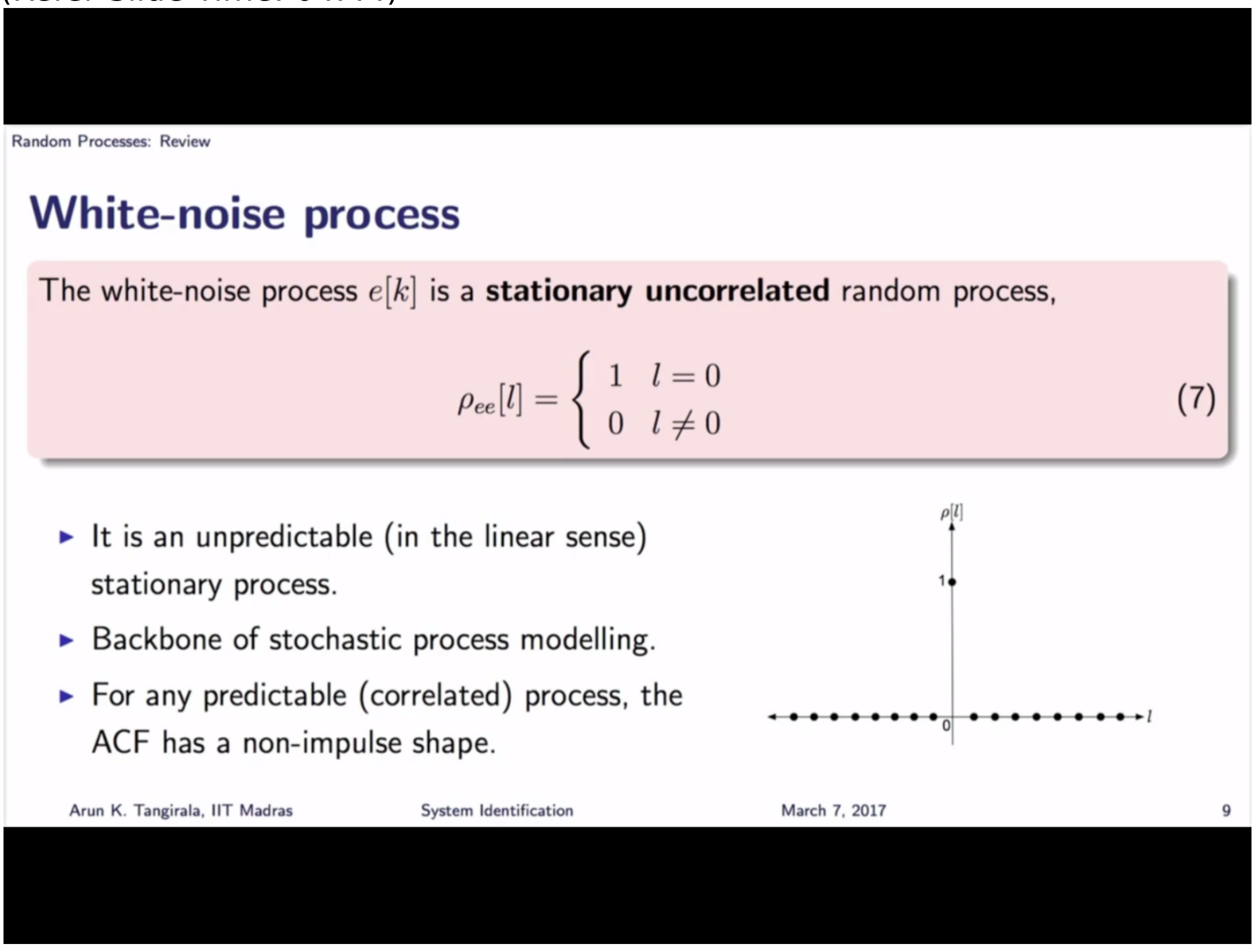
A very good morning. We are in the midst of reviewing the theory of random processes, particularly modeling them and yesterday we learned some critical concepts such as stationarity, ergodicity, and also discussed the notion of autocorrelation function, which is at its heart correlation measure. As I keep saying, do not think of auto correlation function as an alien measure or a new measure, it is at its heart correlation, but apply to observations within a random process. And the need for introducing auto correlation, I have already mentioned one or two needs, one is of course to see if there is hope for any modeling. And I don't… keep saying this repeated, we are only interested in linear models and that's why we are looking at correlation based measures. So the first need for introducing auto correlation function is to let's for predictability and then of course there are other very nice things that come out of this definition and the first gift that auto correlation gives is this notion of a white-noise process, alright. And as I explained to you yesterday, white-noise process is a stationary uncorrelated random process, that's a formal way of looking at it, but conceptually what it is, it's it is that ideal unpredictable process in a linear sense. And why do we introduce this concept, one when I am sitting down to model a series, that is a time series, then I need to know whether the series offers any scope for predictability. So in some sense you are benchmarking any given series against this ideal process and you are seeing if it has properties of this white-noise process. If it does have, remember when I say, we check for series being white or not, we are not looking at values, we are looking at properties. So when we perform, what is known as a whiteness test, what we are actually doing is we are testing if the series has, given series has white noise like properties and this whiteness test we conduct at the beginning of modeling, during modeling and also towards the end. Because at every stage we want to make sure that whatever we are beginning with has some predictability and then during the course of the modeling, if the residuals have any predictability left in them and if they do then again you go back and refine your model and the final model that you settle down with should give you residuals that pass the whiteness test. So this whiteness is an extremely important step in any modeling exercise and that's why it's important to be comfortable with this notion of white-noise process. So the white-noise process itself has so many advantages or so many, I would say functionalities, attached to it. The other role of whiteness process… sorry a white-noise process is that it serves as a fictitious input, remember have been saying that. A stationary random process satisfying certain conditions can be thought of as white noise passing through a filter. So that is the other very important role that white-noise process plays in time series modeling. We have not yet discussed the conditions… which such a representation can be given, but we have already, in fact we are already on our way. We have discussed stationarity first… the first requirement is that the series should be stationary for us to be able to represent it as white-noise passing through a filter. Then there is another condition which is on the spectral density and we have not defined what a spectral density is. I just used that terminology yesterday, but we will talk about it a bit later. So let's move on now and I have already spoken of the role of white-noise in time series modeling, when I say time series modeling, it's to be understood as (inaudible) as well. Now as I just said, one of the critical uses of white-noise is... is imagining it to be the fictitious input that is driving the given series. And let me give you now a peak into that, I mean, give you a feel of how it ends up serving as an input. Although I am not going to do this formally, but I am going to do this somewhat informally.

(Refer Slide Time: 04:44)



So let's take a series VK and we know by definition, a random process always has an unpredictable component, but we hope that there will be a predictable part as well. So let's split it into two parts as V hat of K given all the information up to K-1 + an unpredictable component, which is what is the white-noise. Straight away you can see white-noise process appearing as an indispensible component of the random signal, right. Whether V hat exists or not, E will definitely be a part of VK. At this moment you see white-noise as an indispensible component, as an inherent component of VK. Very soon we will show that this same white-noise process, which is inherent to VK can be thought of as driving VK, that is what we mean by input, right. How do you see that. Well, assume now any form for this predictor, see this V hat can be anything, as I have explained to you the notation, V hat of K given K-1 means that I am predicting V, the value of the signal at K, given all the information up to K-1, when I say information, observations. I can do this in infinitely different ways, right I can predict, I can take the average, I can take some mathematical function and so on. Let's take a very simple mathematical function, which is a linear function. So assume that, now let us say that we are going to just use one information, maybe we can use two in… observations in the past. So I am going to work with a predictor of this form. Very soon we will see that predictors of this form correspond to what are known as auto regressive models. You can see the terminology autoregressive there, it is regressing on to itself. So assume, I am not saying all processes will this, please don't make under impression, assume that there exists a process, for which the prediction is of this form, optimal prediction is of this form, it's a linear predictor. Now plug in this predictor expression into this equation here so that you can straight away write this difference equation form for VK. Do you see that, all we have done… there is nothing mysterious about this. All we are said is, we have… we have used the tenets of random process modeling or theory of random processes. We say that random process inherently contains an unpredictable component and if it contains a predictable component, let us assume that this predictable component has a liner form and we have plugged that in and you can see that there is… now VK has a difference equation form to it. So which means by assuming that predictor, this is the predictor or sometimes we will say prediction and this is model. What is the difference between a model and the predictor. It's important to know this difference. Do you see any difference? Okay, not necessarily, okay fine, fair enough, but anything else?

(Inaudible)

No that's not the case, if you… sorry… let's hear one more answer.

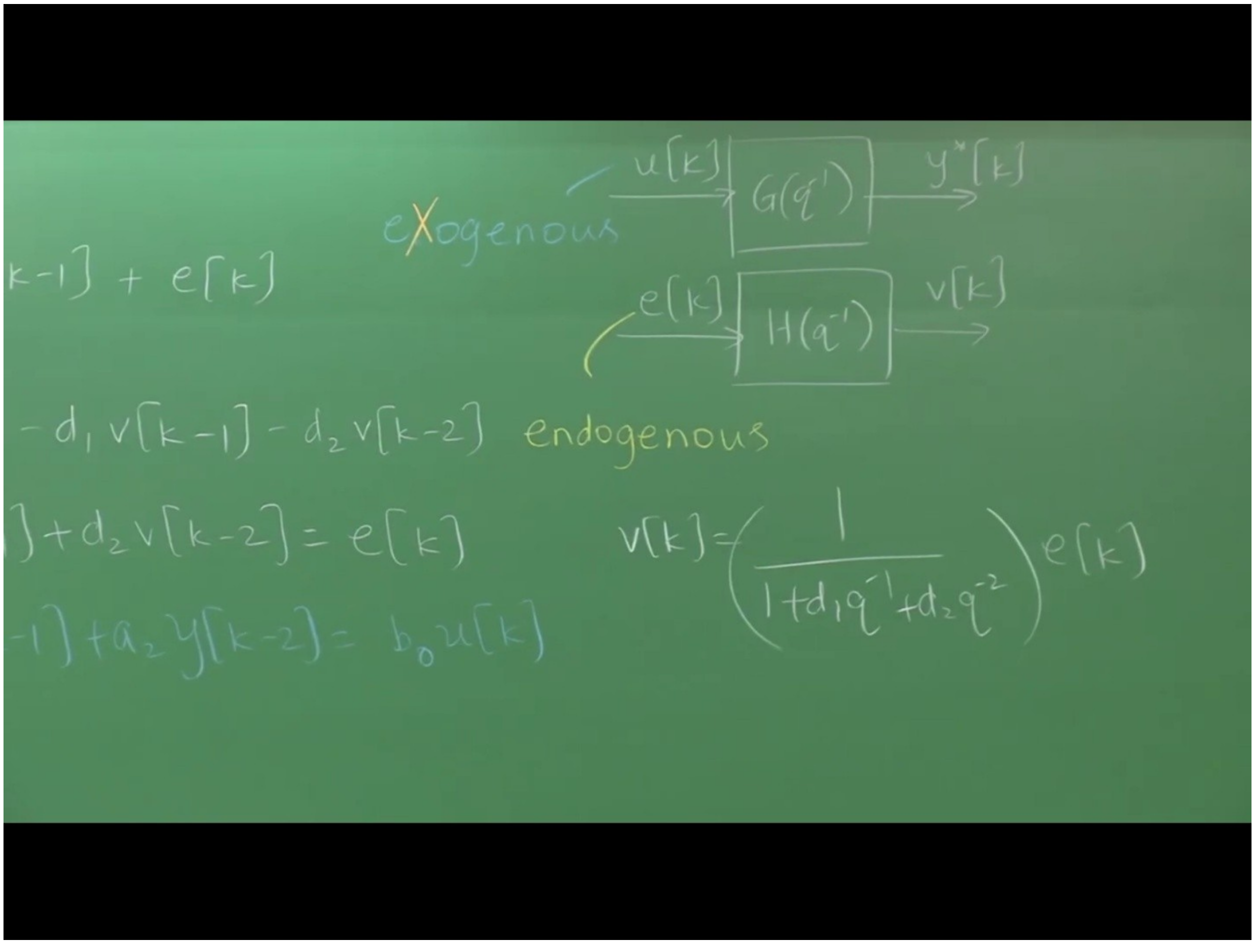
(inaudible)

Okay, good then…

(inaudible)

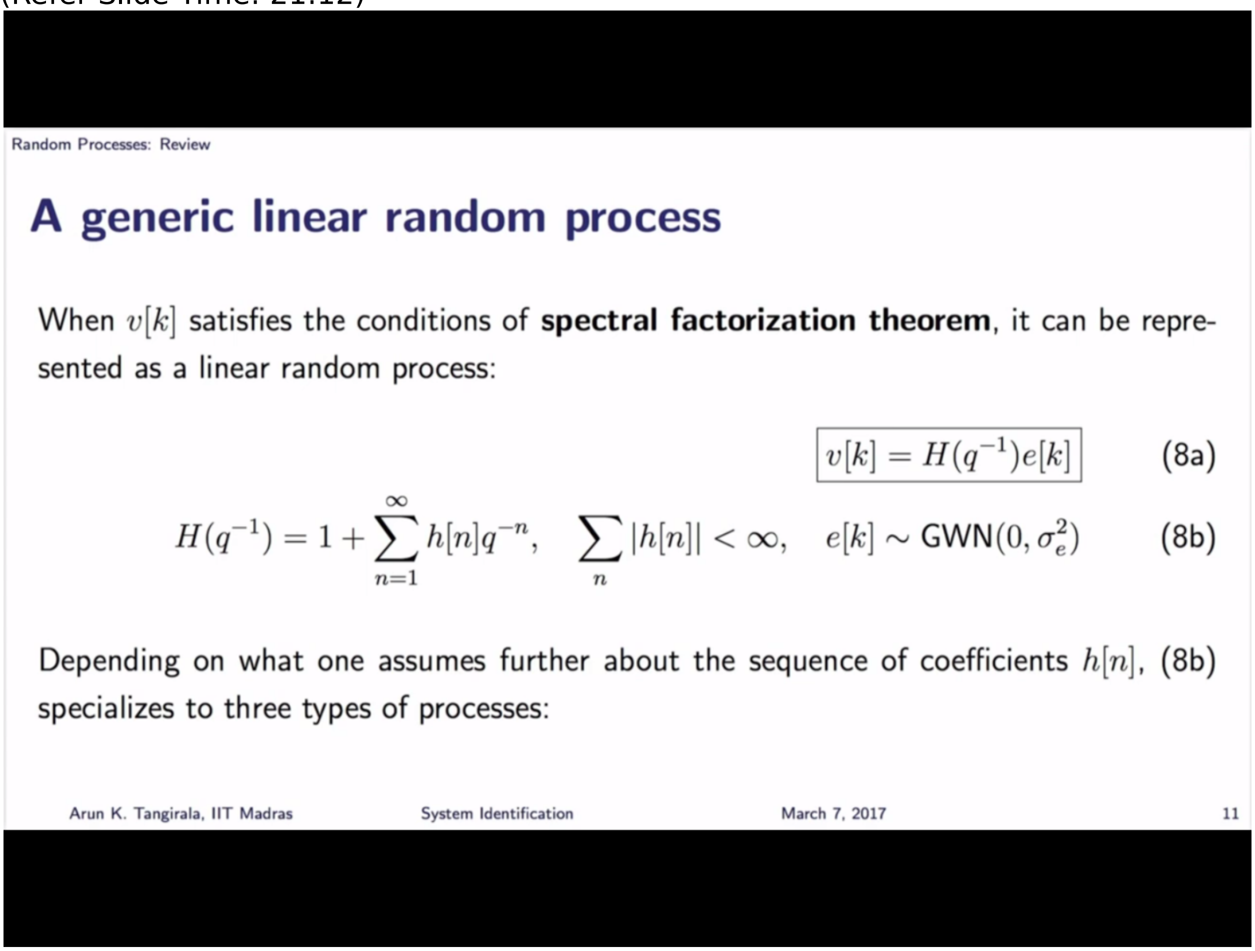
Okay first part was correct. There is a very subtle, but important difference between the model and predictor. In fact model gives you a complete description of the process, okay, in what sense, you are right, so predictor tells me how to compute or how the signal… what will be the value of the signal at a certain instant in future. It could be one step, it could be P steps, doesn't matter. Whereas model tell me both, what will be the prediction and what is being left out, okay so if you… that means in going from here to here is not so straight forward, you have to… this is only giving you one part of the picture, this is giving you the complete picture, in going from here to here, you have to make an assumption of what is being left out, that means what this predictor has not managed to capture. Predictor is only telling you what will be, I mean, it's… it's only estimate of what will be the signal, but it doesn't tell you what it is unable to predict, whereas model is telling you both, from the model you can derive the predictor, but from the predictor you cannot derive the model in a straight forward way, unless you make an assumption that whatever, the predictor couldn't predict is unpredictable, where is the guarantee, right. There is no guarantee that all predictors will leave out an unpredictable component, suboptimal predictors will leave out something that is still predictable. We assume that this is in some sense and optimal predictor, although I have not stated that, in the sense that whatever, is left out of this one-step ahead prediction, that is unpredictable, which is the white-noise. So you should keep this difference intact and remember this. Later on we will again come back to this concept of the notion of a model and a predictor, but the fact is you can derive predictor or prediction expressions given a model, but deriving model given predictions is not necessarily straight forward, unless you make additional assumptions of what is being left out, okay. So now coming back to the discussion here, do you see a striking similarity between this difference equation and the difference equation that you have seen for deterministic LTI systems? Is there is a big difference symbolically or structurally? It doesn't… I mean if you were to write a second order difference equation, you would write, how visible this is? Is this, yeah… So you would write YK+A1 YK-1+A2 YK-2, A = some, let's say B1 UK-1 or may be B naught UK, if there is no delay and so on. So straight away I can give an interpretation to this difference equation that VK is the response of an LTI system driven by EK. Of course I mean, it's a very inform approach. I have to establish a few other things, but at least you get this feeling straight away by looking at the similarity between this difference equation and the difference equation that you have seen earlier. By the way this difference equation is called a stochastic difference equation, because the forcing function is a stochastic signal, that's a big difference between the one that we have seen earlier and that… one that we see on the board right now. Now let us not worry about, under what conditions I can write this, whether this is going to be the form and so on, at least now I am partly convinced and in fact largely convinced that it may be possible to give an LTI representation to a given random process, but not to all random processes, that should be remembered. The random process has to satisfy some conditions. The same goes with deterministic processes, right. The deterministic process has to be linear. When can I write a difference equation form, the deterministic process has to be linear and time invariant and the response should be in a parametric form, right, there are three conditions, but the main conditions are LTI. Likewise you can expect some LTI like conditions on the random process, but we don't use the term LTI, we do use the linearity term, but in place of time invariance what do we see… use? Stationarity. And in addition, because this EK is stochastic, I have to impose one more condition, which is on the spectral density, but otherwise, the LTI assumption carries forward to the stochastic world as well. Linearity carries forward. I have not defined linearity yet, because it was easy to define linearity in the deterministic world, since there is an input and output, I can define linearity very easily, but for a random process like VK, which I don't know is being drive by god knows what input, I cannot define linearity straight away, unless I think of some fictitious some input. Now if I think of the EK as a fictitious input, I can define linearity. I can define Stationarity. I mean stationarity is defined, but notion of time invariance, everything comes in, correct. So I can write here for example, a transfer function form to VK and write here VK as 1 over 1+D1 Q inverse + D2 Q inverse square operating on EK. That's your H of Q inverse, right the 1 1 over 1+D1 Q inverse + D2 Q inverse square is your H of Q inverse. So this is a decent way of getting an idea as to how random processes being end-up as the… as the output of LTI system driven by white-noise input. So look at the beauty. In fact long ago, I would say about seven years ago, when I gave this perspective, initial days I wouldn't give this perspective, but gradually when, you know, when you teach for a long time, you get perspectives. So when I gave this perspective, there was one gentleman who came in the class, this is really surprising as to an inherent unpredictable component turns out to be the input that it is driving VK, but that is the feature of self exciting processes, random processes, because we do not know the causes, you assume that they are self exciting, there is something within the process that is pushing the process forward and that something here is EK, right. So it's as if you are pulling out a wire that is something within a process and you are just pulling it out. It is not truly external to the process, it is just for our own notation and understanding, we write this. However, one should not be drawing complete parallels with this representation or Y or Y star, doesn't matter. The top… in the top one U is exogenous to G, in the bottom that is for the random process E is endogenous to the random process. Although I have shown EK to be coming, you know as if some external think is coming in and exciting H, it isn't, it is a part of VK… it is a part of VK that is driving it. We don't show that here in this block diagram, but you have to understand. So this E is endogenous, where the input here is exogenous. These are the terms that are used in the time series and (inaudible) literature. And it's important to have here this as capital X, so that you will understand why this arcs and R max come in later on. What we have here is an autoregressive model, that's it AR2.

(Refer Slide Time: 17:46)



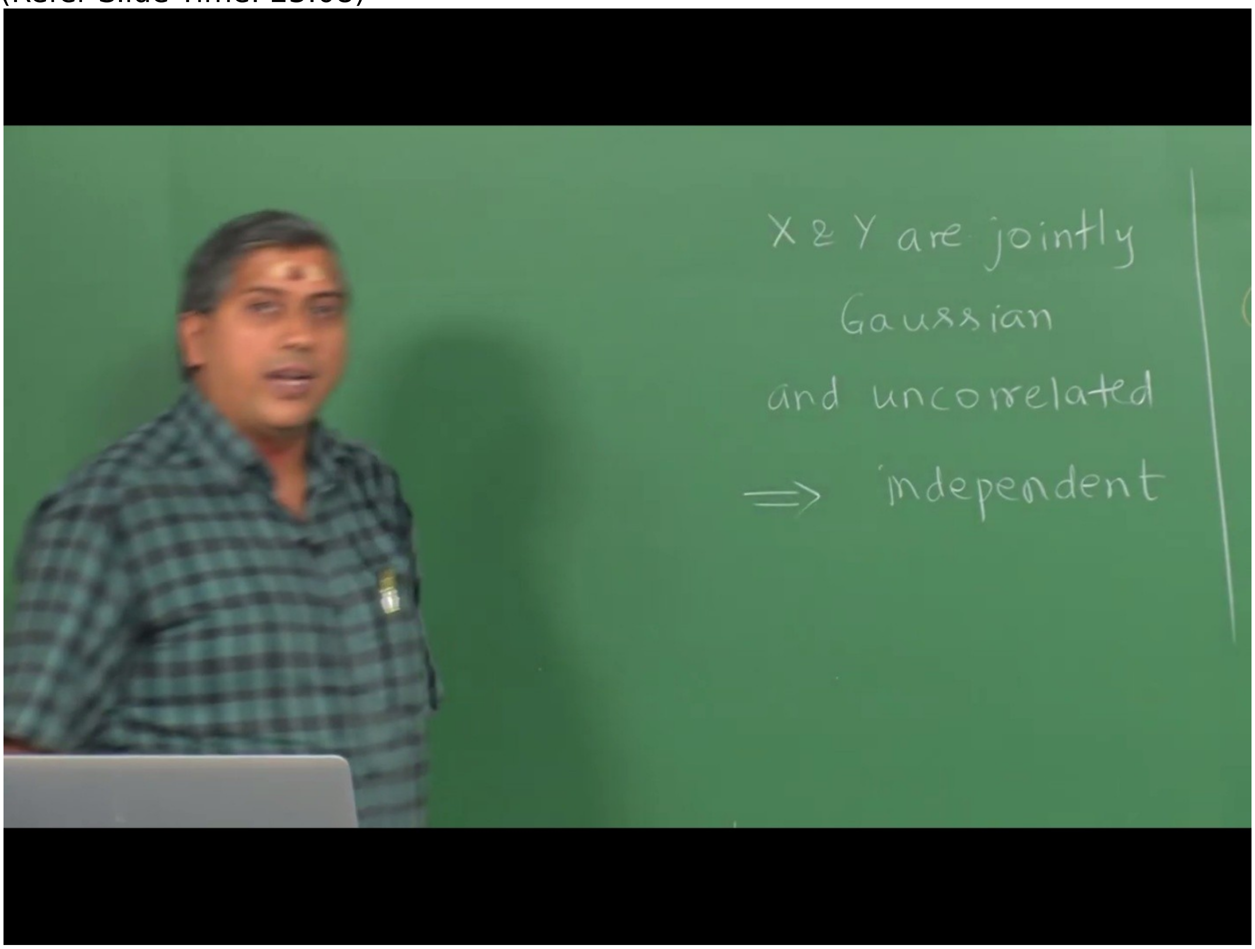
So now we move on to this notion of linear random processes, where I generalize this idea. At this moment, I am not telling you what class of random processes are friendly to this kind of representation, but let's, there exists a large class of process and the idea here is now we represent VK as EK driving an LTI system. Remember very well that this EK is endogenous to VK. The moment I say that there is this situation, I can define a linear random process to begin with. In fact a linear random process is one that can be given this convolution form of representation in addition we require that, you see this condition here, what does this condition remind you of? In deterministic processes, this was a condition for stability, here this is a condition you can show for stationariry. Because one of the key requirements of stationarity is that the process should not blow. It's not necessary and sufficient condition, but it is a necessary condition. There are many processes, whose amplitudes do not blow up, but can be non-stationary, but a necessary condition is that... a necessary condition is that the amplitude of VK should not blow up. There is a chance, right because you are doing an infinite summation there. Now having guaranteed that EK is stationary by definition, because EK is white-noise, correct and white-noise is stationary. So putting together these two conditions that EK is stationary and the coefficients are absolutely convergent, you can show that VK is stationary. So notice now a big difference between the notion of linearity in the deterministic world and in the random world. In the deterministic world, straight away I can define what a linear system is because there is an external input and an output. All I think off is a system has a functional mapping

and I just turned to linearity definition and mathematics, but here I have a problem right from beginning, because VK is assumed to be self-exciting and there is no external… although there is I do not know and therefore I cannot define linearity up front, unless I think of some fictitious input that is driving VK and that fictitious input happens to be VK, alright I have cut shot a lot of formalization here. Just to give you a feel now that VK is being driven by EK. Now that I think of VK being driven by EK, straight away I can define linearity, I can define what are the conditions for stationarity and so on, okay. And also you must notice that I have already fixed the first coefficient to one, we have talked about that for uniqueness. And EK by default here is Gaussian white-noise. There are a few texts and may be in early developments, which are… where things are assumed to be, sorry EK is assumed to be IID. The difference between white-noise and IID, what is IID? IID is identically and independent… independent and identically distributed process.
(Refer Slide Time: 21:12)



White-noise only requires uncorrelated nature to be satisfied. Independent process, IID process has to satisfy two conditions one is independence, which means there is no relation between any two observations. That means there is no hope at all for billing any model. White noise says, there is no hope for a linear model. Identically distributed, which means that every observation falls out of the same distribution. So it’s a lot more stringent. EK being IID or any process being IID is a lot more stringent than a white-noise requirement. However, Gaussian white-noise process alone has a property that it is also an IID process, okay. White noise process is, a white-noise requirement is less stringent then IID requirement. That means not all white-noise processes are necessarily IID, but all IIDs definitely are white. The only exception is Gaussian. All Gaussian white-noise processes are IID processes, why because what does whiteness mean, any two observations are uncorrelated. In addition, we are saying Gaussian, which means these two observations have a joint Gaussian distribution. You can show that for any two random variables, if X and Y are jointly Gaussian, this is a standard result in the theory of random variables and uncorrelated, I have said this earlier also, then they are also independent, that means their joint p.d.f. is factorizable as well, that's a fairly straight forward result to show.

(Refer Slide Time: 23:08)



You start with the joint p.d.f., that is joint Gaussian p.d.f. and make use of the fact that they… they are uncorrelated, which means the covariance matrix of X ray and Y is going to be diagonal, then you can straightaway factorize the joint p.d.f into product of two Gaussian p.d.f.s. So that means you have proved independence, but this is only true for Gaussian white-noise, so it is safe to say EK is Gaussian white-noise, you don't have to say it is IID, okay. so we will assume throughout the course that the by default, unless otherwise, stated, white-noise is Gaussian white-noise, any questions? So this is a definition of a linear random process and we are now introduced to this linear LTI representation in the same… now in the same way as you were introduced to the LTI representation of deterministic process.