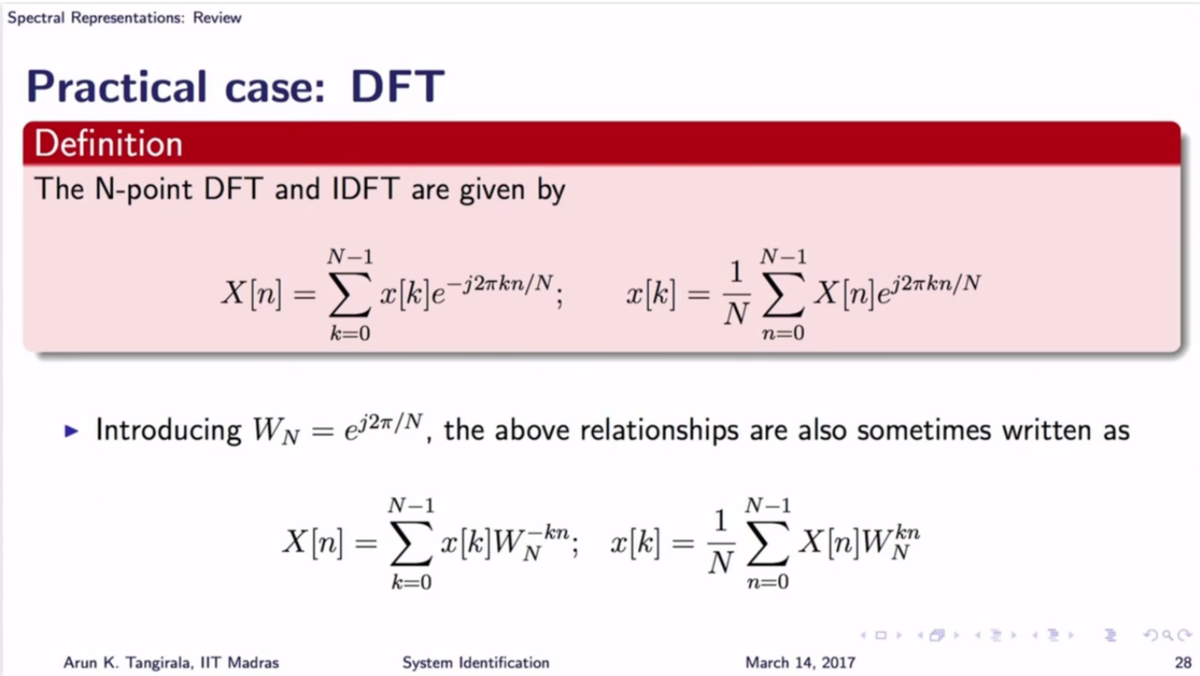
CH5230: System Identification

Spectral Representation 3

Now the one that connects everything in practice is this DFT. So just in transitwe are talking of DFT. What is DFT, we have already spoken about it. So I'm going to, it is nothing but your finite sample and sampledDTFT. We have talked about this.

(Refer Slide Time: 00:29)



And very often you will see this quantity called periodogram being used.Before it I talk about periodogram. What is a consequence or what is an implicit assumption behind using DFT?When I am computing DFT of a finite, I mean DFTis for a finite-length signal. What is the assumption? So the signal outside the interval of observation is periodic. Correct? So in other words we are assuming the signal to be periodic which means that I can only think of a power spectrum. I cannot think of a power spectral density, but when you look at commands in MATLAB or any other software package you will see comments like PSD.

People will use the term power spectral density. How can you think of power spectral density, when I am using DFT,because DFT assumes the signal to be periodic.And moreover you're only computing at a set of frequencies. Are you computing or a continuum? You're not.Computation part a side, the important assumption that DFT makes is that the signal is periodic.Which means theoretically when you're using DFT you can only think of a power spectrum. You cannot, because you've already seen.For periodic signals the notion of density doesn't exist.However to keep things simple, we define an empirical power spectral density and that is what is called a periodogram. This periodogram like your kilogram and so on.Milligram. It tells you how much each period or each frequency is contributing to the power. It's a weight measure. And this was introduced somewhere in the 1890s by Schuster. 1898if Iam correct.For analysing meteorological phenomena, nearly a hundred years after Fourier introduced Fourier transform.

Okay.So what is this periodogram? It's an empirical power spectral density it is not a true power spectral density. What does it do? It simply reports a power per frequency.What is a power, power contribution? The power contribution is mod Cn square.

(Refer Slide Time: 02:50)

Spectral Representations: Review

Periodogram: Heuristic power spectral density

The power spectrum $P_{xx}(f_n)$ for the finite-length signal $\mathbf{x}_N$ is obtained as

$$\implies P_{xx}(f_n) = |c_n|^2 = \frac{|X[n]|^2}{N^2} \qquad (14)$$

A *heuristic* power spectral density (power per unit cyclic frequency), known as the **periodogram**, introduced by Schuster, (1897), for the finite-length sequence is used,

$$\boxed{\mathbb{P}_{xx}(f_n) \triangleq \frac{P_{xx}(f_n)}{\Delta f} = N|c_n|^2 = \frac{|X[n]|^2}{N} \qquad \mathbb{P}_{xx}(\omega_n) = \frac{1}{2\pi N}|X[n]|^2} \qquad (15)$$

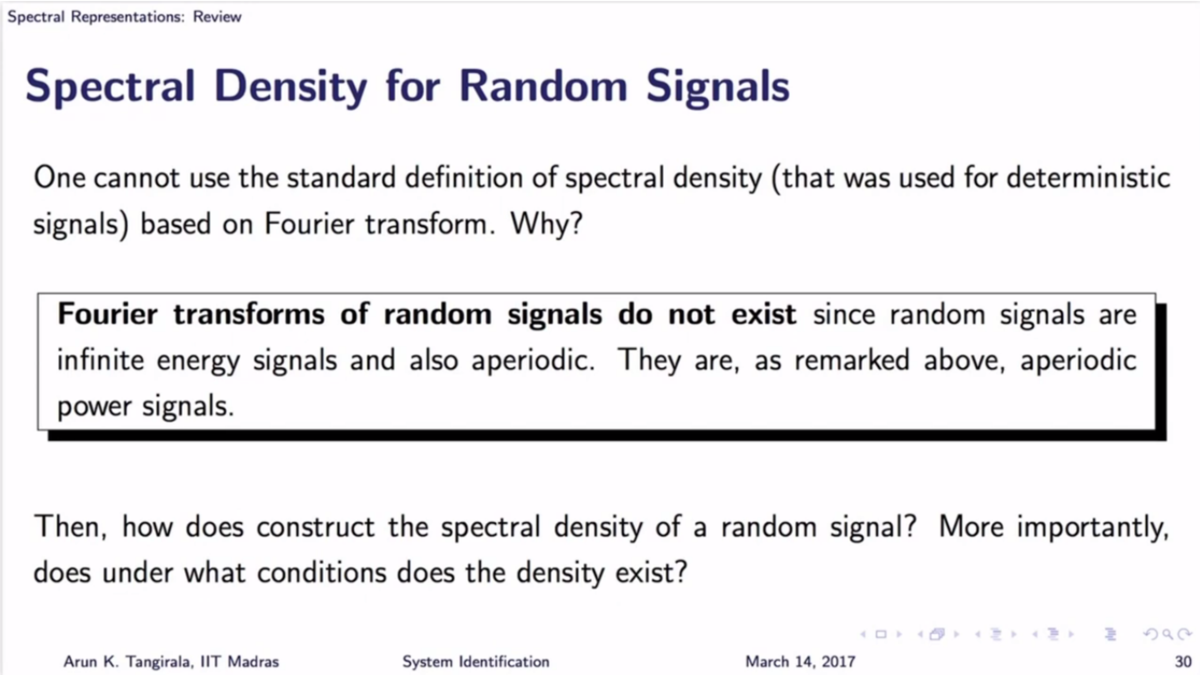
Note the units of frequency in each definition.

Arun K. Tangirala, IIT Madras System Identification March 14, 2017 29

By the way do you recognize this DFT having the same expression as that of a discrete-time Fourier series, right? Except that what is missing here? So X of n is this. But if you look at the Fourier series expansion of xk, what would you do to compute the Fourier coefficient? You would have a 1 over N. So that is the only difference between the Fourier series coefficient and the DFT coefficient. The DFT coefficient is Cn. Sorry. Is X of n. The Fourier coefficient is Cn.What is the difference? Cn is or Xn is n times Cn, right? So X of n is N times Cn. So this is DFT. This is DTFS coefficient, sorry this one coefficient. This is DFT coefficient. So I make use of that and define this periodogram here. Mod C n square, right. So which means C n is X of n by N. Since we work with DFT coefficients the periodogram is redefined in terms of DFT coefficient that's all. So in MATLAB signal processing tool box there is a command called periodogram. It exactly it uses this definition and you have to be careful.Since we are talking of power per frequency you have to be careful the units of frequency.

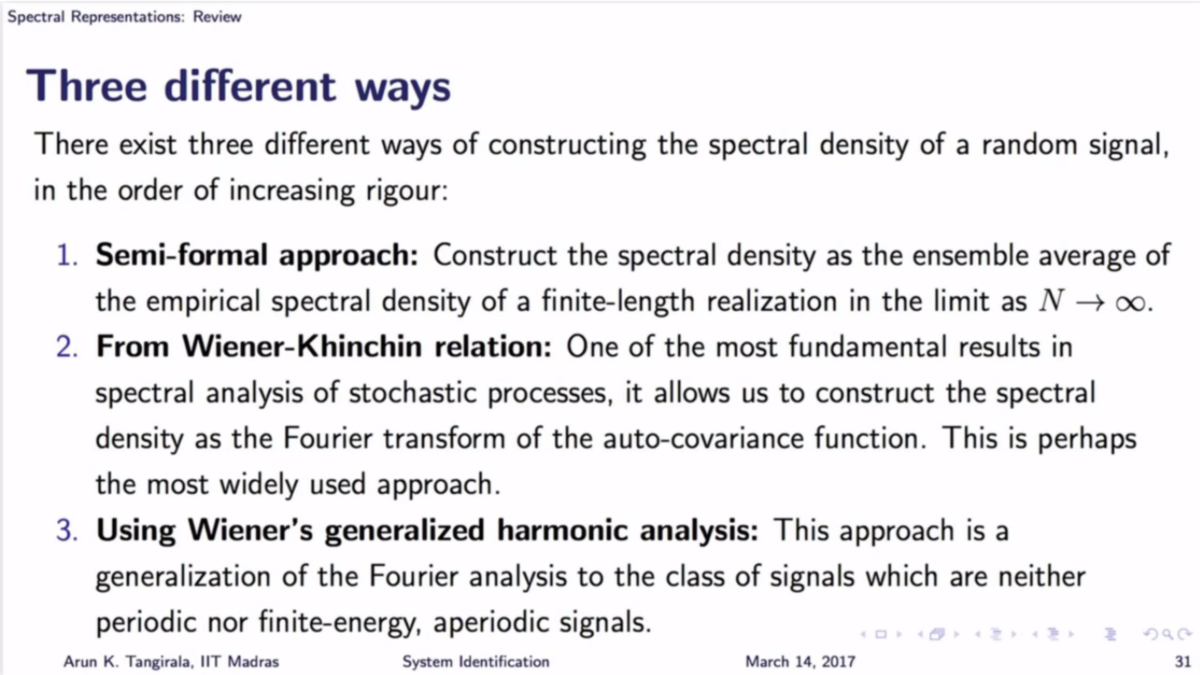
On the left hand side at the bottom, I have given you periodogram when it is expressed in as power per cyclic frequency. If it is power per angular frequency it would be this.And you should, just go and look up in MATLAB periodogram help, whether it returns power per cyclic frequency or per angular frequency. Just as a simple exercise.This should be a habit that you should get into any routine built in routine written by someone. Else you should be thorough with what that person is has assumed. Many do not even know. They just don't even know what that periodogram is about whether it is giving you power spectrum, whether it is giving you power spectral density, is power spectral density is sensible, is it meaningful to be used with, you know, with DFT and so on. But now you are all enlightened and so am I, right? So periodogram is an empirical power spectral density. It is computed from DFT in this following way. You have to be careful with respect to the units. That's all. Okay. So you can also write a code to write periodogram.Now, you know, you just have to take the DFT and write this. That's all. You don't have to rely on any other routine for that matter. The only difference is many packages will give you so-called one sided periodogram or two sided periodogram. What is one sided periodogram? Because it is symmetric, it will take everything on the left and add it to the right, except at zero frequency. So that when you take the area the over all power is obtained. Whereas two sided is as is, whatever you get naturally. So we'll just quickly for 5 minutes talk about spectral density for random signals and that will kind of bring a close. We have already talked about spectral random signals. We know Fourier transformsdon't exist, but we may be able to define a spectral density.

(Refer Slide Time: 06:30)



Now it turns out that the three different ways in which you approach this notion of spectral density. I mean, what I'm giving you is an extremely condensed. What do you say treatment of the problem, presentation of the problem.People had struggled for decades to actually come up with this notion of spectral density. So there are three different ways in which you can arrive at the definition of spectral density. We cannot go the standard route, right, like we went for deterministic signals, because Fourier transforms do not exist. So questions are asked okay, okay, fine, Fourier transform do not exist. Can I think of a density, power spectral density, if I can think of it, what does it mean?When does it exist? These are the questions that were asked.

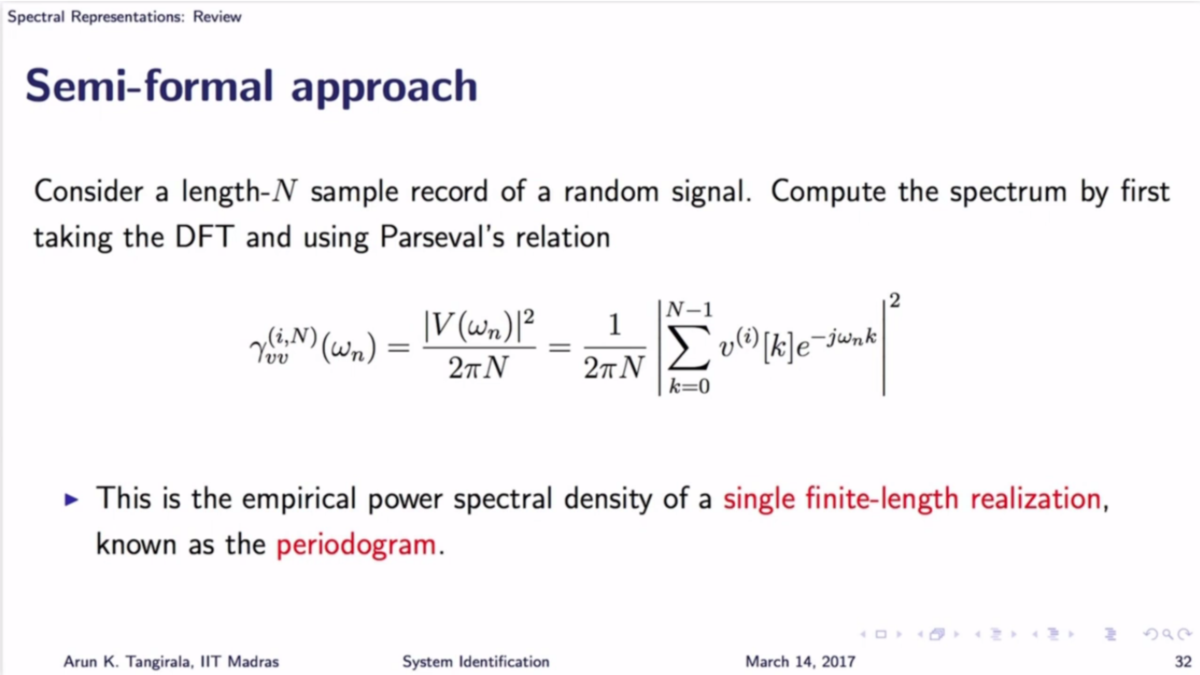
(Refer Slide Time: 07:20)



So the standard approach that you will see in many textbooks today is start with a periodogram. Okay. And one of the most important things that you have to remember is, now that you're moving from the deterministic world to the random world, no longer can you rely on a single realisation.Because you are going to make some comments about the entire process, the random process about all realisations you have to make a comment that is what if you take mean, it is looking at all realizations, if we take auto-covariance function it's looking at all realisations. Spectral density should also be a property of the process rather than a single realisation.Now how we approach that is a very nice way, very simple, just takes a minute to understand. That's why I had to introduce a periodogram to you.So the, what you do is you say I have a finite-length realisation of a random signal of length-N.

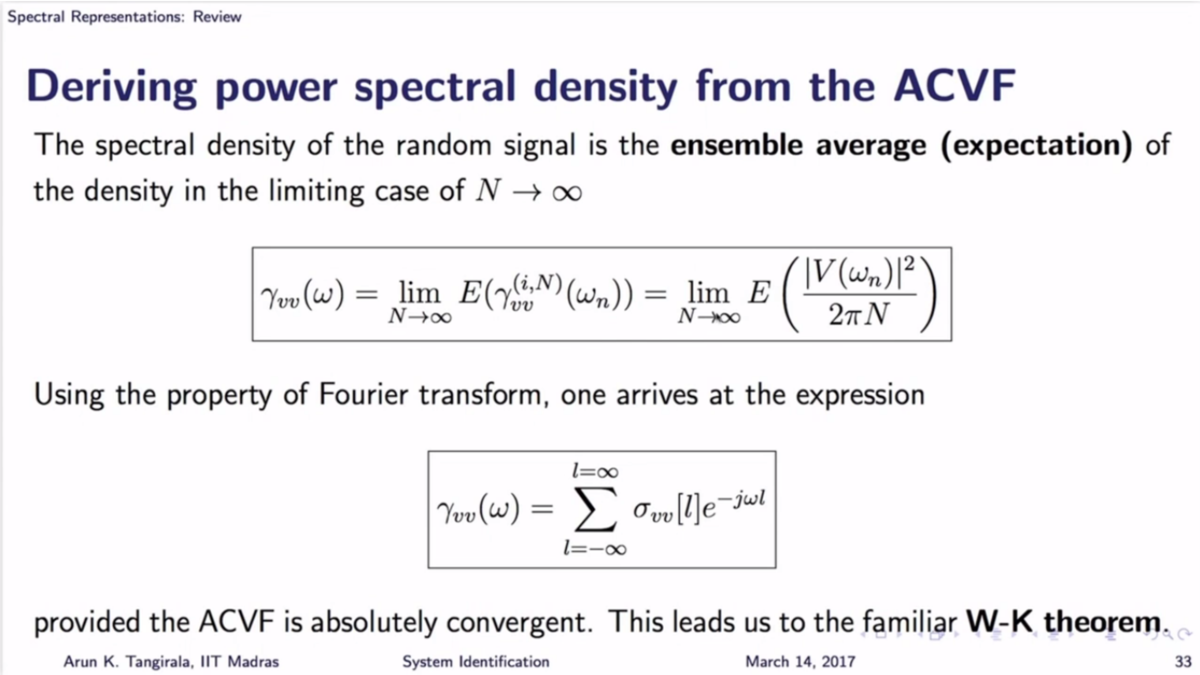
Okay. So what is, suppose I take the DFT of it and then I construct the empirical power spectral density. We know there are already power signals.What would be the periodogram? That is what would be the empirical power spectral density this quantity here, right. That part is clear. But this is a finite-length realisation. Does it represent the entire random process? What are the two things that I have to do? It's finite-length. I have to look at the finite-length realisation and then what else do I have to do? I have to average it across all realisations.

(Refer Slide Time: 08:51)



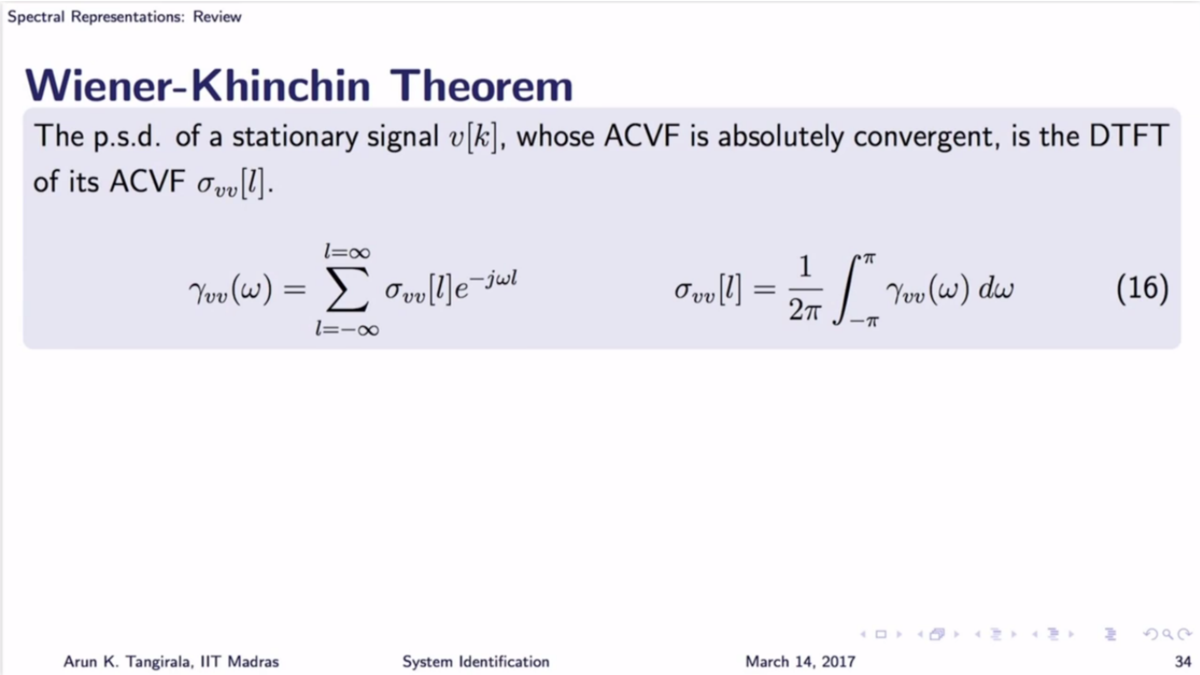
Because any statistical property for a random process is an average across all realisations, correct. So there are two operations that I have to perform. One led the limit and N go to infinity. The length of the realisationgo to infinity and then what?Take expectation, excellent. That's exactly how the power spectral density is defined. The power spectral density for a random signal, this is a semi-formal approach but it works.All you're doing is the inner part here is periodogramof the eighth realisation of length-N. You take the expectation and then let limit N go to infinity. So you average across all finite-length realisations and then let N go to infinity.Now,it turns out that I am cutting down all the math here. It turns out that the power spectral density is nothing, when you apply this definition and work out the math, which takes about 5, 10 minutes, but I'm not going to go over that you can refer to my textbook or any other textbook.

(Refer Slide Time: 09:59)



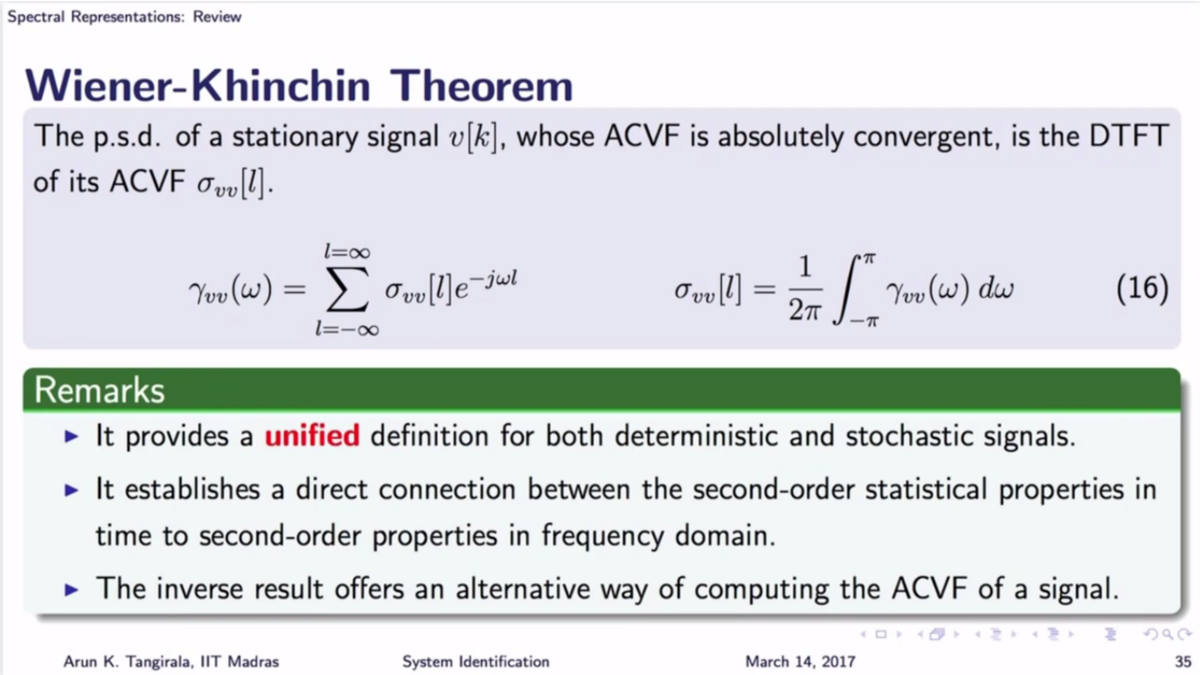
It turns out that the power spectral density is the Fourier transform of the. Which Fourier transform, discrete time Fourier transform. Well, we don't say discrete time here. It's lag. But that doesn't matter time is very generic. It's a Fourier transform of the auto-covariance function.Provided when will this exist. Auto-covariance function is absolutely convergent. That is what is a famous Wiener-Khinchin relation, which says that the power of spectral density is the Fourier transformof the auto-covariance function, if it is absolutely convergent. And I can also write therefore the auto-covariance function as the. There's a mistake here. It should have been e to the J omegal in the right hand side expression sigmaof l should have been. I will correct that in the slides. Sigma of l is 1over 2Pi integral minus Pi to Pi,gamma of omega e to the minus. Sorry, e to the j omega l d omega.

(Refer Slide Time: 11:07)



Okay. When I evaluate this expression at lag zero, what is the result that I get? What is the result that I get at lagzero? By the way if you write in terms of f you should omit the 1 over 2 Pi. That's a, that's understood. So, what do you get at lag zero, when you write that integral? The area under the power spectral density is the variance.Is the power, variance is also the power. Area under the power spectral density is the power, which is also the variance of the signal. So I have a new interpretation for the variance of the signal of a random signal. The variance of a random signal is a measure of the power it in fact, it is the power of the signal, which is very nice.Now the beauty of this relation is it unifies in some sense the deterministic and the stochasticworld. Because we saw a similar relation for the deterministic world, correct.

(Refer Slide Time: 12:23)

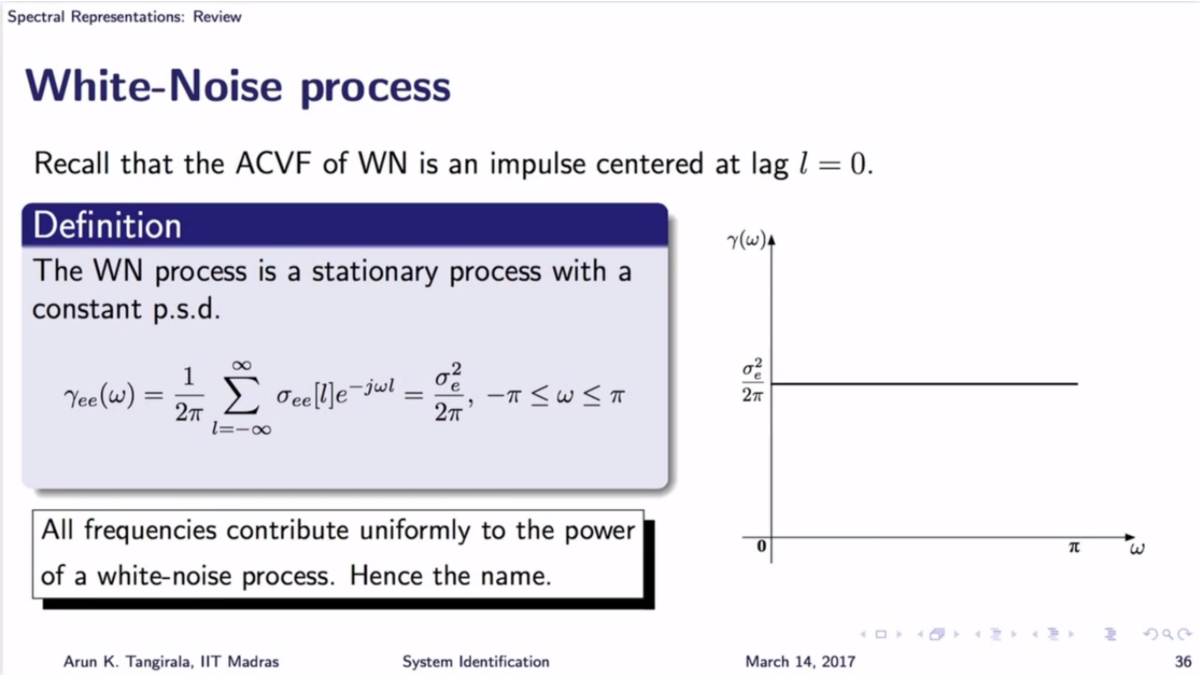


And now comes the mystery and revealed to us as to y white noise is called white noise. Okay. y is it that simply use a Wiener-Khinchinrelation by the way one caution in many textbooks. They say that the power spectral density is actually defined using Wiener-Khinchin theorem. No, that is wrong. The definition of false spectral density doesn't stem from Wiener-Khinchinresult. It's a beautiful relation. It relates power spectral density to auto-covariance function, but the actual definition of power spectral density comes from what is known as Wiener generalized harmonic analysis, which is definitely out of bounds for us in this course. Okay. We'll not get into that, but I'm just cautioning you do not think that this Wiener-Khinchinrelation is defining the power spectral density for you.It is relating the power spectral density and the auto-covariance function. This relation you can use as long as you are convinced that the auto-covarianceis absolutely convergent you can use it. So we'll use it for White noise process. For the White noise process is auto-covarianceabsolutely convergent. Yes or no? How does it look like?

Impulse.

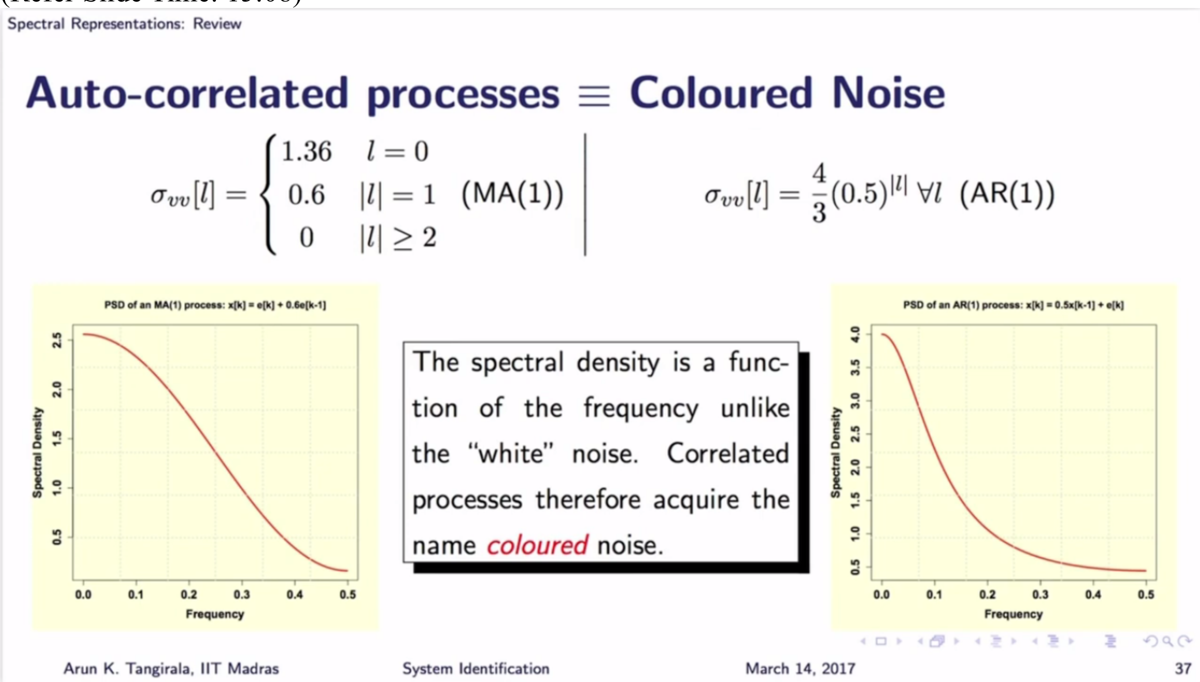
Impulse, that's it. So that means my, the power spectral density of the White noise process is going to be constant at all omega, right? And you see this sketch on the right hand side. That means all frequencies contribute uniformly to the power of the White noise process. Exactly like in your white light, right?White light has all frequencies present in it and that's why engineers called it as White noise process.

(Refer Slide Time: 14:03)



It was not a name given by the statisticians. It was a name given by engineers, because there's an ocean of power spectral density and all of that was found extremely useful by engineers. Therefore correlated processes will give you coloured process,because some frequencies are contributing more to the power than the other frequencies like your coloured light. We know, right. So I'm just showing you examples here for an MA1 and AR1. All right. Where again you plug in this auto-covariance function and then take the Fouriertransform, that is go back to this expression. By the way there is a mistake here in this expression, correct. There should be a 1 over 2Pi here in front of the submission, I will correct that. Okay.So that's it. Any spectral density of any random process if it is not flat, that means it is coloured. That means what? I can build a model, because it's correlated, right?

(Refer Slide Time: 15:08)



It's not white. That means there is a hope of building a model. Can I look at the spectral density and say if it is AR or MA, sometimes yes,but not always that you have to go back to ACF and PAC. Okay.Final thing that I want to say is we have learnt how to obtain power spectral density from auto-covariance. But that is a non-parametric way of obtaining the power spectral density. You can also obtain power spectral density like we brought in a relation for energy spectral density remember in terms of the transfer function or F or F. Likewise, if I have h as a transfer function the power spectral density of v is simply modh of e to the minus j omega square multiplied by the power spectral density of white noise, but power spectral density of white noise as a constant. So effectively the power spectral density of the random signal has the same shape as mod X square. That's all. So this is used in so-called parametricspectral density estimation. What we mean by parametric spectral density estimationis, if I give you a signal now, you have two options or two routes for estimating spectral density. What is the first route? Estimate the auto-covariance function. And then take the Fourier transform. What is second route?Build atime series model. Get a time series model AR, MA, ARMA whatever and then use this relation.There are advantages to both. The first one is called anon-parametricapproach, where I don't fit a time series model. I don't assume any structure. The second approach I'm assuming a certain structure and fitting a model that is called a parametricspectral density estimation approach, which gives you very smooth estimates of power spectral density. Ideally that is what is preferred.You want smooth estimates of the power spectral density.The one that we talked about earlier has its own challenges, but one risk is if you estimate the model incorrectly. Then your power spectral density can go for a toss. So you have to be careful. All right. So that's it. I just want to conclude. Of course, you know, this is just an expression showing you how to construct the power spectral density of an ARMA1, 1. And by know, you know, a linear random process is a filter. So this is a sample and I just want to conclude by mentioning this cross-spectrum and coherence, we'll talk about it a bit later. But this is just a straightforward extension of the WIENER-Khinchin relation.Wiener-Khinchin relation says spectral density is the Fourier transform of auto-covariance. A similar result exists cross spectral density between two random signals is nothing but the DTFT of the Cross covariance function.This has the same interpretation as the cross covariance function. The cross covariance function allows us to relate two random signals in time.Cross coherence function allows us to relate two random signals in frequency. The users are the same. You can use that to estimate delay. You can use this to figure out what is the nature of the filter that is the estimate the transfer function and like we wrote a relation earlier you can write across spectral density is exactly what we saw earlier for deterministic signals.

(Refer Slide Time: 18:57)

Cross-spectrum and Coherence

The cross-spectral density tests for linear relationship between two series, as the CCVF:

$$\gamma_{yu}(\omega) = \text{DTFT}(\sigma_{yu}[l]) = \sum_{l=-\infty}^{\infty} \sigma_{yu}[l]e^{-j\omega l}$$

- It is a complex quantity!
- $|\gamma_{yu}(\omega)|$ gives the strength of common power at that frequency
- $\angle\gamma_{yu}(\omega)$ **(phase)** is useful in estimating delays in the system
- Important result:

$$\boxed{\gamma_{yu}(\omega) = H(e^{-j\omega})\gamma_{uu}(\omega)} \qquad (19)$$

And finally like we normalize cross correlation, across covariance to get cross correlation you have coherence, which is a normalized version of cross spectral density,but do not be mistaken. Only the cross spectral density is a Fourier transform of cross covariance. Coherence is not theFouriertransform of cross correlation. What this means is.When you want to compute coherence you're the first compute cross spectral density and then normalize.Don't come cross correlation and normally and then say take the Fouriertransform and call it coherence.And there exists a beautiful result. The magnitude of this is called coherencewith says a system is LTI. So we have come in full circle now.A system is LTI if and only if the coherence is unity at all frequencies.

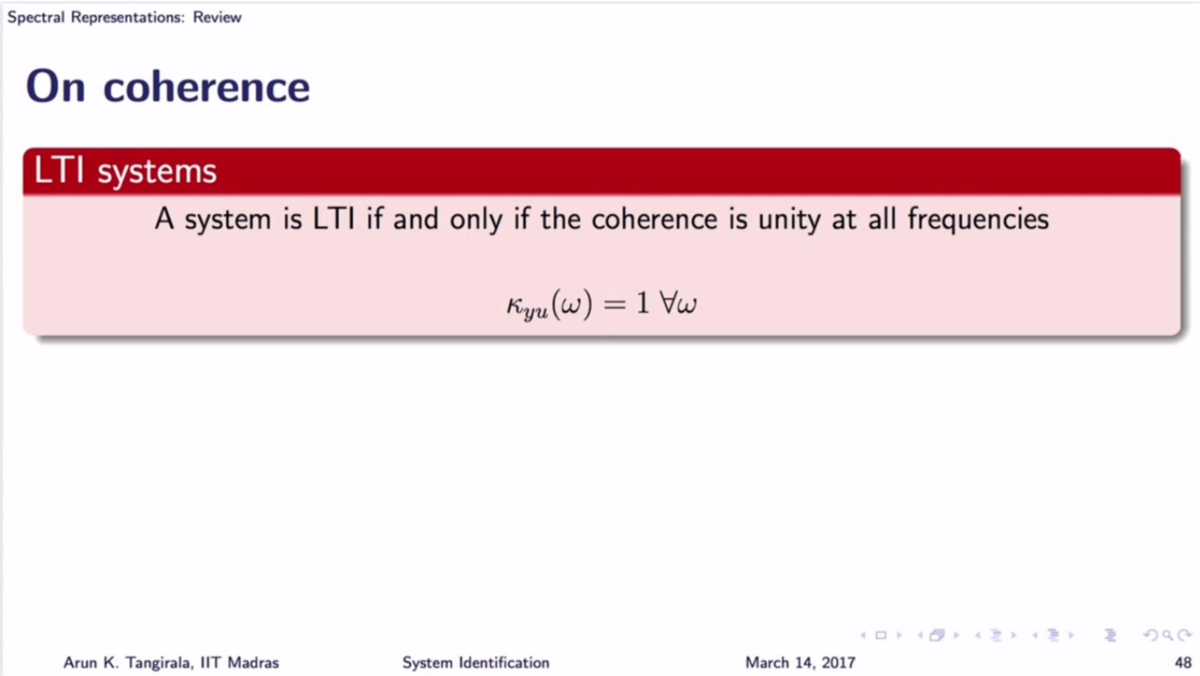
(Refer Slide Time: 19:36)

Coherence

Just as the normalized CCVF, namely, CCF is used to test linearity, a normalized CPSD, known as **coherence function**, is used to test linearity:

$$\boxed{\kappa_{yu}(\omega) = \frac{\gamma_{yu}(\omega)}{\sqrt{\gamma_{yy}(\omega)\gamma_{uu}(\omega)}}}$$

The magnitude of coherence function is **coherence**.



Now this coherence we have defined in terms of power spectral density.For pure deterministic systems you can define coherence in terms of energy spectraldensity. There is no problem. It's the same thing. So the result say is a system is LTI if and only if the coherence is flat. Do you remember in the liquid level case study I mentioned this? I showed you the coherence plot and we talked quite a bit about it. Why the coherence dipped below unity and so on. Same, why does a correlation dip below unity, either because you may have noise and or you have non-linarites in the system. Same story here, coherence but the beauty with coherence is that you do not have to know the delay or anything you can just take the signals input and output compute the coherence plot it as a function of omega and then you will be able to say how linear, whether a linear model will suit the system on. Okay.So tomorrow, from tomorrow we'll get back to the fully deterministic stochastic world. And now our journey is to put these two together. Learn how to make predictions and learn how to estimate the parameters. So that's the rest of the journey for us and society. One thing we have not done is, we have not studied the random processes in the state-space domain. Okay. That will worry about later on. We are still in the input output world. Okay. So we'll see you tomorrow.