Estimation of frequency response functions

Part 1

All right. Welcome to the lecture on estimation of frequency response functions. So until now we have learnt how to estimate non-parametric models in the time domain, namely impulse response functions and step response functions. We looked at least squares methods and also some regularized versions of it. Now as usual as we have done in the early on in this course where we studied non-parametric models in the frequency domain, we studied at that point in time this frequency response function typically, we denote this with as you recall G of e to the j omega, where omega is the input frequency. And we have already. So this is the kind of models now that we learn how to estimate. And maybe it's a good idea for you to recap, what the advantages are what benefits does one have in working with frequency response functions as against impulse response functions? In some applications for example, it is natural to use frequency response functions in many mechanical engineering particularly vibration dynamics or in some electrical engineering applications, frequency response functions are very natural choice of model or descriptions for the process.

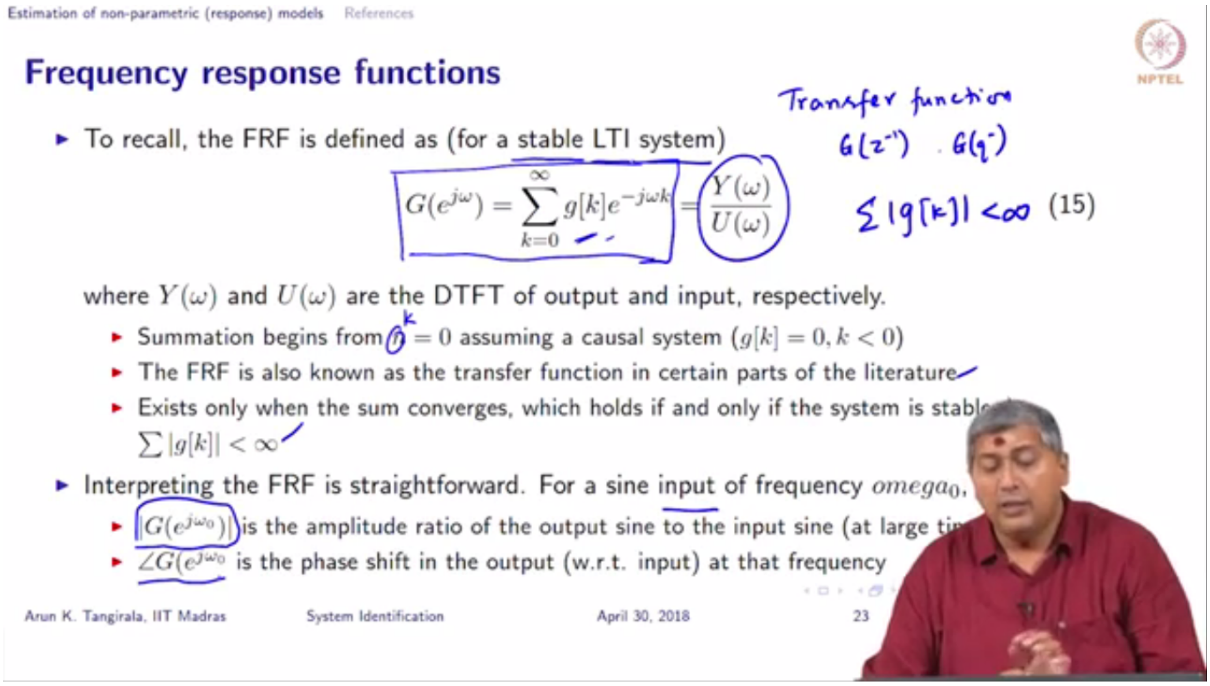
Okay. So let's recap. What is the definition of the frequency response function. Because we know very well that for identification we always start with the definition and then if required we do some tailoring, we have done that with the impulse response functions as well. So here is the first definition that we learnt, which is that the frequency response function is the discrete time Fourier transform of the impulse response sequence and a very important assumption to recall that the notion of FRF is valid only for stable LTI system. Now at this point, I should also caution you on some nominal later confusion that you may see in the system identification literature, which is that very often these FRFs are called transfer functions. We are not new to this terminology, but for us a transfer function typically is represented as G of z or G of z inverse or a transfer function operator is represented by G of q inverse.

Nevertheless for some, not such an unreasonable reason, the FRF is also called the transfer function. But please remember that what they're referring to is the frequency response function. So this is the first definition that we learnt which is valid only for the stable linear time invariant system. And the stability has got to do with the convergence criterion for the DTFT of the impulse response. We know that the-- again it's just a recall, recap that this DTFT convergence if and only if the impulsive response is absolutely convergent, which is indeed the condition necessary and sufficient condition for the system to be BIBO stable. On the other hand, we looked at another definition which is that the frequency response function is the ratio of the DTFT that is discrete time Fourier transform of the output to the input. So this is a different definition. This again comes from the same convolution equation, if that both definitions of the FRF, if you recall come from the convolution equation.

But remember different definitions allow us to come up with different ways of estimating the FRF depending on what we know. So suppose, I have already estimated the impulse response sequence or I have somehow the knowledge of the impulse response sequence. Then I don't need to look anywhere. I just, I can use this definition here to compute the discrete, sorry the frequency response function. On the other hand, I generally in any input output experiment, I would have response of the system to some user designed input, some arbitrary input. In which case I have two options from the given input output data, I estimate the impulse response function and then I use this definition which is kind of a roundabout way of estimating the FRF or the other way is to compute the Fourier transform of the given output to the Fourier transform of the given input. Of course, it's not as simple as it sounds. There are a couple of challenges. One is a finite sample challenge and the other is the fact that there is noise in the measurement of Y.

There is yet a third method of computing the frequency response function, which is called the parametric way. I will briefly talk about that later on. At the moment we are talking about non-parametric models. When we talk of non-parametric models we don't fit a particular structural model. We just want to use the data with the minimal assumptions of linearity and time invariance and that's it. Okay. So just to now put our observations in perspective the summation here. Of course, begins from k not n assuming a causal system. And as I've already mentioned the FRF is also known as a transfer function in the, in certain parts of the literature. And the third point, we have already observed. And also it may be a good idea to recall the interpretation of FRF. It's pretty straight forward. What it tells us is if I inject a sinusoidal input into the system of a certain frequency. Or if I excite the system with a sinusoidal input of a certain frequency, then the magnitude of the FRF tells me, how the amplitude of the injected sine wave is going to be modified by the system at large times. So the main theory behind the use of FRF is that for the stable linear time invariant system. If the input a sinusoidal input of a certain frequency is injected, then given sufficiently large time that is after the systems and transitions have died down. The output also is a sinusoidal waveform with the same frequency. No other system, a non-linear or a linear time varying system has this characteristic necessarily, but only for a stable LTI system the output at large times is a sinusoidal waveform of the same frequency, but with an altered amplitude and a changed phase. These are the only two differences and that amplitude ratio or the altered amplitude can be computed by from the magnitude of the FRF. In other words, if I compute this amplitude ratio I will be able to know what is. And if I know the amplitude of the input then I can compute the amplitude of the output. And the argument or the angle of the FRF it's a complex valued number remember. Gives me the phase shift. So these are the two things. This also allows us to come up with a way of estimating the FRF and we'll talk about that.

(Refer Slide Time: 07:42)

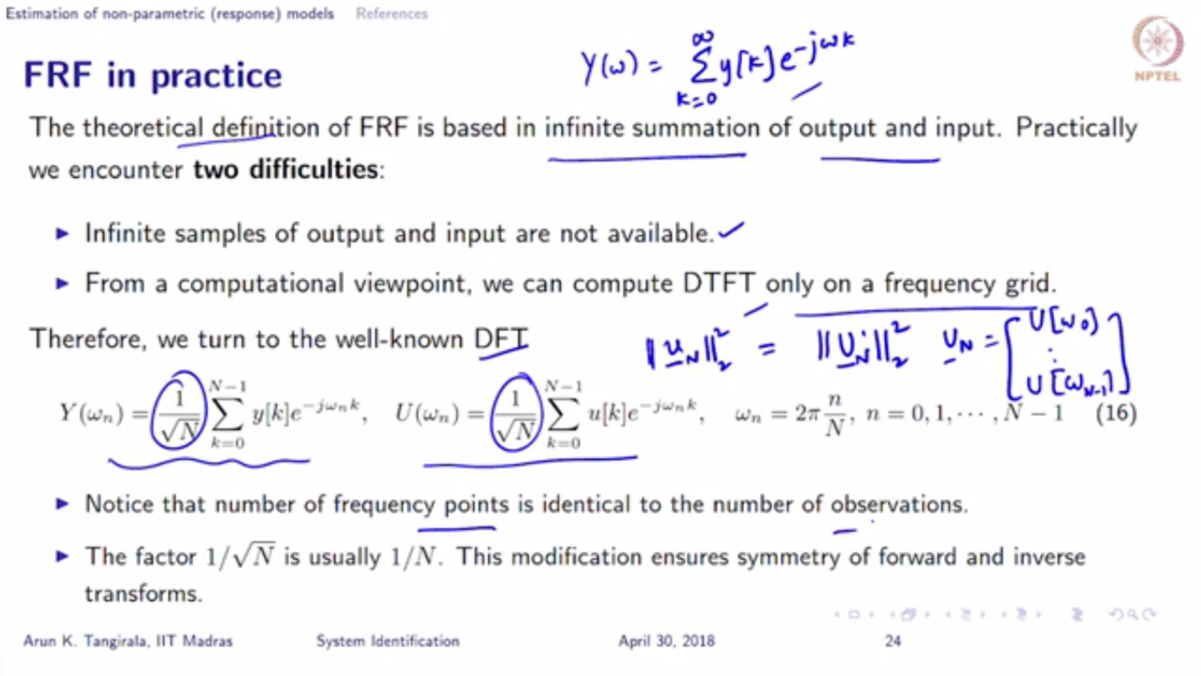


So we'll come to that very soon. But first let us actually look at what are the things that we may have to change when we use the definitions of FRF in practice. So just now we saw that the theoretical definition is based on infinite summation of the output and input. What I mean by infinite summation is? You must recall that y of omega is sigma yk e to the minus j omega k and k runs from 0 to infinity assuming that the output did not exist before at negative times. And likewise for u of omega. So if I want to really use that definition, I need an infinitely long output and of course, the input or I need a mathematical expression for the output and input. I need a close form expressions. In practice, I don't have either of it. In practice what I have is input output data. I have the user or me would have designed some input excited the system, observed the output, even under noise free conditions. Let us assume no noise for now. I have only finite samples. So I can't use that definition as is to compute the FRF. So we have a couple of difficulties, right. So one is that the infinite samples of the output and input are not available. And from a computation viewpoint, even if I had now, suppose I come to that from a computation a viewpoint, we can only compute DTFT only at on a frequency grid. So this is something again, we have spoken about earlier, where we introduce this notion of empirical transfer function. So this is all a recap of what we have looked at earlier. But maybe it helps revisiting them particularly since many of us are not so comfortable with frequency domain concepts. So for that reason we normally work with finite sample versions of the Fourier transform. So we work with what are known as a DFT, the discrete Fourier transform which is both discrete in time as well as in frequency.

The frequency becomes, access becomes discretised, because of this fact that we can only compute at a set of frequencies not over a band of frequencies. So this is the DFT of the input. Normally you may not see 1over route Nin many standard signal processing techs when it comes to DFT, does 1 over a route Nit's just an easy factor to remember, because when I look at the inverse DFT. If I use a 1 over route N here, a1 over route N also appears in the inverse DFT. The inverse DFT is used to recover the time domain signal. But the standard DFT will not have this1 over route N in the forward, in the sense the transform equation. But we'll have a1over N in the recovery or the inverse of the synthesis equation. So 1 over route N is only there to make sure that you can easily remember, but a more important reason is that. It preserves the square to norm in both time and frequency domain. What we mean by that is. I have uk, I have N observations of that, right. And I can think, sorry that may not write this way. Let me write u as a vector, which is of size N. And this vector essentially u N is stacked values of the input, u0, u1 up to u N minus 1.

This is nothing but the big UN, the square to norm of the big UN. What is a big UN? The big UN is very easy to understand. The vector is simply U at omega 0 up to U at omega N minus 1.So the size of big UN is also N by 1. So it's a two norm preserving transform and therefore it's a unitary transform. If you don't have a 1 over route N, then you have to make adjustments for this parseval relation there. Okay. So just to recap now, we work with DFTs instead of DTFTs, but in that process do we make a compromise. Of course, yes if you can recall the ETF, we will recap that shortly as well. In passing you should remember the number of frequencies is equal to the number of observations as well.

(Refer Slide Time: 12:27)



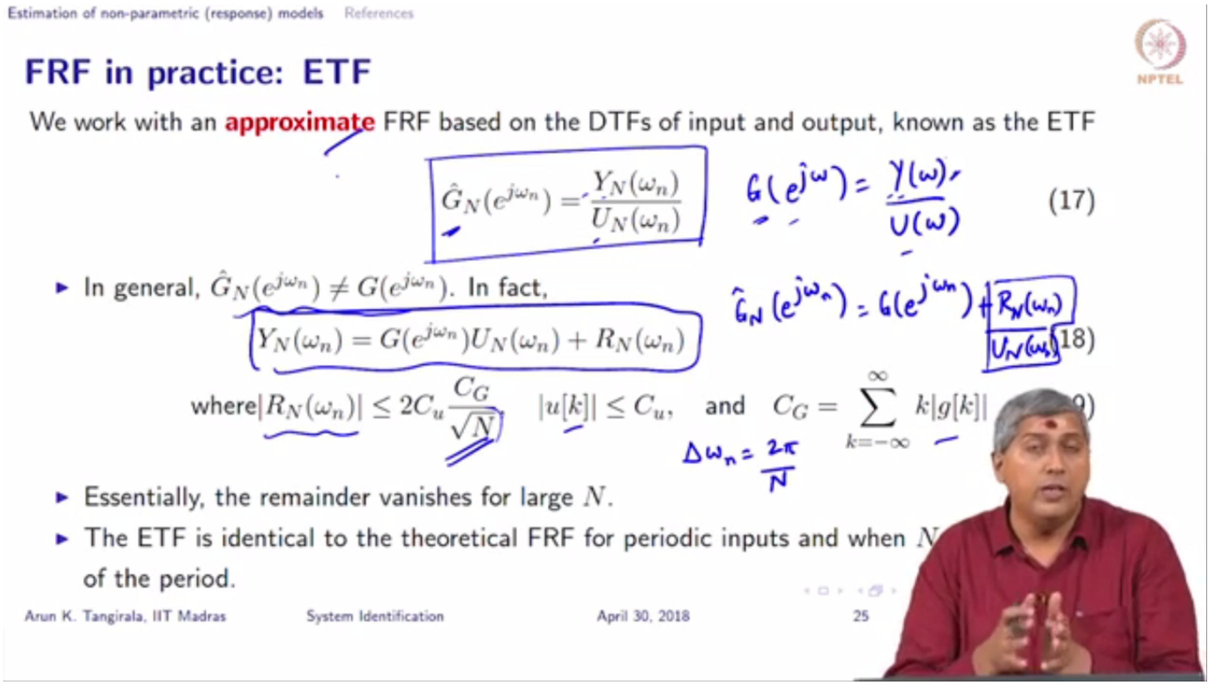
So with that background we must recall, we introduced this ETF, empirical transfer function. This is under noise free conditions. We are still not talking about noise yet. You compare this with the original definition theoretical definition, which is based on DTFT. So this GN hat, it's called a hat, because I'm approximating the numerator with a finite sample version and the denominator with the finite sample version. Therefore there is no need for this quantity to be equal to this. There must be some error, right. Because I am Y of omega here is of involves an infinite summation. And therefore we have a hat here saying that it's an approximation. This comes from estimation theory, the convention.

So in general GN hat, which is the ETF is not equal to G. In fact, you can show that GN hat is such that this relation is satisfied. In other words, GN hat, if I were to divide both sides with UN, I can say, GN hat j omega n equals G of e to the j omega n plus RN by UN. I simply divide both sides with UN. So this is the remainder term. Not reminder. Many people confuse reminder with the remainder. Okay. And that a makes a big difference. So here we are looking at the remainder, but the entire thing that we are talking about is a reminder of what we have learnt. Okay. So this is the kind of a not so good news that the ETF that we have introduced which is based on finite summation is not exactly equal to the FRF, but there is an error term which is the one that I put in the box here. But the good news is that this remainder time is bounded about by some constant and multiplied by 1 over route N. Okay. Those constants depend on the magnitude of the input and the nature of the impulse response. And they do not depend on a sample size. But there is a sample size dependence as well and the good news is, it's good news because as N goes to infinity the remainder goes to 0. Right. Because this magnitude is upper bound is on the magnitude of the remainder.

So which means naturally and that is obvious that as N goes to infinity, the number of terms that you include in YN is going to be being finite which will match with Y. And the frequency spacing that we have that is delta omega n, which is 2 Pi by n, now goes to 0 as N goes to infinity, which means now you're getting back to continuous access for the frequency. So what we learn is that the ETF, which is based on the ratio of DFTs of the output of the input, is an approximation to the FRF. And the remainder term is bounded about by a constant multiplied by 1 over route N and the remainder goes to 0 as N goes to infinity.

So this ETF is the one that perhaps we should work with in practice to obtain an estimate of the FRF. But there is a challenge in practice. In practice what will happen, additional challenge is that now the data that is output will contain measurement error. At least if not some other effects of disturbances. But at least Y will have a measurement error, which will introduce a further error in the calculation of the ETF. Okay. That is what we will call as therefore the ETFE. So in practice what we will be able to construct is an estimate of the ETF. Look at the strange situation we are in. We wanted to estimate the FRF, but then we realized that we can only work with ETF. And now we are going to be in a situation where we will estimate the ETF not the FRF. Okay. So is there a way out. Is that a good estimate that's what we're going to discuss. And that's for the remainder of the lecture.

(Refer Slide Time: 16:57)



So when we're estimating the FRF from noisy data, there are many ways. Before we jump to ETFE we made recall the interpretation of FRF based approach that is now, we should look at that approach. Remember we said the FRF is nothing but it's a function that tells me how the system responds to a sinusoidal input. So great. All I need to do is exceed the system with a sign input of a certain frequency, wait for a sufficiently long time. Okay. So we waited. And then we let the output waveform immerge, right. And then note down the amplitude ratio of the output of the input. So I just look at Ay by Au, Au I know Ay I shall record. And this will give me an estimate of the magnitude of the FRF at that frequency. Let's say that frequency is omega 0. So Ay is a function of omega 0, because I may change the amplitude for, use a different amplitude for each frequency. And the system will also respond differently for each frequency. And then I can also look at the phase difference between the output and input. And all I need to do is look at the phase difference between a phase of y minus phase of u and that will give me angle of G at that frequency.

This is somewhat similar to generating a step response or an impulse response based on the definition.

What do you mean by that? An impulse response is a response to system to the impulse input. So I can inject an impulse input and generate the response. But we have talked about the futility of such an approach. It may be when there is a heavy amount of noise, when there's large noise then the impulse response may go for a estimate obtained through that experiment is not going to be such a great estimate. And therefore we decided to estimate impulse response indirectly from arbitrary input output data. But do we have a similar situation here. If I use a sinusoidal input to excite the system and I wait for sufficiently large times, long time record the amplitude of the output and the phase of the output relative to the input, I get the magnitude, I get the angle, I can put them together to get the G of e to the j omega sounds very simple and it is a very simple experiment indeed. And people do use this. So then why this U and cry about the ETFE and other estimates and so on.

Well, the most I would say important shortcoming of this approach is that it can be very laborious. Particularly for systems which have a broadband FRF. So I suppose the system has an FRF that looks like this. First-order, so this is the magnitude I have drawn, what's this omega here. So I have to estimate FRF at very closely spaced set of frequencies which means, remember how many experiment. Now you can calculate imagine how many experiments I may have to perform. So it can be pretty laborious for systems that have broadband or you can say that have a FRF that are smooth over the range of frequencies, band of frequencies. So the logistics part of it is what prevents us from using this approach for a large class of systems. Now having said that, I did mention earlier that there are a few, in fact a lot of mechanical systems, which are called model systems. Meaning they operate at only certain frequencies and typically the interest is only in the behaviour of the system at those frequencies. Maybe 5 or 6 frequencies in which case it's beautiful, it is actually very simple to excite the system with those 5 or 6 frequencies, record and you will get very good estimates of the amplitude ratio and the phase. It is not as affected by noise as your impulse response estimation, where you directly inject an impulse.

So this method is in fact used in many mechanical systems. So don't dismiss this method as being very futile or useless. It is used in practice, but as I said from a both computational and experimental viewpoint. Maybe it is better to have a method that works with arbitrary input and output data because what if we did not have access to the experimental setup and someone else did the experiment, then we have no choice, we'll have to work with the given input output data. Therefore we turn to the ETFE kind of approaches. So here now what we do is we take the DTF, sorry DFT of the measurement. So this is DFT of the measurement, because now we acknowledge the presence of noise. Earlier when we discussed we said it's a noise free case that is an ideal situation, practical situation output condemns noise.

So this YN is a DFT of the measurement. And as usual this UN is a DFT of the input. But notice that there is a double cap here, because earlier itself when there is no noise the ETF was an approximation to the FRF. So we only placed one hat. Now what we are doing is, when we deal with, when we are using noisy data we are estimating that ETF which is a hat, which already has a hat and therefore this double hat. So this a very peculiar notation, but it's clearly telling you that you are not estimating the FRF, if you are estimating the ETF, which itself is an approximation to the FRF. Okay. So how does a measurement noise effect? Now as usual when we talk of estimators you must recall, we talked about consistency being the top requirement. That means has N goes to infinity. The estimate should converge to the true value. Then we can talk about efficiency and other things. But first and foremost requirement for an estimator is consistency. Are those properties satisfied? The expressions for consistency of ETFE, that is variance of ETFE and so on. Can look a bit intimidating, but the objective is not to memorize those expressions. The objective is to analyse those expressions.

(Refer Slide Time: 24:12)