Course Name: INTELLIGENT FEEDBACK AND CONTROL

Professor Name: Leena Vachhani

Department Name: Multi-disciplinary primarily for Mechanical, Electrical, Aerospace and Chemical engineering streams

Institute Name: Indian Institute of Technology Bombay

Week - 01

Lecture - 02

Hello, so we are back in for intelligent feedback and control videos. This video is talking about process model and experimental methods to identify a process model as well as looking into some approximations and how can we simplify the model itself. So what is the process model that we are talking about? We are talking about linear dynamical systems. And these linear dynamical systems have small deviations from the equilibrium.

So we are looking into modeling these deviations so that we can model this appropriately. Typically, any system has a nonlinear characterization and whatnot, but at the operating range of values where the nominal value is our equilibrium is where we expect that there are small deviations across it, which can be represented as a linear dynamical system. So the representation needs static model identification, dynamic model identification. One can look forward for other methods like feature-based methods. What we have to understand is the difference between the static and the dynamic model that we are talking about here.

When we have a system which we say that, okay, this is a system which is a kind of a blackboard, but I have the input and an output associated with it. And for our case, when it's a control system terminology, we say output has been represented as the variable Y and the input is a control command or the input signal to be given to the system. Now, if there is a direct relationship between Y and U, for example, my Y is following some input value by just a multiplication of gain, where K is a constant value. So, this is what is your static model. But this k need not be simply a constant, but it can be also some function of, some polynomial function of u.

So, for example, this is represented as $kn.u^n + k_{(n-1)}.u^{(n-1)}+.....+k1.u + k0$. So this particular identification of the coefficients of this particular polynomial can be considered as the static model identification. System is currently right now considered under steady state because that's where your dynamical system behavior is not prevalent and the identification of the static model is therefore done in the steady state. For any system, if I consider, one has to also understand whether your system itself is a static system or a dynamical system. So static system, as we said, is going to have a relationship something like this and a direct relationship which we can approximate as a polynomial or of course one can have a different representation or different approximations can be done.

So in order to identify a static model, one has to look for the various different input values and its output values to be considered. The system can have the static as well as a dynamical behavior. The dynamical behavior of the system is represented by differential equations. And these differential equations then is represented in the form of transfer function or state space representation and so on and so forth. And identification of this dynamical model is then the next task that we use to completely describe the linear dynamical system.

We will leave out the feature-wise model ways, which we'll take it up in week four assignments, week four videos and so on. So coming to the static model or static process characterization, what we have to look at is the steady state values of output y for different values of the manipulating variable U as we spoke. Now, this can, if we plot this particular U versus Y curve, so U is the input and Y is on the y-axis. And if we have certain data points given here, and if I'm able to fit a linear curve over here, then your slope of the curve is your static gain. Now, at the same time, we can have the large gains associated with this particular input-output relationship.

If this is large gain, that also indicates difficult control problem. So, we have not yet identified dynamical model. But by looking at the static model gain itself, this large gain is indicating a difficult control problem. We can resort to some other method instead of PID control method right away over here. It will also tell you that how rigorous your model identification process should be.

If your system is having large gain, you might resort to getting into rigor model identification method in order to have more closer control answers to it. So in order to identify this gain, we have experimental method can be something like this. We will consider keeping the input at a constant value. We will then record steady state output value. And then we will change input to cover entire range of inputs.

So this way, what we are looking at is we are plotting this again, this u and y value. For a particular U value, we will identify Y. But this particular output value Y that we have got is a steady state value. It's where the transients have died out. Then we will consider getting some other input U, plotting the output Y, getting the value Y. And this way we will try covering the values of U in the operating range. This operating range of values will then give you the static gain value.

Now, whether this is a line fit or a polynomial fit or whatnot is something depends upon how the data points have come out while experimenting between, while collecting the measurements of the output Y with respect to time. Now, what happens is that when you change this value of U and getting the output value Y, you are waiting for the system to come to the steady state. Every time you change the value of U, you wait for the system to come to steady state, collect the value of output Y. Now, this is a slow process of identification, model identification, and that too you are still identifying the static model. So what has been suggested as an alternative method is you use closed loop system. So here we will do, we will consider output as a constant and correspondingly what is the measured input is.

So what has, in this process what we will do is, we will gain is the, since this is already a closed loop system, the control is already applied, this, we already have an intuition that the control is going to be easy. But to understand this more closely, let's see how it comes out. Let's say this is your plant given by G(s). The transfer function of it is given by G(s), and you have used some kind of a proportional controller or PID controller in order to control this particular plant. By doing some small variations into the PIID control, just by small tuning and so on, you are able to control it to a satisfactory extent. So, how do we collect the data here?

Instead of the previous method, which was the open loop method, you were changing the input and you are getting the output. Here, we will get the output value y when this comes to a constant. And that's when you will collect the value of the output Y and this particular input U. So in order to change now, you got this one value of output Y, you collected the value of input U. So then you are doing nothing but changing these output values Y.

y versus u plot, you change the value of y, you collected the value of input u, you plot this. Now change this y, you got some other u. So it's the opposite that you are getting the, with respect to the y-axis values, you are getting the x-axis values. Fair enough? So in order to change this particular output value y, you will change this particular set point value. Now this set point value is changed, you will get some output value, constant value output and note down the U.

As compared to the open loop method, this is going to be faster because more or less this is a negative feedback methodology where you have a better time constants and so on and so forth. And of course, since it is controlled, the steady state is going to remain at steady state without the disturbances, even if there are disturbances into the system. Looking at identifying the dynamic model, there are two ways. One is the time-based method and the other is the frequency method. Now, the time-based method is identifying the dynamical model with the help of the transient response.

Now, what happens is in order to get the transient response, our underlying assumption is that the system for which we are identifying the model is an LTI system, and therefore, the superposition holds. If the superposition holds, how are we taking the advantage of the superposition? We have studied Fourier series. This Fourier series is nothing but telling us that any band-limited signal can be represented as a constant or a step input and the series of frequencies and their harmonics. So this arbitrary input that we are applying it can be represented in terms of the response of certain simple signals.

So response of any arbitrary signal, for example, I have some arbitrary signal applied as an input to the control system. So this is my input U and this is the time scale that we are talking about. So this particular signal should be able to represent as a step input plus the sinusoids of basic frequency and the harmonics of it. This Fourier series concept is telling us that my basic signals are step signal and a sinusoid signal or the derivative of the step or the integral of the step. So in a way we are saying that these basic signals we consider and the combination of these signals is representing the arbitrary signal.

So response to the arbitrary signal is going to be the superposition of the responses of these basic signals, which are step, pulse, impulse, sinusoids, ramp. All right. So that way, we also know that the response of output Y to a unit step signal S(t) is nothing but the convolution of the impulse response and the arbitrary input U(t). All right, given this, we can go for getting the frequency response, looking into what frequency response gives us. With the help of the frequency response, can I identify dynamical models?

Okay, I'm going back to the previous slide. In order to identify this particular dynamic model using the transient response, now we can apply the step input. And the response to the step input now needs to be collected. If it's a second order system, there's a likely that you'll get a response something similar to the one drawn here. Now, this response could be different.

It could be for a first order system, something like this. These responses which we are drawing here is going to give us an idea what the dynamical model or the transfer function of the system could be. This we'll talk about later. The other way of identifying the dynamic model is frequency way. Now in the frequency way, we know that the system transfer function is given by the magnitude and the phase.

In order to identify the magnitude and the phase of the phasor output of a given input, we apply the sine wave. For an input sine wave of a known frequency and the magnitude, the output is also a sine wave with same frequency, and with some phase difference and the magnitude that you can collect it in order to plot the Nyquist plot. Now if I plot Nyquist plot, I can also look forward for plotting the Bode plot. Bode plot is giving me gain and phase margins. And with the gain and phase margins or the gain and phase plots, I can write the dynamic model.

With the help of the Nyquist plot, I should know the ultimate point, which is the lowest frequency for which the phase shift is minus 180 degree. With these identified values, we will be able to identify the transfer function of the system, which is either first order or second order system or a combination of these. So the advantage with the frequency response is we will not have, we will have a very low measurement errors because input is sinusoid, the output is sinusoid and we are relying on the measurements on of not a single point but the entire sine wave. And that's the reason we will have lesser measurement error as compared to the transient response, which is depending on the response that we get for the step input, which could be erroneous when we are taking the measurements with respect to time. Whereas the frequency way of identifying the model has a disadvantage that the system must always start from rest.

So here we will have to, the system is starting from rest, we are applying the sinusoid, recording the sinusoid values, we're in the steady state, making that this is switched off, once again applying a different sine wave, different frequency of the sine wave and recording the output. This way, every time one has to look forward that the system is starting from the rest. Whereas in the transient response way, one can look forward for applying one step response, step input, making sure that the system is coming at the steady state, then giving another bump of the step and look forward for collecting the output and so on. If you want multiple measurements just to avoid the measurement errors and so on. All right.

Given that we are identifying the model, it is always better to have some idea about how the system transfer function is, whether it's a first order system, second order system, second order with multiple poles and whatnot. If that is what we want to identify the structure of the model, then what we can do is looking at the step response of the system, we can categorize the system in terms of following categories, whether the system is stable or not. whether it is oscillatory, which happens with the spring elastic action or the concentration control of recirculating fluids, whether the system is unstable, whether the system is with transportation delay or a non-minimum phase. So this kind of categorization can be easily obtained by applying a step input and looking at the response. And the different kind of responses that comes up here is what can be seen here.

For a step input, if the system is stable, either if it's a first order system, it is exponentially rising and reaching to the final value. If it is a second order system, we know that the sinusoid or the oscillatory response behavior will die out and is a decaying sinusoid and reaches to the final value. If the system is oscillatory, then it's at the marginal stability point and this will instead of the decaying of sinusoid, it is going to have a sustained oscillations and so on. If it is unstable, then this sinusoid is growing, growing uh growing in the magnitude or if it's a first order system then it is exponential when you apply a step input. The system has a transportation delay then your step response is going to be if it's a first order system then it is going to start with a delay

So this particular delay needs to be accounted for. Similarly, for a second order system, there is going to be a delay in getting the responses and so on. In case of the non-minimum phase, your step response is going to be such that it is first decaying, it is going in the opposite direction and then reaching to the final value after some time. So the opposite, you have applied a positive bump where the response is going in the negative side. So this particular behavior where the output response is opposite to the input applied is what comes up when your zeros are on the right hand side.

It is not because of the delay that you have reached to this particular final value, but because of the non-minimum phase characteristics, you got this final value reached after a particular time instance. So in order to describe the dynamics of the system, this kind of a dynamics of the system where there is a transient before reaching to a particular steady state value, we need to have certain additional parameters. And so one can look forward for adding those additional parameters by saying that, okay, for example, it's a two parameter model. So, or it's a first order system, which we can identify from step, which we can categorize based on the step by looking at the step response of the system. We can say that if it's a simple first order system way, then I need two parameters, which is K and the time constant of the system.

Now, this time constant we represent it as an average residence time. Now, this average residence time is given by A0/K, where A0 is this particular integral, which is representing nothing but if it's a step response of a first order system, this is going to exponentially rise and reach to the final. The step is the response of the system is going to exponentially rise and reach to the final value. So this particular A0 is nothing but the area under this shaded area that is what is represented here. So this is your S of infinity value, which is as we approach infinity, this is what the output final value and how the S(t) is approaching.

So this particular integral gives you nothing but the area under the shaded curve. This is a better measurement way because I can find the time constant by also just appropriately fitting a line over this particular step response. But fitting this line could be definitely erroneous. So it is a better idea to consider this area identification and then finding the time constant of the system. K is coming from the static model itself, static model identification itself.

So in general, I should be able to write this particular average residence time in terms of integral 0 to infinity tg(t)dt upon integral 0 to infinity tg(t)dt, which is nothing but the gain K. So this particular average residence time is the rough measure of time it takes for the step response to settle, which is nothing but the description of a time constant that we talk about. But in practice, none of the system is going to be exactly like a first order system. And that's the reason we are saying that this is an average residence time that we are talking about. Now, for a two parameter model, one can also have the system representation as a integral and a time delay.

Now, this integral is given by can be can be said as the gain 'a' and the time delay 'L'. Now, identification of 'a' and 'L', since the response is some step response is something like an exponentially rising part, we should be able to identify this particular time delay from where the system start rising and 'a' is nothing but the slope of this line. Integral with dead time is also a basis for the popular Ziegler-Nichols tuning method, which we can see in the later part of the video in this week. There comes a simple exercise that you can look forward. For example, my system is having the eighth order system with multiple poles at -1.

Just plot a step response of this. And of course, with some method, we came up with approximating the system as a two parameter model as the first order system or as a integral with a dead time. If you plot the step response of this and this and compare it with the true system, see how much the approximation is matching with the step response of the true system. You'll find that the matching is of the order of 90%. If the matching is of the order of 90%, then it's better to work with the approximate systems and just with the help of the PID tuning, you will be able to control the system.

The idea here is to simplify the control methodology. If my system is a very high order system, I'm able to approximate it with a first order or second order system, or a time delay system. It is better to represent it that way and apply the PID control to achieve the control objectives. Simplification always makes the system more robust in terms of the implementation and so on and so forth. But the key here is whether I am able to approximate the model or not.

This is something we will look into next video as well. So finding the average residence time is also a good idea to have some idea on how can I find this particular average residence time analytically. We said that this average residence time is given by this particular formula, which is the ratio of the area under the step response part that we have seen and the gain K. If I look at the transfer function, the laplace transform of the transfer function is with respect to the impulse response of the system, we can say it is $e^{-st}g(t)$. At s is equal to 0, we get this particular value. Similarly, when we take the derivative of the transfer function with respect to s, then what we have is $e^{-st}tg(t)dt$.

Its value at s is equal to zero, we get this value. So identifying the average residence time now becomes pretty easy because we have the average residence time given by $-g'(0)/g(0)$. By the way, when we were taking this particular derivative with respect to s, we forgot about the minus sign over here. And this minus sign is now appearing over here. All right, fine.

So this way, this also gives a idea about the average incidence time of the cascade system and how they are related. For a cascade system, in terms of the transfer function way, the overall transfer function of a cascade system is given by G1 times G2. So average residence time, we now know that it's simpler way to identify the average residence time if I know the transfer function of G1 and transfer function of G2. So we can find this particular way by getting the derivative of G(s) with respect to s, which comes out as

G1(s)G2'(s) + G2(s)G1'(s). And average residence time can now be given by considering this formula and substituting s=0.

What turns out that just by simple manipulation is that the total average residence time is nothing but the average residence time of the first system, G1(s), plus the average residence time of the second system G2(s). Or it's a summation of any individual average residence times in the cascade way. Similarly, for the three parameter model, we have the characterization given by the first order time delay system. There are three parameters to be identified, K, T, and L. K is the static gain, which comes from the static model identification.

Now we have to look at T and L values. The step response of the system can be given by s(t) over here, whereas now the apparent average residence time over here turns out to be L plus T. So this particular way of representing this particular average residence time has a wonderful idea behind it that if I consider a normalized time delay, which is given by L/(L+T), with the help of this normalized time delay time, I'll be able to say that whether the system can be easily controlled or it's going to be difficult to control. Difficult to control as in the fine tuning of the PID controls will require a lot of effort as compared to the PID for the values for which I'm saying that it is easy to control. At times, it could be a possibility that PID control is not able to achieve the control objectives for the systems which are difficult to control, and we will have to resort to other methods for the control method.

So in this way, experimentally, I am able to understand whether the system will be able to, we will be able to control with the help of PID or not. So here comes the use of this normalized dead time now. Basically the system with a large time lag, PID is usually difficult to tune. We can intuitively understand the response will take its own time to reach to its steady state value or the transients are delayed by some time and its negative feedback. The negative feedback is already not reaching to it and that's the reason the PID will not be able to reach to the control objectives.

All right. So with respect to the time constant and the time delay, if this particular ratio is definitely going to be between 0 and 1. If the tau value or the normalized dead time value is near 0, then it is easy to control because my time constant is much higher than the time delay of the system. Whereas, if the dead time or the time delay of the system is much more than the time constant of the system, means the time delay is much larger than the time the system is taking it to reach to the steady state value, then the system becomes difficult to control. And for that, the tau values are almost near the values of 1.

So with the help of identifying the normalized time itself, we are able to say that we should go for which methodology for the control. Can I go with the simplified methods like PID? The system is easy to control, yes, definitely go for pid no need to go for very rigor control methods, but if it is difficult to control then, give a chance with pid for by simplifying the methods, approximating the models and so on, but the rigor control objectives can then be satisfied with the help of other methods This set of ideas of the average residence time have been covered under these references. One can look forward for these references for more information.

In the next video, we will be looking at the model reduction methods, and that will help you to understand how to approximate the system. Thank you.