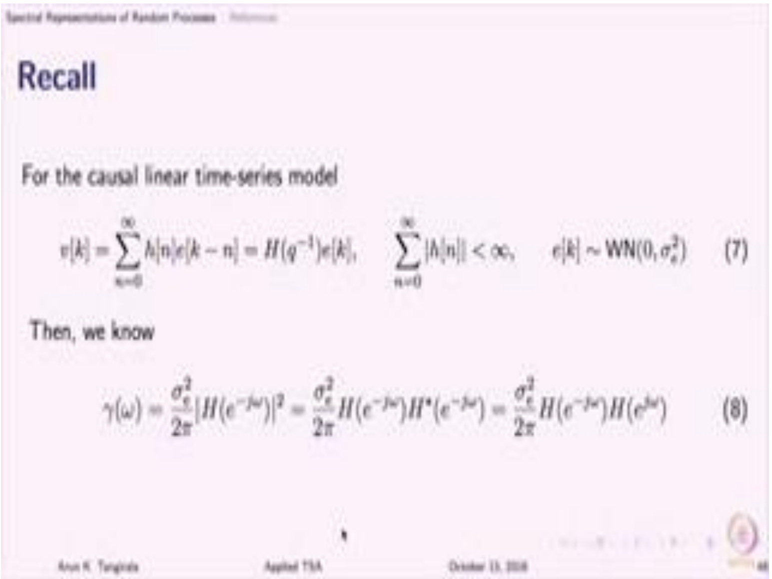
Applied Time-Series Analysis
Prof. Arun K. Tangirala
Department of Chemical Engineering
Indian Institute of Technology, Madras

Lecture – 77
Lecture 34B - Spectral Representations of Random Processes 7

The final straw in the hat is Spectral Factorization. It is nothing new; we are not going to learn anything per say but what we are definitely going to learn is the conditions under which I can build a linear model for a given random process. So, now we are asking the most fundamental question we have discussed already AR models, MA models, ARMA models, we have talked about spectral densities and so on particularly we are talked about I am focusing on AR, MA and ARMA models, but we have evaded the question under for what class of processes I can construct such time series models. And of course, I have briefly talked about it but now is a time to go a bit deeper.

(Refer Slide Time: 01:04)



So, if you recall here the questions now that we are asking is when can I build a linear model. And if you recall the linear model, this is essentially your linear model you should be able to express v k as a linear combination of past present and future shock waves subject to the condition that the coefficients are absolutely convergent and e k is a white noise.

Now, we know already from spectral representation that the spectral density of such a process is what we have seen already; sigma square e by 2 pi times mod h of e to the minus omega square. So, if you look at it the problem of fitting a time series model can be recast as a problem of factorizing the spectral density into three factors. What are those three factors? Essentially actually two factors we do not talk of three factors, but let us talk say three right now.

Student: Sigma square e.

Sigma square e 2 pi will ignore fine that is a matter of choice of omega, sigma square e.

Student: H of (Refer Time: 02:15).

H of e to the minus j omega and h star its conjugate, correct. Normally we do not include sigma square e as the factor; we say that that is a multiplicative constant. We refer to h of e to the minus j omega as a spectral factor. Now obviously, the issue here is that this factorization is not unique. There is something that we know one of the factors is a conjugate of the other, which is good which actually reduces this ambiguity about non uniqueness, but otherwise we I have already discussed this right we have said look if h and sigma square e; if I set is a solution then h by alpha and alpha square sigma square e is also solution.

In other words I can multiply the spectral factor h of e to the minus j omega by some alpha in this factorization and divide sigma square e by alpha square. Why alpha square? All I am saying is if I adjust a spectral factor h by factor of alpha then I have to adjust h star also remember, and that is why I need to adjust sigma square e by alpha square. So, there is this non uniqueness, and we fixed that non uniqueness by requiring that; how did we fix this non uniqueness?

Student: First coefficient.

First coefficients equals 1 which means now I not permitted to multiple h of e to the minus j omega by alpha, I have already fixed. And that is how we fix the non uniqueness. But now let us ask when is this spectral factorization possible, when is it possible to find a spectral factor, what is so special about this spectral factor? How is the spectral factor define h of e to the minus j omega?

Student: Sequence is.

It is what? It is the Fourier transform of.

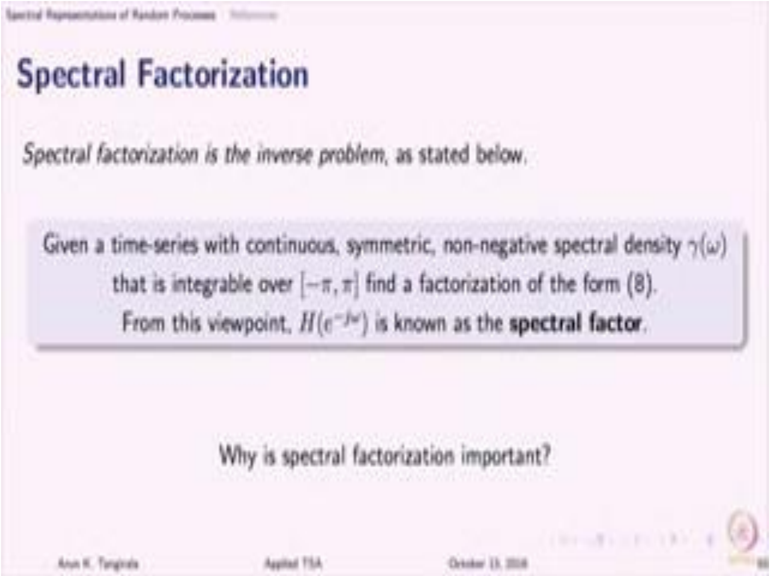
Student: Impulse response.

Impulse response sequence, correct; by the way here we are not talking of periodic processes remember, why are we not talking of periodic processes because we know that the ACVF should be absolutely convergent and so on for this kind of model to hold. So, we cannot bring the periodic processes into this (Refer Time: 04:45). We are only talking of aperiodic processes; that are why I said we are specifically talking of AR, MA, ARMA models and so on.

So, h of e to the minus j omega agreed is a factor but it has something special about it, it is a Fourier transform of impulse response sequence. So, that is an additional requirement, I cannot obtain any factor, I should obtain a factor such that it is a Fourier transform of some sequence and that sequence should be absolutely convergent. So, there are so many things that are required of this factorization. If you are not following what I am saying it is easier to think of fit in this way, fitting a time series model amounts to factorizing the spectral density behind the scenes.

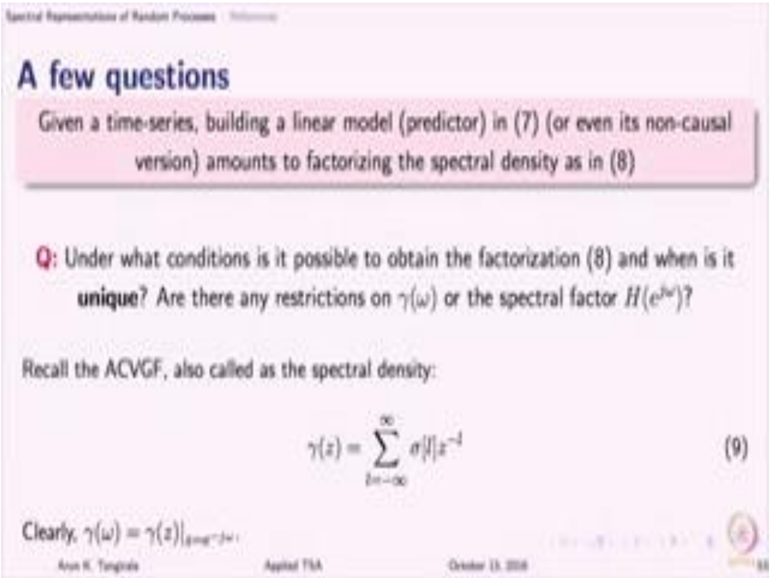
You may not be factorizing the spectral density we do not, when you fit an AR, MA or ARMA model we never talked about spectral density factorization at all, spectral factorization we always talked about fitting the models, but what we are saying right now is fitting a time series model amounts to factorizing the spectral density. This was a profound finding in the theory of random processes. Only when the discovered these connections they where people where able to come up with answers as to when you can fit a time series model. That is why this is called as spectral factorization result.

(Refer Slide Time: 06:12)



So spectral factorization is the inverse problem we have talked about it given a continuous symmetric non negative spectral density gamma. Find a spectral factor. And we already discussed what is spectral factor is.

(Refer Slide Time: 06:27)

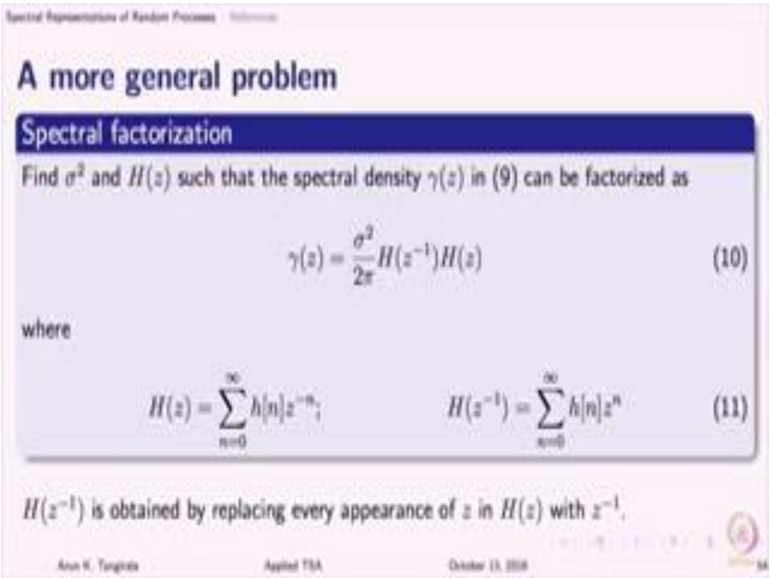


So, why spectral factorization, I have already talked about it. Now the question that we want to answer is under what condition is it possible to obtain the factorization, and when is it unique, and are there any restrictions on gamma. Can I factorize any spectral density? First I should be guarantee that gamma of omega is a spectral density. And we have already talked about the qualifying characteristics of a spectral density, this symmetricity, evenness and so on, periodicity.

So, what we will do is by and large we will avoid the proofs and so on, we will go straight away to the answers to this question, and the proofs are fairly involved. In fact, by going through the proof you can actually show that a random stationary process whose ACVF is a absolutely convergent or for which then the spectral density exist it is possible to express it as a linear combination of white noise. I will upload that proof later on and it is a slightly involved, but I will upload the proof I will try to make it as simple as possible.

So, it is a very nice elegant proof to show that indeed for a random process whose ACVF is absolutely convergent which means a spectral density exist, it is indeed possible to show to express that random process is a linear combination of white noise. The white noise process takes birth from that proof in fact that is how the concept of white noise itself you can think was born.

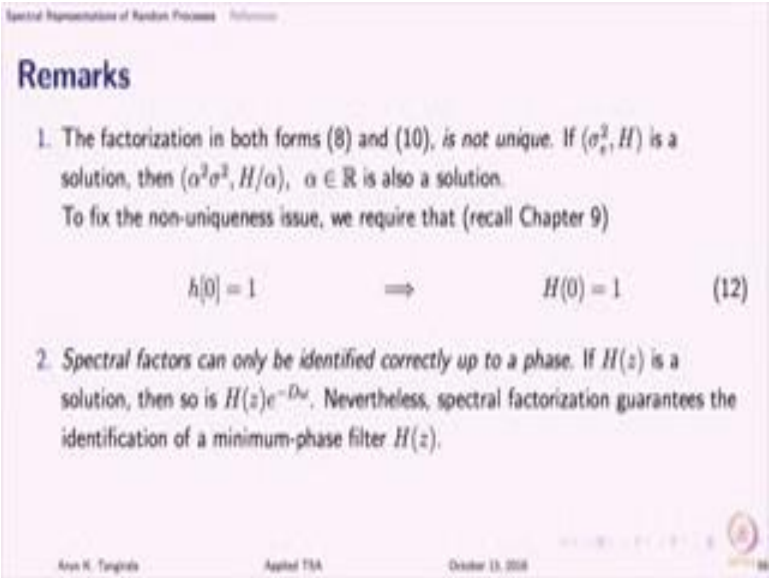
(Refer Slide Time: 08:13)



So, we will avoid all this derivations, I will also avoid the major I will only point the major things. One of the things that you will come across in general when you read this proofs is the generalization of gamma of omega to gamma of z. For those of who are familiar with linear system theory there is a Fourier transform and there is a z transform. The spectral density is a Fourier transform of ACVF. And here we are talking of gamma of z which is a z transform the ACVF. It is nothing but your auto covariance generating function, seen this before nothing special.

The proofs will involve gamma of z, and that is why I talked about it. So, we have talked about this uniqueness of the solution.

(Refer Slide Time: 08:59)



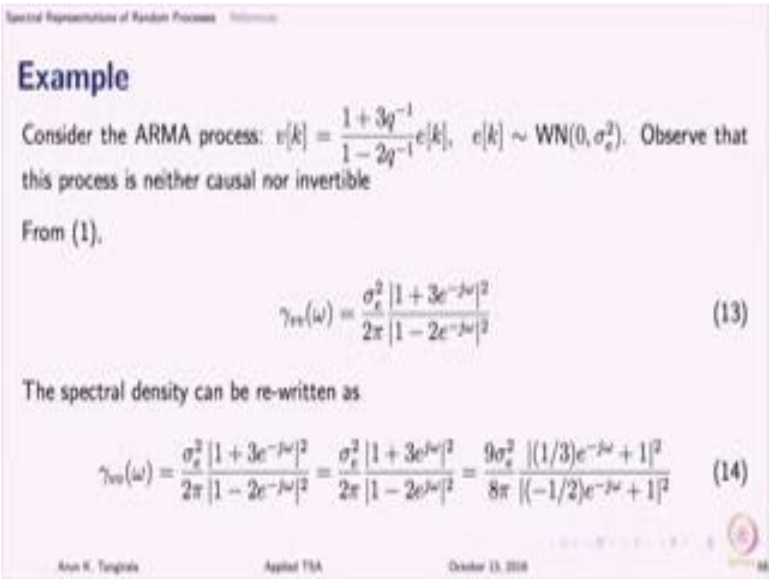
We required that h of 0 is 1. The other point that you could have now taught of or imagine already is that in that factorization I can always adjust the phase of h of e to the minus j omega. What this means is I can always multiply h of e to the minus j omega with some e to the minus j phi, why is that because the complex conjugate will cancel that out; which means that I can always fit a time series model only up to a phase ambiguity there is always a phase ambiguity in that.

In other words if the shock wave there was a delay between the shock wave the fictitious shock wave white noise and the process that you are looking at you will not be able to identify that delay. That is a classic problem with time series models. And you will not be able to identify because that delay will manifest as phase and therefore it is not possible to resolve the phase in any time series model. I am talking of univariate time series models beware. When I am given a single series, I can only figure out the relation between this fictitious shock wave and the given signal, but not the phase. The rest of it can be figure out.

There is not the case in regular system identification or when I am given the input, when I am given the input the phase is fixed because then I have fixed input here the input is not fixed in some sense only its statistical properties are fixed. So, that is something to remember. And the other point that we have also realized earlier is if h of e to the minus j omega is the solution, h star of e to the minus j omega is also a solution. We have talked about this invertibility if you recall; we have said remember the example that we went through MA 1 we had c 1 and 1 over c 1 as solution and then we picked the factor that is invertible.

So, just to give you an example; suppose I have this ARMA process.

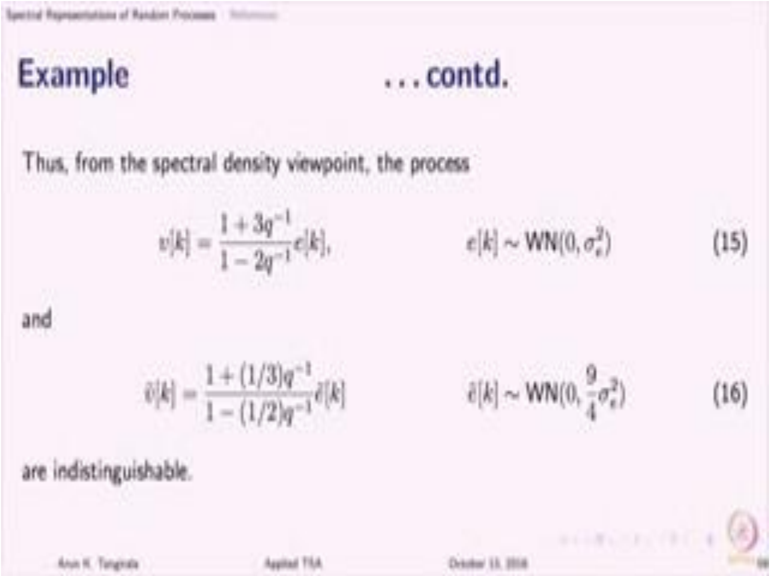
(Refer Slide Time: 11:09)



This ARMA process you see is neither invertible; is this invertible, it is not invertible nor is it stationary or in other words nor it is causal. Amazingly, you can actually when you write the spectral density here. See the existence of spectral density is not conditioned on the invertibility or causality there I can always find remember a stationary representation, but in a non causal sense.

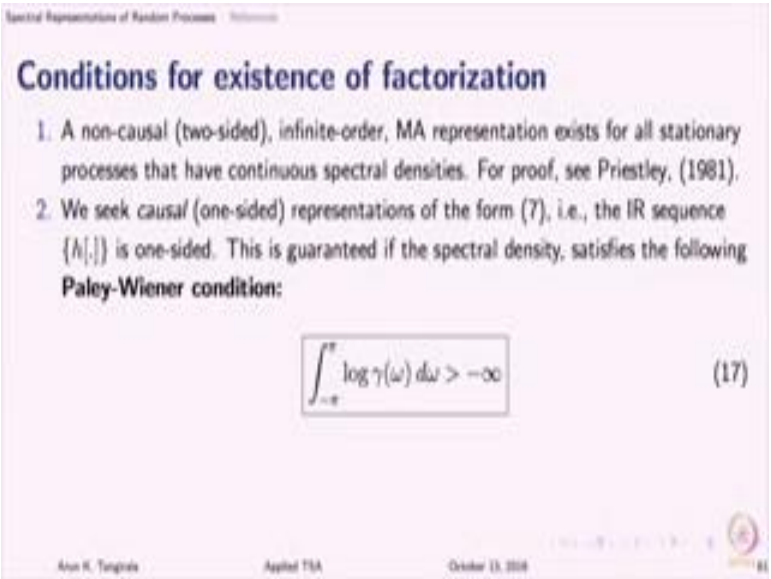
So, here is a spectral density and then I can always rewrite the spectral density or refractor in such a way that I have a new spectral factor or you just algebraic jugglery there we are rewriting now gamma of omega in such a way that I have a new kind of spectral factor.

(Refer Slide Time: 12:09)



As a result this is alternative representation. So, I have 1 plus 3 q inverse over 1 minus 2 q inverse as a model. And I can actually find 1 plus 1 over 3 q inverse and 1 minus half q inverse as a model. So, you can actually find factor that is both invertible and causal. Just a simple example you can go through this example, but what is a difference that you see that sigma square e is different now. That is a difference now. Now that we have adjusted for causality and invertibility the variance of the driving force has changed.

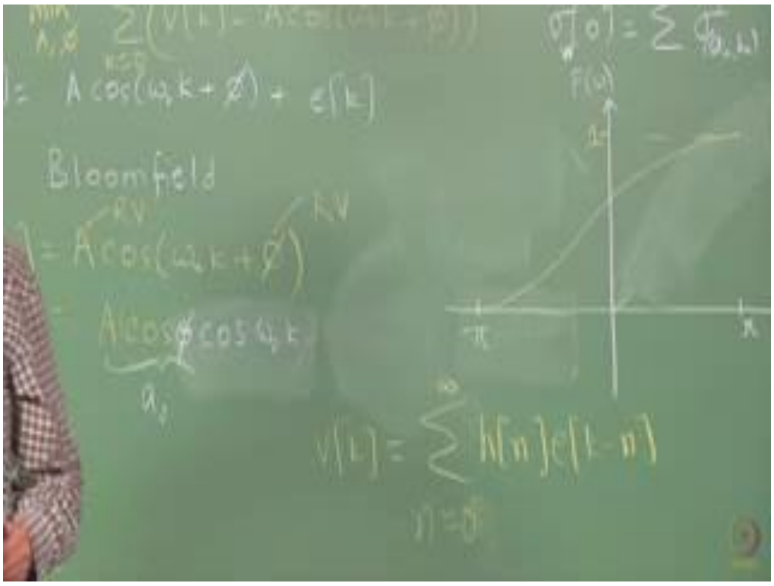
(Refer Slide Time: 12:54)



So, let me now summarize some of the main conditions for existence of factorization and then we will close the discussion on spectral factorization. So, first of all a non causal two sided; first we are talking of non causal because our linear random process model if you recall allows both future and past shock waves to effect this signal. So, first we answer that question. When is it such a representation possible? The turns out that it are possible for all processes that have continuous spectral densities. There is no major restriction per say on that, but what is more important is causal representation. So, we are starting from non causal and then we want causal what we mean by causal is we only want past and present shock waves to effect this signal. We do not want models that relay on future shock waves to explain the present. Ultimately, what is modeling all about? It is all about basically trying to explain the given signal in terms of some shock waves.

So, what we are saying is if it is only a matter of explaining the signal in terms of past present and future then there is no major issue, as long as the process as a continuous spectral density you should be Ok. But, if I want a model that is useful in practice, a non causal model is not useful in practice because I need future shock waves to make a forecast.

(Refer Slide Time: 14:41)



What I want is a one sided model; I want a model that has this form n running from 0 to infinite. When is this kind of a model possible? What restrictions one has to place on the spectral density? And the restrictions are that this particular condition known as the Paley-Wiener condition should be satisfied by the spectral density. If I am only worried about this kind of a representation here two sided then there is no restrictions, any process that has continuous spectral density you should be able to do it. But, this result says if I want this kind of a model which is useful in practice then additionally the spectral density has to satisfy this condition.

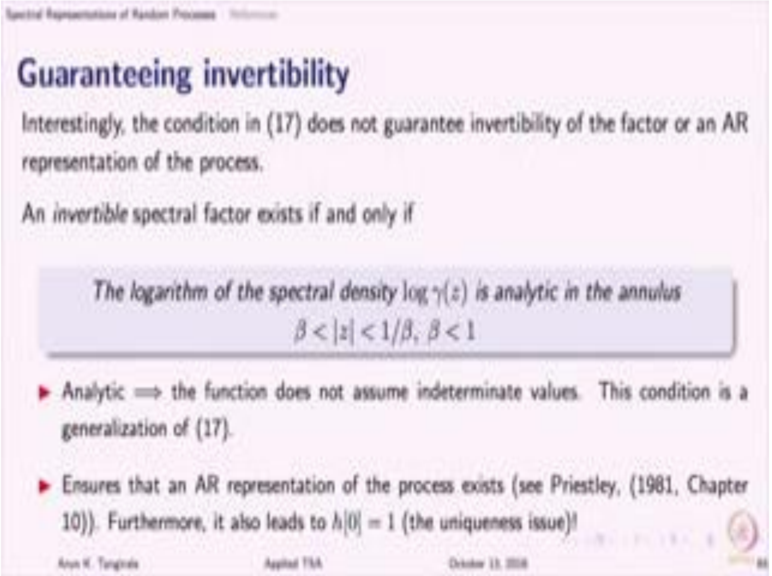
So, the integral logarithm should be greater than minus infinite between minus pi and pi. What else is a requirement? Is this enough? Or do we need more conditions? Are we ok with just this kind of models or we need any other? Remember the condition that h of n should be absolutely convergent wholes that are guaranteed. What else do we need? What kind of a model is this? Is it a moving average or a not aggressive model?

Student: (Refer Time: 16:17).

Moving average right, but I should also be guaranteed that an auto regressive model be constructible; I should be able to construct an auto regressive model. What this condition states is you will be able to construct a moving average model of infinite order provided the spectral density satisfies this. Do not worry about whether I know in practice? I do not know in practice you or I will never be able to verify whether for a given process this condition is satisfied or not. In general but may be for some it will be obvious, but most of the time you will never know whether this Paley-Wiener condition is satisfied, you only hope and pray that it is satisfied and proceed.

What else is required? So, if I want to be able to fit an auto regressive model what else should be guaranteed, invertibility of this model should be guaranteed. If I have to be able to build a stationary AR model then invertibility should be guaranteed in which case the logarithm of the spectral density should be analytic in that annulus.

(Refer Slide Time: 17:26)



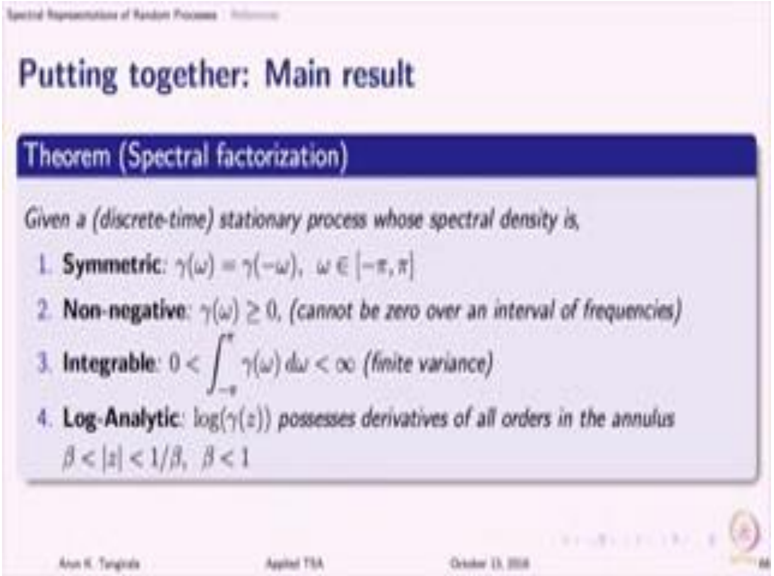
It is just a condition that you just have to know that there exists a condition further on the spectral density for you to be able to obtain an invertible model; which means, what we mean by this is for you to be able to construct an auto regressive model.

So there are three: one is most general when you had minus infinity to infinity where this spectral density any process with continuous spectral density you should be able to build such a model. And then secondly we said know I want useful models. Now tell me whether I can build this model for any process; the answer is no. Any stationary process that as continuous spectral density and which satisfies the Wiener-Paley condition should be able to do it.

And then thirdly, we want to be able to construct an auto regressive model, and that means that I should be able to construct an invertible model and invertibility is guaranteed when the logarithm of the spectral density log gamma of z; now we are not talking of log gamma of omega that should be analytic in this annulus. Analytic essentially means that this function gamma of z should not take on indeterminate values; that is what analytic means it comes from functional analysis.

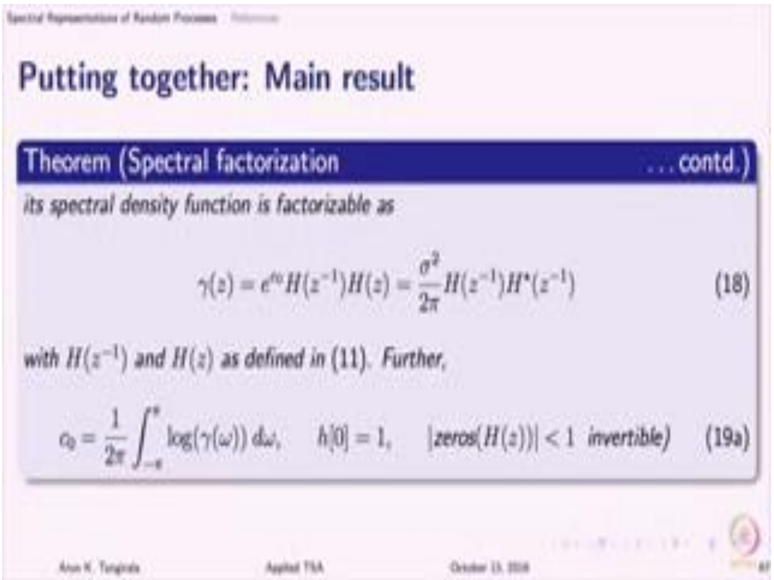
So, this condition guarantees that I can build auto regressive models. So, these are the three main results.

(Refer Slide Time: 19:06)



So, if you put together if you have a discrete time stationary process whose spectral density, of course by spectral density itself we mean symmetricity integrable; why integrable because variant should be finite non negative and of course periodicity is not what I mention it should be also included further now we have included a log analytic a requirement.

(Refer Slide Time: 19:40)

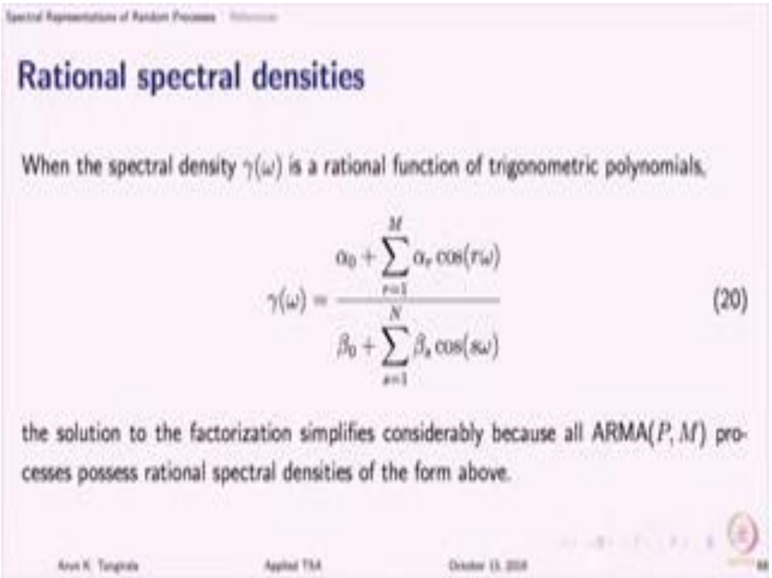


If there exists a stationary process whose spectral density satisfies all these conditions then you can actually performs spectral factorization. Means, you can fit a time series model.

So, first we learnt what a time series models, what are the features of time series models, now we are answering the question when is it possible to fit a time series model. Can I fit this for any process? The last question that we want to ask is when the spectral density as a rational form, until now we have not said anything about the form of spectral density we only said if the spectral density satisfies all those conditions I will be able to fit a model of this form. But now if we specifically say I want to fit an MA model or an AR model of a particular order then that means I am parameterizing h and I am parameterizing ACF and I am also parameterizing spectral density.

We have already seen for ARMA models the spectral density as this form.

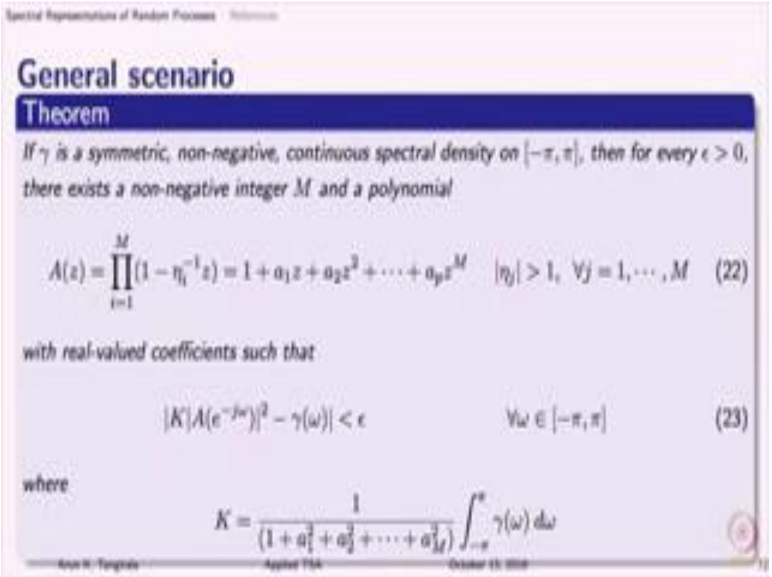
(Refer Slide Time: 20:40)



In this case only I can fit a time series model that will exactly describe the process. Now we want to ask the question what about a general gamma of omega that satisfies these conditions? It satisfies this, but it does not have this form. And I still want to fit a time series model. I am still going to go head, you cannot hold me. I am going to fit a time series model, because that is what is needed for forecasting.

Then what? Am I making a big blunder by fitting a time series model to a process whose spectral density does not have this rational form? The answer is this I am going to skip that example we have already talked about it yesterday.

(Refer Slide Time: 21:32)



The answer is that if you fit; if this is the form of the theorem which says that if you fit an auto regressive model of a sufficiently high order then as long as a spectral density meets, of course it is already understood the spectral density meets all the requirements; regardless of whether your spectral density as that you know what form it has you will be able to approximate the spectral density with an arbitrary degree of accuracy.

To summarize if you do not like any of this theory and so on; to summarize for a process whose spectral density exist and satisfies some important conditions you can fit a suitable time series model that will explain or there will be able to forecast or predict the given process with a fair degree of accuracy. It does not talk a forecasting, but at least it says it talks about spectral density, it is very careful it does not talk of forecast. But good news is you can fit a time series model. How well you will be able to fit depends on now the length of the data that you have, the kind of estimation algorithm that you are going to use, how you choose your order, what is the methodology that you follow and so on; that determines a kind of model that you build.

So until now we have talked a lot about theory and in bits and pieces of estimation, like we have talked about Yule Walker's method, sometimes method of moments and so on.

And occasionally we have talked about variability of parameter estimates then we talked about covariance and so on, but largely we have talked about theory, what are the possible models, what is meant by spectral density, how do you describe a random process in both time and frequency domains, when is it possible to build time series models and so on. So, you can say now enough is enough you have learnt a lot about theory time to turn to practice. Of course, in our I have shown you how to do it. By the way the ARMA spec as I said from the TSA package will allow you to construct the theoretical spectral density for a given ARMA process you should just try it out its very simple to use like your ARMA ACF.

So now, we should turn to estimation and ask tell me how to estimate a good model. What are the things that I have to worry about in estimation? Those are the things that we are going to talk about, I will just spend about 3 or 4 minutes and then of course we will carry on to the next week. So, this kind of concludes theoretical discussion on the theory and now we step into the theoretical discussion on practice. So, there is still the theory of practice you cannot say simply because its practice I will do as I place, it is not possible you have to guarantee that you will obtain good estimates of model parameters, you will obtain so consistent estimates, efficient estimates, you have to guarantee unbiasedness. So, many things you have to guarantee when you use an estimation algorithm, otherwise it is your own personal choice that you can leave with an be happy, but you cannot convenience this scientific community at large; that is the main point.

So, there are a bunch of a references I will not talk of coherence right now will talk about it a bit later, but that kind of concludes the discussion that I wanted to have on univariate time series analysis. Just want to briefly talk about estimation, and then will continue with next week. So, we have just now said and talked about when can be fit a model and so on and now we said we need to focus on how to fit a model. You have talked about types of models when is it possible theoretical conditions and now we are talking of the how.

And that is essentially your estimation. And when you step into the world of estimation you realize that estimation is perhaps the oldest exercise that you can ever have. No, you could have ever come across because may be from the time humans have started to think and analyze situations around inferencing or estimation is a natural thing that would have occur; I mean when I am looking at on and trying to gage how far I have to go, what is the width of the some room are, what is the width of a river that I have to cross. So, all of this is estimation; accept that now today we have a formal way of estimating, but otherwise we keep estimating every moment (Refer Time: 26:19) estimating how many minutes more to go before this class and so, so many things that you estimate.

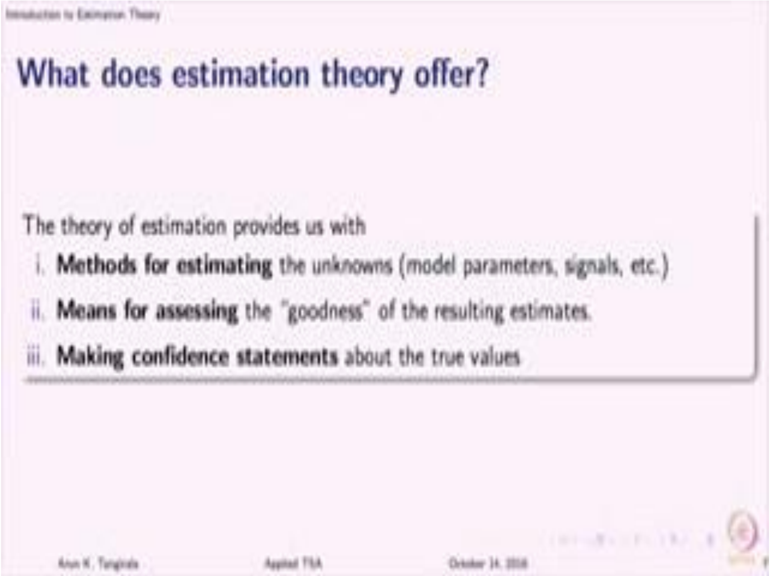
So, one runs into different estimation problems and one of the classic problems that you run into is parameter estimation. And when we mean parameter here not only model parameters, but also parameters of pdf. Do not just think of models parameters all the time. The parameters of the pdf for example, if I look at a Gaussian pdf then I have mu I have sigma and all of those. Or I can think of estimating signal properties, what do we mean by signal properties, again mean variance spectral density all of this is estimation problems of signal properties.

Then we have a state estimation problems or signal estimation; I give you measurement and I want you to estimate the underlining signal. For example I give you sinusoid embedded in noise, I want you to recover the sin for me that is a classic signal estimation problems which was posed extended later to a state estimation problems a Kalman, I do not know how many of you know Kalman recently passed away in July this year.

So, Kalman essentially extended Wieners ideas of signal estimation to state estimation; Wiener of solving more or less a study state problem where the signal does not change with time. And Kalman say are no signal can change with time. And let me actually put in the model for the signal, evaluation model for the signal given the model for the evaluation of the signal how do I estimate; that is essentially your state estimation problem.

So, any time series analysis or any analysis for that matter has the estimation at its heart, you cannot evade at all. Therefore you should be extremely well worse with estimation and that is why we take time off now to spend to discuss on estimation. And they device that performs estimation for us is called the estimator, sounds like terminator (Refer Time: 28:30) but it is called the estimator, it is a virtual device. In fact, the senses that we user are also estimators. And what does estimation theory offer us three things.

(Refer Slide Time: 28:45)



And it is very important to understand first the estimation theory offers us methods for estimation. In fact, most importantly I would say measure means for assessing the goodness of the estimates, but let me order it in this way it gives me methods for estimating. How do I estimate the mean? You have talked about mean, variance, spectral density, covariance, auto covariance and so on, model parameters. What are the different methods available?

Secondly, once I have estimated estimation theory shows us how to assess the goodness of the estimate. And the classic one of the example that I always give a suppose I have a piece of land and I want to build a house, some day you will a do that may be you are already doing I do not know, but some day you will do it. And that time you would approach few builders who will give you estimates, cost estimates of constructing the house there all giving you estimates, but then at the back of your mind your assessing the goodness of such an estimate; how reliable is estimate and so on.

And there as a different criteria that you have in mind, but one of the most important criteria is if I have to ask the builder again if I ask today and I ask tomorrow right or I call two hours later where the market prices would not have change much; the builder should not give me a completely different price. Now at this point in time the builder may say it will cost you about 30 lakhs to build a house and you call one hour later and say is no sir it cost you about 45 lakhs; that is too much variability which is what we call as variability. You want to keep the variability low. You may say 30.5 which you are with it you must not have factored something or something must have drastically changed over the 2- 3 hours time.

But, you cannot expect at this scale of the variability of the process, given the scale of the variability of the process your estimate should not change too much. That is one of the most important criteria. Then there is another criteria called consistency where I give infinite observations, I give a complete realization will you be able to estimate the parameters accurately. So, there are some properties. And finally, in making confidence statements that is after having estimated and after knowing how good the estimates is we want to be able to say something about the truth. That is the most important thing. Why are we actually estimating? Because we want to say something about the truth; so when the builder gives you an estimate what would you ask will say give me a rough interval? Will you be satisfied with the point estimate? First answer that comes out is a point estimate, and then the next answer that you will seek is an interval estimate.

So, ultimately we have to go to the interval, we may begin that point estimate, but we have to go to the interval. What is this interval? This interval contains hopefully the truth. The builder may say ok it may cost you somewhere between 30 plus or minus 5. Do you think finally the actual cost will be in that interval; is it possible always with hundred percent confidences? It slightly that the actual cost may be outside this interval, but there is a degree of confidence and the builder would attach to that interval say- sir 90 percent your cost will be actual cost will be between 30 plus or minus 5. That means, there is a ten percent chance at the reality is outside this. That degree of confidence gives you an assurance that you can go with this builder. So, that is the basic idea.

So, we are going to learn formally how to; we will begin with assessing knowing how to assess the goodness of an estimate and then learn how to estimate itself. Why do we choose that order because, when I discuss a method or estimation I need to know how good the method is and for that I need to know the matrix of goodness. So, that is why will begin with the very simple estimation problems next week and then quickly move on to sorry assessing the goodness of the estimates followed by methods of estimation. So, we will meet next week.

Thank you.