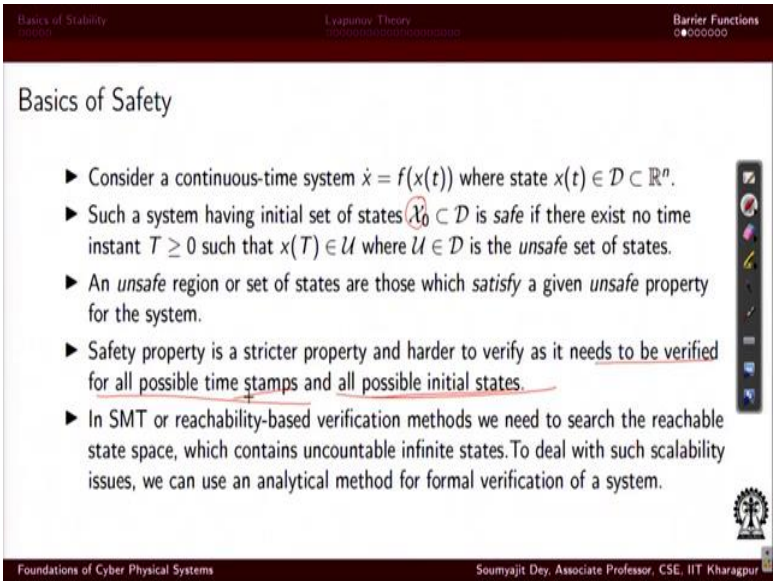
Foundations of Cyber Physical Systems
Prof. Soumyajit Dey
Department of Computer Science and Engineering
Indian Institute of Technology – Kharagpur

Lecture – 42
Lyapunov Stability, Barrier Functions (Continued)

Hello and welcome back to this lecture series on Cyber Physical Systems. So, in the last lecture we discussed about Lyapunov functions and stability. So, in this lecture we will be talking about this new concept called Barrier Functions.

(Refer Slide Time: 00:37)

(Refer Slide Time: 00:37)



So, this is a method to kind of reason about the safety of such autonomous systems. So, let us first understand the basics of safety. I believe we also discussed about safety when we did hybrid automata-based modelling of systems and reachability. So, suppose you have a continuous time system like this and for this system you have this equation x dot equal to f(x) of t and x(t) is belonging to this domain D.

Now, this system ah has initial set of states. ah Let that be x(0) and it is safe if and I mean I mean, let us say that is safe. Right. But the question is when do I say that this system is safe? There should be some notion of I mean unsafe also. So, let us say that inside this domain on which the system is defined, there is another set U. ah So, U is ah U is more like a ah subset. So, this is not belong to these are subset symbol here. Ok

So, you have a subset of D which we call the set of unsafe states. That means, ah it may or may not be possible for the system that even once because it is autonomous, it will be moving around, it will be changing its location in the state space. ah So, it may or may not, ah get into those set of states which are in U. right. And we say that well that is unsafe because they may be not, I mean they may be hazardous for the operation of the system. Ok

Now, the question is how do I define ah the system to be safe. So, what we say that well let the set let the system start in a set of initial states given by X naught that means it is a set of initial states. That means, ah it can start anywhere inside, ah any because X naught. All these are is in a continuous space. Right. So, X naught is also so, a set containing an infinite number of ah possible initial states of the system.

And we say that well this is the set of safe initial states provided as long as the system starts from this Ah from any state which is included in X naught, there does not exist any future time instant t. Such that X at T will be a point located inside you. So, let us understand what we are doing. We are. We are defining something in your state space as unsafe. So, ah we picked up let us say so, this is your domain and then let us say we said that well this is a region which I want to avoid.

Then what is my set of safe initial states? We will say that this is my set of safe initial sets Ah states, as long as if I can start from anywhere here. So, let us say this point x is where I am ah at t naught Ok. Initial time. I have a guarantee that in future, at no point in future, I should land

up inside U. ok So, as long as I have a set X naught so that I can start from anywhere inside this set and in future, there is no time point when I land up in U.

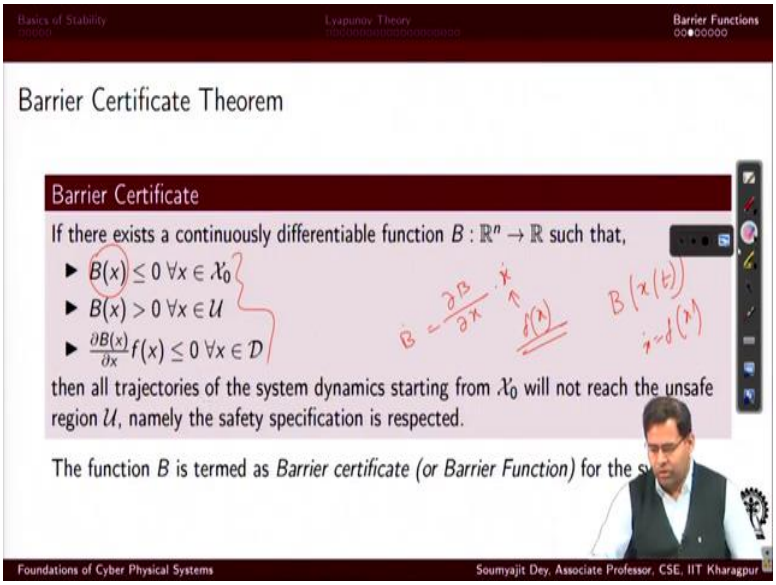
I will say that this set of initial states is safe. Ok. So, ah an unsafe region is something ah which will satisfy some given unsafe property of the system. So, there will be some definition. Let us say for a vehicle, this air to fuel ratio is bad. Right. So, there is an unsafe region in the state space. The vehicles controller should be designed in such a way that overall, it never happens that you get to that kind of air to fuel ratio. Something like that. Ok.

Now the safety property is actually, a stricter property and harder to verify and it needs to be verified for all possible time stamps and all possible initial states. So that is an important thing. You see we are talking about that well we will say that this is safe. If I can show that I can start from anywhere inside this and starting from anywhere in the inside this in future wherever I go, I never land up in U. right.

So, it needs to be verified for all possible time stamps and all possible initial states. Right. So that that is the problem for all possible X naught in this capital X naught and all possible values of T is greater than 0. Right. So, in reachability-based verification or SMT solvers we need to search this reachable space which contains such an uncountable infinite state space, and then the question is how do I deal with that.

So, we will like to have some analytical method for proving that some system is safe indeed and it will never get inside an unsafe region.

(Refer Slide Time: 05:30)



One such method is what we call as barrier certificate. So, let us understand what is the barrier certificate. It is a very simple thing if you have understood Lyapunov function it is just like an extension of that. So, this is also a continuously differentiable function, mapping to a scalar value in the reals. ok So, what it does is you have, you have certain properties of this function which have to be true.

The first property is for all elements in the initial set of states this function must be negative or 0. So, for all x which is inside the set of initial values ah V(x) is negative and for all x which is inside that unsafe set of states, B(x) must be positive. So, notice the trick you are given a system and you are given the set of initial states and you are given the set of unsafe states. Let us say using some closed form expression.

Now, your job is to find the function which is negative for any point which is there inside this initial set of states and it is positive anywhere inside this set of unsafe states. And apart from this for all other positions, I mean, apart from this in general, for all points in this domain of the system, ah the derivative, the time derivative of this function is negative. So, you can see that it is something just like the Lyapunov function I mean.

But with some differences there is a notion of unsafe state where the function is going to change, it is sign, otherwise the function is negative and the time derivative of the function. If you take time derivative of B it is basically the space derivative of B times the time derivative of x which is f(x). right. So, the time derivative of B is negative everywhere in the in the state

space. Right. So, if you can find such a function then you can have a guarantee that every trajectory of the system dynamics which will start from X naught.

So that is the trick, the trick is in finding the function. And then you can claim that well ah I have found a function Ah such that ah see this is this function. If I, if I am able to initialize my system inside X naught this function is negative by definition and ah everywhere in the inside the state space, the rate of change of this function with respect to time ah that is negative. Right. So, it is a decreasing function that means ah if it is initially negative.

And it is a decreasing function, it is not possible for this function to ever be positive. And hence I will conclude that since it cannot ever be positive, it will my system that means my system projector is never going inside U because if my system trajectory has gone inside U, then B(x) would have been positive but if I can figure out the function in such a nice way that you satisfies point one and it satisfies point three.

Then I see that well ah there is no scope of the function to ever be positive and so, ah the system I can say will never actually move inside U. ok So, there is a trick to find out a function which will satisfy all these quantities. If we can figure out that function then we can clearly say that all trajectories of the system dynamics which start from X naught they will not reach the unsafe region. Ah

That means in a way it will we can say that the safety specification is respected. So, this is called a barrier certificate or barrier function for the system. So, suppose you are given a system and you are and it has a I mean complex algebraic formulation or something like that and you are told that well you prove that this is a safe system and this is my unsafe set of states.

And can you show that this will never my system will never go into that unsafe state. If you can design a function like this OK ah then it is like a certificate which is telling that well you see, since I have found this function which satisfies all these requirements, this system that you are talking about will never go into the unsafe state. Now, again just like Lyapunov function, you need to understand the link between this function and the system.

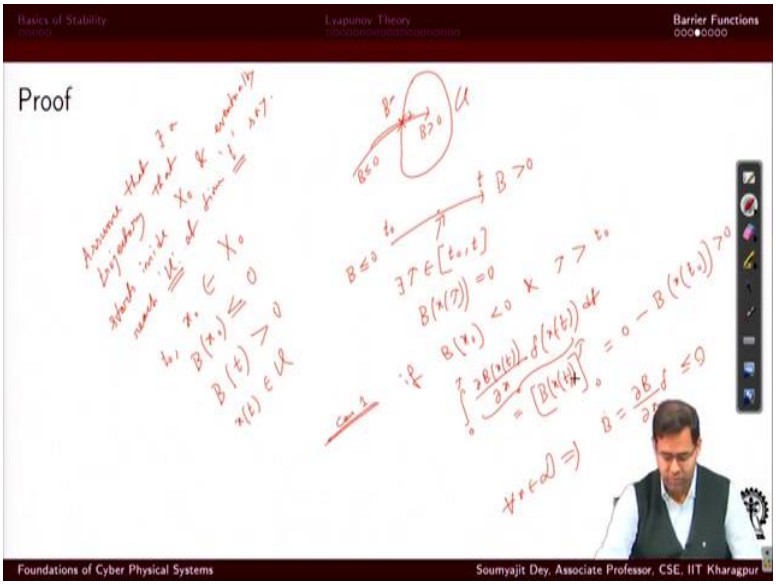
So, the link is again by this that when we say time derivative of B it becomes the space derivative of B times x dot and that is f(x). That means the function is defined over x and x is

changing position with time. That means you are practically evaluating the function only on the trajectory of the system x dot equal to f(x), x dot equal to f(x) is a trajectory ah which is changing x with time and you are always evaluating the function on this trajectory.

As long as this function satisfies this requirement, you are starting with the function being negative and it is rate of change is negative. So, the function will remain negative and that means in a way since the functions evaluation is on the trajectory. The trajectory will never move into U because had the projected of the system moved into U. Since you are evaluating the function over time again and again.

And you are now going to evaluate it in such a position where the function will be positive. Ok. But as long as that does not happen, if you need 1 and 2 then well your system does not really go into the set of unsafe states.

(Refer Slide Time: 10:55)



So, there is a nice elegant theorem, ah which defines ah barrier function. Now Let us prove that well why this should happen. Why I can claim that? I think the idea is very intuitive but again we can give a formal proof. The claim is that well if this function exists then ah the system is safe. So, let us try so, we will try this proof by contradiction. So, we will assume that such a trajectory exists ah for which the system can indeed start in this set of initial states and the system will rear into an unsafe state.

So, it will, it will start at X naught and eventually reach this U at some time t say. So, let us say at time t naught I am at X naught which is inside this ok. So that means the function B at X

naught right, it is negative because X naught belongs to ah this capital X naught (12:54) initial state. Right. But say I reach U at time t then B at this time t is positive. Right. Because x(t) is a member of the unsafe state U, unsafe set of states U. right.

So that means ah say B is a continuous function. right If you look at our definition, these are continuously differentiable function. So, you have safe and you have unsafe right either say you are unsafe. So, if you are moving like this, there must be a position here where you cross it over and if you see we are saying that well when I started, I was negative and then suddenly, when I reached here, I became positive and here I was negative.

So that means there is some position somewhere here right somewhere here, where B equal to 0. This moves to be greater than 0. That means there exists some position where so, let us say this was my initial point. right ah This is ah t naught and this is t. So, there exists some tau right ah which is a member of this interval t naught 2 t. ok. And we can say that B at x of tau that is equal to 0. So that is where the function becomes 0 and then again become positive. Right.

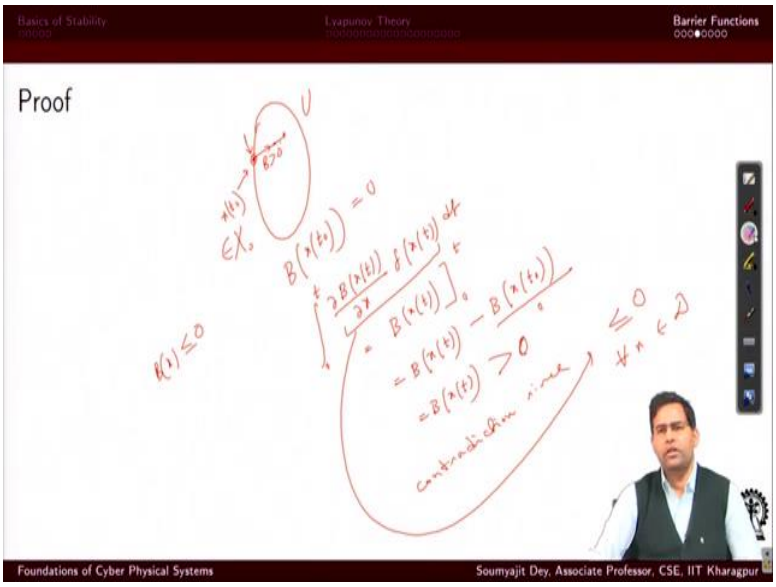
So, ah we consider two cases, tau is greater than this initial point some point here and let us compute this integral. So, this is like ah if you see this thing, ah this is nothing but a time derivative of B right because it is equal to space derivative of B times x dot which is f(x(t)). right So, I can always write this as where t is evaluated at 0 at tau. Now, at tau I have said that B is 0, so, minus B at X t naught. Right

Now, we know that this is less than or is equal to 0. Right. So, we will have this as positive. But this, in that case, ah is kind of contradicting the third condition. Right. Ah Because if you see the third condition says that the integrand which is this, the integrand is always negative for all positions in the domain. So, for all X in the domain we have that B dot which is given by del B del x into f ah that is this. Right.

So, there is a third condition now, if this third condition is true, ah then this integral ah has to be negative but we, whatever we get is positive. Right. So, this is like a contradiction. Now, there can be one condition which is that the initial state is at the boundary of the certificate itself. Ok. That means you have because if you see that you may have a question that well ah we had this. right ah

Let us see this is B x is less than or is equal to 0. Right. And then it become positive inside and somewhere in between it has been 0. Right

(Refer Slide Time: 18:00)



So now, ah let us consider this condition that ah let us assume the initial state ah is not outside, not properly outside this unsafe region but it is somewhere here only at the boundary. Ok. So and why is this important? Let us understand. Earlier, we said it is here and we went into here where inside once I get inside, it has to be B will become positive. Ah That means somewhere in between it became 0.

Now, let us say that well ah it may happen that it becomes 0 right on the boundary. Ok. So, if I mean ah it may I mean this this equation also holds in the boundary situation that will X does not start outside but X starts in the boundary itself. Ok. So, this is your position at initial time. So, suppose ah that is the situation. Now if that is true then B of x t naught or let us say this is called X naught, this itself is 0. I mean it is not negative or less than or is equal to 0.

This itself has to be 0 because the point is the moment I cross the boundary, it has to be positive. It has to be strictly positive. So, if I am on the boundary then this is the last point where I can be 0, after this I have to be positive. right So, ah and as long as I am also a member of the initial state ah set of initial states, my condition is ah B at a x has to be less than or is equal to 0. Right. Now, due to the continuity condition, the moment I cross over ah I have to be positive. I have to write so, this is where I will be 0. Right. So now, again, if we compute the integral that is essentially B itself because this is what is the time derivative of B. And this is 0 and this is

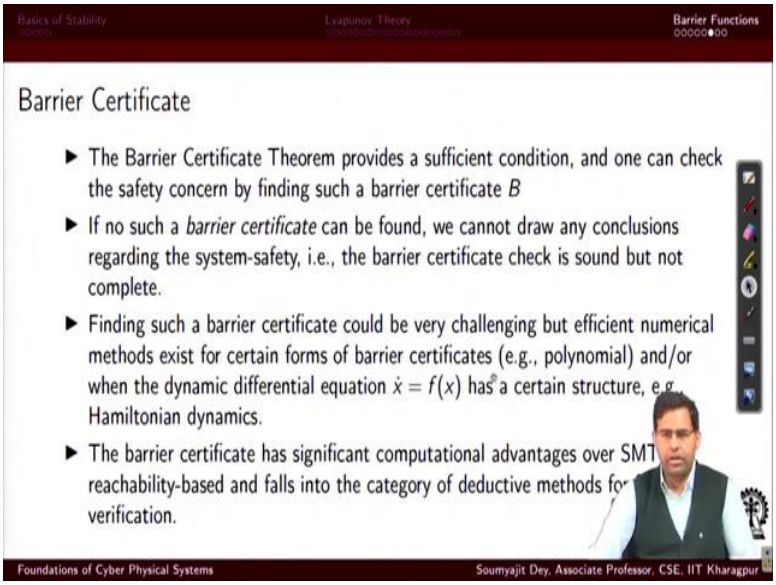
where I am located at time t which is inside U, so, this is positive. Right. Now, you see this again is contradicting the third condition because what we are integrating is negative. Right

So, if we are integrating something negative ah entirely right and then it cannot be this I mean the integration out. output cannot be negative. Like I mean it cannot be positive here. Right. uh Since this thing is less than or is equal to 0 right for all X belonging to my domain here. right So, the state trajectory also ah could not have gone to unsafe. So, in any way we are trying to do it, we are getting a contradiction. Ah

So that means ah this theorem must be true, that well if those conditions happen, ah then the system must not be able to go to an unsafe state. And I believe it is it has been very clear right from the outset. Right. ah If you just look at the theorem, you can always intuitively claim that we have the function ah to be negative or equal or is equal to 0 and if I get inside the state space of unsafe region, ah we must be fully positive, strictly positive.

Now, if initially I am 0 or negative and my time derivative says that I can only decrease in any way, I cannot ever get inside this unsafe region. Right. So that so that is about the proof. It is a simple intuitive as that.

(Refer Slide Time: 23:07)



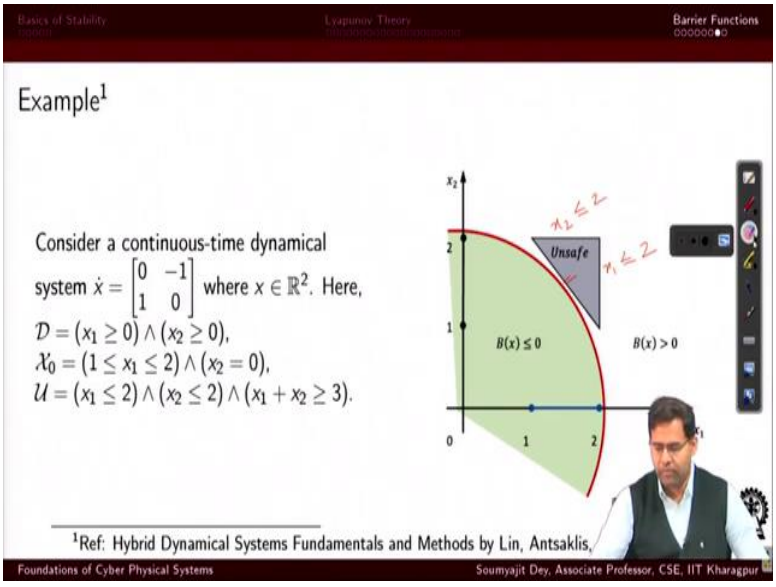
So, fine. ah what is the usefulness of this? So, this provides you a sufficient condition and one can check the safety concern of a control system if, by figuring out a barrier certificate. However, you, if you cannot figure out a certificate, you cannot draw a conclusion like this.

But ah I mean so this is ah I mean as long as you can figure out this certificate ah you are fine that is a sufficient condition.

So but the issue is how easy is finding that certificate? It can be challenging ok. But there exist efficient numerical techniques for funding so, findings some certificates. That means you can, you can guess the polynomial and you can give that polynomial all these constraints of various certificate and ask a solver to solve this for you. It may or may not be possible for the solver. Because so, there exist this efficient numerical methods for certain forms of certificates like this polynomial fitting. Ok

So, you get you, you create a polynomial form with unknown coefficients and you ask a solver to figure out what should be the coefficient. So, the polynomial, so that all these conditions get satisfied. And so that may be possible in that way. Ok. Now, these are some advantages with respect to SMT and other techniques because these are deductive method. If you can have the have this function synthesize then you can just directly say that you are safe. Ok

(Refer Slide Time: 24:33)



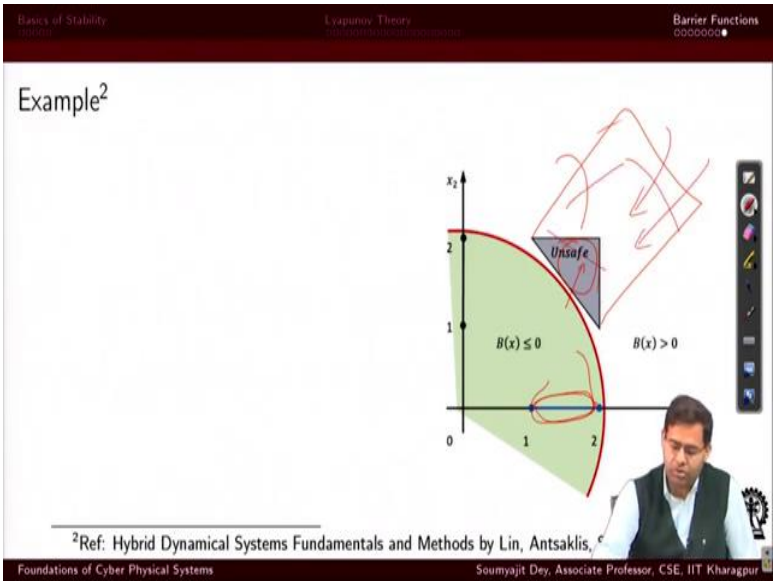
So, here lets take this example. So, you have a continuous time dynamics like this and let us say your initial region ah is a. I mean this is your domain. So, it is the is the is the is the positive ah I mean quadrant ah with x and x_1 and x_2 both being 0 in the greater than or is equal to 0 in the phase plan. Ok And you have an initial region, the initial region says that well ah the value of x_1 is between 1 and 2. Ah So, ah this is the initial region.

That means that well and x_2 equal to 0. So, this is the initial region marked in blue. So, x_2 equal to 0. That means you can you can be starting at any of the points on this line, starting from this point 1 and ending at 2. Ok And you have an unsafe region. So, unsafe region specification is using this triangular form. Because it is an intersection of 3 for 3 Ah linear equations. Each equation will give you one line here right linear in equations.

So, x_1 less than equal to 2. So, x_1 ah less than equal to 2 is what, ah so that means x_2 can be anything so, you have this line. That means anything because if you go beyond this x_1 is greater than equal to 2. So, everywhere in inside this line and then x_2 less than equal to 2 then everywhere on the on this side of this line, which so, this is your, if I just mark them for you x 1 less equals 2. This is x_2 less equals 2 and the other one is this line x_1 plus x_2 greater than 3.

Because if you are on this side of the region, you will have x_1 plus x_2 less than equal to 3. So, they together give you this triangular shape and this is let us say, your unsafe region. Ok. So, ah fine, how do I show that the system is safe or unsafe? ok

(Refer Slide Time: 26:28)



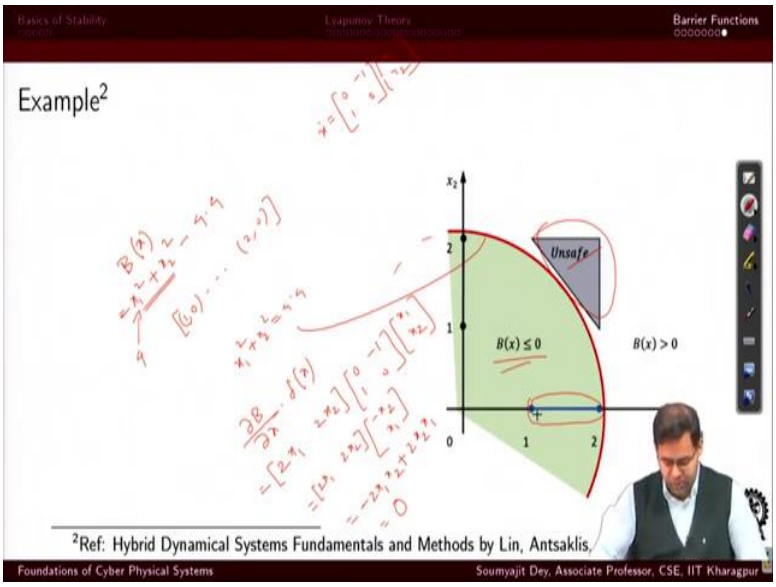
So, let us let us try and figure out. So, my question is that well if I start from anywhere here, is it possible that I get inside this? So, I have to design a function suitably so that the function the moment it gets in here it becomes positive and in my initial region the function is negative and the rate of change of the function with respect to time is always negative. So that is my requirement. Right

Now notice one important thing here. It does not mean that your function should be positive only in the unsafe region. It can be positive at a upper, at a higher approximation of the as a at an approximation, an outer approximation of the unsafe region also. Ok. Then because then all it means that is that it will you are restricting your system to go beyond not only the unsafe region but so, what I am meaning is that let us say I do not get a function which is exactly positive here.

But ah I mean which is which is exactly positive here and it is negative everywhere else but maybe I get a function ah which is positive, Ah in this region and apart from this region everywhere it is negative. Suppose I designed that function. Is that ah is that a proof that my system is safe? Definitely, yes, right because then again, I am able to show that well I have I have this function once I once I can figure out a function which is positive here.

And negative everywhere else, along with other conditions. That means my system can start in the initial region and it will never reach anywhere inside this and since it will never reach anywhere inside this, it will also not reach anywhere inside this right. So that is how we can we are defining the certificate then.

(Refer Slide Time: 28:31)



My point is it need not be that I am only positive in the unsafe region I can be at ah bigger region also which contains the unsafe region. So, with this idea in mind, let us create a barrier function which is something like this. But ah well if I take something like this, ah what is my value at 1, 1? Well it is fine ah am I positive somewhere let us see. What I want is anywhere in my initial region I have to be negative. Right

So, ah my question is B(x) has to be negative, at least in the initial region which is this. right So that means, if I have ah x this is my x_1 so, x_1 can be 2. Right. So, even with x_1 equal to 2, I will have this as 4 right but I still need to be negative. So, ah let us say I make it something like 4.4, something like that. Right. So, you see what happens? ah What is the nature of this function? Ah Then, if I take my initial region which starts from ah 1, 0, x_1, x_2 up to 2, 0, the max value is 4. Right

And so that means B(x) is negative everywhere here. Where else is B(x) negative? That is the important question. So, we can see that B(x) will be negative, ah in this entire circle. If I just extend it like this, ah with this radius right because if you see these are, this is like a equation of a circle. right So, what I have is x_1 square plus x_2 square Ah I mean as long as the I draw a circle with radius 4.4 everywhere inside this it will be negative, on the circle it will be 0 and outside the circle it will be positive. Right

So that is that is what is happening. So, on the circle I will have this equal to 4.4, inside this it is negative and outside it is positive. Right. So, ah so that is about it. ah Now I can show that this function B(x) equal to x_1 square plus x_2 square minus 4.4. This is negative for all the initial states, ah which is true. right ah Because if you put the value the max is 4 right is negative. And if you now take ah B(x)'s value ah here, ah um I mean and technically like I said that if I just take this segment here.

This is the segment where it will be negative in the state space and outside this circumference here ah it will be positive. Right. Now, let us check the other condition which is ah del B del x times f x. So, by the way, this systems Ah. This yeah, this was the systems dynamics. Let me write it here again. This was my system dynamics right and ah. So, if you just use so, this is your x dot right and if you take del V del x then it is $2x_1$ plus $2x_2$. Sorry, ah that is so, when I am taking it, it is basically a partial derivative. Right

So, it will have two components. Right. So, it will be in the first component when I take del B del x_1 so that is $2x_1$ and the rest of it is 0. And when I take del B del x_2 ah it is $2x_2$. right So, this is it and x dot and x_1 and x_2 is also here. Right. So, if you multiply this here, you get minus x_2 in the first row and you get x_1 here. So, overall, what you get is minus $2x_1$ x_2, plus $2x_2$ x_1 so that is minus x_1 x_2 ok yeah Sorry, ah if I look at the last line. Yeah

So, they are actually, cancelling each other, so, this is 0. So, if we see ah the derivative is 0 for all states. Now, if we look at our condition which is my requirement. So, del the derivative of B ah with respect to time which is del B del x times, f x can be less than or is equal to 0. Right. So, if it is equal to 0 that is fine. So that means there is no change in the barrier functions value over time that is all it means. Right

So then as we can see that if I choose my barrier function like this. Sorry, if I choose my barrier function as something like this. I get this kind of a circular region inside which I have my set of initial states and this is negative or is equal to 0 is negative actually here. And beyond the circular region it is positive and this beyond the circular region at I have a subset which is my set of unsafe states.

So, definitely here also it is positive and the rate of change of this function is 0. That means this barrier which is right now negative everywhere inside, is never going to be positive. Because the rate of change is 0. Right. I mean it is not never going to change anywhere. So then it is fine. We can actually, say that well ah this does not change that means the system dynamics that we define is such ah that I have a I have a formal guarantee that this system, ah is never going to reach that set of unsafe states.

So that is a nice formal technique ah to talk about ah stability of such systems. So, with this we will end our discussion here, ah in in this week. Only one thing I will like to mention in our Lyapunov functions and similar formulations, we have talked about these different theorems but you may observe that everything has been in the continuous time domain. They are all applicable in the ah discrete time domain also with some differences.

For example, like ah when I talk about the derivative of the function and we say that it is negative, instead of that we will be writing a difference equation with V x k plus 1 minus V x k instead of saying V dot x is negative. We will be saying V x k plus 1 minus V x k this quantity, the corresponding difference is negative. So, similar theorems also exist, ah for the discrete time Lyapunov functions and we are not covering them here but is just for your information here. Thank you with this. We will change the lecture here.