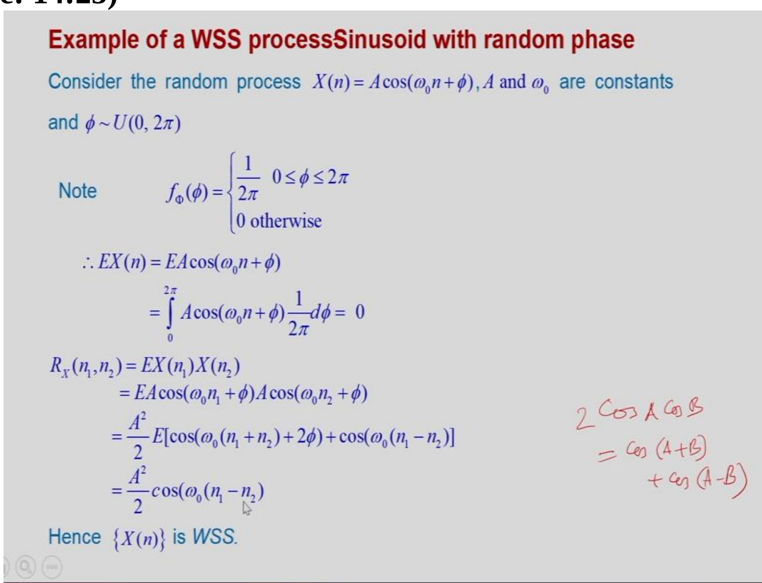
We can write that is joint CDF of random variable Xn 1 Xn 2 up to Xn k at point x 1 x 2 up to x k. So joint CDF at this points is same as joint CDF random variable now we will shift the time axis by an amount into the amount is so X of n 1 is X of n 2 is up to X of n k + h. If we consider this set of random variable instead of this. Then also this joint CDF will remain same. So that means is random variable we are shifting by the same amount Xn 1 has become X of n 1 + h and is X n 2 has become X n 2 + h.

And under such conditions the CDF at a set of points will remain invariant and this is the strict sense stationarity now to analyses strict sense random process is highly complex will look for a weak form of stationarity. So that is the week sense stationary random process or wide sense stationary random process. A random process Xn is called a wide sense stationary process WSS that is if for all h n 1 and n 2 number 1 E of X n = E of X n + h. So to instant of time you have the average value same.

That means average value must be constant, R X of n 1 n 2 is same as R X of n 1 + h n 2 + h this is for all h, now we can put h = - n 1 then R X of n 1 n 2 will be equal to now this h will be, if we will put minus n 1 this 1 will become 0. And here it will become n 2 - n 1. So R X of n 1 n 2 will become R X of 0 n 2 – n 1. So if we want this condition then this is equivalent to the statement that R X of n 1 n 2 is a function of lag n 2 - n 1 only.

It does not depend absolutely on n 1 and n 2 it depends only on the time difference n 2 - n 1 so such a process therefore for a wide sense stationary process the mean E of Xn is constant and the autocorrelation function R X of n 1 n 2 is a function of n 2 - n 1 only.

(Refer Slide Time: 14:25)



Will give an example of a WSS process that is a sinusoid with random phase consider this is the sinusoid Xn = A cos omega 0 n + pi. In omega 0 are deterministic constant suppose A the amplitude and omega 0 is the frequency these are known only phi is unknown and phi. If we model phi as uniformly distributed random variable .This is the symbol for uniformly distributed random variable and it is uniformly distributed between 0 to 2pi.

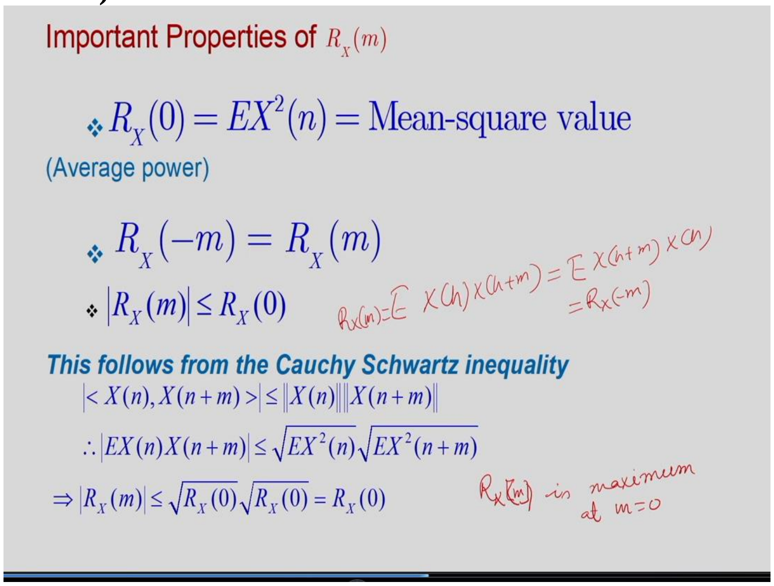
So, this phase part is a random quantity which is uniformly distributed between 0 to 2pi so part we can find out the PDF of this random variable because this is the random part. So PDF will be therefore since it is uniform it is 1 / 2pi for phi lying between 0 and 2pi and 0 elsewhere. Now how to find out E of Xn because I did a function of this random variable phi therefore we can we find a expectation of this function of random variables that is range of random variable is to 2pi.

Therefore this is integration 0 to 2pi of A cos omega 0 n + phi into 1 / 2pi d phi, and this is a cosine function in a period therefore integration will be equal to 0. So what we see that mean is constant. Now let us find out R X of n 1 n 2 this is equal to E of Xn 1 Xn 2. That is by definition and that is equal to E of A cos omega 0 n1 + phi into A cos omega 0 n2 + phi. Now there is a product of 2 cosine terms we can use the relationship that is cos of twice cos A cos B = cos of A + B + cos of A - B.

So if we use this identity then we will have this is equal to A square by 2 into expectation of cos of omega 0 n 1 + n 2 + 2pi this and these are added plus cos of these and these are subtracted cos of omega 0 n 1 - n 2, so this part now this is the cosine function if we integrate over 0 to 2pi you will get 0. So therefore expectation of this part is 0 and this part does not involve pi. So there it is a constant.

Therefore what we will get A square / 2 into cos of omega 0 n 1 - n 2, so what we have seen that this autocorrelation function is a function of the time lag n 1 - n 2 and its mean is constant therefore Xn is a WSS process .

(Refer Slide Time: 17:52)



Now a WSS process is characterized by the autocorrelation function R X of m. This is a very important quantity for WSS random process because it is characterized in terms of R X of m only. So what is R X of 0 if we put R X of 0 by definition it will be E of X n into X n that is E of

X square n and this is the mean square value that means this is the square term you are taking the average therefore it is this is the mean square value and this is also interpreted as average power.

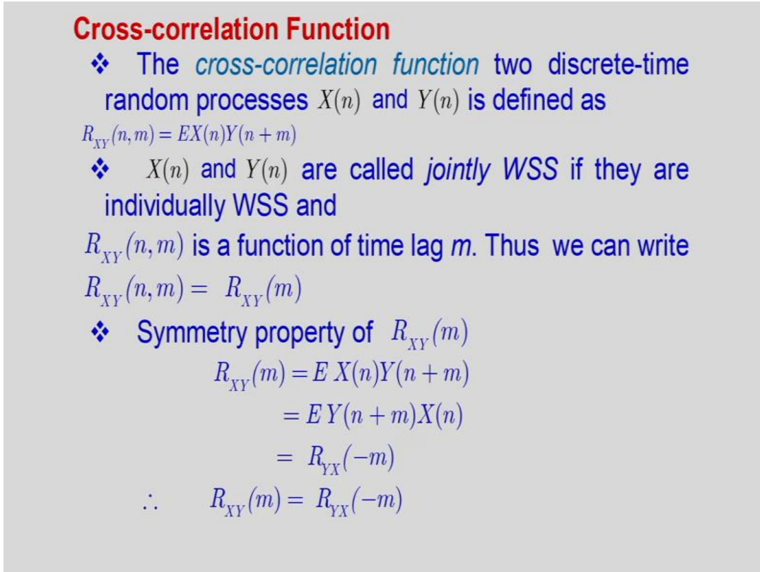
Because this average power E of X square n is the power delivered on 1 ohm resistance unit resistance, so that way E of X square n denote average power. So to interpretation this is A square mean value and this is also the average power of the random process. And R X of m is an even function because what we can write is that E of X n into X of n + m this is same as if I interchanged the order E of X n + m into Xn.

Now this quantity by definition this is my R X of m. And by definition this quantity will be this minus this is the argument therefore this will be R X of minus m. So what do we get is that autocorrelation function is an even function of m. Then magnitude of R X of m at any m is less than equal to R X of 0. So this we can prove using the Cauchy Schwarz inequality that we know. So inner product of X n and X of n + m if we take the inner product and then magnitude of that will be less than normal X n into normal X of n + m.

So inner product is expectation magnitude of E of X n into X of n + m that will be less than equal to normal of X n is square root of E of X square n normal X of n + m is square root of E of X square n + m therefore this quantity is autocorrelation function magnitude of R X of m is less than equal to square root of R X 0 this is equal to R X 0. And similarly this 1 is also R X 0 because this is the product of X of n + m into X of n + m.

So there is lag will be 0 so that way both square root together is R X of 0. So what we have seen that magnitude of R X of m is less than equal to R X of 0 that means the autocorrelation function has a maximum at the origin at m and m = 0 there is a maximum. So R X of m is maximum at m = 0. There are other properties also I will not discuss them.

(Refer Slide Time: 21:33)

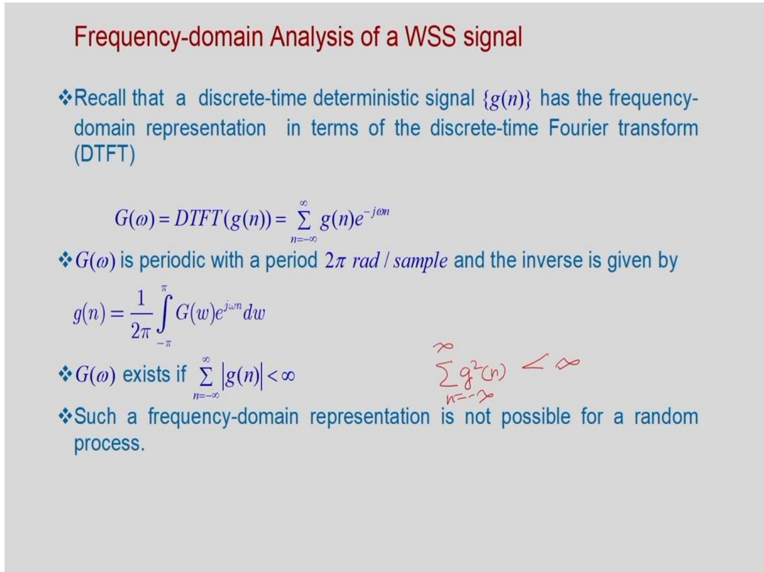


We can consider also multiple random processes for example Xn and Yn. So if Xn and Yn are 2 random processes then their correlation structure will be given in terms of cross correlation function and this is by definition R XY nm that is cross correlation argument n and m it is the expectation of E of Xn that is expectation of X n into Y n + m. So cross correlation R XY at point nm is defined as E of Xn into Y n + m, so that we can define cross correlation function. Here also I want to make some simplistic assumption about the correlation between Xn and Yn.

And we define what is known as the jointly wide sense stationary random process. Therefore jointly WSS process Xn and Yn are characterized by their individual autocorrelation functions and the cross correlation function R XY nm that is equal to R XY m. Now we know that autocorrelation function is even symmetric similar to that there is a symmetry property of R XY m. R XY of m = E of X n into Y n + m this is by definition.

And this now because these are 2 real number I will exchange the order, So E of Y n + m into X n now definition of cross correlation function we take the difference of second augment and first augment. So that way this will be now R YX of - m. Therefore R XY of m = R YX of - m. This is this symmetry property of cross correlation function. So that way if these signals are jointly WSS then we can characterize them with the help of autocorrelation function and cross correlation function.

(Refer Slide Time: 23:59)



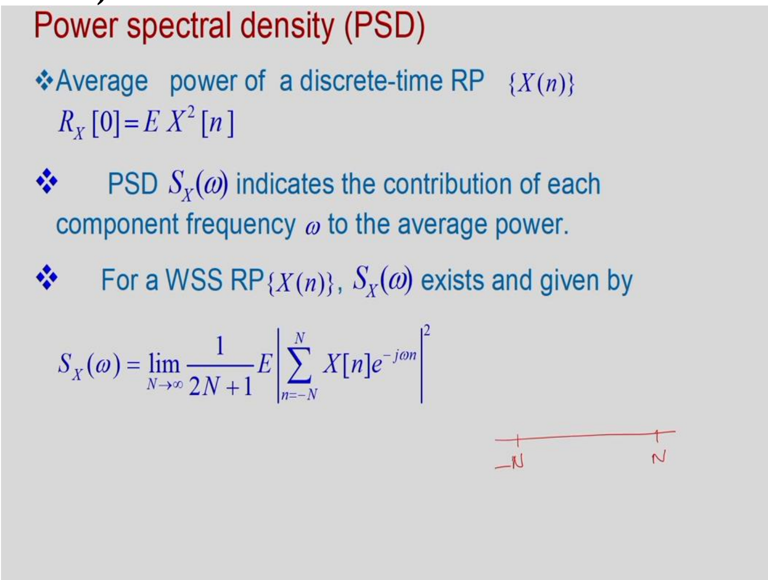
Frequency domain analyses of a WSS signal. We know that in deterministic signal processing the frequency domain analysis plays an important role. We will see if similar representation is possible for a WSS signal. Recall that a discrete time deterministic signals gn has the frequency domain representation in terms of discrete time Fourier transform DTFT. So for example if gn is the signal then its DTFT is given by summation gn e to the power - j omega n and n going from minus infinity to plus infinity.

So this is the definition of discrete time Fourier transform and it is denoted by G omega and it is easy to see that G omega is periodic with period 2pi radian per sample. This unit of omega is radian per sample. So G omega is periodic. Therefore we can find the inverse by this relationship gn = 1 / 2pi integration minus pi to pi of G omega e to the power of j omega n d omega. So will integrate over d omega will get gn. gn and G omega form a Fourier transform pair. Now G omega the Fourier transforms exists if this sum exists.

That is sum from n is equal to minus infinity to infinity of absolute value of gn is less than infinity. In other words gn is absolutely summatable in that case the G omega exist and we can also write a simpler relationship that is instead of mod gn we can write gn square. In that case we will have the condition summation gn square n going from minus infinity to infinity that is less than infinity this expression is the energy of the signal.

So if energy of the signal is finite then it DTFT exist. So this the condition for existence of Fourier transform discrete time Fourier transform of a discrete time signal. But such a frequency domain representation is not possible for a random process because this condition that the energy of the signal is finite not valid for a random process. The definition of stationarity implies that the signal never decays and therefore each energy will be infinite.

(Refer Slide Time: 27:02)

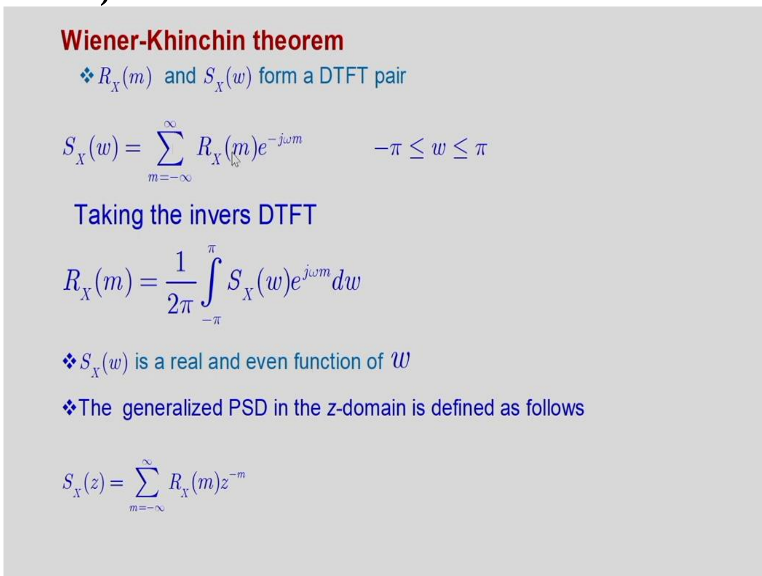


But the power spectral density PSD of a WSS process exists. So let us see what is PSD first average power of a discrete time random process X n is that R X 0 = E of X square n. PSD S X omega indicates the contribution of each component frequency omega to the average power this power spectral density is average power per frequency. So this is the contribution of each component of frequency omega to the average power.

So, that way if we integrate S X omega then we will get the average power. So, integrate E of S square n is equal to integration of S X omega d omega this way we can define S X omega. And now how to define S X omega, So S X omega is equal to limit and tends to infinity 1 / 2N + 1 into expected value this quantity. This is the DTFT of Xn n varying from - N to + N. This is the DTFT of the signal in a window from - N to + N.

So we are considering a window from - N to + N and we are taking the DTFT of the signal within dissidence. Now for a WSS process this quantity exists actually though Fourier transform does not like this quantity exist expected value of this quantity divided by 2N + 1. If we take limit tends to infinity that will be called that the power spectral density. So we have defined power spectral density.

(Refer Slide Time: 29:19)



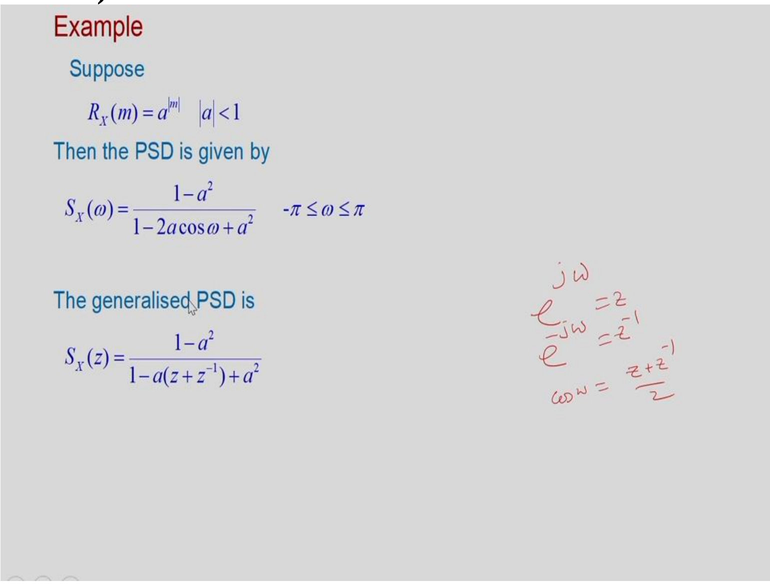
Now we defined what is known as Wiener Khinchin theorem that is the relationship between autocorrelation function and power spectral density R Xm and S X omega form DTFT pair. So they are Fourier transform of each other. So S X omega is equal to that this is this discrete time Fourier transform R X of m into e to the power - j omega m summation of m is equal to infinity minus to plus infinity.

So this is the definition of power spectral density. Now this power spectral density is periodic function of omega therefore we can take the inverse DTFT and we get the autocorrelation function by this integration. So it is in the 1 / 2pi into integral minus pi to pi of S X omega e to the power j omega m d omega. So that way we can find out the power spectral density and from power spectral density we can find out the autocorrelation function from this definition autocorrelation function is an even function and.

Since from this definition power spectrum is a real function, it is a real function because the DTFT is a complex function power spectrum is only real function. And since autocorrelation is even function therefore this part is even and since autocorrelation function is also real and therefore S X omega also must be even so, with this argument autocorrelation S X omega is real and even function of omega. Now it is helpful in signal processing if we have the generalized power spectral density in terms of z transform.

So, is the S Xz transforms of the autocorrelation sequence as S Xz is summation m going from minus infinity to infinity of R Xm into z to the power - h and we can get back R X of m in using this investment formula? So that way that is a complex number now instead of e to the power j omega only we have used z.

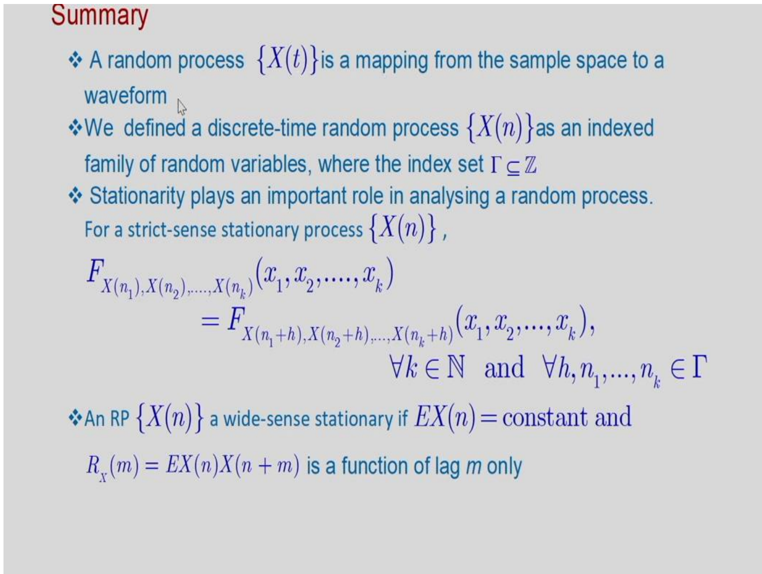
(Refer Slide Time: 31:53)



That will give 1 example suppose R X of m is equal to a to power mod of m mod of a is less than 1 then only it will be become auto correlation function then PSD we can apply the formula of power spectral density that power series expansion. So we will get S X omega = 1 - a square divided by 1 - 2a cos of omega + a square. This is the power spectral density and this power spectral density.

We can convert into the generalized PSD by putting suppose instead of e to the power - j omega will put that is equal z to that it is minus j omega is equal to z inverse. So that if I delete both then cos omega will be z + z inverse divided by 2, so that way we get this relationship so that we introduced power spectrum and generalize power spectrum power spectral density is also called power spectrum.

(Refer Slide Time: 33:05)

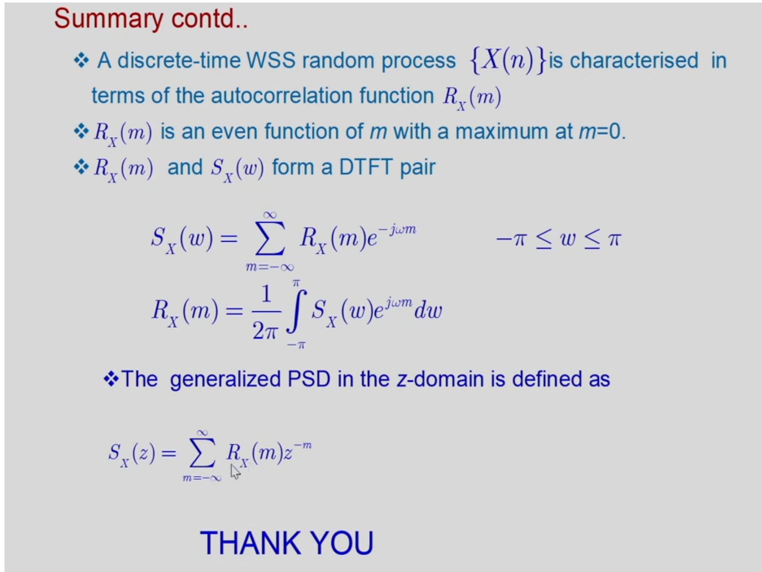


And let us summarize a random process Xt is a mapping from the sample space to a waveform. So this is one important point we define discrete time random process Xn as an index family of random variables where the index set tau is a subset of z we can take z as the index set that this setup integers. Stationarity plays an important role in analyzing a random process for a strict sense stationary random process this CDF at instant n 1 n 2 up to n k is same as the CDF at instant n 1 + h n 2 + h up to n k + h.

So, CDF is invariant under (())(34:04) it time exists. And this is true for all k and this is also all placement of h n 1 n 2 up to n k. So that way this is a very complicated expression which we have to apply to test whether a random processes is strict strictly stationary or not. So to realize this condition we defined it a random process X n is it is a wide sense stationary if E of Xn is constant, and R X of m that is the autocorrelation function that is by E of Xn into X n + m.

And it is a function of lag m only. So it does not depend on and it depends on only the time difference and then we call it a WSS random process we gave one example of WSS random process.

(Refer Slide Time: 35:02)



Also a discrete time WSS random process Xn is characterized in terms of autocorrelation function R X of m because mean is constant anyway. Therefore this autocorrelation function R X of m characterizes a discrete time random process. R X of m is an even function of m with maximum at m = 0. So that we have proved that R X of m at the maximum at m = 0 that we proved in Cauchy Schwarz inequality. Then this is the winner kenchin theorem for discrete time random processes that autocorrelation function and power spectral density form a DTFT pair.

That is S X of omega is the discrete time Fourier transform artist form of the R X m sequence. Similarly R X m can be form from the PSD by this expression and in signal processing the generalized PSD is useful. So it is given by z transform that z transform of R X n sequence and when we put z = e to the power R X omega we get the power spectrum that is S X omega we can get. We also gave one example to illustrate how S X of z that can be obtained from S X of omega. In the next lecture we will discuss the output response of a linear system when the stationary WSS signal is applied as an input. Thank you