

Numerical Methods and Computation

Prof. S.R.K. Iyengar

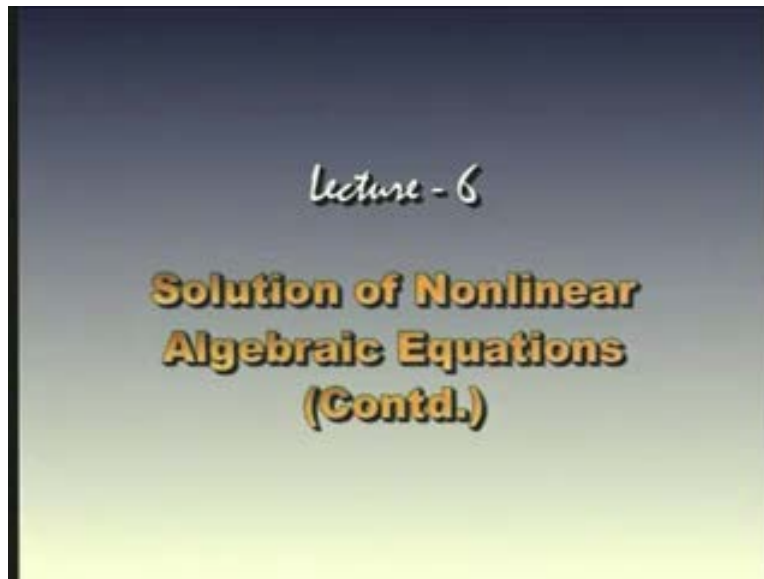
Department of Mathematics

Indian Institute of Technology Delhi

Lecture No # 6

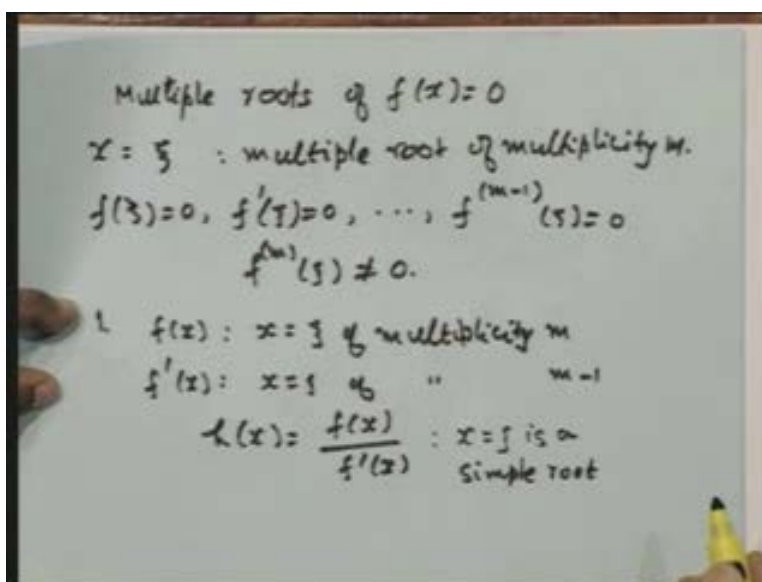
Solution of Nonlinear Algebraic Equations (continued)

(Refer Slide Time: 00:00:56 min)



In our last lecture we were discussing about the finding the multiple roots for an equation $f(x)$ is equal to zero.

(Refer Slide Time: 00:03:38 min)



What we are really looking for was the multiple root of fx is equal to zero. We know the definition of multiple root. Let us say x is equal to x_i which is a multiple root of multiplicity m . So let us take this multiplicity also. It is a multiple root of multiplicity m . Then we know that f of x_i would be zero, and then its derivative is also zero; f prime of x_i is equal to zero and so on up to m minus 1th derivative. It is equal to zero and n^{th} derivative is not equal to zero. This is the basic definition of the multiple root of multiplicity m .

Now to find this particular root we shall first of all attempt whether we can use the methods that we have derived for simple roots by some modification or some other criteria; such that we can use those methods without having to make much changes in the methods that we have. Let us first of all attempt in this particular form. We know that if f of x has got a root, x is equal to x_i of multiplicity m and then its derivative f prime of x has got the same root x is equal to x_i of multiplicity m minus one. Therefore we would like to consider a function which is a ratio of these two and define h_x as fx upon f dash x . Then this particular ratio i.e. this new function that we have defined as h_x has got a simple root at x is equal to x_i . Therefore for this function x is equal to i is a simple root. What does this imply? This implies that since x is equal to x_i is a simple root, I can now apply the all the available methods that we have discussed for finding the simple root; to find the simple root of h_x is equal to zero and hence the multiple root of fx is equal to zero which is the multiplicity m .

(Refer Slide Time: 00:06:13 min)

Newton-Raphson method

$$x_{k+1} = x_k - \frac{h(x_k)}{h'(x_k)}$$

$$h(x_k) = \frac{f(x_k)}{f'(x_k)}$$

$$h'(x) = \frac{d}{dx} \left[\frac{f}{f'} \right] = \frac{(f')^2 - f f''}{(f')^2}$$

$$x_{k+1} = x_k - \left(\frac{f_k}{f'_k} \right) \frac{(f'_k)^2}{(f'_k)^2 - f_k f''_k}$$

$$= x_k - \frac{f_k f'_k}{(f'_k)^2 - f_k f''_k}$$

Let us try to get the Newton Raphson method and its application to find the multiple root. Now what we are saying is, we are applying the Newton Raphson method of $h(x)$. Therefore the iterative method would be x_{k+1} is equal to x_k minus h at x_k divided by h' at x_k . Now I need these two quantities to put here. So h at x_k will be simply equal to f at x_k divided by f' at x_k , so that is from the definition of $h(x)$. Now differentiate h' . So let's get out derivative of this that will be d upon dx of f upon f' . So let us differentiate it. So if I differentiate you will get f' into f' i.e. f'^2 minus $f f''$ divided by f'^2 . So that is the simple derivative of f upon f' . Therefore I can now evaluate h' at x is equal to x_k and then substitute it over here. So I can put this here and get it as x_{k+1} is equal to x_k minus f of k by f' of k ; that is our value of h at x_k . In the denominator we have got h' . So this will go up as f'^2 divided by f'^2 minus $f f''$, which we can simplify in one more step i.e. by cancelling one of f' . So I can write this as x_k minus f evaluated at k , f' at k from here, and the denominator is the same denominator f'^2 minus $f f''$ at k . Therefore I can obtain the simple root of $h(x)$ is equal to zero using this Newton Raphson method modification and therefore this gives me the multiple root of multiplicity of the original equation.

(Refer Slide Time: 00:08:41 min)

Handwritten notes on a whiteboard:

$x \approx x_k \approx \xi$ is an approximation to the multiple root.

Cost: Evaluate f_k, f'_k, f''_k

Order: 2

Secant method

$$x_{k+1} = \frac{x_{k-1} f_k - x_k f_{k-1}}{f_k - f_{k-1}}$$

Therefore we can say that x is equal to x_k approximation. This is ξ . It is an approximation to the multiple root. Therefore we obtain the root without any further modifications in the Newton Raphson method. Now if this is always so we do not need any other method, but if you look at the cost of the computation; the cost of the computation needs to be evaluated, f at k , f' at k , f'' at k . So we have got this f_k , f'_k and f''_k . So I need to evaluate three evaluations to get the Newton Raphson method and the order of the method is same because we have evaluated the simple root of hx is equal to zero. The order of convergence is two. So the order of convergence still remains as two and only thing that we have done is one extra computation. If this extra computation is allowed and we do not have any problem, we can definitely use this particular method; but however we shall give an alternative method wherein we do not need evaluation of three, we need evaluation of only two, provided you know the order or multiplicity in advance.

Now this method can also be applied for secant method. We can apply this secant method. Let us see how you put the same thing into the secant method. Let us take the way in which we had written the computational way of writing the formula. We have written it as x_{k+1} is equal to $x_{k-1} f_k - x_k f_{k-1}$ divided by $f_k - f_{k-1}$. This is the computational form of secant method that we have used earlier. So therefore I need to only insert into this the values of f_k , f_{k-1} ; this is equal to x at k minus one.

(Refer Slide Time: 00:11:51 min)

$$\begin{aligned}
 x_{k+1} &= \frac{x_{k-1} f_k - x_k f_{k-1}}{f_k - f_{k-1}} \\
 &= \frac{\left[x_{k-1} \left(\frac{f_k}{f'_k} \right) - x_k \left(\frac{f_{k-1}}{f'_{k-1}} \right) \right]}{\frac{f_k}{f'_k} - \frac{f_{k-1}}{f'_{k-1}}} \\
 &= \frac{x_{k-1} f_k f'_{k-1} - x_k f_{k-1} f'_k}{f_k f'_{k-1} - f_{k-1} f'_k}
 \end{aligned}$$

Order: 1.618 f_k, f'_k

Now I need to write this for the evaluation of the simple root for $h(x)$ is equal to zero. So I would now use the method as x_{k+1} is equal to x_k minus h evaluated at x_k minus x_{k-1} divided by h_k minus h_{k-1} . Now we have written everything in terms of the function $h(x)$. So let us substitute the value of what is h_k and what is h_{k-1} . So that will simply give us x_k minus f of k minus f prime of k i.e. the numerator minus x_k , f of k minus one is f_k minus one upon f prime at k minus one and then we shall divide the whole thing by h_k minus h_{k-1} , so f_k minus f prime at k minus f_k minus one f prime at k minus one. Now it is just a matter of simplification over here. We can see that from the denominator first of all we can write this as f_k into f prime k minus one minus f_k minus one f prime of k . So I first simplified the denominator; when it goes to the numerator we can see f prime k f prime k minus one cancels with the f prime k from this. Therefore I would not repeat that particular factor. So the numerator would be simply x_k minus one from here, f_k from here and f prime at k minus one minus x_k f_k minus one into f prime of k . So that is the simplification of the numerator. So this will be the secant formula as applied to $h(x)$ is equal to zero. Therefore the order of the method is the same as 1.618. Therefore rate of convergence would not change. In the secant method also we are now using f prime whereas in the original secant method for finding a root of $f(x)$ is equal to zero we have been using only f_k, f_{k-1} . So we have got only one evaluation at any particular stage but here we need f_k . We need f prime at k plus one and f prime k . Therefore at any stage I need to evaluate two evaluations i.e. your f_k is required and f prime of k is also required.

(Refer Slide Time:00:15:26 min)

$$\begin{aligned}
 & 2. \text{ } m \text{ is known (multiplicity)} \\
 & f(\xi) = 0, f'(\xi) = 0, \dots, f^{(m-1)}(\xi) = 0 \\
 & f^{(m)}(\xi) \neq 0. \\
 & f_k = f(x_k) = f(\xi + \epsilon_k) \\
 & = \frac{\epsilon_k^m}{m!} f^{(m)}(\xi) + \frac{\epsilon_k^{m+1}}{(m+1)!} f^{(m+1)}(\xi) + \dots \\
 & = \frac{\epsilon_k^m}{m!} f^{(m)}(\xi) \left[1 + \frac{\epsilon_k}{m+1} F + O(\epsilon_k^2) \right] \\
 & F = f^{(m+1)}(\xi) / f^{(m)}(\xi).
 \end{aligned}$$

So we have now again increased one function evaluation as the cost. Therefore it will be more expensive than the normal secant method. As I said earlier if the evaluation of f' at x_k is not very costly, we can definitely go about using the secant and the Newton Raphson methods. However if the multiplicity is known; let me assume that the multiplicity m is known. Now I would like to use the definition of multiplicity as we have written earlier. Let us repeat it once more; the f of x_i is equal to zero, f' of x_i is equal to zero upto $m-1$ th derivative; all of them are zero and n th derivative at x_i is not equal to zero. So this is the definition of the multiplicity at the root x is equal to x_i . I would like to have a look at what happens in the error analysis as we have done in the Newton Raphson method. Let us start with the what is the definition of f_k . Definition of x_k is, f of x_k and x_k is the exact solution plus the error. Now I would expand this in Taylor series. We are just following the steps that we have done for showing the error analysis of the Newton Raphson method. When I expand it, I will use this information. Therefore when I write f of x_i the first term is zero and second term is $\epsilon_k f'$ of x_i i.e. zero and so on. All the terms, the first m terms are zero because of this information. The first non vanishing term in the expansion is containing $f^{(m)}$ of x_i . Now if I write that one, the first non vanishing term is $\epsilon_k^m / m!$ $f^{(m)}$ of x_i plus (the next term is) $\epsilon_k^{m+1} / (m+1)! f^{(m+1)}$ of x_i . Before I add the denominator let us first simplify this expression. Take out whatever is possible for us. This is the common factor here. So I will take out the factor $\epsilon_k^m / m!$ $f^{(m)}$ of x_i also I will take it out and write this as one plus (I have taken out $\epsilon_k^m / m!$ $f^{(m)}$ of x_i) I have ϵ_k left out here; denominator is $m+1$ and this ratio let us just call it as F . Let us call the ratio F as the ratio of $f^{(m+1)}$ of x_i divided by $f^{(m)}$ of x_i . So I have just taken the common factor here and then I have written this remaining part as one plus $\epsilon_k / (m+1) F$; F is this ratio and whatever we are writing here are second order terms and higher order terms.

(Refer slide Time: 00:19:35 min)

$$\begin{aligned}
 f'(x_k) &= f'(\xi + \epsilon_k) \\
 &= \frac{\epsilon_k^{m-1}}{(m-1)!} f^{(m)}(\xi) + \frac{\epsilon_k^m}{m!} f^{(m+1)}(\xi) + \dots \\
 &= \frac{\epsilon_k^{m-1}}{(m-1)!} f^{(m)}(\xi) \left[1 + \frac{\epsilon_k}{m} F + O(\epsilon_k^2) \right] \\
 \frac{f(x_k)}{f'(x_k)} &= \frac{\epsilon_k}{m} \left[1 + \frac{\epsilon_k}{m+1} F + \dots \right] \left[1 + \frac{\epsilon_k}{m} F + \dots \right]^{-1} \\
 &= \frac{\epsilon_k}{m} \left[1 + \frac{\epsilon_k}{m+1} F + \dots \right] \left[1 - \frac{\epsilon_k}{m} F + O(\epsilon_k^2) \right] \\
 &= \frac{\epsilon_k}{m} \left[1 + \left(\frac{1}{m+1} - \frac{1}{m} \right) \epsilon_k F + O(\epsilon_k^2) \right]
 \end{aligned}$$

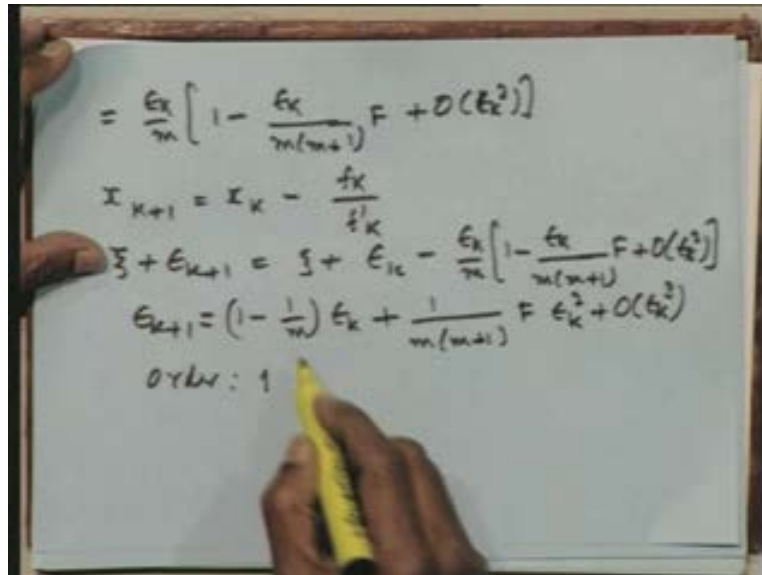
Now the same thing we would like to do for the denominator also. Now if I take the denominator, the denominator will contain f' of x_k . So this is f' of ξ plus ϵ_k . Now again expand this by Taylor series. Again I would be using the information that all these derivatives are zero upto $m-1$ derivative and this will be the non vanishing term. Therefore the first non vanishing term will be containing $f^{(m)}$ derivative but this will be ϵ_k to the power of $m-1$ by $(m-1)!$ and $f^{(m)}$ of ξ plus the next term is ϵ_k to the power of m by factorial m $f^{(m+1)}$ of ξ plus so on. Now as we have done in the previous case let us now take out the common factor here also; ϵ_k to the power of $m-1$ by $(m-1)!$, $f^{(m)}$ of ξ is again a common factor. Now whatever is left out is, one ϵ_k here, m here; this ratio is exactly the same $f^{(m+1)}$ of ξ is same as F and order of ϵ_k square. Now I will take the ratio of f of x_k upon f' of x_k . Now let me just put this slide back once more; we have here $f(x_k)$, the factor outside is ϵ_k to the power m by factorial m $f^{(m)}$ of ξ and the factor here is ϵ_k to the power of $m-1$ by $(m-1)!$ $f^{(m)}$ derivative of ξ .

Now if I take the ratio of these two you can see that $f^{(m)}$ of ξ cancels, ϵ_k to the power of m minus one cancels and $(m-1)!$ cancels. So whatever is left out is simply ϵ_k in the numerator here and m in the denominator, so remaining parts get cancelled out of this particular ratio and what we have here is one plus (I am writing this particular term) ϵ_k upon $m+1$ into F plus higher order terms; in the denominator is this particular factor which I will take it up and write this as ϵ_k by m to the power of -1 F to the power of -1 . So I have taken the denominator up.

Now let us open it up or simplify it further. I will retain the first term as it is, ϵ_k by m plus one of F . Then I expand it out by binomial expansion; one minus ϵ_k by m into F plus order of ϵ_k square times. Now let us multiply it out and simplify it, so what I will have here is ϵ_k by m into one and ϵ_k term is one upon $m+1$ from here, minus one upon m from here, ϵ_k into F plus order of ϵ_k square. This is ϵ_k square; this is

epsilon k square and this is epsilon k square. So whatever is left out are all the remaining terms of order of epsilon k square.

(Refer Slide Time: 00:23:09 min)



$$= \frac{E_k}{m} \left[1 - \frac{E_k}{m(m+1)} F + O(E_k^2) \right]$$

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

$$\xi + E_{k+1} = \xi + E_k - \frac{E_k}{m} \left[1 - \frac{E_k}{m(m+1)} F + O(E_k^2) \right]$$

$$E_{k+1} = \left(1 - \frac{1}{m}\right) E_k + \frac{1}{m(m+1)} F E_k^2 + O(E_k^3)$$

order: 1

Now let us simplify it further. So I would therefore have this as epsilon k by m into (now here I am simplifying this, therefore I will have here m minus m minus one). So I will have one minus epsilon k by m into m plus one into f plus order of epsilon k square. So this is what I have from the simplification. May be if you just want to have both the terms together; I will bring this slide down once and we can just have a look at what we have. This is one minus epsilon k by m into m plus one into f k square.

Now let us put this into our Newton Raphson method and see why the difficulty was arising. So the method was x_k plus one is equal to x_k minus f by f' .

Therefore we have the error as x_i plus epsilon k plus one is x_i plus epsilon k minus (now we have just now obtained the ratio - this and this; this is ratio) therefore this is epsilon k by m into one minus epsilon k by m into m plus one into f plus order of epsilon k square. Now let us simplify; x_i cancels off from here, so the left hand side is epsilon k plus one and on the right hand side I will connect epsilon k from here, so one minus one by m into epsilon k. This is plus, we can write down this quantity; one upon m, m plus one f epsilon k square plus order of epsilon k cubed. So we just collected the terms and wrote this as one minus one upon m into epsilon k; this is plus one upon m into m plus one f into epsilon k square plus one.

Now the error for the Newton Raphson method; the right side is epsilon k. Therefore order is one. Therefore Newton Raphson method as applied to finding a multiple root would drop down the order by one from quadratic convergence to linear convergence only. As I said this can easily be experienced when you are actually doing the computation. When you know that the method should converge fast, after few iterations what would happen in your computation when look at the results is, you find that after even large number of iterations the error is reducing it very slowly. If you know the exact solution of the problem and then see the error, the error is reducing

by only a factor of one decimal place or so. Therefore this slow convergence will imply that you are possibly going at a multiple root. Now we have shown that for a multiple root Newton Raphson method has only order one; then what do we do? We can do a very simple modification to the Newton Raphson method.

(Refer Slide Time: 00:25:38 min)

$$x_{k+1} = x_k - \frac{\alpha f_k}{f'_k} \quad \alpha = \text{constant}$$

$$\epsilon_{k+1} = \left(1 - \frac{\alpha}{m}\right) \epsilon_k + \frac{\alpha}{m(m+1)} F \epsilon_k^2 + O(\epsilon_k^3)$$

choose $\alpha = m$

$$\epsilon_{k+1} = O(\epsilon_k^2) = C \epsilon_k^2$$

$$C = \frac{\alpha}{m(m+1)} \frac{f^{(m+1)}(x)}{f^{(m)}(x)}$$

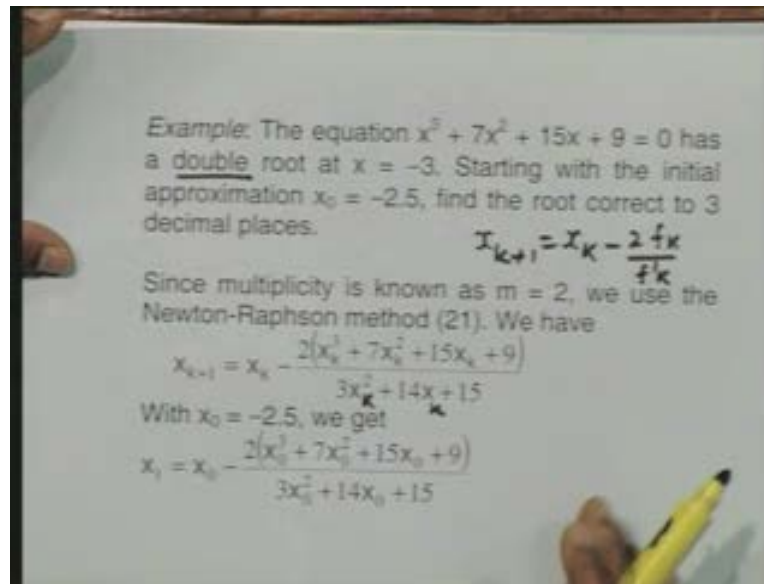
Order: 2

That simple modification is; suppose I start the Newton Raphson method with x_{k+1} equal to x_k minus alpha. Alpha is a constant. So I would try to start a method like this. Now let us just go back to this slide once. Why I have introduced an? I have introduced an alpha for this, because this ratio is f_k by prime k . So let me go to slides behind; we have taken the ratio of f by prime k at this quantity which you have finally brought it to the form of this as f_k by f prime k . Now if I put an alpha over here, what I am doing here is; I am bringing alpha over here. If I am bringing an alpha over here this is going to be an alpha here, that means this error expression would therefore look like, (I would just write the same thing straight away because it is trivial quantity) one minus alpha by m epsilon k plus alpha m into m plus one F epsilon k square plus order of epsilon, epsilon k cubed. Therefore if I introduce this alpha over here, I am introducing an alpha here in this bracket, one minus alpha by m and I am multiplying a by alpha in this one.

Now you can see that we started with alpha as a constant. If I now chose alpha is equal to m , then epsilon k vanishes. Now we shall say, choose alpha is equal to m . If I choose alpha is equal to m , then epsilon k square is order of epsilon k square or we can call it as C epsilon k square, where now C is this quantity alpha by m into m plus one; if you open it out (what we had written it earlier) f_m plus one x_i by f_m of x_i . So the order has immediately gone up to two and now we have got the rate of convergence of this method as two. That means again we have got back our quadratic convergence by simply using the order of multiplicity that is known to us in advance in the Newton Raphson method here. If we are able to know it therefore we are now avoiding evaluation of one more function. As we have seen earlier we wanted evaluation of f at x_k , f prime at x_k and f double prime at x_k for the Newton Raphson method to have the required second order convergence. So here I don't need any extra evaluation. The cost is the same;

identically same as the cost as the earlier method for simple root except that we know in advance the order of multiplicity. So that factor can be brought in here and then use that as the required Newton Raphson method, the order of convergence will be remain same. Of course if you do not know the multiplicity of the root, we have to follow the other way of getting the root.

(Refer Slide Time: 00:28:35 min)



Now let us take a simple example of this. Now I take this equation x cubed plus seven x square plus fifteen x plus nine is equal to zero. It is a double root. It has double root at x is equal to minus three. I start the initial approximation x_0 is equal to minus 2.5 and the find the root corrected to three decimal places and I would be using the Newton Raphson method.

Now the first thing is, the multiplicity given to us; it is a double root. Therefore m is equal to two. Now since m is equal to two is given to us I would now write down my method as x_{k+1} plus one is equal to x_k minus two times f_k by f' prime k . So I would now be using that x_{k+1} plus one is equal to x_k minus two times f_k upon f' prime k . Now let us substitute the values of f of x_k , f' prime x_k i.e. x_k plus one is x_k , two f of k is from here, that is x_k cube plus seven times x_k square plus fifteen times x_k plus nine. We are differentiating with respect to x . So I will have three x square plus fourteen x plus fifteen. I have differentiated this; this is three x square plus fourteen x plus fifteen in the denominator.

(Refer Slide Time: 00:29:03 min)

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

$$f(x) = 3x^3 + 14x + 15$$

$$f'(x) = 9x^2 + 14$$

$$\text{With } x_0 = -2.5, \text{ we get}$$

$$x_1 = x_0 - \frac{2(x_0^3 + 7x_0^2 + 15x_0 + 9)}{3x_0^2 + 14x_0 + 15}$$

$$= -2.5 - \frac{2(-0.375)}{-1.25} = -3.1$$

$$x_2 = -3.1 - \frac{2(-0.021)}{0.43} = -3.0023$$

$$x_3 = -3.0023 - \frac{2(0.000001)}{0.000001} = -3.5000001$$

In three iterations we have got an accuracy of 5 decimal places.

Now we start the iteration. It is given to us the initial approximation as minus 2.5.

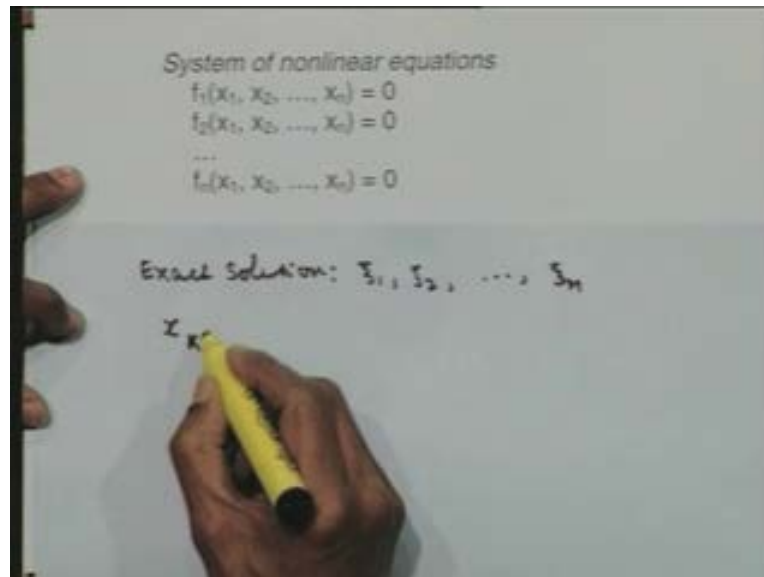
So I can substitute it here; k is equal to zero, I have the right hand side as x_0 , two times x_0 whole cube, seven times x_0 square, fifteen times x at zero plus nine divided by three x_0 square plus fourteen x_0 plus fifteen. So we just substitute the value of minus 2.5 in the expression and then evaluate it from here. I substitute it over; this is minus 2.5 and I evaluate the numerator, two times minus 0.375, the denominator is minus 1.5 and that is equal to the value minus 3.1. So minus 2.5; the iteration has gone to minus 3.1; I substitute x_2 is equal to x_1 minus two times everything evaluated at x_1 numerator and denominator is also evaluated at x_1 . So I have here minus 3.1 minus two times of what I evaluated; I evaluate this at x_1 i.e. at the value minus 3.1, I get minus 0.021 and the denominator is 0.43. I simplify the whole thing. I get minus 3.0023. I repeat the same thing by taking this as x_2 two and numerator evaluated at x_2 ; denominator evaluated at x_2 . So I have got here minus 3.0023 minus two times the numerator by denominator and this is minus 3.5000001.

Now we can see that in three iterations we have got an accuracy of five decimal places. Indeed it would be very interesting for you to just take it as an exercise and solve the same problem by Newton Raphson method without inserting the factor two. You will be able to know just after three iterations with your computations that we are converging very slowly in this one. So hence if I do not know the multiplicity that is given to us, I would go to the alternative way of finding the second derivative also and then using the Newton Raphson method.

Now very important extension or application of the solution of a single nonlinear equation is a system of nonlinear equations. Why it is very important here is, in most of the practical applications the mathematical model of the physical system is either ordinary differential equation or a partial differential equation and in most cases the process are nonlinear processes. Therefore the differential equations that are produced are nonlinear ordinary differential equations or nonlinear partial differential equations. Now we use some numerical methods; finite difference method or finitely methods to solve this equation and we produce a system of

nonlinear algebraic equations which are the solutions of nodal points; the points of intersection of the mesh. Therefore what we have is thousands of nonlinear algebraic equations and how do we solve it? It is possible for us to extend directly the methods that we have discussed now for solving the system of nonlinear equations.

(Refer Slide Time: 00:32:09 min)



So let us take the system of nonlinear equations as this. So I have got n equations as unknowns. Now as I said it is not unusual to get few lakhs of equations, particularly in finite difference methods. If you have got a three dimensional problem, in the three dimensional problem it is not difficult at all to get the thousands or even few lakhs of nonlinear equations. Now one comment before we give the method is that, we have already seen that if Newton Raphson method is to converge it is necessary to have some good initial approximation; we cannot be very far away from the exact solution of the problem. Now the case becomes much more important or discussion becomes more important in the case of system of nonlinear equations. Wherein let us say you have got ten variables and if one variable or two variables are given very accurately but other variables given are very far away from the exact solution. By the time those have been modified the solutions which are close enough could be spoiled. Therefore it is necessary that all these values, through some other procedure or from physical considerations, one must have an idea of what could be the order of magnitude or what the type of solution could be, so that it can be used as your starting solution.

(Refer Slide Time: 00:35:51 min)

Exact solution: $\xi_1, \xi_2, \dots, \xi_n$
 $x_k^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}$: Any approximation
 $\xi_1 = x_1^{(k)} + h_1, \xi_2 = x_2^{(k)} + h_2, \dots$
 $\xi_n = x_n^{(k)} + h_n$
 $f_i(x_1^{(k)} + h_1, x_2^{(k)} + h_2, \dots, x_n^{(k)} + h_n) = 0$
 $i = 1, 2, \dots, n$

We shall follow the same procedure as used in deriving the Newton Raphson method earlier in one of our lectures. So let us take the exact solution of this particular problem. Let us say the exact solution is x_{i1}, x_{i2} and x_{in} . So this is exact solution of this problem. We have used the Taylor series expansion to get the Newton Raphson method. One of the ways of deriving is using the Taylor series method. We shall use that method to derive the method for the nonlinear equations. Now let us take any initial approximation; x_1k, x_2k and x_nk . So this is any approximation. I would now add to these suitable quantities, so that they become the exact solution, which means I will add x_{i1} is equal to x_1k plus some h_1 , x_{i2} is equal to x_2k plus h_2 and so on. I will put $x_{in}k$ plus h_n is equal to x_{in} . So I am adding the suitable quantities so that I get the exact solutions. Therefore if I insert this here, x_{i1}, x_{i2} and x_{in} are exact solutions. Therefore when I substitute, it is satisfied identically. So if I take any equation, any i^{th} equation, this will read as x_1k plus h_1 x_2k plus h_2 plus x_nk plus h_n is equal to zero. I have substituted these values in this.

So i is going from one two three n .

What I now do here is, I expand this in Taylor series; i.e. the function of n variables and retain only the first order terms, that means I would now write this as f_i evaluated as x_1k, x_2k, x_nk plus; from earlier expansion this will be the h_1 into partial derivative with respect to x_1 evaluated at the point, plus h_2 partial derivative with respected to x_2 evaluated at the point and so on. So this will be having h_1 into Δf_i upon by Δx_1 . All of them are evaluated at k , so you can put it as "at k " if you like. Then we will have this as h_2 partial derivative of f_i with respected to x_2 evaluated at xk plus so on and h_n partial derivative of f_i by x_n evaluated at k is equal to zero.

(Refer Slide Time: 00:39:57 min)

$$f_i(x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}) + h_1 \left(\frac{\partial f_i}{\partial x_1} \right)_k + h_2 \left(\frac{\partial f_i}{\partial x_2} \right)_k + \dots + h_n \left(\frac{\partial f_i}{\partial x_n} \right)_k = 0, \quad i = 1, 2, \dots, n$$

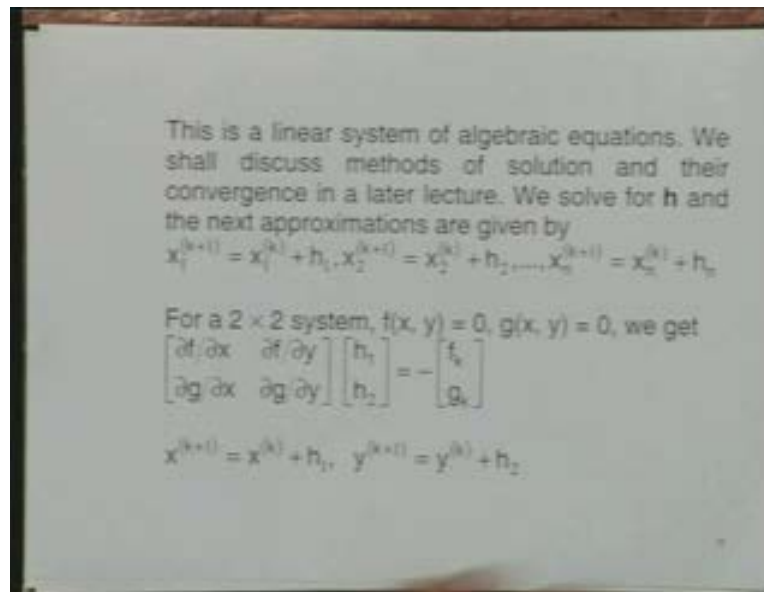
$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \dots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix} = - \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix}_k$$

$$\underline{J} \underline{h} = -(\underline{F})_k$$

\underline{J} : Jacobian matrix of the system.

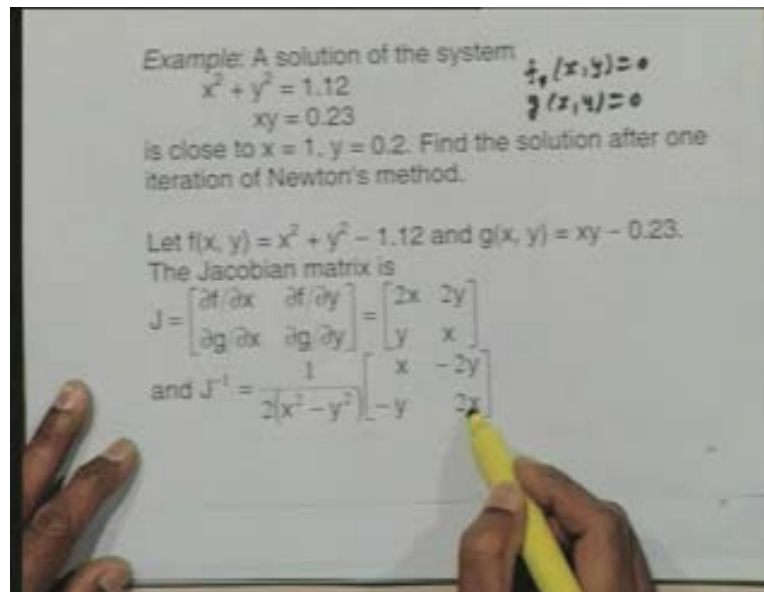
Now I have a collection of these equations one, two, three, n all these equations. I take this. This is a quantity that is known to us; x_1^k, x_2^k, x_n^k is a previous iteration. So the computed values are all known to us. So I can take this to the right hand side and then I will write this in the matrix format for each one of them. So let us take i is equal to one, two, three, so on. So I will have Δf upon Δx_1 , Δf of upon Δx_2 , Δf_1 by Δx_n , f_2 by Δx_1 , Δf_2 by Δx_2 , Δx_n . So we are putting in the matrix format. So this will be Δf_n by Δx_1 Δf_n by Δx_2 . This is Δf_n by Δx_n . Multiply this by h_1, h_2, h_n . Now I write this particular thing; this is taken to the right hand side, therefore I will have this as minus f_1, f_2, f_n , all of them evaluated at the iterate k i.e. k^{th} iterate. This matrix if you denote this by J ; this by vector h , this is a matrix h ; this is the right hand side, let us call it has f ; f is evaluated at k . So I can put it as f evaluated at k . This J is called the Jacobian matrix of the system. So it has a name called the Jacobian matrix of the system. Now I think we better put a suffix here also, because Δf_i by Δx_1 all of them are evaluated at x_k . So these are all evaluated at x_k . Therefore these are all numbers, these are all constants. Therefore given at any iteration x_k , we have a constant coefficient matrix multiplied by the vector and the right hand side is known; therefore what this has produced us is a system of linear of algebraic equations. So this is a system of linear algebraic equations. Now therefore we need methods for solving the system of linear algebraic equations, which is a topic which we shall be taking in the later lectures as to how to solve a linear system of algebraic equations.

(Refer Slide Time: 00:41:07 min)



Let us now consider a particular case of a two into two system and show how we can obtain the solution; we shall then apply it to get the solution in an example. Now let us take simply a two by two system and see how we can solve a simple problem. Let us take a two by two system, take the system as $f(x, y) = 0; g(x, y) = 0$. For this what we will have here is Δf by Δx , Δf by Δy and the second row is Δg by Δx , Δg by Δy ; the vector is h_1 and h_2 ; right hand side is f_k, g_k ; f is evaluated at x_k, y_k and g is evaluated at x_k, y_k . What will be the next approximation? We should have mentioned there. The next approximation will be, as you can see x_{k+1} is equal to x_k plus h_1 ; x_{k+2} is x_k plus h_2 ; x_{k+n} is x_k plus h_n . So the remaining part is the same as the Newton's method. So this is really an extension of Newton's method to a system of nonlinear equations. So once we determine this h_1 and h_2 from here we can go back, evaluate these partial derivatives again with respect to k ; I can as well put this as a suffix with k . I will evaluate everything at the current iterate again, solve the system, get the increment and then again repeat the same thing. So that means for each iteration, we shall be solving one linear algebraic system of equations. The order of the system depends on the number of the equation that is given to us. Here we are solving two by two system; in the general case we are solving the n by n linear algebraic system where we have one particular system for each iteration.

(Refer Slide Time: 00:44:16 min)



Now let us take a simple example on this. We are given x square plus y square is 1.12. I have taken a simple example in which you will be able to eliminate it, but we will not eliminate it. Just to illustrate it we have a simple example; x square plus y square is 1.12, xy is 0.23. The solution of this problem is close to 1 and 0.2, so that means we are given initial approximation to the problem. Find the solution after one iteration in the Newton's method. So what I need to build up is the Jacobian's matrix of this first. So I build up the Jacobian. Now we will write this as $f(x, y)$ as the first equation, so we take everything to left hand side. Even when we are writing this, we will have to write this as f_1 of (x, y) is equal to zero or g of (x, y) is equal to zero. So these are the two equations; therefore I will bring this right hand side to the left hand side and define it as f . For the second equation I bring 0.23 to the left hand side, define it as $g(x, y)$, therefore $f(x, y)$ will be x square plus y square minus 1.2 and $g(x, y)$ is equal to (x, y) minus 0.23. Then the Jacobian matrix will differentiate this partially with respect to x ; differentiate this partially with respect to y ; then differentiate g also partially with respect to x and y . So I have here two x , I have here two y and the derivative of this will be y and x . Therefore the Jacobian matrix is this.

Now since it is two by two system what I have done here is, I have inverted it; because once I invert it, I do not have to repeat the solution of the system of linear equations. I just have to evaluate what it is. So this is a two by two system. So I will find the determinant of this i.e. two x square minus two y square that goes into the denominator. I will write down the co-factors for this. Co-factor for this is x ; co-factor for this two x ; co factor for this and its transpose also, I take it as minus two y and minus y here. So I get this as the inverse of this particular matrix J . Then once I write this as the inverse of the matrix, I can now write down my system of equations.

(Refer Slide Time: 00:45:21 min)

The image shows a handwritten derivation of the Newton-Raphson method for solving a system of nonlinear equations. The equations are written on a piece of paper. The first equation is $h_1 = -J^{-1}(F)_k$. Below it, the Jacobian matrix J is given as $J = \begin{bmatrix} x & -2y \\ -y & 2x \end{bmatrix}$ and the function vector F is $F = \begin{bmatrix} f \\ g \end{bmatrix}$. The next line shows the calculation of h_1 and h_2 for $x_0 = 1$ and $y_0 = 0.2$. The Jacobian matrix is evaluated at $(1, 0.2)$ as $J = \begin{bmatrix} 1 & -0.4 \\ -0.2 & 2 \end{bmatrix}$. The function values are $f(1, 0.2) = -0.08$ and $g(1, 0.2) = -0.03$. The correction vector h is calculated as $h = -J^{-1}F = \begin{bmatrix} 0.0354 \\ 0.0229 \end{bmatrix}$. The new values are $x_1 = x_0 + h_1 = 1 + 0.0354 = 1.0354$ and $y_1 = y_0 + h_2 = 0.2 + 0.0229 = 0.2229$. The final line states: "We can check the error in satisfying the given equations".

$$h_1 = -J^{-1}(F)_k$$
$$\begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = -\frac{1}{2(x^2 - y^2)} \begin{bmatrix} x & -2y \\ -y & 2x \end{bmatrix} \begin{bmatrix} f_k \\ g_k \end{bmatrix}$$

With $x_0 = 1, y_0 = 0.2$, we have

$$\begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = -\frac{1}{2(0.96)} \begin{bmatrix} 1 & -0.4 \\ -0.2 & 2 \end{bmatrix} \begin{bmatrix} -0.08 \\ -0.03 \end{bmatrix} = \begin{bmatrix} 0.0354 \\ 0.0229 \end{bmatrix}$$
$$x_1 = x_0 + h_1 = 1 + 0.0354 = 1.0354$$
$$y_1 = y_0 + h_2 = 0.2 + 0.0229 = 0.2229$$

We can check the error in satisfying the given equations

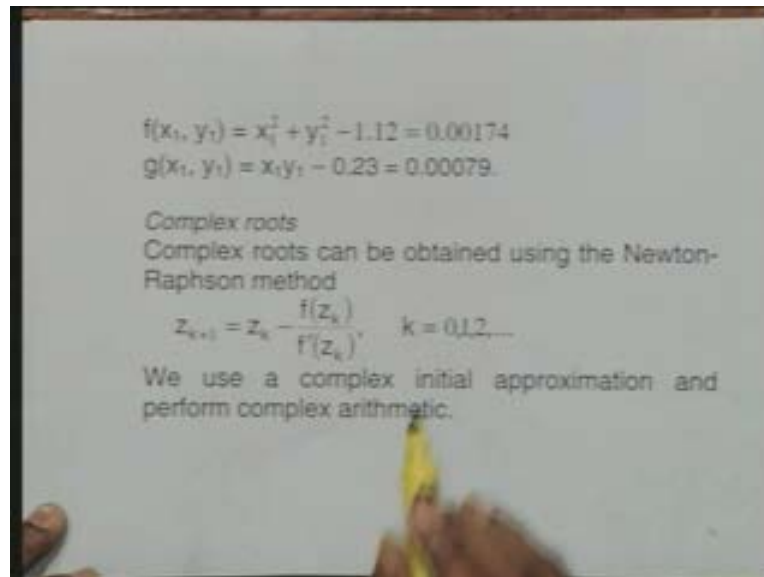
So h_1, h_2 are there. Now what we are writing here is the next step. We are writing this vector h is equal to J inverted of $(F)_k$. So this means h is equal to minus J inverted of $(F)_k$. Now we have h is equal to h_1, h_2 . We have just evaluated the J inverted. So I substitute it here; one upon two x minus x square minus y square and the Jacobian is this, we have just written it. Now I will substitute the initial values x_0 is one, y_0 is equal to two, then it is a simple matter of computation. So this is one divided by two times. I have computed what is x square minus y square; set the value of x as one; two x is equal to two; minus $2y$ minus 0.4 ; minus y is minus 0.2 . Now this is the evaluation of f at x_k, y_k ; that means I have evaluated $f(x, y)$ and $g(x, y)$ from this and written these values as minus 0.08 and minus 0.03 . Now just multiply it out and divide it out by this quantity and produce this as $0.0354, 0.0229$ and we have x_1 is equal to x_0 plus h_1 , y_1 is equal to 0.2 which is your initial approximation given to us plus 0.0229 that is 0.2229 . This is the solution of this particular problem. The initial approximation was given to us as 1 and 0.2 and now we have got the new solution as 1.20354 and this.

Now it is possible for us to check the error and satisfy the given equation, that means we can just put these values in the equation and see whether f is sufficiently close to zero or g is sufficiently close to zero. The values are f at x_1, y_1 is equal to x_1 square y_1 square minus 0.00174 and $g(x_1, y_1)$ is this. So when we started it initially the values of f and g were here; this is your f at x_k, y_k ; g at x_k, y_k . The value of minus 0.08 is now reduced to 0.00174 and 0.03 is reduced to 0.00079 . So that means we are talking of how the equation is satisfied. So we are now testing whether the equation has been satisfied sufficiently. Now we are able to get the accuracy faster; probably you need just another three or four iteration to get an accuracy of four or five decimals places. Therefore it is possible for us to solve a system of nonlinear algebraic equation also by simple extension of the Newton's method.

Now earlier we made a comment that of all the methods that are available the Newton Raphson method is the most commonly used even today for any scientific and engineering applications.

The reason we have given was, the reasonable order of the method and the computational cost. Now if you look at these nonlinear systems, you will find that the extension of secant and other methods will be much more complicated and much more difficult and therefore the Newton Raphson method is very valuable for the system of nonlinear equations also.

(Refer Slide Time: 00:49:07 min)



$$f(x_1, y_1) = x_1^2 + y_1^2 - 1.12 = 0.00174$$

$$g(x_1, y_1) = x_1 y_1 - 0.23 = 0.00079$$

Complex roots

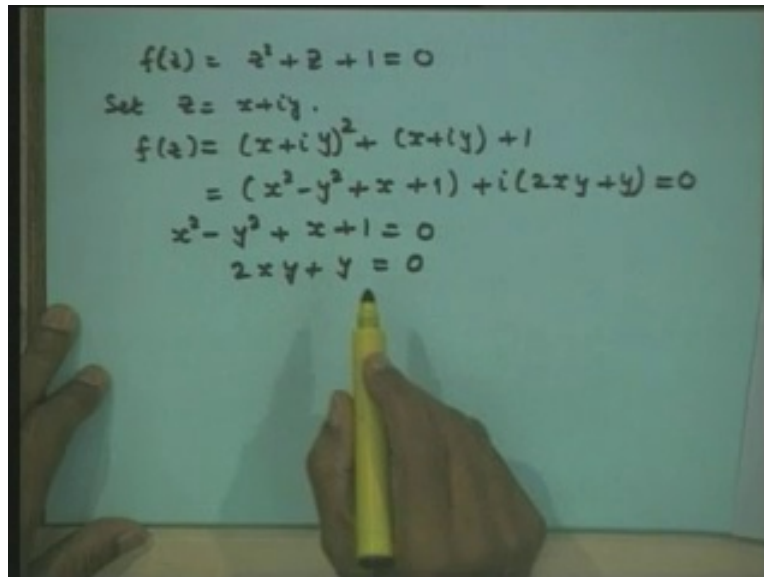
Complex roots can be obtained using the Newton-Raphson method

$$z_{k+1} = z_k - \frac{f(z_k)}{f'(z_k)}, \quad k = 0, 1, 2, \dots$$

We use a complex initial approximation and perform complex arithmetic.

We all along talked of the real roots; a simple root and a multiple root. Now how do you get the complex root; say you want to solve $x^2 + 1 = 0$, so there is no real roots; they are all complex roots, how do you get the complex roots? All the methods that we have done, all of them work; except that you have to do complex arithmetic. You have to start with the complex initial approximation and do the complex arithmetic. Every method that has been done so far can be used. Therefore that is what we have. For example, I can write down the Newton Raphson method as $z_{k+1} = z_k - f(z_k) / f'(z_k)$; k is equal to 0, 1, 2 and so on. Now the only thing is that, a real approximation cannot lead to the complex. So I need to have the complex initial approximation and I do the complex arithmetic. So you can define in advance the computation as complex and then go start with initial approximation and perform it; and the order of convergence retains as the same, whether it is a real root or complex root there is no difference. Therefore we do not need any special method for obtaining the complex root. However we can derive an alternative method for the solution. We shall illustrate this method through an example.

(Refer Slide Time: 00:49:59 min)


$$\begin{aligned}f(z) &= z^2 + z + 1 = 0 \\ \text{Set } z &= x + iy. \\ f(z) &= (x + iy)^2 + (x + iy) + 1 \\ &= (x^2 - y^2 + x + 1) + i(2xy + y) = 0 \\ x^2 - y^2 + x + 1 &= 0 \\ 2xy + y &= 0\end{aligned}$$

Let fz is equal to z square plus z plus one is equal to zero be the given equation. Set z is equal to x plus iy . Then we get f of z is equal to x plus iy whole square plus x plus iy plus one is equal to x square minus y square plus x plus one plus i into two xy plus y is equal to zero. Now setting the real part zero and imaginary part to zero, we get the two equations; x square minus y square plus x plus one is equal to zero two xy plus y is equal to zero. This gives a system of two nonlinear equations in two variables x and y . We solve this system by the Newton's method that we have derived earlier. However in this case the computation may take more time. We shall choose this procedure if we cannot perform the complex arithmetic and using complex initial approximation. Therefore the general technique in this particular case would be that substitute z is equal to x plus iy in fz is equal to zero; separate the real and imaginary parts, put them equal to zero, solve them as a system of two simultaneous nonlinear equations in two variables of x and y . That gives the alternate method of the Newton's method that we have discussed earlier. We shall stop it for today.