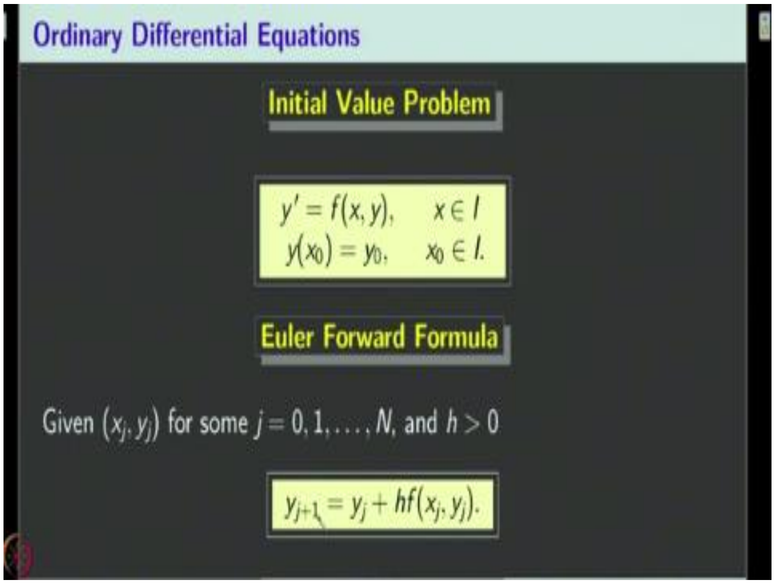
Numerical Analysis
Prof. S. Baskar
Department of Mathematics
Indian Institute of Technology, Bombay

Lecture - 56
Numerical ODEs: Euler Methods (Error Analysis)

Hi, we are learning numerical methods for initial value problem of a first order ordinary differential equation. In this, we have learned the Euler method for approximating a solution of a first order initial value problem in our last class. We will continue our discussion on Euler method in this class and study the error involved in the approximate solution computed by the method.

(Refer Slide Time: 00:50)

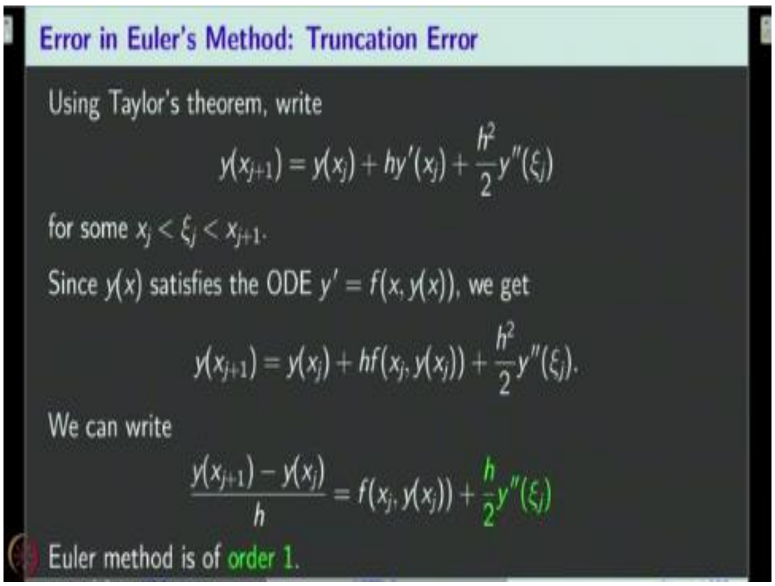


Recall, that we are interested in solving the initial value problem $y' = f(x, y)$, where we are given an initial condition $y(x_0) = y_0$, for some x_0 in the interval on which the problem is posed. In the last class, we have studied Euler method for approximating a solution of this initial value problem at some grid points. We had two formulas, one is the Euler forward formula which is given here and another one is the Euler backward formula.

In this class, we will study the error involved in the approximate solution obtained using the Euler forward formula. The error analysis for the Euler backward formula can be carried over in a similar way. Let us recall the Euler forward formula first. Given a point (x_j, y_j) and a parameter $h > 0$, the Euler forward formula can be used to obtain the value y_{j+1}, which is an approximation to the solution y of this problem at the grid point x_{j+1}.

This can be done for each j = 0, 1, 2 so on up to some N number of grid points. Observe, that we are given x_0 and y_0 in the problem itself, which is the initial condition. Once we fix h and generate the grid points $x_0, x_1, \cdots, x_n$, say in the interval I, then we can obtain the approximation to the solution y of the given initial value problem at this grid points using this formula.

(Refer Slide Time: 03:19)

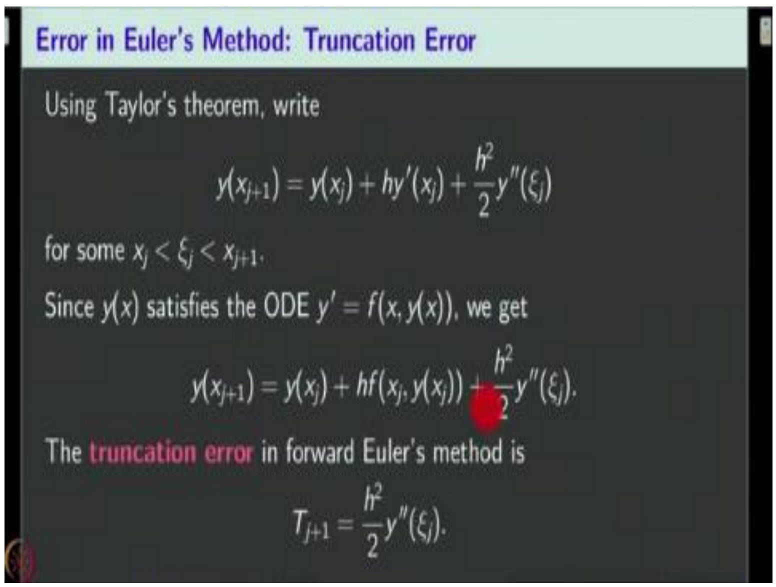


To derive the error involved in the approximate solution y_{j+1}, when compared to the exact solution $y(x_{j+1})$, we first use the Taylor's theorem about the grid point x_j and write the solution at x_{j+1} as $y(x_{j+1}) = y(x_j) + hy'(x_j)$. This is the Taylor's polynomial of degree 1, plus you have the remainder term $\frac{h^2}{2}y''(\xi_j)$.

Since y satisfies the given first order ordinary differential equation $y' = f(x, y)$, we can substitute y' in this representation by $f(x, y)$ and that gives us $y(x_{j+1}) = y(x_j) + hf\left(x_j, y(x_j)\right)$ + the remainder term. Now, what we will do is, we will bring this $y(x_j)$ to the left-hand side and divide both sides by h and thereby this h will go and you will be left out with only $\frac{h}{2}$ in the remainder term. Thereby you will get this equation where the left-hand side is the approximation for $y'(x_j)$ and that is equal to $f\left(x_j, y(x_j)\right)$.

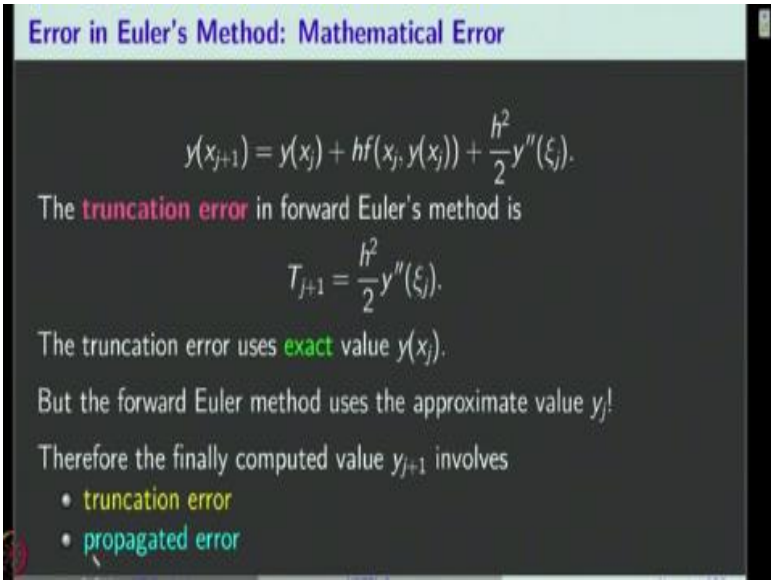
Up to here, this is what we have taken as the equation and now instead of having y', we now have the corresponding finite difference formula and therefore we have this error. And if you carefully look at this error, you can see that this error goes to 0 as h goes to 0. And what is the order in which this term goes to 0? You can say that this is of $O(h)$. In this case, we say that the error term is of order h that is $O(h)$ and because of this we can see that Euler method is of order 1.

(Refer Slide Time: 06:20)



Remember, the remainder term is given like this and the remainder term as it is going to 0 with order 2. But because we have divided both sides by h, in order to make this left hand side to be an approximation to y', we got to lose one order here and therefore Euler method finally happens to be of order 1. But what is the truncation error involved in this? The truncation error is nothing but the remainder term in the Taylor's expansion. And therefore, we call $\frac{h^2}{2}y''(\xi_j)$ as the truncation error.

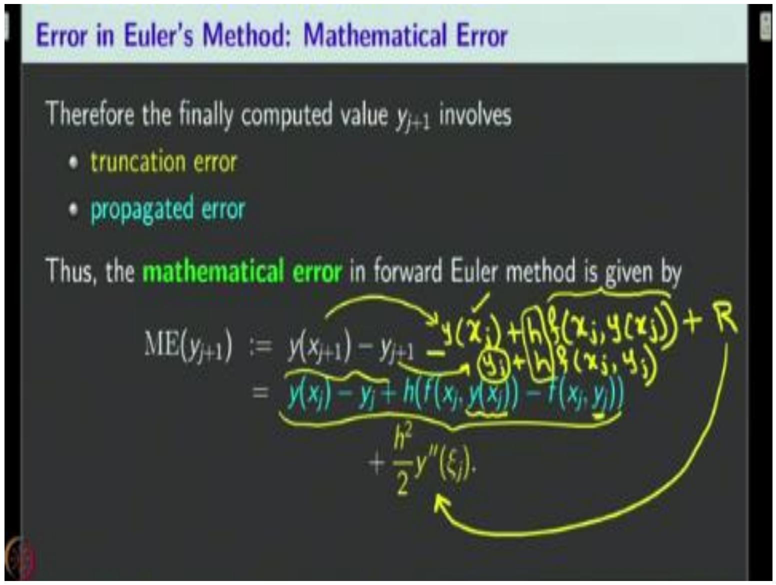
(Refer Slide Time: 07:12)



Note that, the truncation error is obtained when we use exact solution $y(x_j)$ while computing in the Taylor's expansion. But in the Euler formula we use the approximate value y_j while computing y_{j+1}. This shows that there are more levels of error involved in computing y_{j+1}. One is, of course the truncation error and the other one is called the propagation error. Therefore, the mathematical error involved in the forward Euler method has two components.

One is the truncation error and another one is the propagation error. Recall, that we came across a similar situation in our discussion on performing arithmetic operations with floating point approximations. Here we are facing a similar kind of situation.

(Refer Slide Time: 08:23)

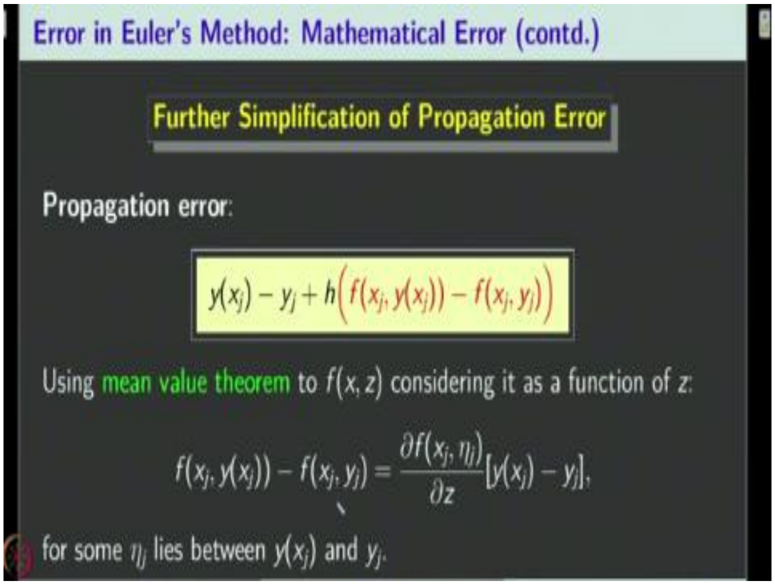


As I said, the mathematical error in the forward Euler method defined as the exact value minus the approximate value involves two levels of errors. Namely the truncation error which we have already derived and is given by $\frac{h^2}{2} y''(\xi_j)$ and the propagation error now given by this expression. How do we get this expression for the propagation error? Let us see from the Taylor expansion we have seen that $y(x_{j+1})$ is written as $y(x_j) + hf\left(x_j, y(x_j)\right)$.

Remember we had y' here but then we use the equation to replace y' by f, plus we had the remainder term which is sitting here and then you have minus the Euler forward formula which is $y_j + hf(x_j, y_j)$. So, this is what I am writing here precisely, you have $y(x_j) - y_j$, which is coming from here plus $hf(x_j)$, exact value that is coming directly from your Taylor expansion of the exact solution $-f(x_j)$, approximate value of the solution at x_j that is y_j coming from your forward Euler formula.

Plus of course the remainder term which is coming from your Taylor part is sitting here. And finally, we have written the mathematical error involved in y_{j+1} as the propagation error plus the truncation error.

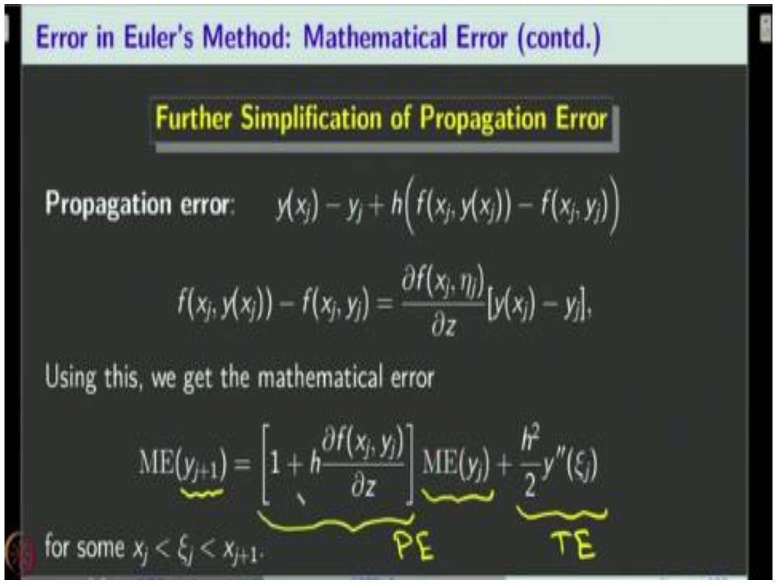
(Refer Slide Time: 10:53)



We can further modify the expression of the propagation error and bring it to a more clear form. Let us now do this. Recall the propagation error is given by this expression from the previous slide. We will now use the mean value theorem for f with respect to the second argument. Let us denote the second argument by z, just for the notation clarity. By mean value theorem we can find an η. We will denote it by η_j here, lying between the points $y(x_j)$, which is the exact value of the solution and y_j which is the approximate value at x_j computed by the Euler forward formula.

And thereby we get $f\left(x_j, y(x_j)\right) - f(x_j, y_j)$ is equal to now, we have to differentiate only partially with respect to the second argument times $y(x_j) - y_j$. This is precisely the mean value theorem for one variable but here it is applied only to the second argument of the function f by fixing the first argument x_j. Let us use this expression in the mathematical error to get mathematical error involved in y_{j+1}.

(Refer Slide Time: 12:33)



If you recall, this is nothing but the propagation error plus the truncation error. Remember we are just writing the truncation error as it is, without disturbing it. Now we are just writing the propagation error which is originally given by this expression in a slightly different way, using the mean value theorem. What we are doing is, just observe that this is nothing but mathematical error involved in y_j.

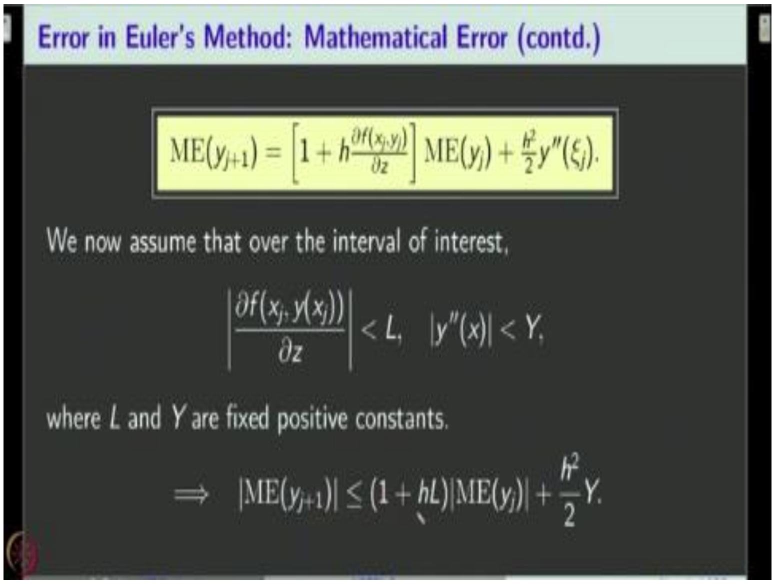
That is what we are writing here, one into mathematical error involved in y_j is nothing but the first term in the expression of the propagation error. Whereas the second term in the propagation error is now written like this and here also you can see that the mathematical error involved in y_j is sitting here. Therefore, this is also mathematical error in y_j that is what we are writing here h times this h is already here $\frac{\partial f(x_j,y_j)}{\partial z}$ which is coming from the mean value theorem times $ME(y_j)$.

Therefore, the mathematical error involved in y_{j+1} can be written as, something into the mathematical error in y_j, that is the propagation error, plus the truncation error. Here you can observe that the mathematical error in y_{j+1} is equal to this into the mathematical error in y_j + the truncation error, that is, the mathematical error at every grid point includes more or less the mathematical error coming from the previous grid point plus a new error.

That is the truncation error is something new that is getting accumulated at this step. If all these terms are positive, say, then the mathematical error will keep on increasing because mathematical error at the present grid is something which is coming from the previous grid plus new error. So, it may keep on increasing as we go on with the grid points. If you recall in our previous class, we have observed in the numerical solution of an example that as we go on with the grid points the error was increasing gradually.

Now, we can see clearly the reason for such a behaviour in the numerical solution. This is because every time the mathematical error in the present grid is nothing but the mathematical error from the previous grid may be multiplied with some number greater than 1, times surely a new level of error is added here. So, that is quite interesting. Let us try to obtain an estimate that is the upper bound of the absolute value of the mathematical error.

(Refer Slide Time: 16:21)

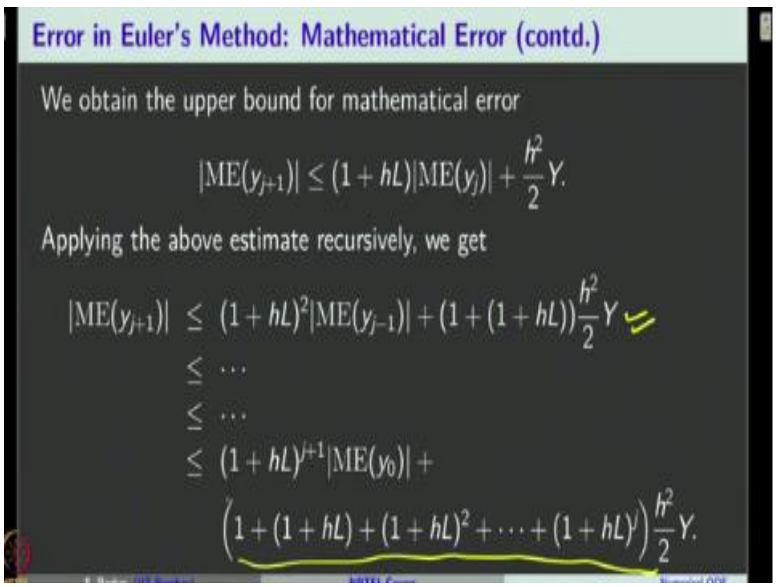


For this let us assume that $\left|\frac{\partial f(x_j,y(x_j))}{\partial z}\right|$ is less than L at the grid point x_j, that is this term is bounded by L, is what we are now assuming and also, we assume that $|y''(x)| < Y$, that is we are also assuming an upper bound for this term where L and Y are fixed positive constants.

Now what we will do is, we will take the modulus on both sides of this equation and then use these upper bounds in the appropriate places, to get an upper bound for the absolute value of the mathematical error in y_{j+1} and that is given by 1 which is already there plus h into, now instead of this term since we have taken modulus, we will put the upper bound here that is L times $|ME(y_j)| + \frac{h^2}{2}|y''(x)|$. Now we will put the upper bound here. Perhaps we will have to write it as strictly less than or we should write an equal sign here.

I am sorry for this error but we will keep this in mind. Now we see we got this inequality for the mathematical error in y_{j+1}. You can observe that the same inequality will hold even for the mathematical error in y_j. You simply have to put *j* instead of *j* + 1 then this will become y_{j-1}, in order to get the same type of inequality for $ME(y_j)$.

(Refer Slide Time: 18:31)

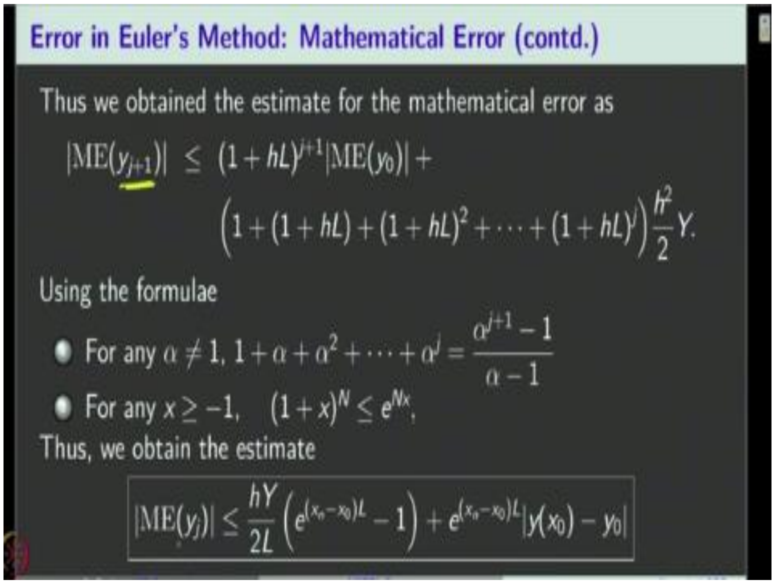


Now we will apply the same inequality for $ME(y_j)$ and then we will do this recursively. For instance, if you apply this inequality for $ME(y_j)$, you can observe that you will get this expression on the right hand side. How will you get it? This is less than or equal to, you already have 1 + hL, now you will have $(1 + hL)ME(y_{j-1})$, plus $\frac{h^2}{2}Y$, and that is for this term plus you already have $\frac{h^2}{2}Y$.

So, you combine all this, you will get this expression on the right hand side, when you put the same inequality to $ME(y_j)$. Now we will keep on putting this idea again and again. For instance, now you can put this inequality that is this inequality in place of $ME(y_{j-1})$ and get an expression after rearranging the terms, that will be something more than this. And you keep on going with this idea till you reach y_0 at this position.

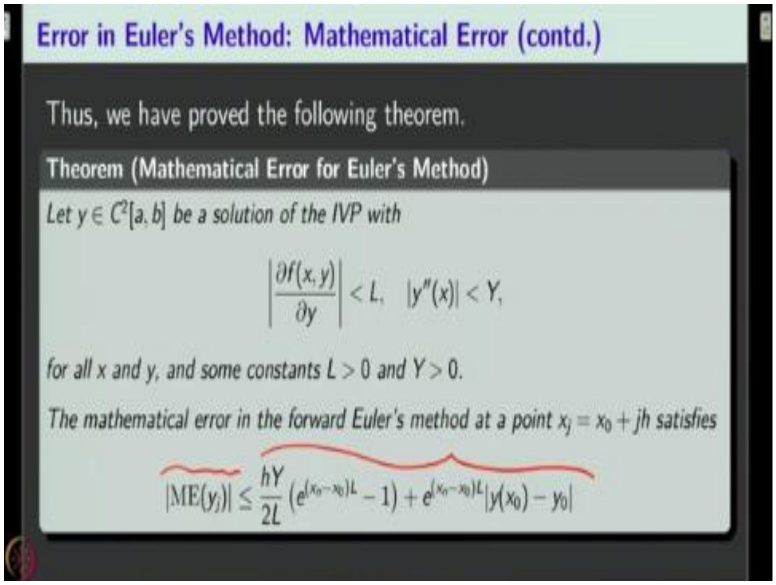
And at this stage, the right hand side expression will be given like $(1 + hL)^{j+1}ME(y_0)$ and the second term will look like this. You can just write it and see two or three times if you do this exercise, you will see how this pattern is forming and from there you can write this expression on the right hand side.

(Refer Slide Time: 20:44)



So, we got this estimate so far. In fact, we can simplify the right hand side further using some well-known formula, namely this one, which you can use for this term and we can also use this inequality, which is well known in the place of this term, and finally we can get the mathematical error involved in y_{j+1}. I have just written it as y_j, that does not matter because you can see that now after putting these formulas into this expression and after simplification, we can get this as the upper bound and that is clearly independent of y_j.

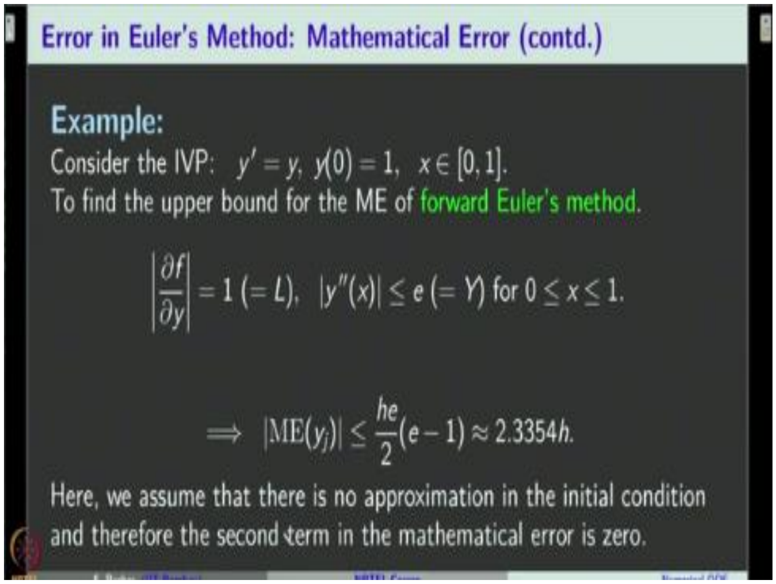
(Refer Slide Time: 21:49)



We will state the result, we have derived so far in the form of a theorem. As you can see that we need a bound for y'', so we have to assume that y is a C^2 function, in the interval in which we have posed our initial value problem. Then we can see that the mathematical error involved in y_j can be bounded by this quantity. Let us try to understand this upper bound. We may assume that the second term is 0 because $y(x_0)$ is actually given to us as the initial condition.

We also see that the upper bound depends on $x_n - x_0$. Thus, if $x_n - x_0$ is large, the upper bound will be very large, irrespective of any j, that we are computing. In that way, we can see that this estimate is an overestimate and often the actual error is much smaller, then what is predicted theoretically. However, theoretically this is the estimate that we could obtain so far. We have proved a bound for the mathematical error involved in the approximate solution of the Euler forward method.

(Refer Slide Time: 23:36)

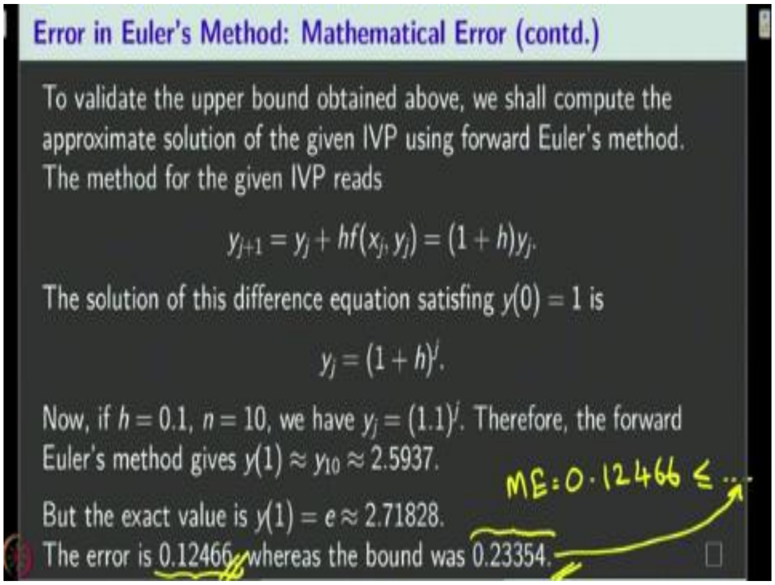


Let us take a simple example, where the ODE is $y' = y$ and the initial condition is $y(0)$ = 1, that is we have taken x_0 = 0 and the interval is taken as $[0,1]$. Let us find the estimate for the mathematical error involved in the forward Euler method, using the previous theorem. For this, first we have to find an upper bound of the partial derivative of *f* with respect to *y*. You can see clearly, that this is equal to 1. And let us find a bound for $y''(x)$ also.

For this we need to know the solution and from there you can see that the upper bound may be taken as *e*. So, this is just for the sake of example we are doing. Then, from the estimate we obtained from the previous theorem, we can see that the mathematical error is bounded by 2.3354h, where $h > 0$ is the discretization parameter. Remember, in the theorem we had two terms, the second term is not appearing here.

Because, we have taken the initial condition as 1 and therefore, we assume that there is no error committed in the second term. Therefore, we have directly taken the second term as 0, that is why there is no second term involved in our upper bound here. The bound we obtained above is purely from the theoretical point of view. Let us see how it works numerically.

(Refer Slide Time: 25:50)



First let us write the formula of the Euler forward method for the present example. If you recall the forward Euler formula is given like this. Here you have to take $f(x_j, y_j) = y_j$, that is what is given as the right-hand side in our ODE. Therefore, if you take $f = y_j$, then the Euler forward formula reduces to this expression. Since $y(0) = 1$ we can see that $y_j = (1+h)^j$. Why it is so? Take y_1, which is equal to $(1+h)y_0$, where $y_0 = 1$ that is equal to therefore $1 + h$.

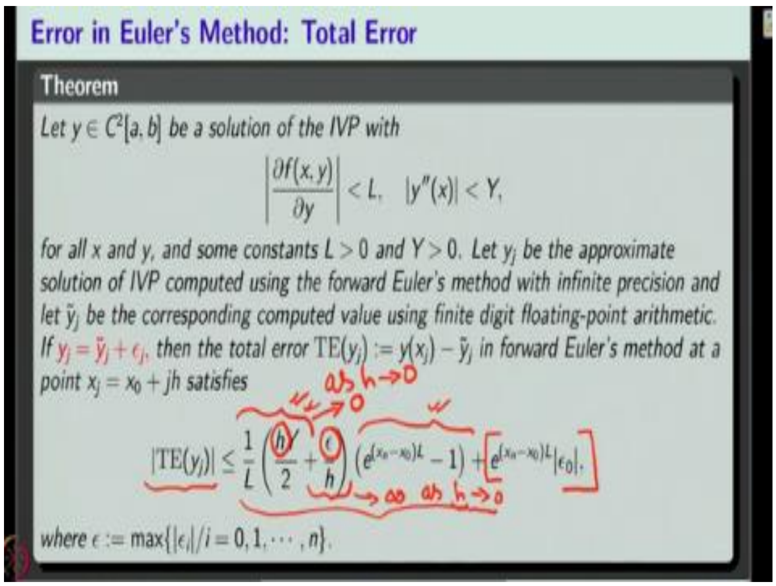
Now if you go to y_2, then y_2 is given by $(1+h)y_1$ and that is equal to $(1 + h)$ into $(1 + h)$ which is coming from here and therefore it will be $(1+h)^2$. Similarly, you can see that $y_3 = (1+h)^3$ and so on. That is why we have written $y_j = (1+h)^j$. Let us take $j = 10$ and we are interested therefore in y_{10}. Let us take $h = 0.1$ and of course $n = 10$. The moment you take $h = 0.1$, then y_j is given by $(1.1)^j$.

Now we are taking $j = 10$, therefore you can see that y_{10} is nothing but $(1.1)^{10}$, and that is given approximately as 2.5937. But y_{10} is the approximate value of $y(1)$, because we are starting from 0 and taking $h = 0.1$ and going 10 grid points. Therefore, at the end y_{10} will be the approximate value of $y(1)$. And what is the exact value of $y(1)$? That is nothing but e and it is given by 2.71828.

And you can now compare the approximate solution and the exact solution and see what is the mathematical error involved in them. You can directly find the mathematical error now, because we have exact solution and approximate solution. You can directly take the difference between them and the error is given by 0.12466 whereas the bound that we obtained from our theoretical estimate is actually 0.23354, because in the previous slide we have obtained that.

So we have this now, you took h = 0.1. Therefore, we have the upper bound as 0.23354 which is theoretically predicted and from the numerical experiment, we see that the error is 0.1246. So, as we expected the mathematical error involved in the approximation is less than or equal to the theoretically predicted number that is this. Therefore, our numerical example is well in agreement with the theoretical prediction.

(Refer Slide Time: 30:27)



Finally let us obtain an estimate for the total error, when we include the arithmetic error in our computation. Let us assume that we obtained $\tilde{y}_j$ instead of y_j due to some floating-point error involved in the computation. The error in $\tilde{y}_j$, when compared to y_j is say ϵ_j, then the total error which is defined as $y(x_j)$, that is the exact value $-\tilde{y}_j$ now. Because we are not obtaining y_j as we have some rounding error involved in our calculations.

Therefore, we only obtained $\tilde{y}_j$, therefore the total error is nothing but the mathematical error which is $y(x_j) - y_j$ + the arithmetic error which is $y_j - \tilde{y}_j$. And that will be your total error. So, this is the mathematical error and this is the arithmetic error. We have already derived an estimate for the mathematical error. Similarly, we can also derive an estimate for the arithmetic error and that can be given with a factor like this.

Remember the mathematical error is this much into this plus this and now the arithmetic error brings in a new term like this. So, this is the new term coming from the arithmetic error. I will leave it to you to derive this. It is not very difficult, once you understand the derivation of the upper bound of the mathematical error, the idea goes exactly the same. Recall, we have also obtained the bound for arithmetic error in certain finite difference formulas.

The idea will go exactly in a same way. But what is interesting for us to observe here is, when you take h tending to 0, you can see that the mathematical error which is this times this one. Of course, we will always not give that much importance to the second term because that will often depend on the initial condition and the error involved in the initial condition. So, we will only bother about the actual computation part.

If you see the mathematical error alone, it is this term into this and if you take as h tends to 0, you can see that the mathematical error is nicely tending to 0. What about the arithmetic error part? You can see that the arithmetic error part tends to infinity as h tends to 0. What it says? If you have even a small arithmetic error in your calculation, that may tend to amplify at least the upper bound here. But the fact is, it will also amplify the total error if you keep on reducing the grid step size h.

So, do not think that you keep on reducing h, you get better and better approximation. Just like we discussed in the finite difference formula, there will always be an optimal h. If you go on reducing h below that optimal h, then your total error will tend to increase. This is the message we are getting from the upper bound of the total error involved in the forward Euler method. Remember the same kind of analysis can also be performed for backward Euler method.

I leave it to you to see that. With this cautious note, we will close this lecture, thank you for your attention.