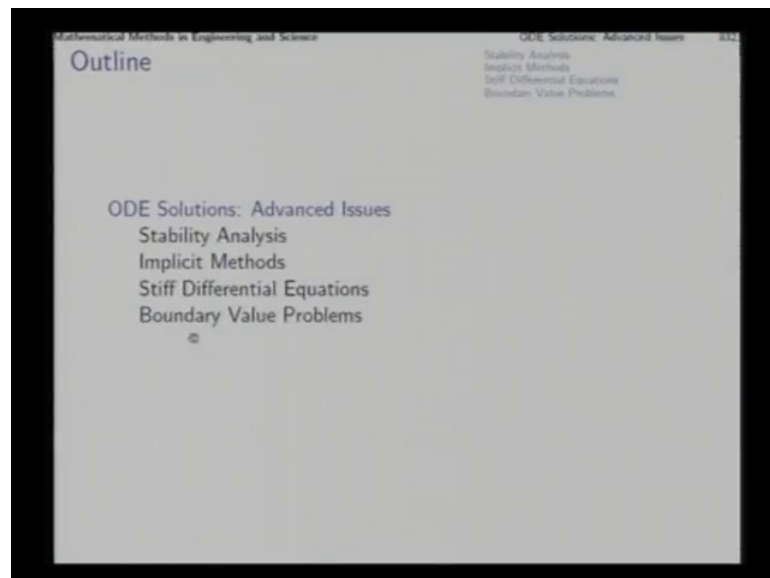


**Mathematical Methods in Engineering and Science**  
**Prof. Bhaskar Dasgupta**  
**Department of Mechanical Engineering**  
**Indian Institute of Technology, Kanpur**

**Module – V**  
**Selected Topics in Numerical Analysis**  
**Lecture - 04**  
**Boundary Value Problems, Question of Stability IVP Solution**

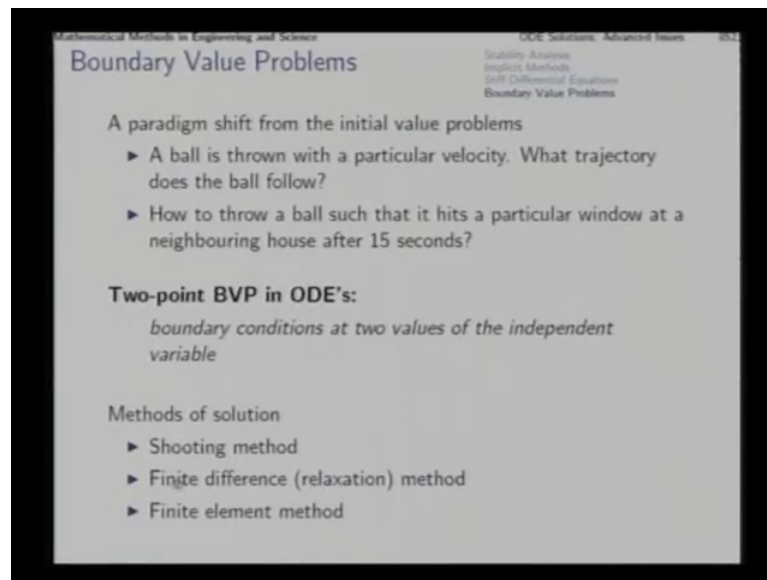
Good morning, in the last lecture we studied numerical method for solution of ordinary differential equation in particular the initial value problems. In the initial value problem at one value of the independent variable all the component of the dependent variable vector  $y$  is given and that signifies an initial value problem. We will have quite a few further issues to discuss in the numerical solution of initial value problems, but today's lecture let us start with a brief introduction to the boundary value problem after which we will return to the advanced issues of initial value problems.

(Refer Slide Time: 01:01)



So, let us start with boundary value problems from the initial value problem to boundary value problems in differential equation there is a complete paradigm shift, which can be visualized with the help of these 2 questions.

(Refer Slide Time: 01:20)



Suppose we ask this question there is a ball is thrown with a particular velocity; velocity means speed and direction combined a ball is thrown with a particular velocity at an instant of time what trajectory does the ball follow that is in future, what will be the trajectory of the ball that is at the starting time all the conditions are given the complete state factor is given position as well as velocity, and we want to know the position and velocity in instants of time in future, this is a typical initial value problems. Corresponding the and in contrast a typical boundary of the problem will look like this how to throw a ball such that it hits a particular window at a neighboring house after 15 seconds; that means, at time 0 we know the position we have to throw the ball from here and at time equal to 15 seconds, we want a particular position that is the conditions are given half at this time instant and half at a future time instant.

So, at the initial time we are giving only the position and not the velocity that is r how to throw a ball on the other hand at the end point we are giving the position there is position at this point and position at that point. So, this signifies a boundary value problem now this is a difficult problem physically as well as computationally physically also you can consider this issue, that if you want to target a particular window where the ball goes exactly after 15 second, for that you need to make a very detail preparation and possibly you have to try 1 or 2 times before perfecting the speed and direction. On the other hand for the first problem physically all that you need is a sophisticated measurement mechanism by which at every time instant you will go on taking the snapshot of the ball.

Now physically also this is a more practical problem comparably this is more interruptible computationally as well we find that solving boundary value problem will require more iterations and will require greater computational resource. A particular case we will discuss here which is called a 2 point boundary value problem, there are kind of problems that is illustrated by this question. So, in this kind of boundary value problems the conditions are given at 2 points 2 edges of the independent variable as here, the boundary conditions at 2 values of the independent variable will signify a 2 pint boundary value problem.

The number of conditions could be dispersed even at more than 2 points for example, we could have given some position at 1 time instant, some other position at time another time instant some velocity values at another time instant that is also possible. So, a specific kind of problems involves 2 point, typically at the 2 end points of the domain. So, that is that constitutes boundary of the domain and that is why it is a boundary value problems, but in general boundary value problem could have conditions given at various points, but we will consider here this class of problems for 2 point of boundary value problems. So, the methods of solution one is shooting method which is conceptually the easiest and then you have finite difference method finite element methods and so on. So, we will consider 2 method here one is shooting and the other is relaxation or finite difference method.

(Refer Slide Time: 05:18)

Mathematical Methods in Engineering and Science

Boundary Value Problems

Stability Analysis  
Integral Methods  
Stiff Differential Equations  
Boundary Value Problems

**Shooting method**  
*follows the strategy to adjust trials to hit a target.*

Consider the 2-point BVP

$$y' = f(x, y), \quad g_1(y(a)) = 0, \quad g_2(y(b)) = 0.$$

where  $g_1 \in R^{n_1}$ ,  $g_2 \in R^{n_2}$  and  $n_1 + n_2 = n$ .

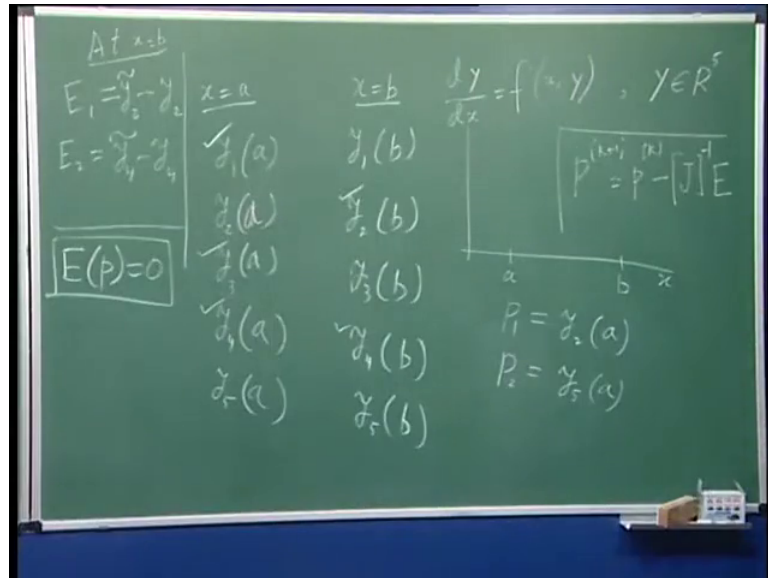
- ▶ Parametrize initial state:  $y(a) = h(p)$  with  $p \in R^{n_1}$ .
- ▶ Guess  $n_2$  values of  $p$  to define IVP

$$y' = f(x, y), \quad y(a) = h(p).$$

- ▶ Solve this IVP for  $[a, b]$  and evaluate  $y(b)$ .
- ▶ Define error vector  $E(p) = g_2(y(b))$ .

In the case of shooting method let us take it let us understand it with the help of an example.

(Refer Slide Time: 05:26)



Suppose at  $x$  equal to  $a$  we have got some value is given and at  $x$  equal to  $b$  we have some other value is given; now let us consider a problem in which the differential equation is this where  $y$  is 5 (Refer Time: 06:03) that means, let us take  $y_1, y_2, y_3, y_4$  and  $y_5$  be the state variable. Now at  $x$  equal to  $a$ ; these are the values if all this values all this 5 values at  $x$  equal to  $a$  are given, then it would be an initial value problem.

Now, in this particular case suppose at  $x$  equal to  $a$  all values are not given, but then at  $x$  equal to  $b$  some values are given total how many conditions we need how many values out of ten values how many values we need 2 consistently solve differential equation, we need 5 values out of these ten values right. So, for the 5 values suppose at  $x$  is equal to  $a$   $y_1$  is given, at  $x$  is equal to  $b$   $y_2$  is given; that means, the value of  $y_1$  is given at  $x$  equal to  $a$  the value of  $y_2$  is given at  $x$  is equal to  $b$  then value of  $y_3$  is given again at  $x$  equal to  $a$  and value of  $y_4$  is given at both end and value of  $y_5$  is given neither at this end nor at that end that is possible the a total of 5 conditions are given.

Now from these 5 conditions and from this differential equation we want to solve the differential equation solve the boundary value problem over this domain right. Now we make this observation that if we knew these 2 values also then it would be. So, nice we would solve it as an initial value problem right. So, suppose we make a guess of these 2

values. So, say we take  $P_1$  as this value which is not given, similarly another guess if we make which is this value which is also not given  $y_5$ , right that 2 values at  $a$  which are not given for that suppose we make some guess. It is like time to hit the target we try to make a guess of the direction and then shoot it, similarly we try to make a guess of  $y_2$  and  $y_5$  at  $x$  is equal to  $a$  and solve the initial value problems. Solution of an initial value problem is equivalent to pulling the trigger and letting the bullet go.

So, as the bullet goes typically first round it does not hit the target it goes somewhere else what is meant here by saying that the bullet does not hit the target, it means that the positions which are prescribed at the end point they are not correctly achieved; that means, with these guesses of  $y_2$  and  $y_5$  at (Refer Time: 09:39) equal to  $a$ , the point in the state space where we reach from that point we can work out with the figure out  $y_1$   $y_2$   $y_3$   $y_4$   $y_5$  at  $x$  equal to  $b$ , and we find that at these 2 points the value does not match the at these point whatever the value we do not know the true value we do not know. So, there is no way to match these 2 values what it should be we know and where we have reach that we get from the solution of the initial value problem.

So, between these 2 we can work out the discrepancies first error is error in this. So, with these guesses for the 2 unknown state variable values at (Refer Time: 10:25) equal to  $a$  with these guesses for them the initial value problem that we have solved gives the value of  $y_2$  here as  $\tilde{y}_2$  and the true value is this at  $b$  right. Similarly the other value that is known for that also we can find out the error right because of our specific choice of these 2 values we have got these 2 specific values of error right. Now the question boils down to finding out the suitable values here such that these 2 errors are 0 in a typical target hitting problem what we do? After we find the error that is compare to the bulls eye whatever shift we have made in our first attempt we make an estimate of the direction change that we should give to the gun so that next round it hits the target correctly.

Accordingly which is the direction of the gun and hit again so; that means, the initial condition whatever was assumed was slightly wrongs we try to make a correction is that right and correction should be made in such a manner that eventually these errors turn out to be 0. Now note that with small changes in  $P_1$   $P_2$  there will be small changes in  $e_1$   $e_2$  right. So, that way we can consider  $e_1$   $e_2$  as functions of  $P_1$   $P_2$ , and then try to solve this system of equations. In this particular case this system of equation is

actually a system of 2 equations  $e_1 = 0$  and  $e_2 = 0$  in 2 unknowns  $P_1$  and  $P_2$  right.

So, we can try our standard Newton Raphson method for solving these 2 equations in this 2 1 and for that we will need the Jacobian; Jacobian means derivative of these 2 errors with respect to these 2 variables; that means, a 2 by 2 Jacobian matrix we will get we will need and for that we need to give small change to  $P_1$  and small change to  $P_2$  and register whatever is the differential change in  $e_1$  and  $e_2$  and accordingly build the Jacobian matrix now; that means, that for every iteration after solving 1 initial value problem in order to find the Jacobian; we will need to solve 2 more initial value problems 1 with a little shift in  $P_1$  another with a little shift of  $P_2$  at  $P_2$  with at  $p_2$ .

So, then with so, many initial value problem solutions, we will get the error value at the end at these 2 values and its Jacobian its state of change based on that we can apply 1 newton raphson step 1 correction and this can go on iteratively eventually hopefully converging to  $e_1 = 0$  and  $e_2 = 0$ , which will mean that now we have converged to the correct values of  $P_1$ ,  $P_2$  that is correct values of these 2 state variables at a point  $x = a$  and the corresponding solution of the initial value problem then will be the correct solution of the boundary value problem as well, because initial values are matched and final values have 0 error.

So, this is the theme of shooting method that is you make trials to shoot the target this 2 and depending upon error at every iteration you try to change the initial values and accordingly per try to perfect step by step your attempts. Now here we took a simplified case in which 3 values were given it could be somewhat more general rather than giving 3 values at this point  $x = a$ , there could be 3 relationships between all the 5 values similarly here rather than giving 2 values it could be 2 relationships among 5 values then also conceptually the problem remains the same because in that case if there are 3 relationship between this 5 variables, then we can find out 3 parameters or rather 2 parameters in terms of which these 5 can be expressed those 2 parameters will be  $P_1$   $P_2$  because 5 variables 3 equations will mean 2 choices in our hand.

So, in terms of 2 parameters  $P_1$  and  $P_2$ , we could express all the 5 unknowns or 5 variables. So, in general the shooting method will look like this, it will follow the strategy to adjust trials to hit a target. Now in a 2.1 variable problem where this is a

differential equation and there are a few relationship given among the state variables at  $x$  equal to  $a$  and another few relationships at  $x$  equal to  $b$ . Now the number of this let us say it is  $n_1$  and a number of this dimension of  $g_2$  suppose that is  $n_2$  then  $n_1$  plus  $n_2$  must be equal to  $n$  because if the dimension of the problem is  $n$  if the state with space is  $n$  dimensional then you will need  $n$  condition less than  $n$  will not be enough more than  $n$  may be inconsistent right.

So, in that case what we do with the help of these  $n_1$  equations we parameterize  $y$  of  $a$  as a function of  $n_2$  parameters that is  $n$  minus  $n_1$  there is  $n_2$  parameters. So,  $n_2$  varies  $p_1, p_2, p_3, p_4$  up to  $p_{n_2}$ . So,  $n_2$  becomes the dimension of  $p$ . So, in terms of those  $n_2$  variables we parameterize  $y$  at this end, right. Now if we make choices for this  $n_2$  values then we get a complete state vector at  $x$  equal to  $a$  with such guesses with  $n_2$  values of  $p$  we make guesses and then we define the initial value problems same differential equation and  $y$  a completely given right with these  $n_2$  guesses of  $p$  made  $p_1, p_2, p_3$  up to  $p_{n_2}$  right now this initial value problem can be solved over the interval  $a$  to  $b$  and we can evaluate  $y$  of  $b$ .

And then we try to put this  $y$  of  $b$  in to this equation right and that will give us  $n_2$  equation which we try to solve. So, error vector is this and we try to solve  $E(p) = 0$ .

(Refer Slide Time: 17:33)

Mathematical Methods in Engineering and Science

Boundary Value Problems

Objective: To solve  $E(p) = 0$

From current vector  $p$ ,  $n_2$  perturbations as  $p + e_i \delta$ : Jacobian  $\frac{\partial E}{\partial p}$

*Each Newton's step: solution of  $n_2 + 1$  initial value problems!*

- ▶ Computational cost
- ▶ Convergence not guaranteed (initial guess important)

Merits of shooting method

- ▶ Very few parameters to start
- ▶ In many cases, it is found quite efficient.

So; that means, that for the solution of this system from the current vector  $p$   $n_2$  perturbations, having perturbation component in the  $e_1, e_2, e_3, e_4$  directions that is 4

direct shifts we can evaluate the Jacobean matrix; that means, that each newton Raphson step will require solution of  $n^2 + 1$  initial value problems 1 for getting  $e$  of  $p$  and another  $n^2$  here to evaluate the Jacobean after that we can execute 1 newton Raphson equation which is next  $p$  is equal to current  $p$  minus that  $\frac{1}{\text{Jacobian inverse}}$  in to current error right. So, this is the typical newton step for this problem.

So, shooting with this method has the advantage that it requires very few parameters to start with only  $n^2$  variables we have to select and then we can start shooting. On the other hand the disadvantage is that there is no guarantee of convergence and the problem is very sensitive to the working of the method is very sensitive to the initial guess; that means, initial guess becomes very important. If the initial guess is far away then the iterations may go anywhere and that is a typical drawback of Newton's method on the other hand it is a good method to start with that is if you have a problem for which you need the solution it is a good idea to first try a shooting method quickly if it works you have your solution in hand on the other hand if it does not work then a more sophisticated method can be tried which is finite difference method.

(Refer Slide Time: 19:35)

Mathematical Methods in Engineering and Science

Boundary Value Problems

Finite difference (relaxation) method  
adopts a global perspective.

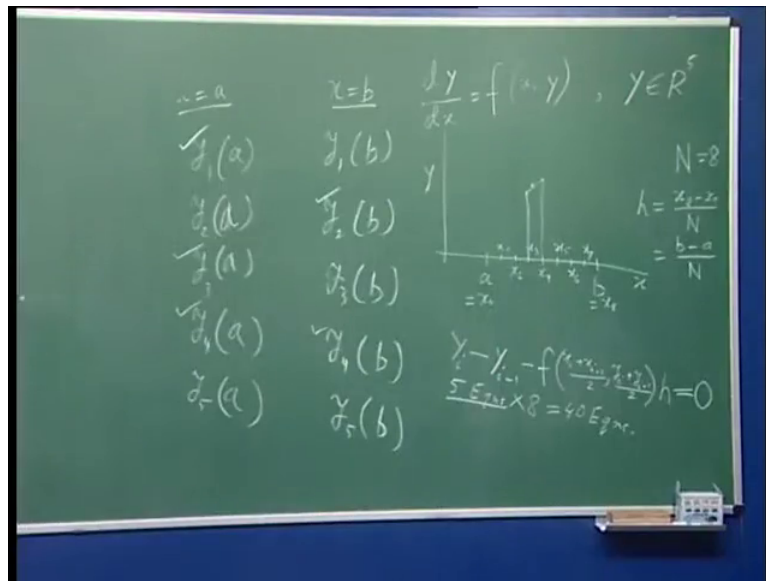
1. Discretize domain  $[a, b]$ : grid of points  
 $a = x_0 < x_1 < x_2 < \dots < x_{N-1} < x_N = b$ .  
Function values  $y(x_i)$ :  $n(N + 1)$  unknowns
2. Replace the ODE over intervals by *finite difference equations*.  
Considering mid-points, a typical (vector) FDE:  
$$y_i - y_{i-1} - hf \left( \frac{x_i + x_{i-1}}{2}, \frac{y_i + y_{i-1}}{2} \right) = 0, \text{ for } i = 1, 2, 3, \dots, N$$
  
 $nN$  (scalar) equations
3. Assemble additional  $n$  equations from boundary conditions.
4. Starting from a guess solution over the grid, solve this system.  
(Sparse Jacobian is an advantage.)

Finite difference method adopted global perspective.

So, for this same problem let us see how a finite difference method will work. In a finite difference method first we will discretize this domain.



(Refer Slide Time: 19:58)



Let us say for the time being that we are discretizing this domain  $a$  to  $b$  into 8 steps. So, 8 intervals this is  $x_0$  and this is  $x_1$ , this is  $x_2$ , this is  $x_3$ , this is  $x_4$ , this is  $x_5$ , this is  $x_6$ , this is  $x_7$ , this is  $x_8$ . So, there are 9 values of  $x_n$  is 8 and there are 9 values of  $x$  starting from  $x_0$  up to  $x_8$  and 8 intervals  $x_0$  to  $x_1$ ,  $x_1$  to  $x_2$  and so on 8 intervals now if we do that then what we can do that for every interval we can write this differential equation in discrete form right. So, then from the differential equation we get what is called finite difference equation.

So, for example, suppose at  $x_3$  if this is the value of  $y$ , now here I am showing  $y$  directly along this axis, but actually that is not a 1 dimensional entity that is a 5 dimensional entity. So, suppose the vector  $y$  vector at  $x_3$  is something and at  $x_4$  it is this. Now the differential equation written for this interval at what point we write the differential equation suppose we try to write it in the at the midpoint then it will look like difference in  $y$  that is  $y_i - y_{i-1}$  difference in  $y$  is equal to  $dy$  by  $dx$  which is this in to difference in  $x$ ,  $h$  here is  $x_8 - x_0$  by  $n$  right or  $b - a$  by  $n$  and this is same as  $x_i - x_{i-1}$ . Now this  $f$  evaluated at which point we are planning to evaluate at the midpoint. So, that will mean  $x_i + x_{i-1}$  by 2  $y_i + y_{i-1}$  by 2.

So, now consider this, this is an equation in  $x_i$   $x_{i-1}$   $y_i$   $y_{i-1}$  and that's it right you can also say you bring the entire thing on this side and how many equations are here;  $y$  has a dimension of 5 in this particular problem. So, actually you have got 5

equations here 5 equations here and like that you will get 5 equations over every interval; that means, in to 8 eight intervals are there you will get 40 equations forty equations you get in these variables the variables the unknowns involve the values of y values of x are known right. So, unknowns involve value values of y.

So, we have got forty equations and how many unknowns are there? There are 9 points at each point there are 5 unknowns; that means, unknowns 45 unknowns forty equations 5 equations less those 5 equations you get from here 1, 2, 3, 4, 5; 5 conditions are given. So, 45 equations and 4 5 unknowns you can solve. Now there are 2 issues if the differential equation is linear, then this will be a linear function of y right and; that means, that this system of equation that you get from here will be a linear system of equations on the other hand if the differential equation is non-linear then you will get non-linear system of equation and for solving them iterations will be necessary this is one issue.

Second issue is of course, in the first issue even if the differential equation is linear and the resulting algebraic equations that you get out of this are also linear, till iterations could be used and at times iterations are found useful and relaxation method typically uses such iterative methods like Gauss Seidel method or Jacobian method to solve the linear system. Now if the system is non-linear then the; if the differential equation system is non-linear, then this algebraic equation system is non-linear and the iterations will be necessary you cannot solve it in one step.

So, you have to use iterative method and for that purpose need Jacobean now one good issue here good point here is that differentially that finite difference equation that you see here will involve variables only from here and here not from here not from here; that means, every bunch of equations like this every bunch of 5 equations that you get from here will involve only a few variables which are close by in around this interval right and; that means, that the Jacobean that you will get out of it will be parse and here you will take advantage of the parsity of the matrix and you will use those kinds of method which work well with parse coefficient (Refer Time: 259) or solution of linear system now let us see outline in a generalized version generalized form first of all finite difference relaxation method adopted global perspective that is it tries to get the solution all over the domain advance and from this point to this point everywhere it tries to adjust the solution together not that it not the not like shooting method in which we try to build

the solution from this end towards that end here in the global perspective all the parts of the domain are given due importance and they are adjusted all together.

Now one more issue when we try to frame this equation using this derivative function here, then for developing this derivatives or the Jacobean there off we will need the values of  $x_i$   $x_{i-1}$   $y_i$   $y_{i-1}$  and so on right and therefore, to begin with even at the very beginning of solution we need to assume 1 solution from this end to that end that may be wrong in general that will be wrong, but then 1 value 1 set of values for  $y$  at every point of the grid is necessary to start the system start the solution process.

So, what we do we discretize the domain  $a$   $b$  with the grid of points  $x_0$  to  $x_n$   $n$  was taken as 8 in this example right then function values at  $x_i$  will mean  $n$  in to capital  $N$  plus 1 unknowns capital  $L$   $n$  plus 1 because capital  $N$  plus 1 points are there in to  $n$  which is the dimension of  $y$ . Now we will replace the ordinary differential equations over intervals by finite difference equations which will look like this right and this will give us  $n$  in to  $n$  small  $n$  in to capital  $N$  scalar equation; that means capital  $N$  vector equation at so many intervals and each vector equation will stand for small  $n$  scalar equations component wise.

Now we will assemble the additional  $n$  equations from the boundary conditions and then this total number of equations will be same as total number of unknowns and then starting from a guess solution over the grid we solve this system and at the end of the process we will get the complete solution. So, this method is for the relaxation method which works based on the finite difference equations.

There is also a finite element method which is much larger than this course will allow us to discuss in that we do not consider finite differences like this, but we try to split the entire domain to small elements finite elements and we try to propose the type of function with appropriate continuity requirement across boundaries and then the details of those functions we try to work out through iteration through the numerical methods. So, finite element method constitute a full course in itself and so, we do not go in to the details of that towards the end of this course when we discuss variational calculus there we will make a small introduction to the basic ideas of finite element methods and now we go back to the initial value problems.

And consider a few important and interesting issues in the solution methodologies of initial value problem. In the previous lecture, we discuss method for solving initial value problems of ordinary differential equation. Now we try to address a few important questions regarding the way the method handles the differential equation we try to make a stability analysis of the method adaptive Runge- Kutta for third order method or for third order fifth order method.

(Refer Slide Time: 30:58)

Mathematical Methods in Engineering and Science

Stability Analysis

Adaptive RK4 is an extremely successful method.  
*But, its scope has a limitation.*

Focus of explicit methods (such as RK) is accuracy and efficiency.  
*The issue of stability is handled indirectly.*

**Stability of explicit methods**  
 For the ODE system  $\mathbf{y}' = \mathbf{f}(x, \mathbf{y})$ , Euler's method gives  

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \mathbf{f}(x_n, \mathbf{y}_n)h + \mathcal{O}(h^2).$$

Taylor's series of the actual solution:  

$$\mathbf{y}(x_{n+1}) = \mathbf{y}(x_n) + \mathbf{f}(x_n, \mathbf{y}(x_n))h + \mathcal{O}(h^2)$$

Discrepancy or error:  

$$\begin{aligned} \Delta_{n+1} &= \mathbf{y}_{n+1} - \mathbf{y}(x_{n+1}) \\ &= [\mathbf{y}_n - \mathbf{y}(x_n)] + [\mathbf{f}(x_n, \mathbf{y}_n) - \mathbf{f}(x_n, \mathbf{y}(x_n))]h + \mathcal{O}(h^2) \\ &= \Delta_n + \left[ \frac{\partial \mathbf{f}}{\partial \mathbf{y}}(x_n, \bar{\mathbf{y}}_n) \Delta_n \right] h + \mathcal{O}(h^2) \approx (\mathbf{I} + h\mathbf{J}) \Delta_n \end{aligned}$$

Is an extremely successful method, but its scope is limited now the focus of explicit method such as Runge- Kutta method is accuracy and efficiency of the method for the purpose of efficiency we use adaption of (Refer Time: 31:25) right. Now there is another important issue in the solution process of ordinary differential equation and that is stability.

Stability of the method not the stability of the dynamic system for which there are differential equation is framed that is another issue which we will discuss much later in the course. Now stability of the numerical method means that if there is an error at 1 point then as the computation progresses does the error grow or does the error decrease if the error grows then we say that the process is not stable.

Now the issue of stability is handled in explicit method quite indirectly through the root of accuracy. Now if we try to make an analysis of the stability then let us consider this differential equation and rather than starting our analysis from r k 4 method that is fourth

order method let us start from Euler's method itself first order method the discussion will be simpler than Euler's method gives us this from  $y_n$ , we get  $y_{n+1}$ . So, this relationship which we saw in the previous lecture now Taylor's series of the actual solution around  $y$  equal to  $x$  equal to  $x_n$  is this  $y$  at  $x_n$  such that the true value of  $x_n$  true value of  $y$  at  $x_n$  plus  $f$  evaluated at this correct  $x$  and correct  $y$  in to  $h$  plus higher order term you see what are the differences here, here it is  $y$  and plus 1 the value of  $y$  at  $x_{n+1}$  given by Euler's method this is similarly the value at  $x$  equal to  $x_n$  given by Euler's method and so on here this is the true value this is the true value and the function is evaluated at the true point.

The true solution of the o d e system. Now if we try to see their discrepancy between these 2 points then this difference that is the value obtained minus the true value that is the error or discrepancy at  $x$  equal to  $x_{n+1}$  right and this is the difference  $y_{n+1}$  minus  $y$  at  $x_n$  that is the value obtained from the method minus the true value that is this plus similarly this minus this. So,  $f$  evaluated here minus  $f$  evaluated here in to  $h$  and then the rest are higher order terms now this is error at error step and what is this function at  $x$   $y_n$  and function; function at  $x$   $y$  of  $x_n$  this is this is the evaluated  $y$  at  $x_n$  this is the true  $y$  at  $x_n$  now from mean value theorem we know that between  $y_n$  and  $y$  of  $x_n$  in the  $y$  space there will be a point in the straight line join in the straight line segment joining, this erroneous  $y$  and true  $y$  in the straight line segment joining this 2 points there will be some point at which the derivative of  $f$  will be such that this derivative in to difference in  $y$  values will give this difference of  $f$  values that we get from the mean value theorem. So, between  $y_n$  and  $y$  of  $x_n$  in the straight line join of this 2 points suppose that point the existence of which is guaranteed by mean value theorem is  $y_n$  bar. So, at that point  $y_n$  we are not making any such discussion for  $x$  because  $x$  is same at these 2 points. So, at  $y_n$  bar then this turns out to be the Jacobean of  $f$  which has the 2  $y$  the derivative; derivative matrix in to  $\Delta_n$ . Now if we combine this 2 and call this matrix as  $J$  the Jacobean matrix, then we will get identity in to  $\Delta_{n+1}$  plus  $J \Delta_n h$ ; that means,  $i$  plus  $h J$  in to  $\Delta_n$ .

Plus of course, higher order term. So, you see we get a relationship between the discrepancy at  $n+1$  plus minus state and the discrepancy at  $n$  th step.

(Refer Slide Time: 35:48)

Mathematical Methods in Engineering and Science ODE Solutions: Advanced Issues 13.11

### Stability Analysis

Euler's step magnifies the error by a factor  $(\mathbf{I} + h\mathbf{J})$ .

Using  $\mathbf{J}$  loosely as the representative Jacobian,

$$\Delta_{n+1} \approx (\mathbf{I} + h\mathbf{J})^n \Delta_1.$$

For stability,  $\Delta_{n+1} \rightarrow 0$  as  $n \rightarrow \infty$ .

*Eigenvalues of  $(\mathbf{I} + h\mathbf{J})$  must fall within the unit circle  $|z| = 1$ . By shift theorem, eigenvalues of  $h\mathbf{J}$  must fall inside the unit circle with the centre at  $z_0 = -1$ .*

$$|1 + h\lambda| < 1 \Rightarrow h < \frac{-2\text{Re}(\lambda)}{|\lambda|^2}$$

**Note:** Same result for single ODE  $w' = \lambda w$ , with complex  $\lambda$ .  
For second order Runge-Kutta method,

$$\Delta_{n+1} = \left[ 1 + h\lambda + \frac{h^2\lambda^2}{2} \right] \Delta_n$$

Region of stability in the plane of  $z = h\lambda$ :  $\left| 1 + z + \frac{z^2}{2} \right| < 1$

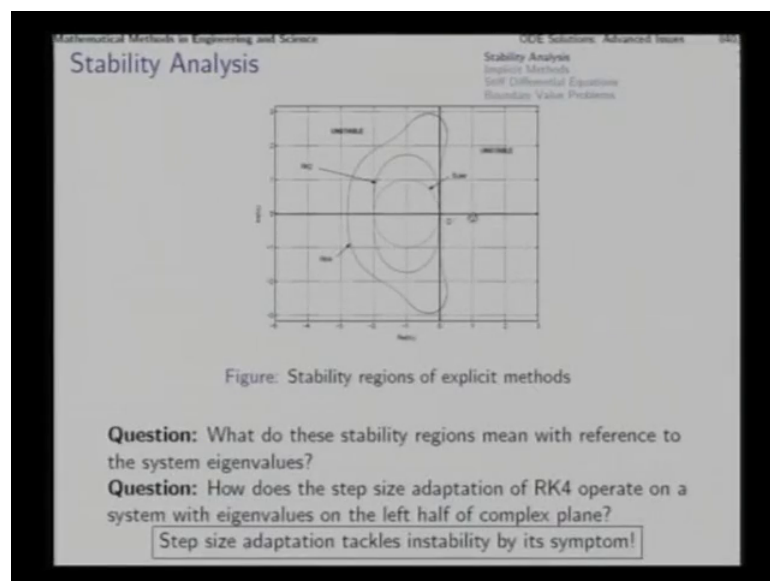
So, that shows that Euler's step magnifies the error by a factor of this now  $\Delta_n$  is a vector  $\Delta_n$  plus 1 is also a vector error in  $y$  at  $x_n$  error in  $y$  at  $x_{n+1}$  now this vector is mapped to this vector with this matrix  $\mathbf{I} + h\mathbf{J}$  and that will give you the magnification produce in the error as a result of this one step then the question is this magnification of size now we are talking about magnification of vectors. So, if the initial vector is in this direction  $\Delta_n$  and the resulting vector is in this direction then we will be talking about magnification in terms of sizes; now is this size magnification size the size of magnification greeter than 1 or less than 1 if it is greater than 1 that will mean that the size of the error will keep on increasing on the other hand if it is less than 1 then that will mean the size of the error will keep on decreasing.

So, using this  $\mathbf{J}$  loosely as the representative Jacobian why we say loosely as the representative Jacobian because as we move from 1 point to another  $x$  will change  $y$  will change  $f$  will change and therefore,  $\mathbf{J}$  will also change in general, but as a representative Jacobian using this same notation  $\mathbf{J}$  we can say that from the error at step 1 as we have taken  $n$  steps and come to the steps  $n + 1$  then. So, many times this matrix has been multiplied of course, in actual practice the matrix in between will keep on changing because  $\mathbf{J}$  is changing.

But using the rough Jacobian this will be the situation, Now for stability we need that this error at  $n + 1$  th step goes to 0 that is it goes on reducing. So, for reducing the

error step by step we will need this matrix to have all its Eigenvalues of magnitude less than one. So, we need Eigenvalues of this matrix to fall within the unit circle around the origin the Eigenvalues may be complex that is why we are talking about the Eigenvalues falling within the unit circle within the unit circle around the origin in the complex plane that is this this is a unit circle and the Eigenvalues Tymhs Wsjddof this matrix we need their magnitudes we need less than 1. So, by shift theorem we know that Eigenvalues of this matrix are related to the Eigenvalues of  $h, j$  because addition of the identity matrix is basically increasing the Eigenvalues by 1.

(Refer Slide Time: 38:46)



So, if the Eigenvalues must fall within the unit circle with origin at with the center at origin for  $i$  plus  $h j$  then the Eigenvalues of  $h j$  should fall within this circle which is shifted by 1 unit this is origin this is minus 1 0. So, within this circle all the Eigenvalues of  $h j$  should fall and that gives us this condition  $1$  plus  $h \lambda$  should be magnitude of that should be less than 1 which will give this limit on  $h$  that is.

Now the real  $\lambda$  real part of  $\lambda$  if it is negative then this will be positive and then it will make sense that is  $h$  is anyway positive. So, then it will make sense and then the method will be stable that is the error will not go on growing, but it will keep on reducing over steps that is if  $\lambda$  a  $\lambda$  has negative real part and if  $h$  is less than this this is what you will get from this figure also. If this unit circle if  $\lambda$  is  $\lambda$  has positive real part then anyway if the Eigenvalue cannot fall within this circle if it has

negative real part and  $h$  is so small that Eigenvalues of  $h j$  that is  $h \lambda$  comes within this then there is a hope that the Eigenvalues of  $h j$  falling in this and the method is stable

On the other hand if you take a take an  $h$  step size which is larger. So, that the Eigenvalues of  $h j$  falls somewhere here, here, here, then the method will be unstable because in that case Eigenvalues of  $i$  plus  $h j$  will have sizes which are larger than 1.

So, the same result this you will get if rather than the matrix vector equation system if you consider a single ODE which is  $w' = \lambda w$  allowing  $\lambda$  to be complex now the way we got this result from the first order method that is Euler's method for the second order Runge- Kutta method, we will get a similar relationship in which this matrix will appear this expression will appear for  $\lambda$  which earlier for the Euler's method was  $1 + h \lambda$  here will be a quadratic vector.

So, there the region of stability will be that this magnitude must be less than 1 which is this elongated figure. So, for RK 2 that is the Runge- Kutta second order method the Eigenvalues of  $h j$  that is  $h \lambda$  should fall within this for fourth order Runge- Kutta method similarly you can work out the expression of  $h n \lambda$  that also will be in the form of an inequality and when you plot it you will find that values of  $h \lambda$  for stability should lie within this and fall for higher order method there will be similar slight expansion still the problem remains that for large zones in the complex plane an instability problem remains that is if the Eigenvalues of  $h \lambda$  fall on this side then the method is unstable even if the values of  $h \lambda$  fall on this side, but depending upon the value of  $h$  if they fall outside this figure here here here then also the method is unstable.

That means for unstable differential equations for which the matrix  $j$  has Eigenvalues with positive real part, the method is anyway unstable and if the differential equation is stable that is Eigenvalues have negative real part, but if the value of  $h$  is large. So, that the Eigenvalues fall here the method is still unstable. So, then how does the step size adaptation of Runge- Kutta fourth order method operate on this kind of system as we discussed earlier the issue of stability in r k 4 or such explicit team is handled indirectly that is as the error goes grows up as the error grows up step by step then after a few steps it is identified that the grow the error growth has been too much and that will be identified through the comparison of 2 estimate as we discussed earlier through the



comparison of 2 estimates it will be found that at the seventh step say the growth error is too much; that means, over this interval we cannot use a single step we will sub divided into 2 steps and then any Eigenvalue which was found here that will be drawn inside like this and it will be found that now it is fine error is out of 2 estimates the error estimate will show that error is not too much error is fine.

But slowly because of the stepwise magnification of the error after a point again the error will grow too much and then due to the error growth the step size adaption will be performed and this goes on and finally, after quite a few iterations quite a few steps the adaptation will be taking place. So, frequently that you will have extremely small steps, but then the errors will be quite large this is the problem of explicit steps explicit method. So, step size adaption tackles instability by its symptom it does not handle the actual melody what will happen.

(Refer Slide Time: 44:40)

Mathematical Methods in Engineering and Science

Stability Analysis  
Implicit Methods  
Diff. Differential Equations  
Boundary Value Problems

### Implicit Methods

#### Backward Euler's method

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \mathbf{f}(x_{n+1}, \mathbf{y}_{n+1})h$$

Solve it? Is it worth solving?

$$\begin{aligned} \Delta_{n+1} &\approx \mathbf{y}_{n+1} - \mathbf{y}(x_{n+1}) \\ &= [\mathbf{y}_n - \mathbf{y}(x_n)] + h[\mathbf{f}(x_{n+1}, \mathbf{y}_{n+1}) - \mathbf{f}(x_{n+1}, \mathbf{y}(x_{n+1}))] \\ &= \Delta_n + h\mathbf{J}(x_{n+1}, \bar{\mathbf{y}}_{n+1})\Delta_{n+1} \end{aligned}$$

Notice the flip in the form of this equation.

$$\Delta_{n+1} \approx (\mathbf{I} - h\mathbf{J})^{-1}\Delta_n$$

Stability: eigenvalues of  $(\mathbf{I} - h\mathbf{J})$  outside the unit circle  $|z| = 1$

$$|h\lambda - 1| > 1 \Rightarrow h > \frac{2\text{Re}(\lambda)}{|\lambda|^2}$$

**Absolute stability** for a stable ODE, i.e. one with  $\text{Re}(\lambda) < 0$

If we do not use an explicit method, but use an implicit method what it means it means that rather than writing the differential equation in terms of the derivative  $\frac{dy}{dx}$  at  $x_n$   $y_n$  we could write it in terms of  $x_{n+1}$   $y_{n+1}$  at the other end of the small interval that is.

Rather than taking the slope from here if we say that what about taking the slope at the point where we would finally, reach at that point the question is that where we will finally, reach that is  $y_{n+1}$  we do not know now so; that means, that here we do not

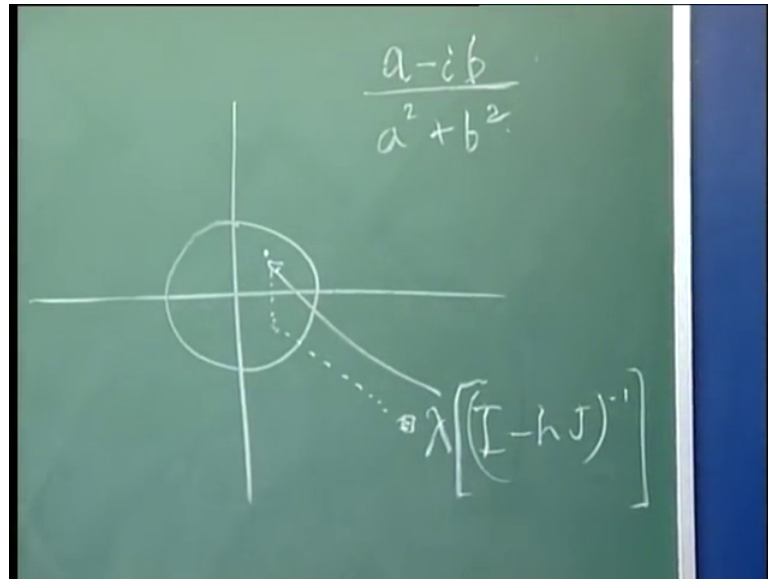
get an expression for the unknown  $y_{n+1}$ , but we get a system of equations in  $y_{n+1}$ .

So, question is we still have  $n$  equations in unknowns. So, we can solve it by Newton Raphson iteration or whatever, but first we ask a question is it worth solving it that is the problem of instability will that be address if we solve this and accordingly find  $y_{n+1}$ . So, let us first try to see whether it is worth solving this in order to evaluate  $y_{n+1}$ . So, for that we try to make a similar analysis what is the discrepancy at  $x_{n+1}$  if we use this kind of formula that will be  $y_{n+1}$  minus the true value that will be  $y_{n+1} - y(x_{n+1})$  in to the function value right the same equation which we used earlier except that now this  $f$  is evaluated at  $x_{n+1}, y_{n+1}$  rather than  $x_n, y_n$  similarly here and that will give us this is  $\delta_n$  and this will be  $h_j$ .

This  $j$  is slightly different from that earlier  $j$ , but still it is the Jacobian matrix at a representative point in to change in  $y$  difference in  $y$  now notice that earlier in the explicit scheme we got  $\delta_{n+1}$  is equal to  $\delta_n + h_j \delta_n$ , now here we are getting  $\delta_{n+1} = h_j \delta_n$ . So, in order to express  $\delta_{n+1}$  in terms of  $\delta_n$  what we do we take this on the other side and then we get  $i - h_j$  here right identity into  $\delta_{n+1} - h_j \delta_n = 0$ .

So,  $i - h_j$  inverted and pre multiplied here will give us this now this is the difference here now for stability of the method we need  $\delta_{n+1}$  to decrease our step that is compared to  $\delta_n$  we want  $\delta_{n+1}$  to be less; that means, the Eigenvalues of this matrix the inverse of  $i - h_j$ .

(Refer Slide Time: 47:49)

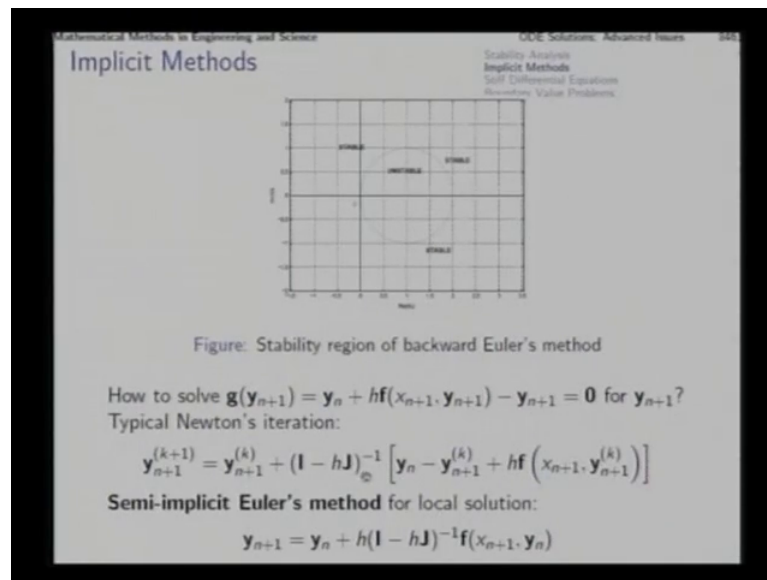


We want to be to fall within the unit circle we want the Eigenvalues of  $i$  minus  $h j$  inverse to be within this. So, for that suppose the Eigenvalue is here. So, for  $i$  minus  $h j$  inverse Eigenvalue of  $I$  minus  $h J$  inverse. So, for the Eigenvalue to fall here for  $I$  minus  $h J$  inverse where should be the Eigenvalue of  $i$  minus  $h j$ , that will be reflection right reflection first by this circle and then by the real axis here. So, if it is a plus  $i b$  then its inverse will be  $1$  by  $a$  plus  $i b$ ; that means,  $a$  minus  $i b$  divided by  $a$  square plus  $b$  square this conjugacy is equivalent to reflection by the real axis and this division is equivalent to reflection by this circle here. That means, the Eigenvalues of  $i$  minus  $h j$  should fall the unit circle Eigenvalues of  $a$   $i$  minus  $h j$  and; that means, that Eigenvalues of  $h j$  should be again accordingly shifted. S

o, if we analyze that there is Eigenvalues of  $I$  minus  $h j$  should fall outside the unit circle and that will give us the limitational  $h$  as greater than this. Now if  $\lambda$  has positive real part then  $h$  larger than this quantity will be required for stability and if  $\lambda$  has negative real part then this is always true because  $h$  is positive and then; that means, it will be always stable so; that means, we get absolute stability for a stable o d e system that is this.

$I$  minus  $i h j$   $i$  minus  $h j$  having Eigenvalues outside this unit circle will mean that  $h j$  must have Eigenvalues shifted i mean outside the shifted circle like this that is this.

(Refer Slide Time: 50:14)



So, if  $h$  falls  $h_j$  if Eigenvalues of  $h_j$  fall within this Eigen within this a circle then the method will be unstable for all the rest of the domain on the complex plane the method is absolutely stable. So, this is the advantage of the backward Euler's method, now we come back to this question should we solve this and if. So, how now since we know that the stability issue will be guaranteed then we know that it is what solving the system to find out  $y_{n+1}$ ; because we are using the slope function at the other unknown end of the interval rather than known end. So, we get this system of equations in the unknown  $y_{n+1}$ . So, we need to solve that.

So, for solving that what we need we need  $y_n$  minus  $y_{n+1}$  minus  $y_n$  that is this there is a differential equation we write it and this is a system of equations in  $y_{n+1}$ . So, we need to solve  $\mathbf{g}$  of  $y_{n+1}$  that is this expression equal to 0 for  $y_{n+1}$  newton step newton iteration will be this there is value at step  $k+1$  is equal to value at step  $k$  in to  $i$  minus  $h_j$  inverse because its Jacobian, Jacobean of this function will be minus  $i$  plus  $h_j$  right; that means,  $h_j$  minus one. So, in the newton step you have  $y_k$  minus Jacobean inverse. So, (Refer Time: 51:57) has been taken care of here that minus has become plus. So, minus  $I$  plus  $h_j$  has become  $I$  minus  $h_j$ . So,  $i$  minus  $h_j$  inverse in to the difference of the function values right.

So, you have got this the error value at  $k$  th iteration. So, that is this this whole thing evaluated at the  $k$  th iteration that is this now this is an iterative method iterative step.

Now in a semi implicit Euler's method what we say is that any way we will use the guess value for this newton Raphson step at the as the value of  $y_n$  the previous value that is known at  $x$  equal to  $x_n$ . So, if we use that as the guess value here and here then the first step in the newton Raphson iteration that you will get will have  $y_n$  here. So, that is same as this. So, this both are and  $y_n$  here.

So, the function value evaluated at  $x_n + 1$ , but  $y_n$  which is known. So, the function is evaluated here and this is the first iteration of the newton step Newtons method now if you utilize only the first iteration, then what you get is called semi implicit Euler's method otherwise the complete implicit Euler's method will require you to have this iteration in convergence. Now similar iterative schemes similar implicit schemes can be developed for second order Runge- Kutta method also. Now these implicit schemes will have the advantage that stability will be guaranteed, but the accuracy is another matter which has to be handled separately.

Now in the next lecture in the beginning we will consider what is called stiff differential equations and see how the application of implicit method helps in handling stiff differential equations and further in that in the next lecture, after discussing stiff differential equations we will raise more fundamental questions regarding the validity of solution of initial value problems, in a physical and practical sense and from that we will develop some deep mathematical results which will give us confidence to use the solution of initial value problems as we solve in the typical problems. So, these 2 issues we will discuss in the next lecture.

Thank you.