STOCHASTIC APPROXIMATION: THEORY AND APPLICATIONS

Dr. Gugan Thope

Department of Computer Science and Automation

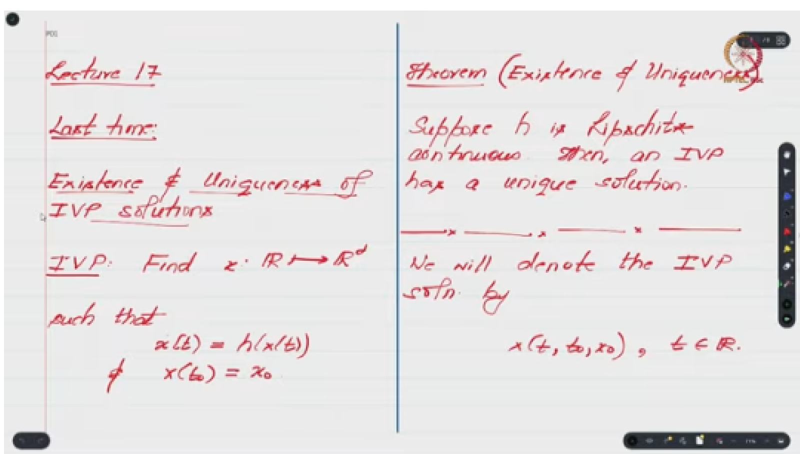
Indian Institute of Science

Week 4

Lecture 17

Continuity of ODE Solutions with Respect to Initial Conditions

Hello and Namaste everyone, welcome to Lecture 17 of this NPTEL course on Stochastic Approximation. So, we are in Week 4 of this course, right? And in this week, we are going to focus on understanding some basics of ordinary differential equations. And let us do a quick recap of what we did in the previous class. In the previous class, we sort of discussed something called the initial value problem and then discussed the existence and uniqueness of solutions to initial value problems. And recall what the initial value problem looks for.



So, we have been given some ODE, we have been given some initial condition, and we want to find a trajectory which, for every time instance T in R, gives something in RD. And the existence and uniqueness theorem said that if your driving function, which is basically this function, is Lipschitz continuous, then the initial value problem indeed has a unique solution. So, notice that this statement makes two conclusions. One is that there exists a solution, and the second conclusion is that this solution that exists is unique. And often, we will denote the solution to an initial value problem by this notation.

So, this notation over here has t as the time index; this t0 will be referred to as the initial time instance, and x0 is the value of xt at time t0. So, this is like the initial time instance, and this is the initial value. So, this is the solution trajectory which, you know, satisfies this initial condition. So, we will use this notation to denote the solution to this initial value problem.

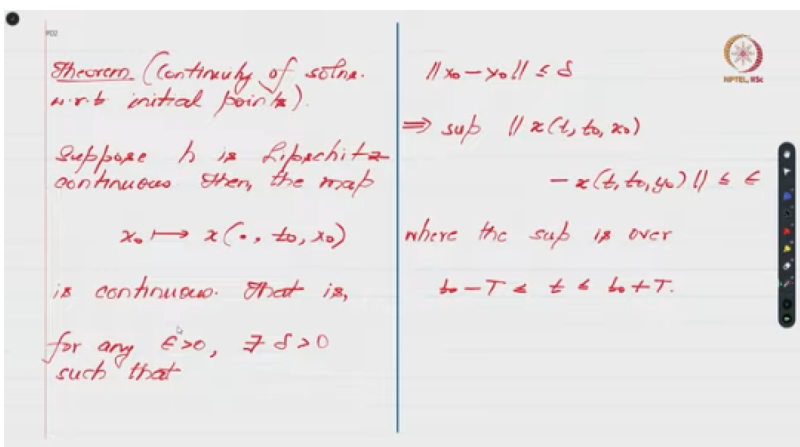
$$h: R \mapsto R^d$$

$$x[t) = h(x(t))$$

$$x\left(t_0\right) = x_0$$

$$x\left(t, t_0, x_0\right), \quad t \in R$$

So, if you recall, the second half of that theorem that we discussed in the previous class spoke about the continuity of solution trajectories with respect to initial points.

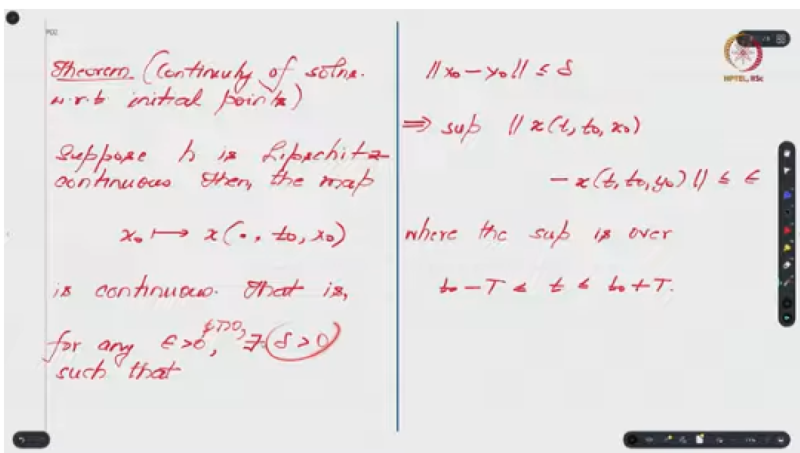


So, what does that mean? It means that if you somehow perturb the initial condition by a small amount, then the solution trajectories will not differ by much. That is roughly what continuity of the solution trajectories with respect to initial points means. Now, this theorem gives a sufficient condition that ensures the continuity of the solution trajectories with respect to initialization, and the theorem reads as follows. It says suppose your driving function H is Lipschitz continuous.

So, recall that this is the same condition that guaranteed uniqueness and existence, right? So, the result says that if your driving function H is Lipschitz continuous, then this map which that, you know, takes as input an initialization point and maps it to this solution trajectory. Okay. So, we will fix T naught.

Right. And the only thing that is changing is this X naught. Okay. So, the map that maps X naught to a solution trajectory that passes through X naught at T naught is continuous. So, what does this mean?

It means that for any epsilon greater than 0 and t greater than 0—so I should say any epsilon and t greater than 0. There exists some delta greater than 0 such that if the two initial points are delta-close, then the two solution trajectories in this supremum metric will be epsilon-close. So, let us understand what this is and what this is.



This is the solution trajectory to the initial value problem which passes through X naught. This is the solution to the initial value problem which passes through y0, right? And we are comparing the distance between them at time instance t, right? And again, notice that both solution trajectories start at t0 and pass through x0 and y0, respectively. And the supremum over here is taken for all t, which is between t0 minus capital T and t0 plus capital T.

$$x_0 \mapsto x\left(\cdot,\ t_0,\ x_0\right)$$

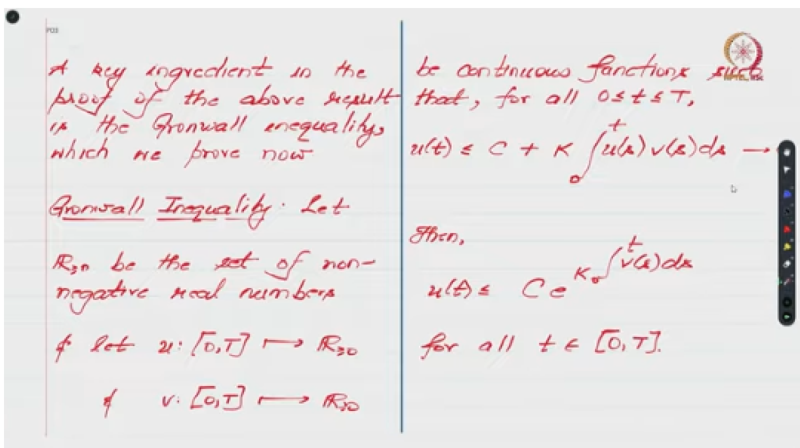
$$\epsilon > 0,\ T > 0,\ \delta > 0$$

$$\left|\left|x_0 - y_0\right|\right| \leq \delta$$

$$\sup sup \left|\left|x\left(t, t_0, x_0\right) - x\left(t, t_0, y_0\right)\right|\right| \leq \epsilon$$

$$t_0 - T \leq t \leq t_0 + T$$

So, the continuity means that you give me an epsilon and t; I will give you a delta such that if the two initial conditions are delta-close, then the solution trajectories in this supremum norm will be epsilon-close. So, we will now see how to prove this result. And to prove this result, a key ingredient is what is known as the Grönwall inequality. Now, Grönwall inequality, you know, on one hand looks very simple. However, you know, throughout the analysis of differential equations and also in the analysis of stochastic approximation algorithms, Grönwall inequality plays a very, very key role.



So, let us try to understand what this Grönwall inequality is. So, let us start with some notations. Let R greater than or equal to 0 denote the set of non-negative real numbers, and for this Grönwall inequality, we need two trajectories on the interval [0, T], right? These two trajectories take values on the non-negative real line. So, notice that u takes as input some value between 0 and t and maps it to a non-negative real number. Similarly, v takes as input some t between 0 and T and outputs a non-negative real number.

So, the statement goes as follows. Let u and v be two trajectories like this, which are required to be continuous. And for every t between 0 and T, u of (t) and v of (t), the trajectory v satisfies some inequality like this. So, u of (t) is less than c plus k times the integral from 0 to t of u of (s) v of (s) ds. Now, c and k are some constants, and this should hold for every t between 0 and T.

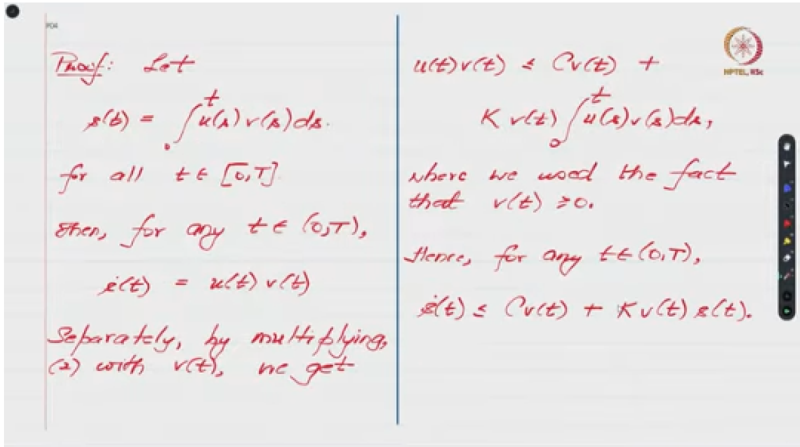
$$U: [0, T] \mapsto R \geq 0$$

$$V: [0, T] \mapsto R \geq 0$$

$$U(t) \leq C + k \int_{0}^{t} U(s)\, V(s) ds$$

So, if such an inequality holds, then the conclusion of the Grönwall inequality is that u of (t) is less than c times e raised to k times the integral of v of (s) ds, and this inequality holds for all t between 0 and T. So, let us first try to decipher what is nice about this Grönwall inequality. It takes as input some inequality and outputs another inequality. So, at a high level, this is what the Grönwall inequality does. So, let us see what the differences are between the two inequalities. The first inequality observed is an implicit inequality.

By implicit, I mean u of t is there on the left-hand side, and you know the u trajectory is there on the right-hand side as well. So, in that sense, this inequality is implicit in nature, whereas here you have an explicit relation—that is, u of t is there on the left-hand side, but there is no u on the right-hand side. So, in that sense, this is an explicit inequality, and that, in some sense, is the utility of Grönwall's inequality—that is, you can take some implicit inequality and convert it into an explicit inequality. Now, let us see how to prove this relation over here. This is a bit of algebra, and I will take you through that, and you will soon see that it follows a very similar trick to what is used for solving first-order differential equations.



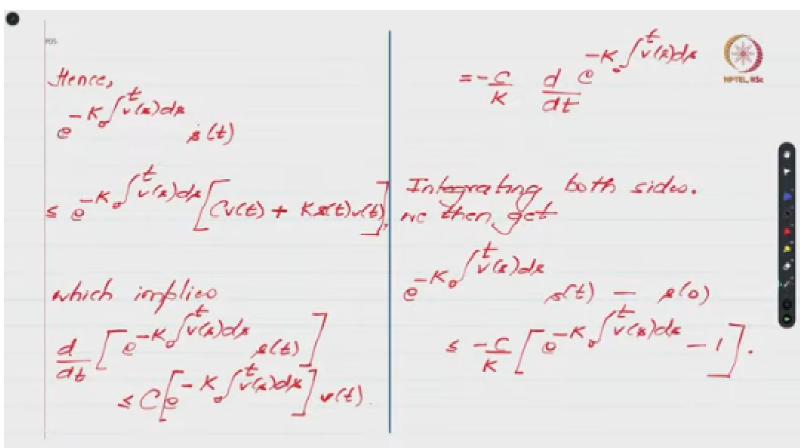
So, what we do is, you know, the expression that we have over here—let us call this S of t. In particular, let S of t be this integral of this product u of s, v of s, ds for all t between 0 to capital T, right? And because u and v are continuous functions and we are looking at the integral, right, from the fundamental theorem of calculus, we know that S is differentiable, and S dot of t is basically the product u of t times v of t. Now, this is some fact that we can observe, and you know we have been given this inequality. So, what we will do is we will multiply both sides by V of t. So, if you multiply both sides by V of t, this expression over here will have V of t, this expression will have V of t, this expression we have V of t, and we have defined this expression to be S of t. So, you can see that u of t times v of t is less than c times v of t plus k times v of t times s of t. So, that is what we have over here.

Is this okay? And notice that the inequality is not reversed. So, whatever inequality you had, you have the same inequality over here, and the inequality holds its nature because of the fact that this V is a non-negative function. So, it is at this place where we are crucially using the non-negativity of the V function. Is this okay?

So, now what we will do is we will sort of replace everything in this expression in terms of what we know or what we have defined over here. So, this left-hand side is u(t)v(t), which is actually ṡ of (t). So, I have replaced this expression over here with ṡ of (t). You have Cv of (t), I have kept it as it is, right? You have Kv of (t), I have kept it as it is, and this integral, right? Is s of (t), and hence I have, you know, replaced that integral with s of (t), right?

So, the integral inequality that we were given before, right? So, this integral inequality that we were given before, by doing suitable manipulations, we have converted into an inequality involving derivatives. So, I mean, the reason is that, you know, if the derivatives satisfy some inequality relation, then the integrals of these derivatives will also satisfy that relation. However, you know, if the integrals satisfy some inequality relation, we cannot say that the derivatives of the functions will also satisfy that inequality relation. So, that is the reason we took this given inequality relation and, you know, converted it into an inequality relation in terms of some derivatives of this s of (t), right?

Now, let us see what we can do over here. Now, observe that here you have a first-order derivative, right, and you have this product v of (t) times s of (t). I am sure you must have seen, you know, tricks to solve differential equations of this form. So, we will use the same trick. So, what we will do is we will multiply both sides by e raised to minus k times the integral of v of (s) ds. So, observe that this thing we have, you know, taken its integral, and we are multiplying both sides by this.

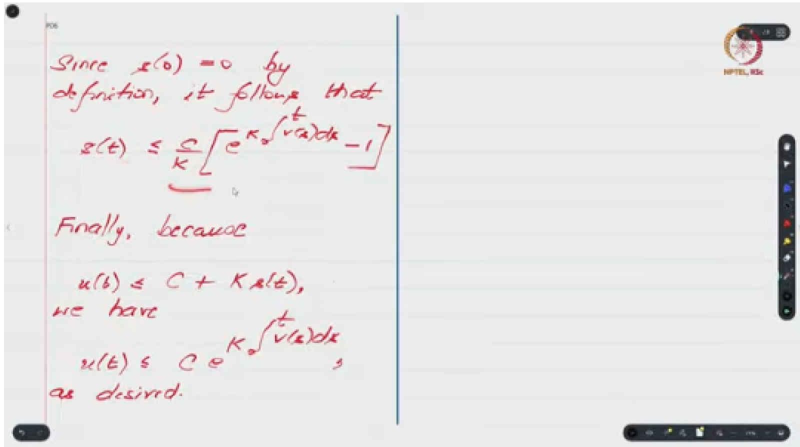


Again, observe that this quantity over here is non-negative; hence, the inequality is preserved. And now, what we can do is we can, you know, take this expression that you have to the left-hand side and whatever you have, you can express it in this way. So, if you see, there is a product of two functions. So, if you apply this product rule for derivatives, then you can see that, you know, the derivative of this product will lead to an expression like this minus an expression which involves the product of these terms. So, that is what you will have, and the right-hand side will basically be left with whatever is over here, which is C times this exponential times V of t, and you know, it is not difficult to see that this expression on the right-hand side can be written in the following way.

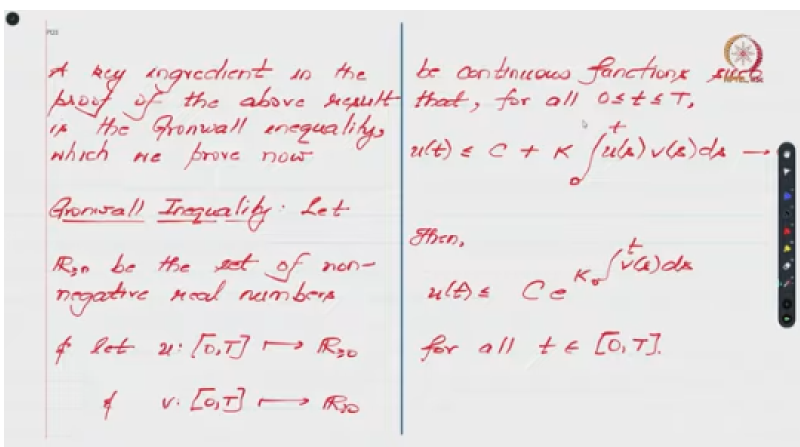
So, you take the derivative of this; this will be, you know, because it is an exponential, you will have the same thing, and then if you differentiate it again, you would have a minus k times V of t. So, this V of t is there, and since there is no minus k, observe that I have divided it by minus k. So, at the end, what we can conclude is that this derivative is less than or equal to this derivative over here. Is this okay? And because two derivatives, you know, have some inequality relation between them, if you integrate both sides, okay, from 0 to t, then that inequality will continue to hold. So, if you integrate this between 0 to t, this is the value that you will have at time t minus the value of this at t equals 0.

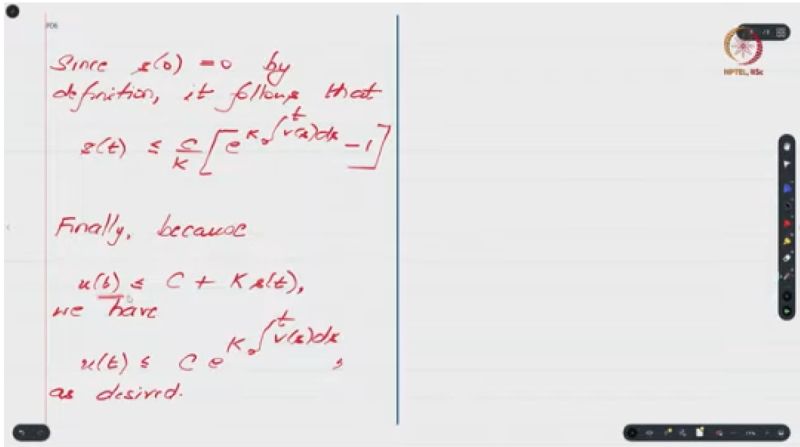
So, when you substitute t equals 0, this expression is e raised to 0. So, this becomes 1; that is why it is not there, and this will become S of 0. Similarly, on the right-hand side, okay, we will end up with minus C over K times the value of this expression at time t minus the value of this expression at t equals 0, which is exactly 1, right? So, we will end up with this inequality over here. Now, recall the definition of S of t. S of t is defined in this fashion.

So, from this one can immediately see that if I substitute t equals 0, this is an integral which goes from 0 to 0 of something, and hence this integral will be 0. So, S of 0 becomes 0. So, I can replace whatever I have over here by S of t is less than. So, I take this expression, multiply it to the right-hand side, and I will end up with something like this. So, now notice that the minus k becomes plus k over here, right? And here earlier I had this minus k. So, if I multiply this e raised to minus k with this, it will vanish, and this negative sign has been pulled inside the square bracket, because of which we end up with some expression like this.

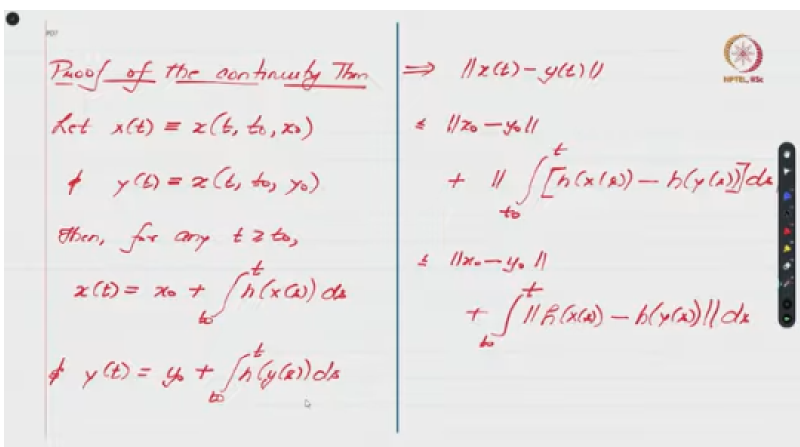


So, we end up. Concluding that S of t is less than this. But our original goal was to obtain a bound on U of t. But if you recall, the given condition was that U of t satisfies some relation like this, and we had defined this to be S of t. And since we now have an upper bound on S of t, that upper bound can be used to derive an upper bound on U of t, and this is exactly what we do. You know, we have an upper bound on S of t, and we know that U of t is upper bounded by this expression. So, if you substitute this upper bound over here, we end up with this inequality as desired.





So, this is the proof of the Grönwall inequality. Again, the Grönwall inequality converts an implicit inequality into an explicit inequality, and that is what we have done over here. So, with this in place, we now go back to the continuity theorem. Recall the goal of the continuity theorem is to show that if the initial conditions differ by a small amount, then the solution trajectories also differ by a small amount over a given time interval 0 to t, right? Both on, you know—sorry, I should not say time interval t, 0 to t—I should instead say the window of length capital T starting from t0, both on the positive side and the negative side.

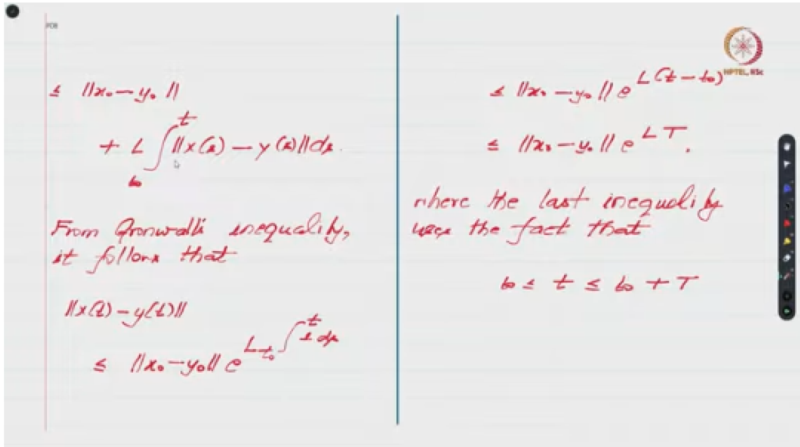


So, let us see how to use Grönwall's inequality to prove this. So, we take two initialization conditions, and you know, let us look at this solution trajectory and this solution trajectory, and we want to look at the distance between them. And for simplicity, to keep the notations and calculations simple, what we will do is we will denote this by X of t and we will denote this as Y of t. This is just for convenience. That is, when the initial condition is y of t, we will denote it as y of t, and when the initial condition is x naught—I should not have said the initial condition is y of t; I should have said the initial condition is y naught—then we will denote this solution trajectory by y t, and when the initial condition is x naught, we will denote this solution trajectory by x of t. So, now we are going to compare the distance between X of t and Y of t, and towards that, observe that because X of t is a solution to the initial value problem, I can integrate both sides and get this relation, right?

X dot of t equals H of X of t; I can integrate that and get a relation like this, and this X naught comes up because at time t naught, your value of x of t is indeed x0. Similarly, one can conclude that y of t is y0 plus this integral. Is this okay? Now, notice that there is an implicit relation over here.

There also is an implicit relation, right? And now what we can do is, you know, take the difference between x of t and y of t and look at the distance between them. Then, by the triangle inequality, the distance between them will be upper bounded by the distance between x0 and y0 and the norm of this difference, okay? So, that is this difference, right? And by a simple calculation, we can take this norm inside, which is referred to as the triangle inequality for integrals, right?

So, we take this norm inside, and we will end up with something like this. And now we can invoke the fact that your h is actually Lipschitz continuous to conclude that. You know the difference between x of t and y of t is upper bounded by x naught minus y naught plus L times the integral from t naught to t of x of s minus y of s, right? So, observe that on the left-hand side you have x of t minus y of t, and on the right-hand side you also have x of s minus y of s. So, you can see that we now have an implicit inequality, okay? So, here this will be your c.



This will be your K, and your U of S is this expression, and your V of S is actually 1.

$$x(t) \equiv x\left(t, t_0, x_0\right)$$

$$y(t) = x\left(t, t_0, y_0\right)$$

$$x(t) = x_0 + \int_{t_0}^{t} h(x(s)) ds$$

$$y(t) = y_0 + \int_{t_0}^{t} h(y(s))\, ds$$

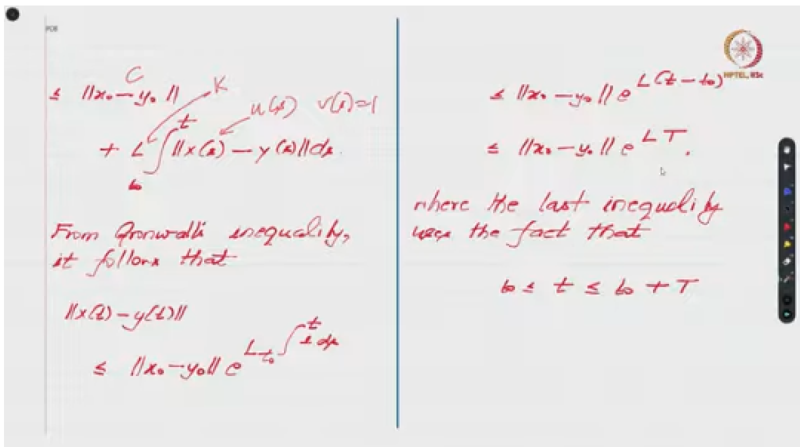
$$||x(t) - y(t)|| \leq \left|\left|x_0 - y_0\right|\right| + \int_{t_0}^{t} [n(X(s)) - h(Y(s))]\, ds$$

$$\leq \left|\left|x_0 - y_0\right|\right| + \int_{t_0}^{t} ||h(x(s)) - h(y(s))||\, ds$$

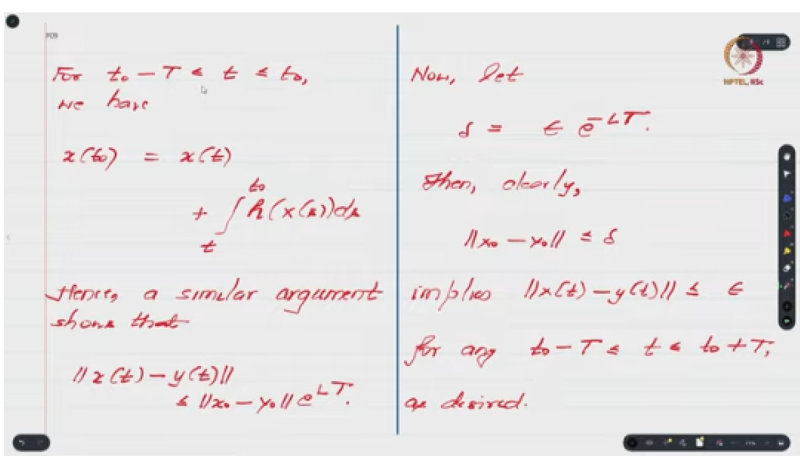
$$\leq \left|\left|x_0 - y_0\right|\right| + L\int_{t_0}^{t} ||x(s) - y(s)||\, ds$$

Is this okay? So, we can indeed invoke the Grönwall inequality now and conclude that x of t minus y of t—this distance—must be upper bounded by the product of C, which in this case is the distance between x0 and y0, times e raised to L, which is playing the role of k, right, times the integral from t0 to t of v of s ds. And since v of s is identically 1, I have put this 1 over here. And, you know, we end up with something like this.

So, this will be like t minus t naught. And because your little t is between t naught and t naught plus t, one can easily see that this expression is upper bounded by the distance between x naught and y naught times e raised to L t. Is this okay? And notice that the right-hand side now does not have any dependence on little t, right? So, one can conclude that this expression is upper bounded by this thing, which does not depend on t, and this is true for every t between t naught and t naught plus capital T.



And similarly, we can show that the inequality also holds for little t, which lies between t0 minus t and t0. We basically repeat the same story again. In this case, one can see that x of t0 must equal x of t plus this integral. So again, this follows from the fact that your x is a solution to this IVP. So, it satisfies x dot of t equals h of x of t. If you integrate it between t and t0, then one can see that x of t0 must be x of t plus this expression.



This just follows from the relation between integrals and derivatives. Right, and one can again see that y of t will also satisfy some relation like this, and hence one can use a similar argument as before to conclude that the distance between x of t and y of t is upper-bounded by this. And because the right-hand side does not depend on little t, one can take the supremum of this expression on the left-hand side for little t between t0 minus t and t0 plus t and conclude that that supremum is upper-bounded by this expression. So now we have to show that, for this fixed t, what is the value of delta so that the right-hand side is less than or equal to epsilon. Well, this is not very difficult. You choose delta, which is epsilon times e raised to minus l t. So, you choose such a small delta.

As soon as you choose such a small delta, if you multiply this expression with this, right? And we can multiply this with this because we are presuming that x0 minus y0 is less than delta, and delta has some value like this. So, because delta has some value like this and x0 minus y0 is less than delta. So, we can conclude that this product will then be less than epsilon, as desired. So, with this, we will conclude today's class.

Let us quickly summarize what we have done so far. Today, we looked at the continuous dependence of solution trajectories to IVP on the initial condition. So, if you fix a window of length capital T and perturb the initial condition by a small amount, then what we have managed to show is that the solution trajectory, at least over this finite interval, will not differ by a significant amount. That is roughly what we have seen. And in the subsequent classes, we will see how both these properties—the existence and uniqueness of solutions and the continuous dependence on the initial condition—can be utilized.

With this, let me stop today's lecture. Thank you. Hope to see you in the next class.