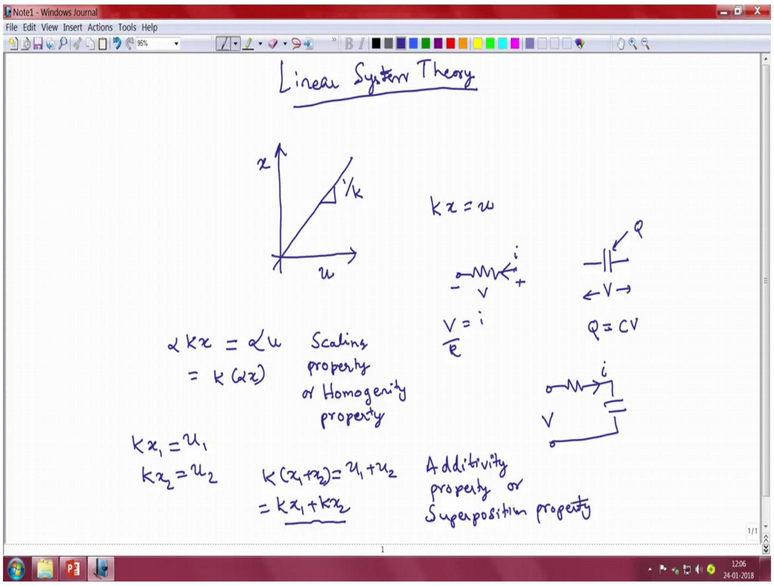
Control System Design
Prof. G. R. Jayanth
Department of Instrumentation and Applied Physics
Indian Institute of Science, Bangalore

Lecture – 02
Linear Systems

Hello. In this video clip we will introduce Linear System Theory.

(Refer Slide Time: 00:21)



So, I shall write the title of this clip here. Now, let us start with the assumption that none of us know anything about linear systems. So, what does one mean, when one talks about a linear system. If you were to ask a person, who is uninitiated in linear system theory, he might look at the name you see the word linear in the name, so probably there is going to be a straight line somewhere. And you have system also in the name. And therefore one would expect that there would be an input and an output. So, perhaps this uninitiated person would think that if there is a straight line relationship between the input and the output of the system, then one could call that system a linear system.

In other words, if one were to graph the input-output relationship, this uninitiated person might assume that then the input u to a system is changed, the output x of the system changes according to a straight line. Perhaps if he is to uninitiated, he might draw a straight line that passes exactly through the origin. Of course this is not adequate as a useful definition for linear systems, but let us start with this as our initial assumption of what a linear system might look like.

So, if the slope of this line is 1 by k, then I can write the input-output relationship for such a system as k x equal to u. Now, before we want to label a system with such a straight line relationship between the input and the output as a linear system, we want to ask ourselves is there any use to labeling such a system as a linear system. And as engineers it would be useful for us, if there are enough examples in our respective disciplines that have this particular input-output relationship.

So, what examples can we think of, where the input and the output is related by a simple proportionality constant k. One possible example is ohms law. So, if I have a resistive element, and there is a voltage V across it then the current i is given by V by R. So, there is a proportionality relationship between the current and the voltage. Likewise, if I take a capacitor and I have some potential difference V across the capacitor, and a charge Q on the plates of the capacitor. Then I have a relationship between the charge and the voltage as Q equal to C V.

Likewise, in the mechanical domain one could think of a spring, where the relationship between the force applied by the spring and its extension is once again a proportionality constant, which is equal to the spring constant. So yes, there are examples of physical systems where the input and output are related by a simple proportionality constant. But, the set of examples is quite small.

For instance, if you take the two examples that I have considered here namely a resistor and a capacitor, and you were to connect them in series. So, I have a resistor in series with a capacitor, and I apply some voltage V across them. Then, and I assume that the current through this series combination is what would be the output then I can show that the input, output relationship would look something like V is equal to d Q by d t times R plus Q by C or equivalently d V by d t is equal to d i by d t times R plus I by C. So, this would be the kind of relationship between input and the output even for a simple combination of a capacitor and the resistor.

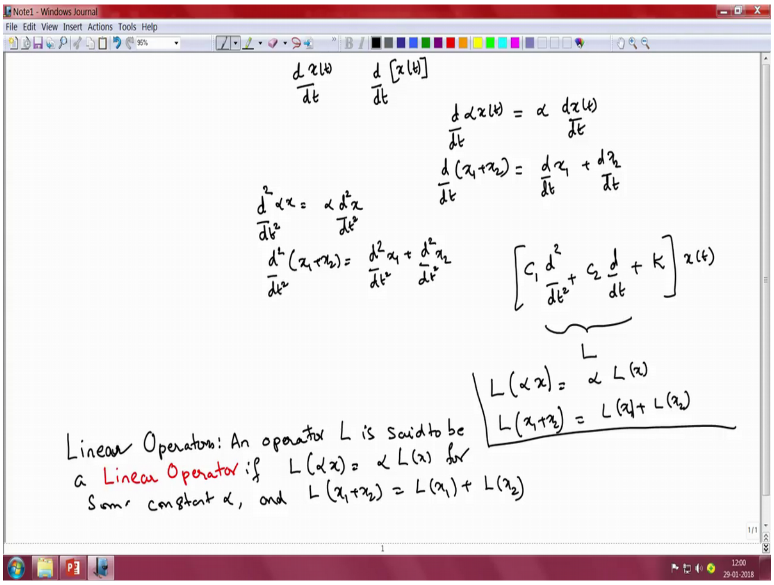
And clearly you can see from this equation that one cannot cast this equation to have a linear relationship between the input and the output a straight line proportionality relationship between the input and the output. So, this simplistic definition of a linear system is not going to help us too much, because we do not have too many physical systems that obey such an equation, but we can still use these equations for inspiration.

For instance, if we revisit this equation k x is equal to u, and assume that k is some kind of an operator. Of course, we know that it is simply a proportionality constant, which scales the input to give me the output. But if you assume that it is an operator a scaling operator if you will, then what are the properties of this operator. One thing you notice is that if the input to the system gets scaled by a factor alpha. So, if I apply alpha u as the input to my system, then the output would be of course, alpha times k x, which I can write as k times alpha x. Or in other words what this equation reveals is that, when the input is scaled by a factor alpha, the output also gets scaled by the same factor. So, this property of this proportionality operator we shall call it the scaling property or homogeneity property. This is one property that is that the proportionality operator possesses.

There is one other property that it possesses. Suppose, I were to apply u 1 as an input to my system the output will be let us say x 1. So, the input and the output are related by that equation there k x 1 is equal to u 1. And if the input is u 2, then the output would be x 2 let us say. Then if I apply an input of u 1 plus u 2 then I know that the output would be k times x 1 plus x 2, which I can rewrite as k times x 1 plus k times x 2. And what this equation reveals is that when I apply two inputs to a system, the response of the system is a summation of the responses to the individual inputs. And this property, we shall call as the additivity property or the superposition property superposition property.

Now, let us go out and search for other operators, which obey these properties namely that of scaling, and superposition. The properties which we have now discovered in case of the proportionality operator k, do they also exist in other operators that is the question, we are trying to answer.

(Refer Slide Time: 08:49)



Now, let us take the example of differentiation. So, let us say we differentiate a signal x of t with respect to time, in other words for evaluating d x by d t. Now, one can view this differentiation as an operator in that the differential operator d by d t acts on the signal x of t. Now, is it possible that the differential operator is also is also one that obeys superposition and scaling, we can verify that by inspection.

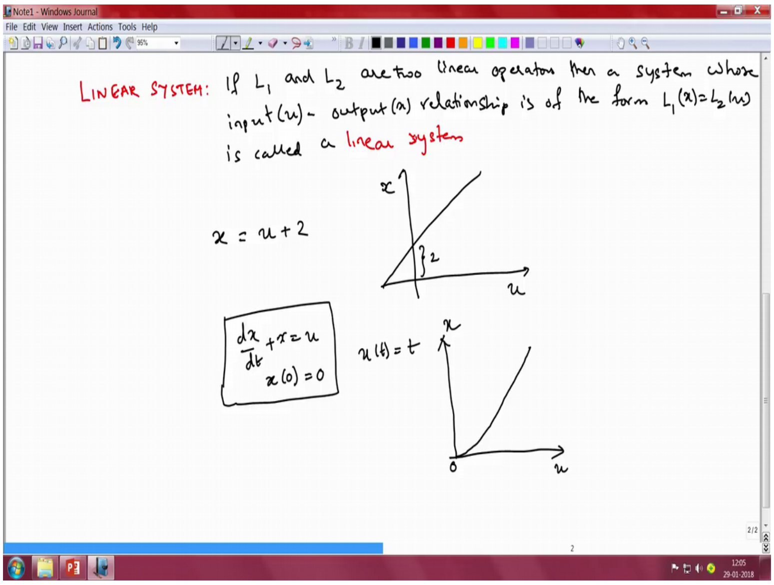
So, let us say I were to multiply x of t with some constant alpha, so that I get alpha of x of t. Then we know that the derivative of alpha of x of t with respect to time is alpha times the derivative of x of t with respect to time. Similarly, the derivative of two different signals x 1 plus x 2 sum of two different signals is equal to the sum of the derivatives of the individual signals d x 1 by d t plus d x 2 by d t. So, indeed we see therefore that the differential operator is one that also satisfies superposition and scaling.

How about second derivative t square by d t square second derivative with respect to time? Once again one can verify that d square by d t square of alpha times x is alpha times d square by d t square of x, and d square by d t square of x 1 plus x 2 is equal to d square by d t square of x 1 plus d square by d t square of x 2. So, the first one shows that it obeys homogeneity; the second one shows that it obeys superposition or additivity. So, we have found two other operators namely a differential operator, and a second derivative operator both of which satisfy additivity and homogeneity.

How about a linear combination of these operators, so for example, some constant C 1 times the second derivative operator plus another constant C 2 times the differential operator plus some proportionality gain k is this something that obeys additivity and homogeneity. Indeed, if you simply plug in and inspect you will discover that, if this operator acts on some signal x of t, and I were to call this operator as L. Then just by substitution, I can verify that L of alpha times x is equal to alpha times L of x, and L of x 1 plus x 2 is equal to L of x 1 plus L of x 2.

Now, having come across a range of different operators that satisfy additivity and homogeneity, let us now define all these operators that satisfy additivity and homogeneity as linear operators. An operator L is said to be a linear operator, if L of alpha times x is equal to alpha times L of x for some constant alpha, and L of x 1 plus x 2 is equal to L of x 1 plus L of x 2. We are now in a position to revisit our definition of a linear system. And this time, we shall define a linear system by employing this definition of a linear operator.

(Refer Slide Time: 13:32)



I shall write out the definition of a linear system here, linear system. If L 1 and L 2 are two linear operators, then a system whose input bracket u, output bracket x relationship is of the form L 1 of x is equal to L 2 of u is called a linear system.

Notice though that this new definition of a linear system, which is based on operators L 1 and L 2 satisfying additivity and homogeneity is not necessarily the same as our initial guess of what might constitute in a, a suitable definition for a linear system. I want to give you two counterintuitive examples to illustrate the fact that this definition is a slightly more abstract one, but nevertheless a vastly more useful one to us as engineers. So, in the first example, let us say the input-output relationship is given by the equation x is equal to u plus 2, where u is the input and x is the output. The input-output relationship can be graphed, and you can quickly see that it would be a straight line with a y intercept of two units.

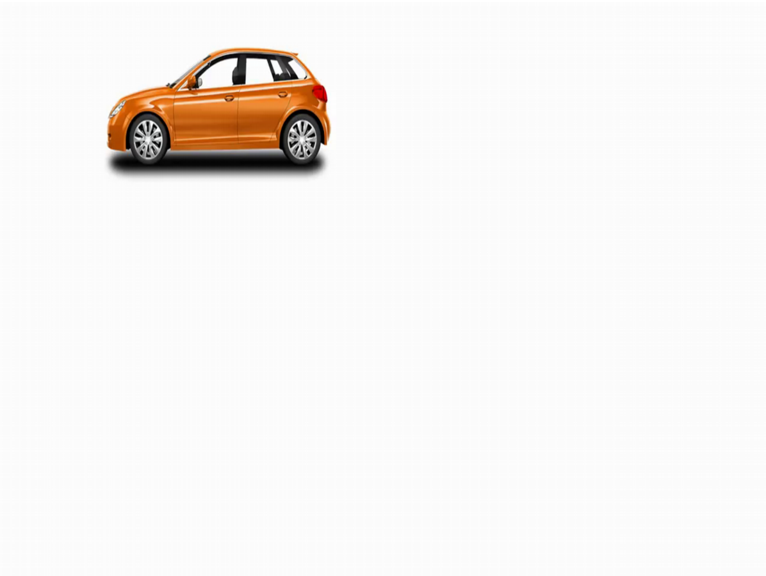
Now, in this case, you see that the input-output relationship is a straight line, but if you were to strictly apply the definition of a linear system to verify whether this input-output relationship constitutes one of a linear system or not, you would conclude that it does not constitute a linear system, because it does not satisfy either additivity or homogeneity. Likewise, if you want to take another example, let us say I have a system of the kind d x by d t plus x is equal to u, and I start with 0 initial condition. And I provide a ramp input to my system.

If I were to plot the input versus, output relationship a few instants after applying the input, the input-output relationship would look something like this. In other words, it would show a pronounced non-linearity in the input-output relationship near the origin. However, if you were to apply the definition of a linear system namely as one whose input-output relationship is governed by an equation of the kind L 1 of x is equal to L 2 of u, you will quickly notice that this particular system d x by d t plus x equal to u is actually a linear system.

So, here are therefore two counterintuitive examples one where the input, output relationship is a straight line, but because it does not have a zero intercept either with the x-axis or with the y-axis, it is actually not a linear system. And in the other case, where the input-output relationship is not a straight line it is a non-linear curve, but when you look at a differential equation relating the input to the output it satisfies superposition and scaling, and therefore, the system can be considered as a linear system.

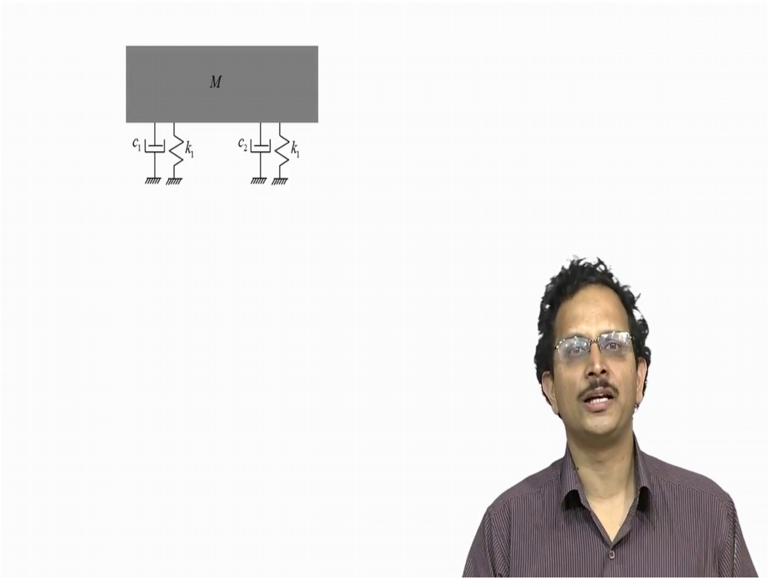
Now, why did we choose to define systems that satisfy superposition and scaling as linear systems. As I said that is because, there is a large number of engineers engineering systems that fall under the ambit of linear systems, if we choose to define linear systems in this particular manner. So, let me know give, let me know validate this claim by giving examples of various physical systems that obey such relationships input-output relationships.

(Refer Slide Time: 18:32)



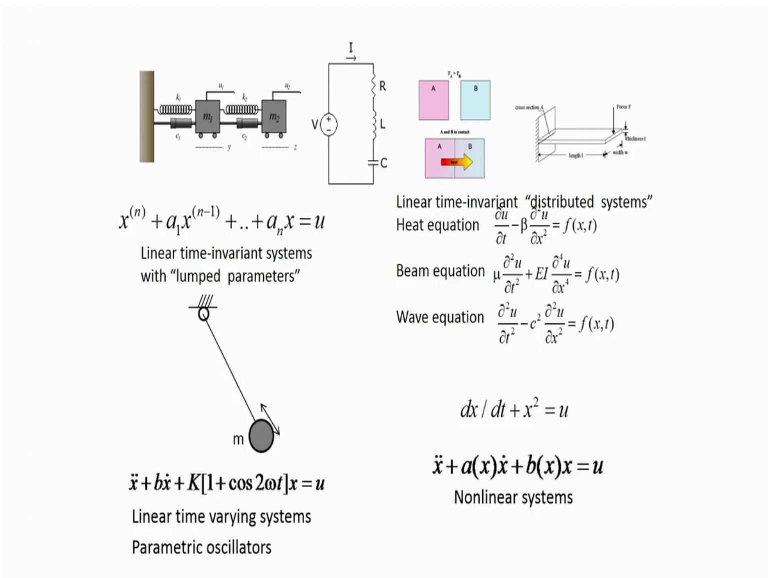
Let us first take the example of a car. Of course, all of us know, but the mass of a car is distributed over its entire body as also is its stiffness and damping coefficient and so on. But for purposes of analyzing how a car might respond to vibrations that it might receive from the ground as it is traveling. So, when one looks at the model of vertical vibrations of a car. For all the practical purposes, one can model, the entire body of the car as one single mass and this mass can be modeled to have been suspended by a combination of springs and dampers.

(Refer Slide Time: 19:11)



In this particular case, the dumping coefficient, and the spring constant are determined by the shock absorbers, and the tires and so, on. And the mass the value of, the mass m is determined by the overall weight of the car. So, one can therefore come up with an approximate model, which captures the important dynamics of a car moving down a bumpy road. By using a simple model, where the mass has been lumped at the center of gravity of the car; and the stiffness, and the damping coefficients have been lumped at the locations of the shock absorbers, and the tires and so on; such a model is called a lumped parameter model.

(Refer Slide Time: 20:06)



It is an approximate model, but the degree of approximation can be improved by using more springs and dampers connected in some particular fashion in order to better represent, the dynamics of the system. Likewise, in the electrical domain, one can think of lumped elements of resistors inductors and capacitors to model an electrical circuit, all of these can represent a wide variety of physical systems. So, in the mechanical domain, in addition to the car whose example I just took this network of springs, dampers and masses can be used to model for example, civil structures such as bridges and buildings and so on. It can be used to model aircraft wings entire aircraft itself and all and so on and so forth.

So, one can come up with good approximate models by stringing together these springs and masses and all of that. And such models, where these masses and springs, and other elements are lumped at specific locations are called lumped parameter models. And for small deformations of the object about its mean position, the differential equation that relates the output of the model to its input would have the form of the kind that is shown in this slide.

Here, x superscript bracket n represents the nth derivative of x with respect to time. So, the general dynamic model would be nth derivative of x plus a 1 times the n minus 1 derivative of x plus so on and so forth, plus a n times x is equal to u. And this you can easily verify represents a linear input-output relationship. Here u is the input, and x is the output. And just by applying blindly the rules necessary to validate superposition and scaling one can verify that it is a linear system.

So, here is an example of a dynamic model, which represents a wide variety of physical systems both in the mechanical domain as well as in electrical domain that will be modeled by using linear time invariant ordinary differential equations and can therefore, be modeled as linear systems. Of course, one need not have to model a system such as a car with distributed mass, and distributed compliance and so on. As a lumped parameter system; one can also consider that the inertia and the stiffness are distributed over the entire volume of this object. In such a case, one has to employ distributed parameter systems, which are essentially systems that are governed by partial differential equations.

So, if we take the example of a thermal system, for instance this room, we would note that one corner of the room would probably be hotter than the other corner. For instance, the top of the room might be hotter in summer than the bottom of the room, and there might be temperature variation along the length of the room as well. And certainly the temperature of the room will also depend on time. So, every part of the room would be colder at the middle of the night than for example during midday.

So therefore, there is a spatial variation of temperature as well as a temporal variation of temperature, and this is captured by what is known as the heat equation. And you can verify that this equation is again one that represents a linear system here f of x comma t represents the input, this input can vary both with time as well as with space, and u here represents the output. Likewise if one considers an aircraft wing that is flexing under the effect of wind load and so on. The deformation of the wing is dependent on the distance of the point, where you are measuring from the base of the wing. And if this wing is vibrating in the breeze, then it also depends on the time at, which you are measuring the vibration.

Therefore, there is once again a temporal dependence of the deformation as well as a spatial dependence of the deformation, and both these together are dependent on the input force which could itself be dependent on space as well as time. And this once again by applying the rules for determining a linear whether it is a linear system or not can be verified to represent a linear system, here f represents the input and u once again represents the output.

Likewise the famous wave equation is also one that represents a linear system. So, we have therefore, lumped parameter systems which are approximate models of physical systems of various kinds we have distributed systems, which once again model physical systems of various kinds all of which fall under the category of what we define as linear systems. In this two sets of examples I have considered, the coefficients of the of the derivatives of x maybe the partial derivatives or the regular derivatives are all assumed to be constant, but that is not really necessary.

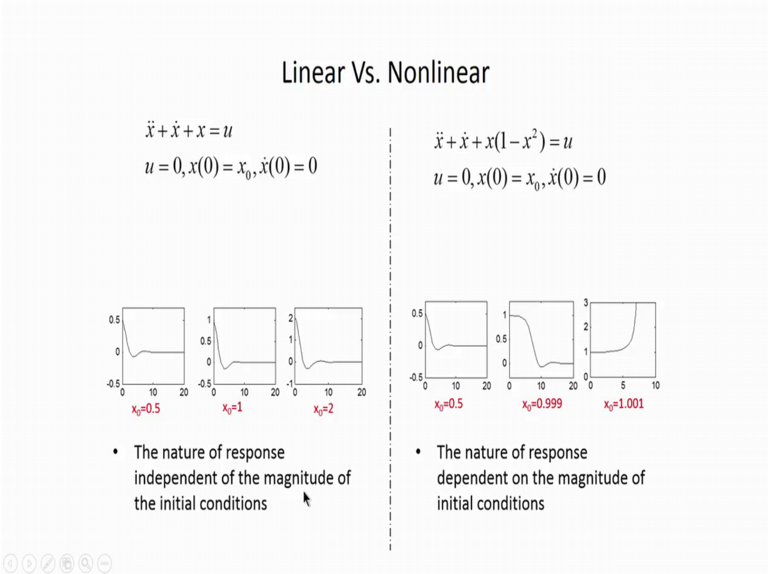
So, for example, if you have a differential equation of the kind that is shown here, x double dot plus b x dot plus k times 1 plus cos 2 omega t times x is equal to u you see that the coefficient of x has a time varying component. Now, one can easily show that even this represents a linear system even though there is a variation in time it represents a linear system and that is so because, it satisfies homogeneity and additivity.

Now, this kind of an equation is obeyed for instance by a child swinging himself or herself on a swing. So, the child sits down. When the swing reaches its extreme position and stands up, when the child, when the swing is at the bottom position, and that way it is able to pump himself or herself, and overcome the losses in the oscillation amplitude due to friction effects. More from a more practical perspective a an input-output relationship of this kind is also seen in the case of ion traps, where time varying electromagnetic fields are used to trap ions within a certain volume. And the dynamic equation of the ions is given by an equation of that looks something like this.

So therefore, the first important reason why our definition of a linear system in the manner that we adopted a couple of minutes back namely L 1 of x is equal to L 2 of u is useful is because, it encompasses a wide variety of physical systems. Now, how does a non-linear system look like it is worthwhile to build our intuition to be able to identify nonlinearities in the dynamics of a physical system. So, here are a couple of examples that are that represent systems that are non-linear.

So, in the first example we have d x by d t plus x square is equal to u. And just by inspection you see that you have a term x square here and therefore, you would one you would suspect that the system is non-linear and it indeed it turns out to be so. Likewise, any other second order differential equation with the damping term or the stiffness term that is non-linear. Or in other words dependent on the variable x or x dot or x double dot can be easily shown to be a non-linear system, because it does not satisfy superposition and scaling. So, the first reason: why our definition is useful because of the ubiquity of such models in nature. The second is the mathematical facility with which one can both analyze as well as design such systems. So, let me give two examples to illustrate the mathematical ease with which one can handle linear systems.

(Refer Slide Time: 28:06)



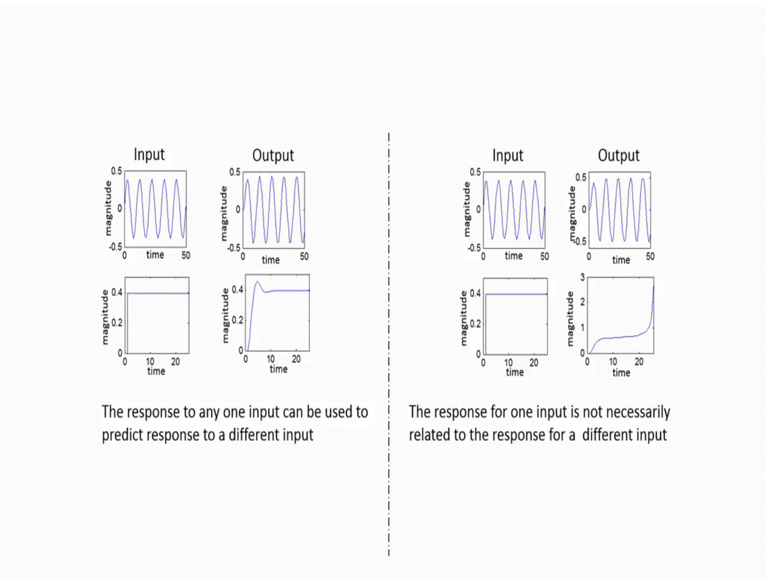
Therefore, in this example, I have considered two systems. On the left hand side, I am looking at a linear system x double dot plus x dot plus x is equal to u. And on the right hand side, a non-linear system x double dot plus x dot plus x times one minus x square is equal to u. In the example that I would consider in this slide, I have set the input u to be equal to 0, but the initial condition x of 0 is assumed to be nonzero some value x naught while the velocity x dot of 0 is also set to 0.

Now, suppose I were to simulate how the, how the state x would change with time if it is released from its initial condition x naught, we see that if it is released from x naught is equal to 0.5, it decays down to 0 after some time, if I were to not change the value of x naught and make it 1. It once again decays down to 0, but with a very similar looking, profile time domain waveform. And if x naught were to be made equal to 2, it once again decays down to 0 in exactly, the same manner as it did for the other two initial conditions.

On the other hand, let us say we were to repeat the same experiment in case of the non-linear system, when we x naught equal to 0.5; in other words, the same initial condition as for the linear system, we see that the way that the system decays to 0 is very similar to that of a linear system. However, if I were to bring the initial condition close to 1, I have not made it exactly 1; I have made it x naught equal to 0.999 very close to 1. We see that already there is some difference the shape of the curve as it decays down to 0 qualitatively looks different from how it appeared when x naught was equal to 0.5.

Now, let us increase the initial condition just a little bit from 0.999 to 1.001. And what we see, what we see is that the solution explodes x of t tends to infinity. And this is one unsettling fact about the behavior of non-linear systems in that the nature of response is dependent on the magnitude of initial conditions. When I say the nature of response I am talking about the kind the appearance of the solution for scaled values of initial conditions, whereas in case of the linear systems, the nature of response is independent of the magnitude of initial conditions. Likewise, in the next slide we shall consider what would happen if we were to force our linear and non-linear systems by using different inputs.

(Refer Slide Time: 30:42)



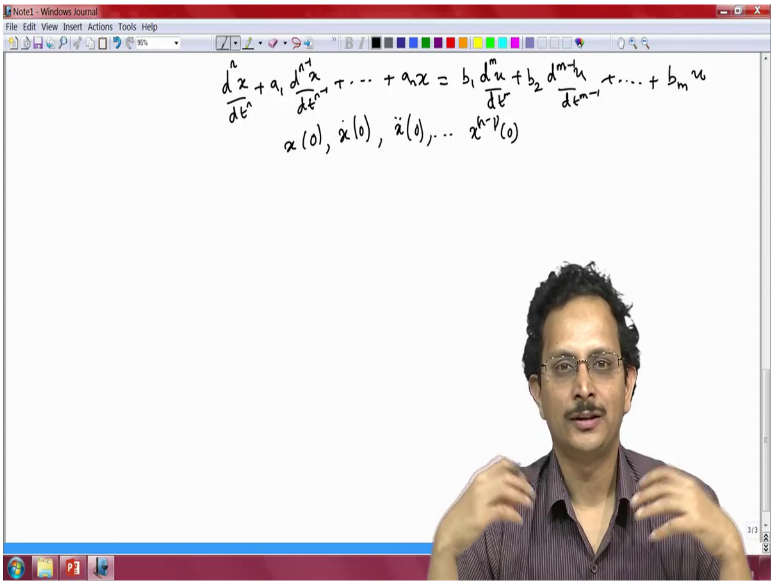
So, on the left hand side, once again I am forcing my linear system with a sinusoidal input. And what I see is that the response is also sinusoidal. It is shifted in phase of course, and also modified a little bit in magnitude, but in the sinusoidal in steady state. What I can conclude qualitatively by looking at this graph is that if I provide a bounded input maybe a sinusoidal input, I get a bounded output a sinusoidal output ok. Likewise, if I provide a step input I get a response that looks something like this once again a step input is one that has whose final magnitude is bounded and the response therefore is also bounded.

On the other hand, let us consider the response of a non-linear system or the same non-linear system that, we considered in the previous slide. So, for a same sinusoidal input, the output looks qualitatively very similar to what was obtained from a linear system, it is, it looks sinusoidal in steady state. However, if I were to apply a step input what I see is that my solution once again explodes. So, here is another unsettling fact about non-linear systems; in that the response for one input is not necessarily related to the response to a different input. On the basis of the response to a sinusoidal input, I might have been led to assume that if I were to apply a step input I once again have a finite response, but that is not really the case.

On the other hand, in contrast, if you look at a linear systems the response to any one input can be used to predict the response to a different input. And this prediction is not just a qualitative prediction of the kind that we did in this slide, but actually we can show that quantitatively if you have the response to any one input you can actually predict, how it might look for a different input.

So, with these two huge advantages in favor of defining our linear systems in the manner that we did namely the ubiquity of physical systems, which obey linear dynamics as well as the ease with, which one can analyze differential equations that are linear time invariant and so on. We shall stick with this as a definition of a linear system. However, from the examples that, we considered a couple of slides back, we saw that linear systems, encompass lumped parameter systems, distributed parameter systems and systems whose parameters are varying with time. Now, in this course however, we shall restrict ourselves to only lumped parameter systems, whose coefficients are not changing with time.

(Refer Slide Time: 33:41)



In other words, I shall write down the equation of that connects input and output for all the systems that would be considered in the course of this these lectures. In general, the input-output relationship would look something like this n-th derivative of x with respect to time d n x by d t up d t to the power n plus a 1 d n minus 1 x by d t for n minus 1 and so on and so forth plus a n x is equal to b 1 d m u by d t power m, m-th derivative of u with respect to time plus b 2 d m minus 1 of u by d t power n minus 1 and so on and so forth plus d m times u.

This we will restrict ourselves in this course to systems whose input u and output x are related by differential equations of this kind. Of course, when we as control engineers, we are interested to know how these physical systems behave in response to certain inputs u and that is what we will be looking at in the next video clip. But before we get to that I want to point out that it is not sufficient to simply state, the differential equation that relates the input and output one also has to specify the initial conditions to be able to derive mathematically, how the output would look like for a given input

So, we shall specify n initial conditions because, on the right hand, on the left hand side, we have an nth order differential equation. So, n initial conditions are x of 0 x dot of 0 which is the time derivative of x at the time p equal to 0, x double dot of 0 and so on and so forth up to x n minus 1 of 0. So, we have this differential equation with these initial conditions. We shall now start out by looking at how one can obtain the response x of t for the specified set of initial conditions and the specified input u.

Thank you.