Elementary Numerical Analysis
Prof. Rekha P. Kulkarni
Department of Mathematics
Indian Institute of Technology, Bombay

Lecture No. # 31
Initial Value Problem

Today, we are starting a new topic, that is, approximate solution of ordinary differential equations. So, we are going to consider two classes of the problem, one is the initial value problem, so it will be differential equation of the first order and we consider boundary value problem, which will be differential equation of the second order. As the name suggests in the case of initial value problem, we are considering ordinary differential equation of first order. So, for the uniqueness, we specify one condition and that is the initial condition.

When we consider second order differential equation, we need to specify two conditions and these two conditions in general, they are specified at the 2 end points of the interval. These boundary conditions, they may be conditions on the function or its derivative or combination of the both.

So, first, we are going to look at the initial value problem, for the initial value problem there are going to be two techniques for defining approximations, one will be the truncated taylor series expansion, y is your unknown function you are trying to find. So, assuming sufficiently sufficient differentiability for this function, one can write down the taylors series expansion, truncate it, and that gives you approximate methods.

So, under these you get euler's method or you get taylors method and a variation of taylor's method are runge-kutta methods. So, these methods they are classical methods. For these methods, we will look at local discretization error, and in case of euler's method, we can also look at the total error. Then, there are methods which are of relatively recent origin, in that you have got y dash is equal to f of x y, so integrate both the sides.

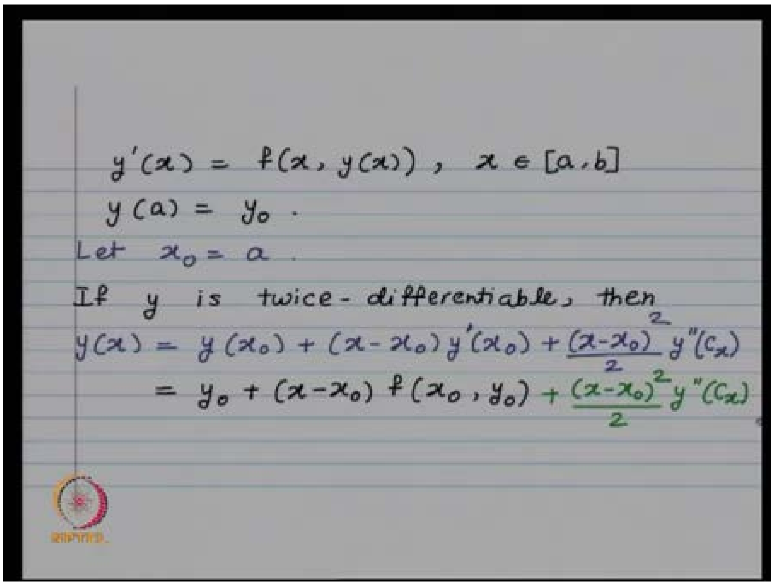
Now, instead of integrating exactly which may not be possible always, you can apply some numerical integration formula and that will give you approximate formulae. So, under these category, the main methods or the important methods they are adams bashforth method and there are predictor corrector formulae, which are one of them is adams moulton method.

So, we will compare various methods, for the comparison what we will do is, we will look at the order of local discretization error and how much work one needs to do. So, how many function evaluations, so these are going to be our criteria. So, our methods they fall into two categories, one is single step methods and other is multi step methods. So, this euler's method, runge-kutta method of order 2, order 4, they come under the category of single step methods.

The adams bashforth method and other methods which we obtain using numerical integration, they are going to be multi step method. Now, as I said that when we compare the methods, what is going to be important is, number of function evaluations, then the order of the error - local discretization error - but then, in case of multi step methods, the stability also comes into picture.

So, we will be considering the stability of multi step methods, when you look at boundary value problem, we are going to consider only finite difference methods. So, the derivatives they will be replaced by numerical differentiation formulae. So, this is a rough plan of what we are going to do for approximate solution of differential equation. So, let us start with initial value problem, I will first state a theorem which tells us about existence and uniqueness of the solution and then, we will define what is euler's method.

(Refer Slide Time: 05:58)



So, here is the initial value problem, y dash is equal to f of x y x, x is varying over a to b and y at a is equal to y0. So, this is the initial condition, what is a unknown is function y, the notation is y dash is equal to dy by dx. So, f is given to us and we want to find y which satisfies this differential equation. So, d is a rectangle and a y a, a was the left end point on which our function y is defined, y at a will be value of our unknown function y. So, we assume that this point is going to be interior point of d then on this rectangle, we will assume some conditions on the right hand side function.

First, let f x y be continuous on d and lipschitz continuous, so that is modulus of f of x y 1 minus f x y2 is less than or equal to k times modulus of y1 minus y2, for all points x y1, x y2 belonging to d. So, d is a rectangle, on which our function f is defined, it is lipschitz continuous. The point a y a is not going to be a boundary point, but it is going to be an interior point of this d, if this is the case, then y dash is equal to f x y, y a is equal y0, it is going to have a unique solution.

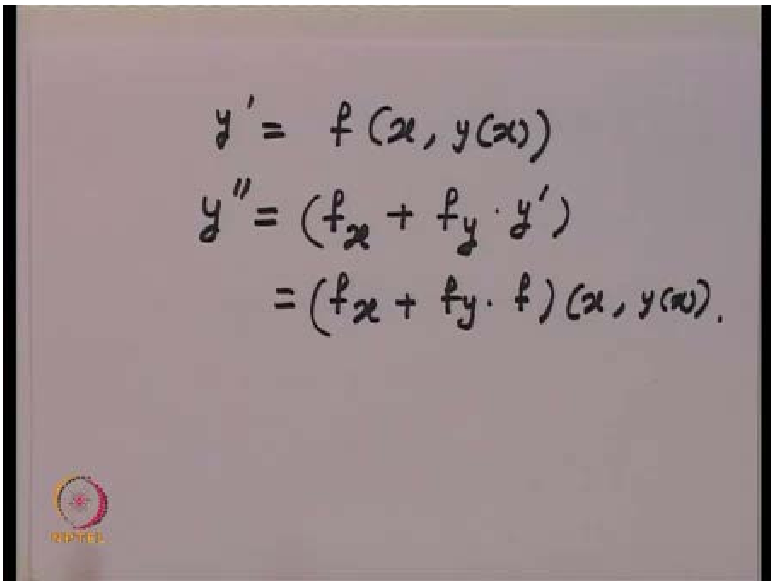
So, this is existence of unique solution for our initial value problem. So, here is an example, y dash is equal to y, y 0 is equal to 1, in this case you can solve exactly. So, y x is equal to e raise to x x belonging to, say, interval 0 to b, the right hand side function is f x y is equal to y, it is continuous on whole of r 2 and modulus of f of x y1minus x y 2 is equal to modulus of y1 minus y2, so k is equal to 1.

So, if your function f is such that it is partial derivative with respect to y exists and it is continuous on d, then this condition will be satisfied. So, this is a weaker condition than existence of partial derivative with respect to y, but that is one of the sufficient conditions, so now we consider approximate solution. So, the first thing we are going to look at is, write down the taylor series expansion for our unknown function y, so we have got… this is the initial value problem.

If y is twice differentiable function then y x is equal to y x0 plus x minus x0 y dash x0 plus x minus x0 square by 2 y double dash, at some point c x between x and x0, y x0 is given to be y 0 plus x minus x0 f of x0 y0 plus x minus x0 square by 2 y double dash c x.

Now, before I proceed, we are saying that function y should be twice differentiable, now I do not know what is the unknown function. So, even if we do not know the unknown function by looking at the right hand side function f, one can deduce differentiability properties of y. Like you have y dash is equal to f of x y, so our function y is once differentiable and it is derivative is the right hand side function f x y. Now, if this function f, if it has got continuous first order partial derivatives; that means, delta f by delta x and delta f by delta y, if they exists, then the second derivative will be given by this.

(Refer Slide Time: 11:22)



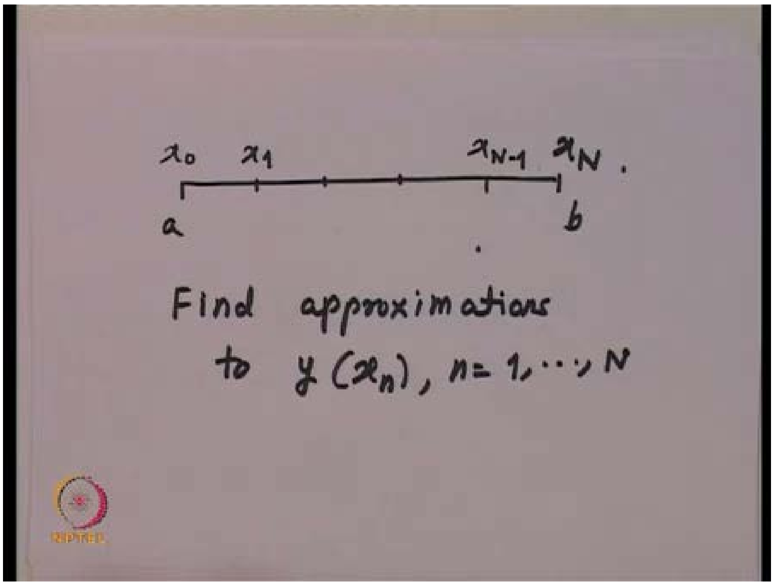
So, you have y dash is equal to f x y x, so y double dash is going to be equal to partial derivative of f with respect to x plus partial derivative of f with respect to y into y dash, because y is a function of x. So, this will be equal to fx plus fy and y dash is f at point x y, so that means, by looking at a function f, I may be able to tell that whether the function is twice differentiable or not. Now, we are writing taylor series expansion, you have got y x is equal to y at x0 plus x minus x0 into y dash, so for y dash we substitute f and then you have got a error term.

So, the first approximation will be truncated, just keep the first 2 terms, so that means, y x will be approximately equal to y at x0 that is given to us, that is y0, plus x minus x0 into y dash at x0, so that is going to be f of x0 y0, so I can calculate this. But as you know, as you go away from x0 the error is going to increase, the truncated taylor series gives good approximation in the neighborhood of your point x0.

So, whatever is this method, it will be valid for the neighborhood of x0, our aim is to find approximation to our function y over the interval a to b, so the interval a to b is fixed, x0 is our left end point a. So, in a neighborhood of a I can find approximation, but I have to go till a, till the right end point b. So, what one does is, divide this interval a b into n equal parts, so x0 is our a and then you go up to point x1. So, y at x1 will be obtained by this truncated formula, then from x1 you go to x2.

So, you remain in the neighborhood, first it was x0 x1 is near x0, so you have got truncated taylor series expansion, using the values at x1 you go to x2. So, like that, step wise you go up to the right end point, so that is the classical euler's method.

(Refer Slide Time: 14:55)

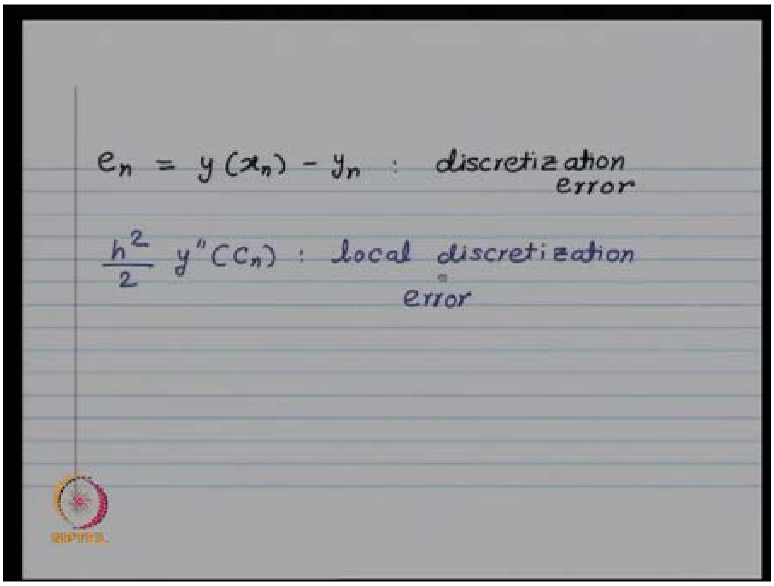


So, here you have got y x is equal to this; the error is, it contains x minus x0 square, so as you go away from x0, this error is going to increase. So, our interval a b, we divide into capital N equal parts, these are equidistant points, so x n plus n1 minus xn is going to be equal to h which is b minus a by n, our small n varies from 0, 1 up to n minus1.

The notation is going to be y at xn is the exact value, y sub n is going to be approximation and at the starting point there is no error, so y0 is equal to y of x0. So, here is the euler's method, y x is approximately equal to y x0 plus x minus x0 y dash x0, so that gives you y x1 to be approximately equal to y0 plus h times substitute for y dash, so it is f of x0 y0.

So, this y x1, it is approximation, so I call it y1, so y1 is equal to y0 plus h times f of x0 y0. And then, you define similarly, y n plus 1 is equal to y n plus h times f of x n y n n is equal to 0, 1, 2 up to capital N minus 1, so we have got our interval a b. In this, we consider some equidistant points, so x 0, x 1, x n minus 1, x n and we find approximations to value of y at x sub n, n is equal to 1 to up to N. And once you have got approximation to y at some finite number of points then, you can do interpolation, you can do spline interpolation to obtain a function which is continuous or which is differentiable.

(Refer Slide Time: 17:40)



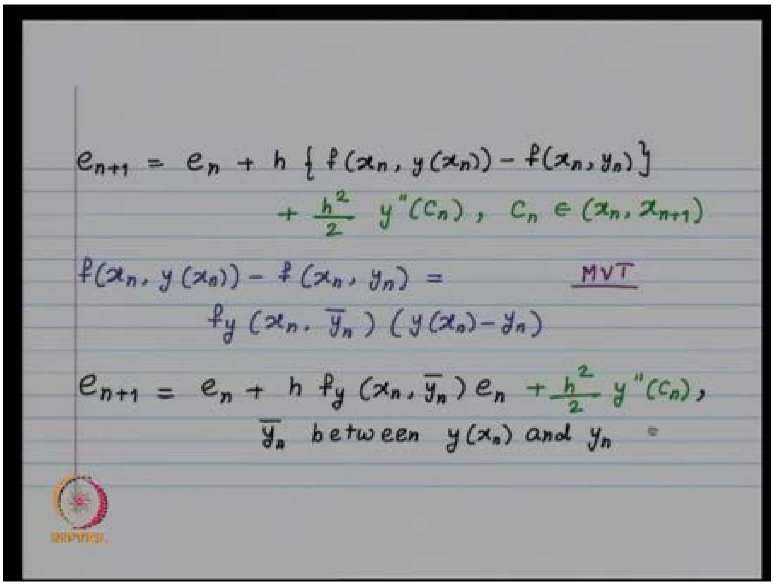
So, now, if you consider y x n plus 1, remember our notation is that y at x n plus 1 is going to be exact value; its taylor series expansion is, y at x n plus x n plus 1 minus x n that is h, so it is h y dash x n plus error h square by 2 y double dash c n. So, this is equal to y at x n plus h times y dash is f, so f x n y x n plus h square by 2 y double dash c n.

Euler's method is y n plus 1 is equal to y n plus h times f of x n y n. So, when I consider the error y x n minus y n, it will have e n plus 1 is equal to e n plus h times f of x n y x n minus f of x n y n plus h square by 2 y double dash c n. This term is known as local discretization error and your y n is also an approximation to y at x n. So, this is going to be, when you have got e n, the error at n stage; the error at n plus first stage will have this term plus this term.

So, when you want to look at the total error, you have to take this into consideration, it is possible for euler's method, but otherwise, we will be just considering the local discretization error. So, here the local discretization error, one says that is of the order of h square, this assuming y to be second time - 2 times - differentiable, this can be dominated by a constant.

So, e n which is y x n minus y n that is known as discretization error, and h square by 2 y double dash c n that is known as the local discretization error. So, now, this is our… we are now going to look at the error in the total error in the euler's method.

(Refer Slide Time: 20:24)



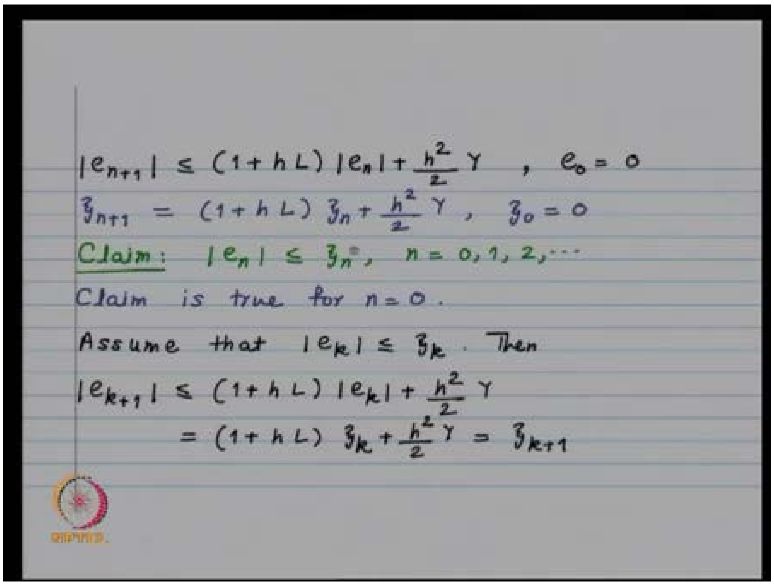
So, look at this expression e n plus 1 is equal to e n plus h times this plus this is the local discretization error. We are going to apply mean value theorem to this term, so you have f x n y x n minus f of x n y n, the first term is the same, so this will be partial derivative of f with respect to y at x n y n bar. So, y n bar will be something in between the 2 into y x n minus y n, so that is the mean value theorem.

So, you have e n plus 1 is equal to e n plus h times partial derivative of f with respect to y plus h square by 2 y double dash c n. So, now, we are going to impose certain conditions, so we are going to assume that the partial derivative of f with respect to y, let us assume it to be continuous, and then we will say that the maximum value of the partial derivative, its modulus is less than or equal to sum l. And second derivative of y, that is y double dash, we will assume that modulus of y double dash of x is going to be less than or equal to capital Y.

So, you think these we are going to try to find the total error, we had discretization error, discretization error is y x n minus y n, we have local discretization error which is… If you do not take into consideration, see you go from x0 x 1, so in the first step your y at x0 is equal to y 0. So, there will be only local discretization error, but when you go from x 1 to x 2, there is already error in y at x 1. So, you have got that error and you have error because you are truncating your taylor series expansion. So, these 2 together, so local discretization error is of the order of h square and we will show that the total error y x n minus y n that is going to be of the order of h.

It is like in the numerical integration, when we looked at composite numerical integration rule, then you had certain power h raise to k for each interval, when you add it up then you lost 1 power of h, similar thing happens here that local discretization error is h square, but then, when you apply it to the whole interval a b your euler's formula, the error is going to be less than or equal to constant times h.

(Refer Slide Time: 23:47)



So, these are our assumptions, partial derivative of f with respect to y is less than or equal to L for all x y belonging to d, d is the rectangle which contains the point a y a in its interior. Modulus of y double dash x is less than or equal to capital Y for x belonging to a b.

Take mod of both the sides, so you will have modulus of e n plus 1 to be less than or equal to mod e n then, modulus of f y x n y n, that is, L and here you have, in fact, one more term, you have got here h f y x n y n bar and this term here, so this I have forgotten, so there is going to be mod e n here. So, this is h, this gives you L and there is mod e n plus h square by 2 into y. Next, our claim is that modulus of e n is going to be less than or equal to h into y divided by 2 L into e raise to x n minus x0 into L minus 1, this is our uniform partition of interval a b.

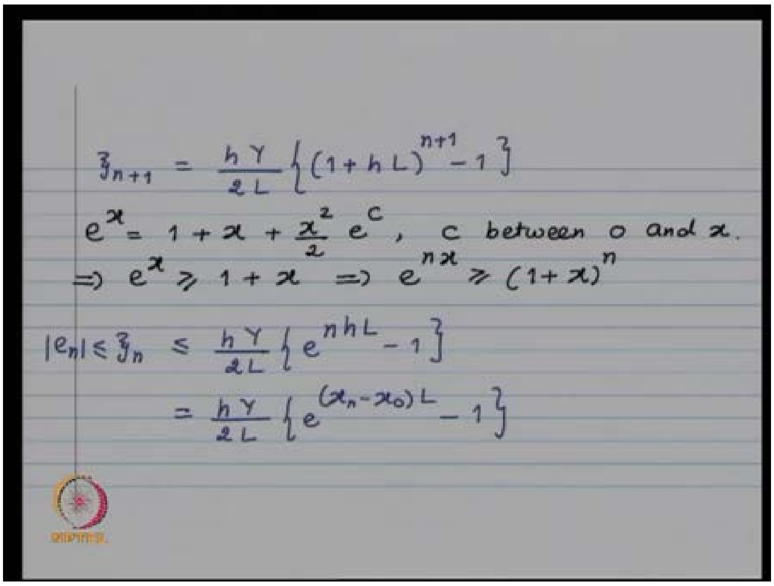
I am looking at this point x n, and error at this point is going to be y x n minus y n. I want to say or I want to show that modulus of e n is less than or equal to this term, look at the term in the bracket, it includes x n minus x0 but there is no h appearing here. So, if I fix x n to be fixed, then I can change my length of the sub interval, I can go from here to here with step size h, then I will need only n steps. If I decide to take step size to be h by 2 in order to reach here, I will need two times 2 n steps.

So, here modulus of e n is less than or equal to constant times h, so that is what I was saying that local discretization error is of the order of h square, but the total error is going to be of the order of h, so this is the result we are going to prove. So, we have got this estimate e 0 is equal to 0, I define a new sequence psi n plus 1 to be 1 plus h l psi n plus h square by 2 y. So, the difference is instead of less than or equal to I am defining it to be equal to, then the claim is mod e n is less than or equal to psi n for n is equal to 0, 1, 2 and so on.

The proof is by induction, it is true for n is equal to 0, because e 0 is equal to 0 psi 0 is equal to 0, and assume the result to be true for k, consider modulus of e k plus 1. This will be less than or equal to 1 plus h l mod e k plus h square by 2 y by induction hypothesis mod e k is less than or equal to psi k, so you will get 1 plus h l psi k plus h square by 2 y and that is our psi k plus 1.

So, we have got, instead of inequality, now I have to work with this sequence. So, now, we have got psi n plus 1 is equal to some formula. So, now, you replace psi n by psi n minus 1, psi n minus 1 by psi n minus 2 and so on. And then, you will get psi n plus 1 in terms of finally, you will reach psi 0 and psi 0 is going to be equal to 0.

(Refer Time Slide: 28:22)

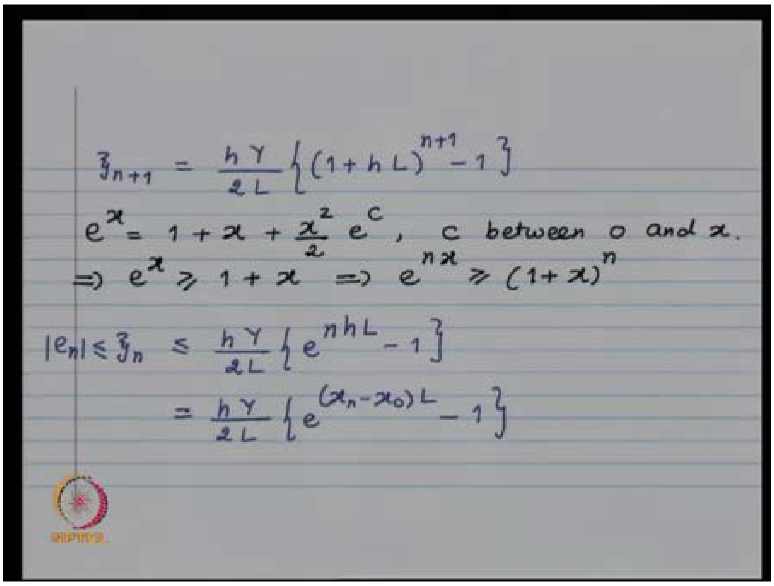


So, let us look what it looks like, so psi n plus 1 is equal to this. Now, use the same formula with n replaced by n minus 1. So, we will have 1 plus h L square psi n minus 1 plus h square by 2 into 1 plus 1 plus h L y. I am substituting for psi n, so it will have 1 plus h L psi n minus 1 and then h square by 2 y, so that is why I have got this h square by 2 y and then 1 plus h l y. When I continue, I go up to 1 plus h L raise to n plus 1 psi 0, see here; when the power is 1, here you have n; when the power is 2, here you have n minus 1. So, some of these two, it has to be equal to n plus 1, so that is why when you have 0 here, here it is n plus 1 plus h square by 2 1 plus 1 plus h L plus 1 plus h L raise to n y, there is always one term less.

The psi 0 being 0, this term is equal to 0, h square by 2, the term in the bracket is equal to 1 plus h L raise to n plus 1 minus 1 divided by 1 plus h L minus 1. You can multiply this and this, then verify that it is the same which will be equal to, now in the denominator you will have h L, so one h will get cancelled. So, you will have h then this y upon 2 L, 1 plus h L raise to n plus 1 b minus 1. Then, this 1 plus h L raise to n plus 1, so I want some compact formula for that, so that is why I look at function e raise to x. This function e raise to x is always bigger than or equal to 1 plus x, because I write down e raise to x as 1 plus x plus x square by 2 and then e raise to c.

So, e raise to x will be bigger than or equal to 1 plus x, so e arise to n x will be bigger than or equal to 1 plus x raise to n. So, using this formula, we obtain an expression for psi n plus 1.

(Refer Time Slide: 31:06)



So, we have got e raise to x is bigger than or equal to 1 plus x and hence, e raise to n x is bigger than or equal to 1 plus x raise to n. Now, mod e n is less than or equal to psi n, this will be less than or equal to h y upon 2 L e raise to n h L minus 1, so n h is nothing but x n minus x0, so you get h y upon 2 L e raise to x n minus x0 into L minus 1.

So, we have now proved that the error in the euler's method is less than or equal to constant times h. So, if I want error to be less than some, say, 10 raise to minus 6, I have to choose my h small enough. So, when you choose h small enough, in order to reach the right end point b, you will need many more steps because our aim is to find approximation to our function y on the interval a to b. So, we are dividing interval a to b into sub intervals of length h. So, if you choose h small then, the number of intervals they will be big, and then like you have to take into consideration the round off errors.

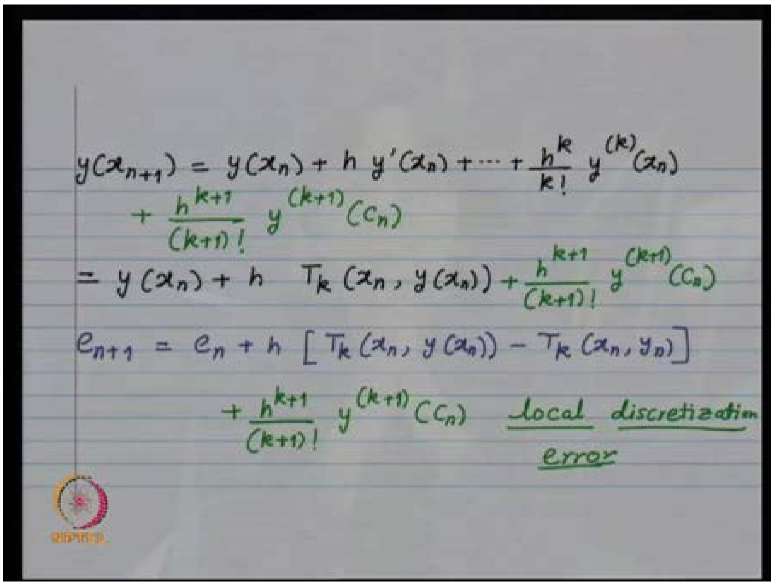
It should not be that your interval a b, you are sub dividing into really small intervals and then by the time you reach b, the round off errors they get added up. So, it will be better to have a method which has got higher discretization error or higher order discretization error.

In euler's method what we did was, we truncated our taylor series expansion by retaining only first two terms. So, why I will retain only first two terms? You retain some more terms, so suppose, you retain, say, when you retain two terms, then your discretization error - local discretization error - was of the order of h square. From now onwards, we will be considering only local discretization error, that means we will ignore. I just want to know from x n to x n plus 1 going what is the error which has occurred, there is error because of the y x n is not equal to y n, so that part we will ignore.

So, you choose more terms and then you will get higher discretization error, but there is always a trade off. When you look at the first two terms, what you had was y x0 into x minus x0 y dash x0; our y dash was equal to f. If you consider the second order y double dash x0, y double dash will be in terms of partial derivatives of the function f with respect to x and with respect to y. So, in the euler's method what you need is only function, now you will need partial derivatives, then if you want to also retain the term y triple dash x0, then you will have a higher order partial derivatives, so it depends. If your function f is something simple for which you can calculate the partial derivatives easily, then the taylors method will be a good choice and if this is not the case, then one has to think off some other way.

Also what one wants is, one wants to have some automatic program, like I do not want, given a problem ok, now let me look at the function, then calculate its partial derivatives by hand, we do not want to do that. So, if the function is given to me, if I have a formula only in terms of the function that will be good and that is what is achieved by the runge-kutta method.

(Refer Time Slide: 35:57)



So, let me first tell you what is the taylors method, we have got y x is equal to y x0 plus x minus x0 y dash x0 plus x minus x0 raise to k upon k factorial y k x0 and this is the error, y dash is equal to our function f. The second derivative will be partial derivative of f with respect to x plus partial derivative f with respect to y into y dash, y dash is nothing but f, so you have y double dash to be this term. Let us calculate one more derivative, so y triple dash your function f x is function of 2 variables, we have f is a function of x and y. Then, we had y double dash to be f x which will be a function of x and y plus f y into f both functions of x and y.

So, when I consider y triple dash; that means, d by d x of y double dash, so you have to do implicit differentiation. So, it is going to be f xx x y, it is the second order partial derivative with respect to x plus f xy - f x is function of x and y y is a function of x, we are taking the derivative with respect to x, so you have to have f x y - into y dash, but y dash is nothing but f. So, this takes care of this, then you will have f y x into f plus f yy and then f square, because 1 f will come from y dash and 1f is here, plus f y into f x, because you are using product rule. And then, you will have f y square into f and then you can assume that the second order mixed partial derivative, they are equal, so f x y is equal to f y x, so then these two terms you can combine.

So, now, taylor’s algorithm of order k is defined f x y plus h by 2 f dash x y plus x raise to k minus 1 by k factorial f k x y. Now, let me explain the notation, when you say f dash that is derivative with respect to x. So, it is going to be equal to f x plus f y into y dash, so that is f, so this f dash f k these are the total derivatives. And then y n plus 1 is equal to y n plus h times t k x n y n, if you consider k is equal to 1; that means, you have got only f x y, so that gives us euler's method. If you retain more terms, then you are going to have a better discretization order, but you will need to evaluate all this derivatives.

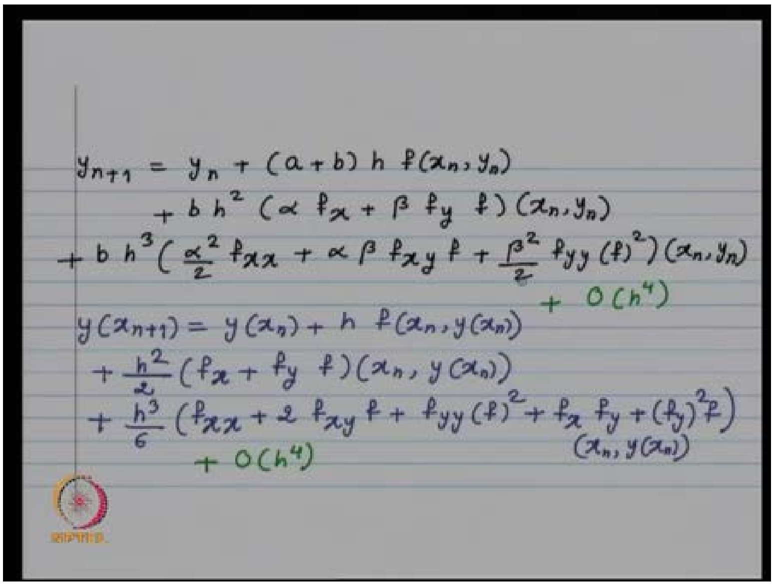
So, this is the taylor's series expansion for unknown function y, where we are retaining k terms. This is the error term y dash is f, y double dash is f double dash, y k is f k minus 1 and hence this is nothing, but y at x n plus 1 is equal to y x n plus h times t k x n y n. Let us go back to what was our t k, so t k x y has f h by 2 f dash h raise to k minus 1 f k and f is y dash, f dash it will have f x plus f y into f and so on.

So, here, this is y x n plus h times t k x n y x n plus this is the error, then e n plus 1 which is y x n plus 1 minus y n. This will be e n plus 1 is equal to e n plus h t k x n y x n minus t k x n y n, what is y n plus 1? y n plus is y n plus h t k x n y n. So, I take the subtraction, y x n plus 1 minus y n plus 1 is e n plus 1, y x n minus y n is e n plus h t k x n y x n minus t k x n y n and then this, so this is the local discretization error.

So, we are going to concentrate only on local discretization error now onwards. So, if you are retaining more terms, you get higher power of h. Now, so you choose your k and you can have various methods, here in the taylor's method one needs to calculate partial derivatives of f. So, we will like to avoid that and that is why now we are going to consider what is known as runge-kutta method of order 2. We are going derive runge-kutta method of order 2 and I am going to state the formula for runge-kutta method of order 4.

So, the idea for runge-kutta method is, that in the taylor series expansion try to match as many powers of h as possible. So, we will start with some formula, like I want to look at formula of the type y n plus 1 is equal to y n plus h times, now I want to evaluate function at two points. So, once I will evaluate the function at x n y n and once I will evaluate at some other point, so f of x n plus alpha h and y n plus something and then we will try to match the powers of h.

(Refer Time Slide: 44:03)



So, let us be more specific, so here you have, this is the taylor series expansion for unknown function y. We want a formula of this type y n plus 1 is equal to y n plus a k 1 plus b k 2, k 1 is h times f x n y n k 2 is h times f of x n plus alpha h, y n plus beta k 1. So, we have got a b, alpha and beta, these are the constants at our disposal. We want to use we want to determine these constants such that 2 agrees with 1, for as many powers of h as possible. So, here look at k 2, it is f of x n plus alpha h y n plus beta k 1, where k 1 is here.

So, what we are going to do is, how are we going to match the powers here for this function of 2 variables. We are going to write down the taylor's series expansion, this is taylor's series expansion for function of one variable, we will write similar taylor's series expansion for function of two variables and try to match the powers. So, we have k 1 is our h times f of x n y n, k 2 is h times f of x n plus alpha h y n plus beta k 1. Let me look at this f of x n plus alpha h y n plus beta k 1, this will be f of x n y n plus alpha h increment in x n multiplied by a partial derivative with respective x evaluated at x n y n plus increment in the second variable beta k 1 partial derivative of f with respective to y evaluated at the same point plus alpha h square by 2 f x x, alpha h into beta k 1f x y and beta k 1 square by 2 and second order partial derivative with respect to y.

Plus there will be higher order terms, here you have got power h square, here you have alpha and beta are going to be constants. You have h and in k 1, you have got h, so this also has square, this also has h square, so it is going to be plus order of h cube. Our y n plus 1 was y n plus a k 1 plus b k 2, so what it will be? It will be a times k 1 plus b times k 2, k 2 has this h here, so it will be y n plus h times a f plus b h times, again see, you are substituting for this here. So, you are going to have y n plus 1 is equal to y n plus h times a plus b into f that is going to be the term which contains h.

So, here what was the term containing h, it was y dash x n y dash is f. So, here the first is y at x n and then, we are going to not distinguish between y x n and y n. So, you have got the same here. Here you have got h into f here you have got a plus b into h into f and then higher order. So, when we collect the terms you are going to have y n plus is equal to y n plus a plus b into h f of x n y n plus this etcetera.

So, in order to match the power of h, you need a plus b to be equal to 1, here you will need b alpha to be equal to 1 by 2 b beta also to be equal to 1 by 2 and then so on. So, we have got four constants to be determined, we will try to determine these constants and we will show that, in runge kutta method of order 2, the local discritization error is of the order of h cube. In euler's method, it was h square, so you are improving from h square to h cube and the price you are paying is instead of evaluating the function once, you need to evaluate function twice.

But you are gaining one power of h in the error and that becomes more significant. So, runge kutta method of order 2 is going to be better choice than the euler's method, it needs only function evaluation. This matching of the powers in more detail, we will see in our next lecture and then I will also state what is runge kutta method of order 4. We will look at some example and then, we will go to numerical integration, so thank you.