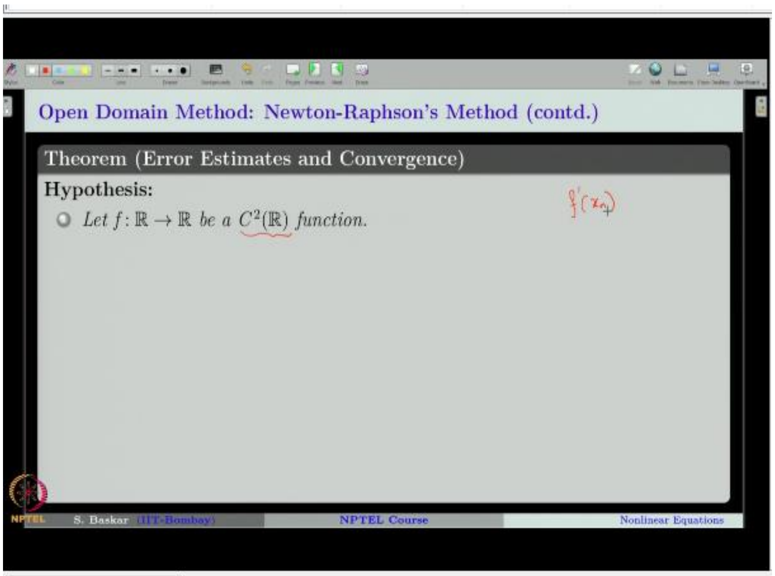
Numerical Analysis
Prof. Dr. S. Baskar
Department of Mathematics
Indian Institute of Technology-Bombay

Lecture-33
Nonlinear Equations: Newton-Raphson's Method (Convergence Theorem)

Hi, in the last lecture we have discussed Newton-Raphson's method for approximating an isolated route of a non-linear equation. We have also stated the convergence theorem of the method in the last class. In this lecture we will prove the convergence theorem. As a first reading you may skip this lecture. However, I advise you to carefully understand the proof of the convergence theorem of the Newton-Raphson's method.

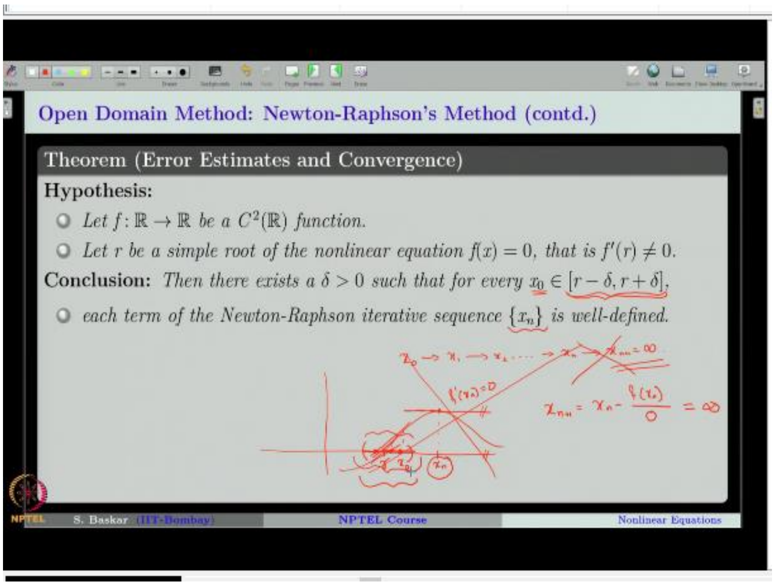
Because the mathematical tools used in this proof is very important for you especially if you want to train yourself as a numerical analyst from the research point of view. Let us get into the theorem. Let us first state the theorem which we have already stated in the last class.

(Refer Slide Time: 01:17)



Let f be as C^2 function, recall to set up the Newton-Raphson's iteration we just need $f'(x_n)$. That is we just need the first derivative of the function f, but for the convergence we need one more order of smoothness of f. That is why we have assumed that f is a C^2 function.

(Refer Slide Time: 01:44)



And also, we will assume that the root r which we are interested in, is a simple root. It means what? It means $f'(r) \neq 0$. If these two conditions hold for our function f then the theorem concludes that there exists a small neighbourhood of the root r. That is what we mean by saying that there existed δ such that $[r - \delta, r + \delta]$ is a small δ neighbourhood of r.

And if you start your iteration that is if your initial guess is taken in this small neighbourhood, remember we have to find this neighbourhood δ, there exist means we have to find this δ neighbourhood, such that if you start your iteration with your initial guess x_0 in this neighbourhood then everything will go nicely. That is what the theorem says. What is mean by everything will go nicely? Let us see.

First thing is each term of the Newton-Raphson's sequence is well defined. What is mean by this? What happens if you started with a x_0 then you went to x_1, then you went to x_2 like that you kept on going. At some x_n you see that your x_{n+1} is infinity then what happens? Then the Newton-Raphson's iteration sequence is not well defined. The theorem says that if you start your initial guess pretty close to the root then this will never happen. That is what it says.

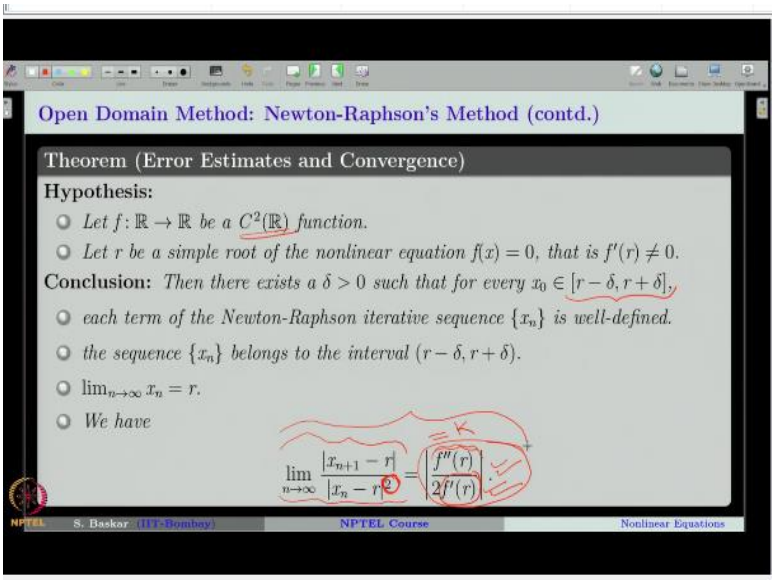
The question is when such situation will happen? Let us see suppose we are working with a function f whose graph is like this; you can see that it is a nice smooth function at least visually you can see that and suppose we are interested in capturing this root r. Now you started with some x_0 say and you started computing x_1, x_2 and so on. At some n say x_n is this then what happens?

Your $f(x_n)$ is this and now x_{n+1} is nothing but the point of intersection of the tangent line with x-axis. What is this tangent line? Tangent line is parallel to the x-axis and therefore it never intersects x-axis. That means you can never get x_{n+1} at all if you face such a situation. This is a geometrical illustration what is happening analytically? You see in this case you will have $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$.

What is $f'(x_n)$ here? That is equal to 0. So, this becomes 0 here and that will give you infinity. This is how the Newton-Raphson's iteration sequence may fail to exist. The theorem says that I can find a small neighbourhood of r as long as I start my initial guess within that neighbourhood then such a situation will never occur. That is what it says. Suppose you see if you restrict yourself to this neighbourhood.

And start your x_0 here then you see your x_1 will be somewhere here and your x_2 will be somewhere here and your x_3 will be somewhere here. Everything will lie in this interval only. That is all your x_n's will lie in this interval only and they all will be well defined. So, that is what we are going to see. So, as long as you start your iteration with the initial guess x naught in a small neighbourhood like this then your Newton-Raphson's iteration sequence will be well defined.

(Refer Slide Time: 06:00)

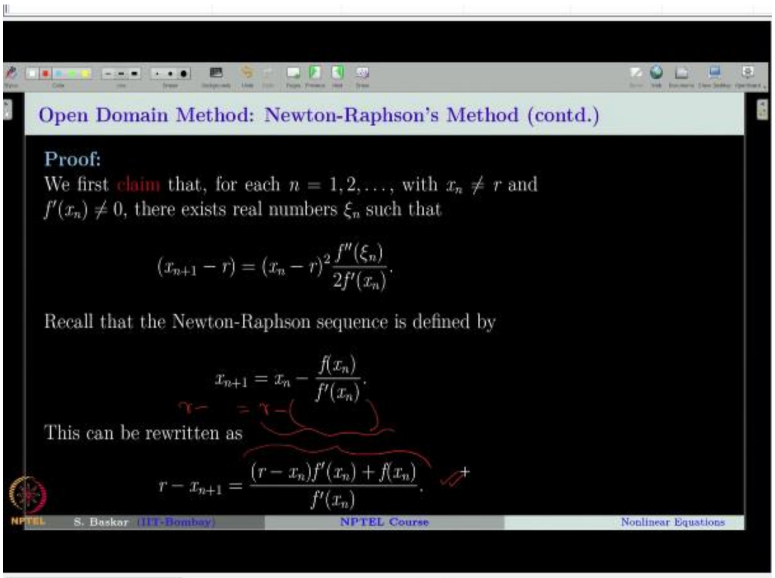


The sequence x_n will remain within this interval. This is what I have shown geometrically in the previous slide and the important point is that the sequence will also converge to the root and finally this gives us the quadratic convergence of the Newton-Raphson's method. It means what this quantity is equal to constant, some finite number. That is what we have to show in order to conclude that the sequence converges with this order.

You can see now here it is 2, therefore the order is 2. Why this term is finite because we assumed $f'(r)$ is not equal to 0 and f'' is a continuous function. That is what we assumed here that f is a C^2 function and we are always restricting to a closed and bounded interval. Therefore, this is a finite quantity, this is also non zero finite and therefore this entire thing will be finite.

So, if you prove this expression, it precisely means that the Newton-Raphson's iterative sequence converges quadratically. Let us see how to prove this theorem.

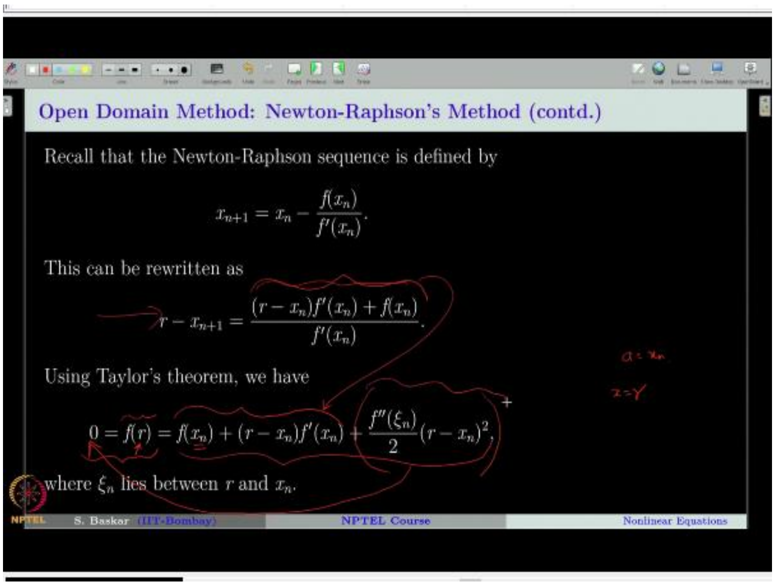
(Refer Slide Time: 07:23)



The proof of this theorem is much simpler than the proof of this secant method's convergence theorem, but the idea goes more or less the same. First, we have to establish this relation for some ξ which is very close to x_n and this we have to establish for each n. Let us see how to prove this expression, it is not very difficult. You start with the Newton-Raphson's method formula and now what you do is you take r - on both sides. So, you will take r - this equal to r - this term.

Now this right hand side can be rewritten like this, it is very easy to see that the right hand side with $r - x_n - \frac{f(x_n)}{f'(x_n)}$ can be rewritten like this. Now let us see how to deal with this expression.

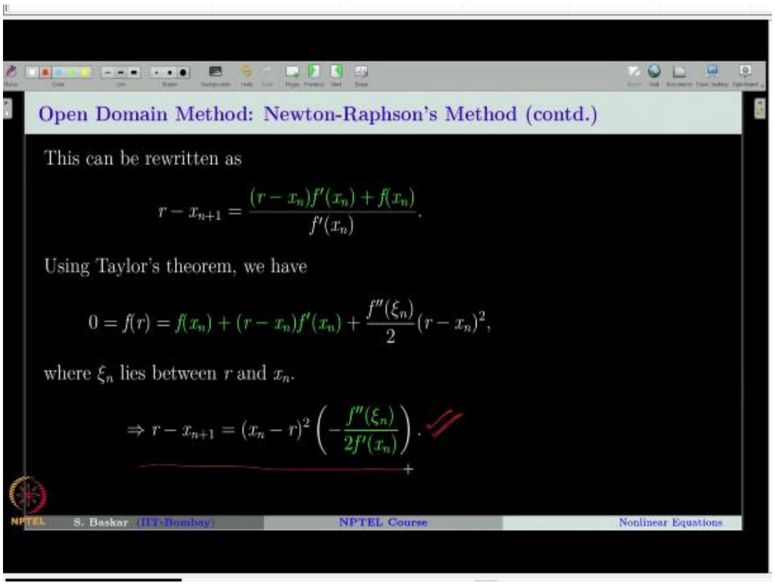
(Refer Slide Time: 08:32)



For that, first we will go to the Taylor's theorem and try to expand the Taylor's theorem around x_n at the point r. If you recall you are taking $a = x_n$ and you are writing the Taylor's formula for some $x = r$ here. What is r? r is precisely the root of the equation $f(x) = 0$. Therefore, the left hand side is 0, right hand side is precisely the Taylor formula and this is the remainder. Now you can see that this term that is the linear term is precisely what you have here.

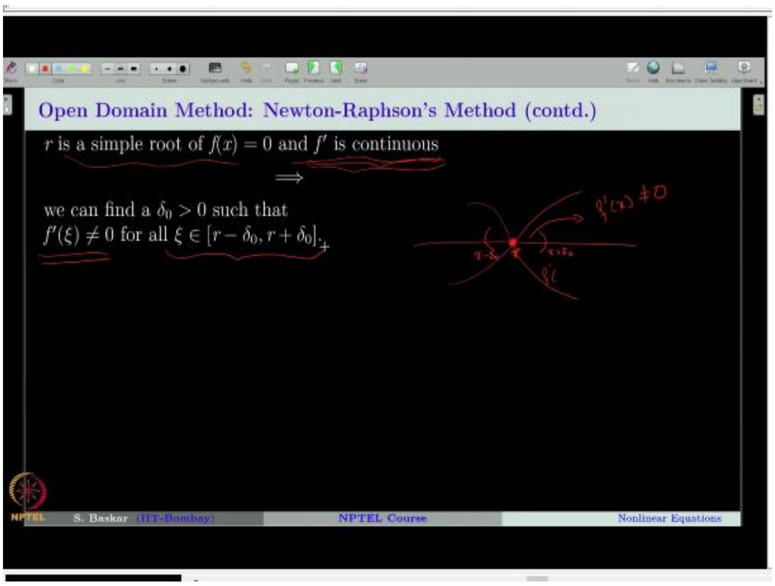
This is quite natural because in the Newton-Raphson’s method we precisely did the linearization and took that expression. That is why this and this are just coinciding. Now what I will do is this term is equal to minus of this term. So, I will take this remainder term to the other side and I can replace the numerator in this expression by minus of this. So, that is what I am going to do. I will replace this with minus of this one. I hope you understand this, where ξ_n lying between r and x_n, which is coming from the reminder term of the Taylor's theorem.

(Refer Slide Time: 10:01)



Once you do that, you get this expression. This is precisely what we wanted to show. Therefore, we have obtained our claim. That is we wanted to prove this and we have derived this expression right from the Newton-Raphson's method. Now we can use this expression to prove all our conclusions. Let us see how to do that.

(Refer Slide Time: 10:31)



First of all, if you recall we have assumed that the root r is a simple root and also if you recall we assume that f' is continuous. In fact, we assume that f'' is also a continuous function. Why we assumed first and foremost we want this Taylor formula to hold. For that we need f to be C^2. We already used that hypothesis that f is a C^2 function. Now we are only using a part of it.

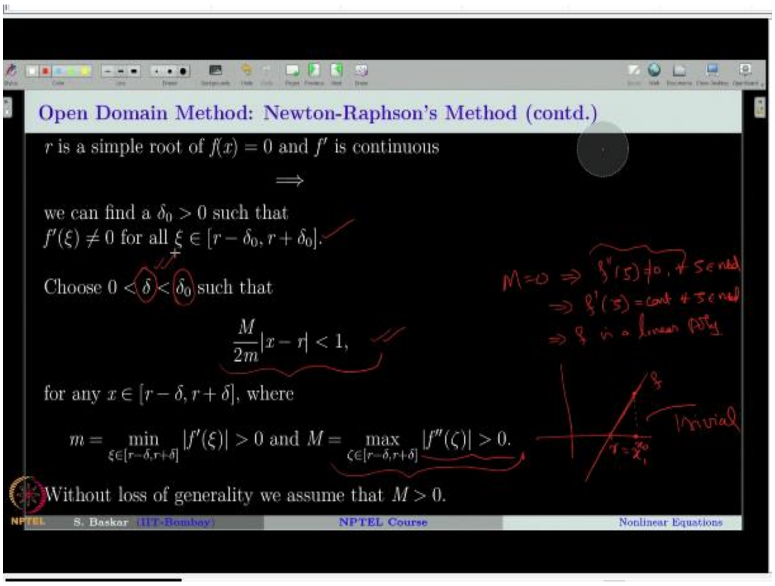
That is f is a continuous function. Therefore, we will combine these two results that is f' is not equal to 0 at the point $x = r$. It means what when f crosses the x-axis at r, $f'(r)$ is not equal to

0. Therefore, it will either cross like this or it may cross like this and so on. At the point r when it crosses it has some non-zero slope. That is what we mean by saying r is simple. Once $f'(r)$ is not equal to 0 then we can find a small neighbourhood of r in which f' remains non-zero.

So, in this interval $f'(x)$ is non-zero. How we get this? This is by the intermediate value theorem. So, intermediate value theorem says that if the continuous function f' is non-zero at r then you can find a small neighbourhood say δ_0 neighbour. That is $[r - \delta_0, r + \delta_0]$, such that f' will remain non zero for all ξ in that neighbourhood.

That is what the intermediate value theorem says. Remember we have to capture a δ neighbourhood, but this is not the neighbourhood that we want.

(Refer Slide Time: 12:32)



We want further a small value δ. From where we pick up, we should choose this δ which is less than δ_0 such that this quantity is less than 1. What is this? M is the maximum of f'' and m is the minimum of f'. So, minimum of f' is always greater than 0, because we are restricting ourselves to this neighbourhood where f' is not equal to 0.

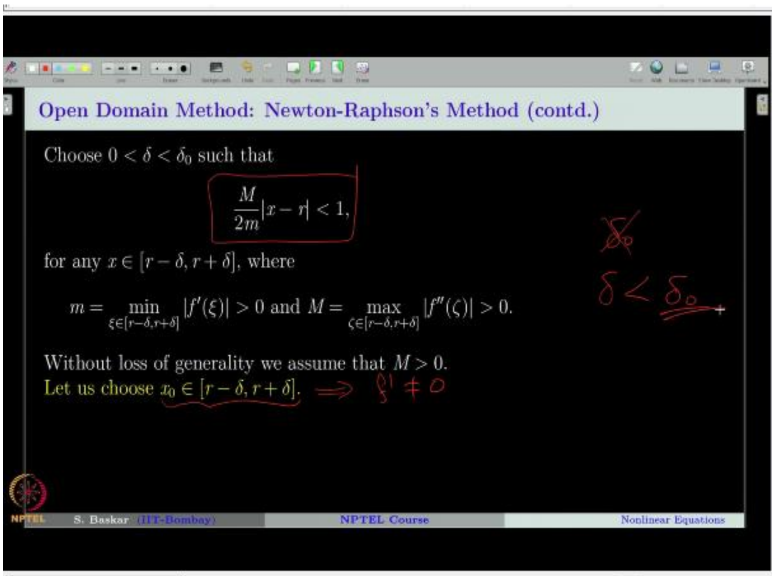
Therefore, this will always be positive, why should f'' and its maximum should be positive? That can be assumed without loss of generality. This may not be true but we can assume this without losing any generality, why just if $m = 0$, it means $f''(\zeta)$ is 0, for all ζ in the neighbourhood. That is this δ neighbourhood. It means what? It means $f'(\zeta)$ is constant for all ζ in this neighbour.

It means what? f is a linear polynomial; it means its f graph is a straight line. So, what it means if you take any x_0 and apply the Newton-Raphson's method. If you recall Newton-Raphson's method will take the tangent line at the point x_0 and see the point of intersection of that tangent line with the x-axis. In that sense you can see that if f itself a straight line, then the Newton-Raphson's method will capture the exact root at the first iteration itself.

Therefore, there is no question of convergence and there is no question of the existence of the Newton-Raphson's iteration or everything become trivial. Therefore, we will not consider the case that $f''(\zeta) = 0$. It means, we will assume that $f''(\zeta)$ is not equal to 0 therefore without loss of generality you can assume that M is greater than 0 and we choose our δ which is less than δ_0 you should also think and observe that such a choice is possible.

I leave it to you to see why such a choice is possible therefore you have to choose like this and the corresponding δ you take.

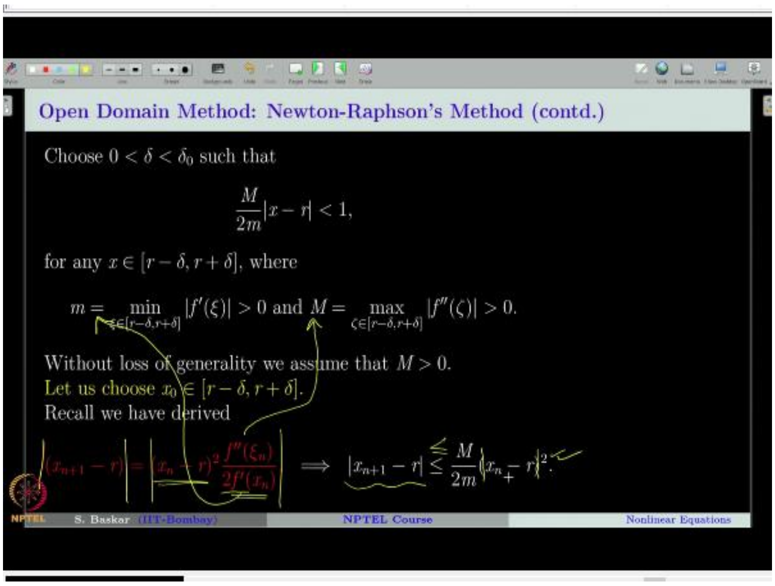
(Refer Slide Time: 15:26)



Now let us see that this δ will work very nicely for us, how? Let us see that. Choose x_0 in that smaller δ neighbourhood not the δ_0, δ_0 we are not going to take, we are going to take the δ which is smaller than δ_0, such that this holds and this will work very nicely for us. Let us check that you choose x_0 in this neighbourhood.

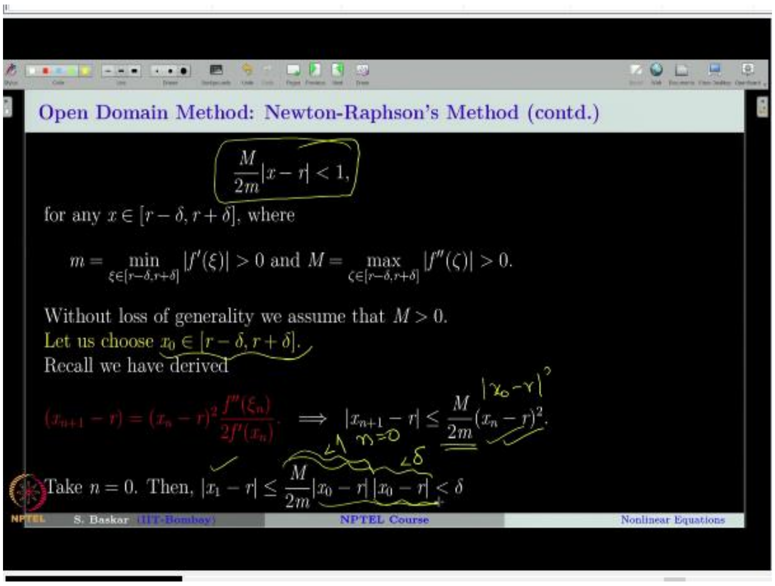
Remember in this neighbourhood f' is not equal to 0. That is something which is already there, because it is there for the bigger neighbourhoods. Therefore, it will be there for this smaller neighbourhood also.

(Refer Slide Time: 16:07)



Now let us take this formula which we have derived as the first claim of our proof and we will use this, how, take modulus on both sides, you can see that this term is less than or equal to now what I am doing is I am taking modulus and putting maximum for this, that is this one and minimum for this. Therefore, this will be less than or equal to M divided by $2m$ into this I am keeping as it is. So, I am not taking modulus here because I am squaring it. So, we got this inequality.

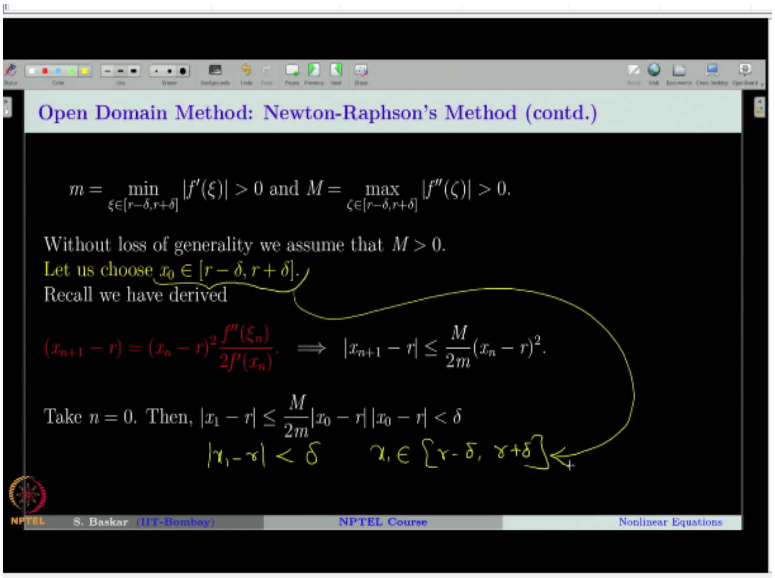
(Refer Slide Time: 16:53)



Let us see what we can do with this inequality. In this inequality you take $n = 0$. Now what you will get? The first left hand side will be $|x_1 - r|$ and that is less than equal to $\frac{M}{2m}|x_0 - r|^2$, that is what I am taking like this and now this is less than δ and this is less than 1 why that is how we have chosen our δ.

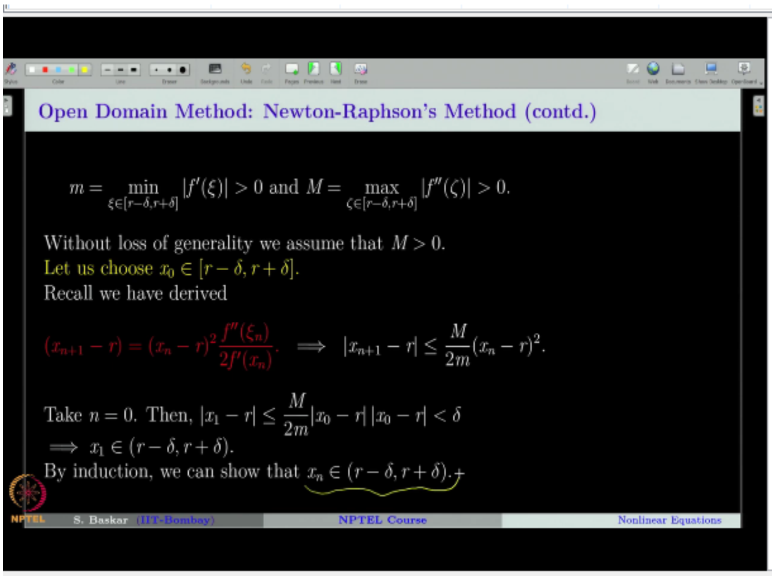
Therefore, as long as x_0 belongs to this interval this first term will be less than 1 and the second term obviously will be less than δ, therefore the both together will be less than δ. What it means?

(Refer Slide Time: 17:44)



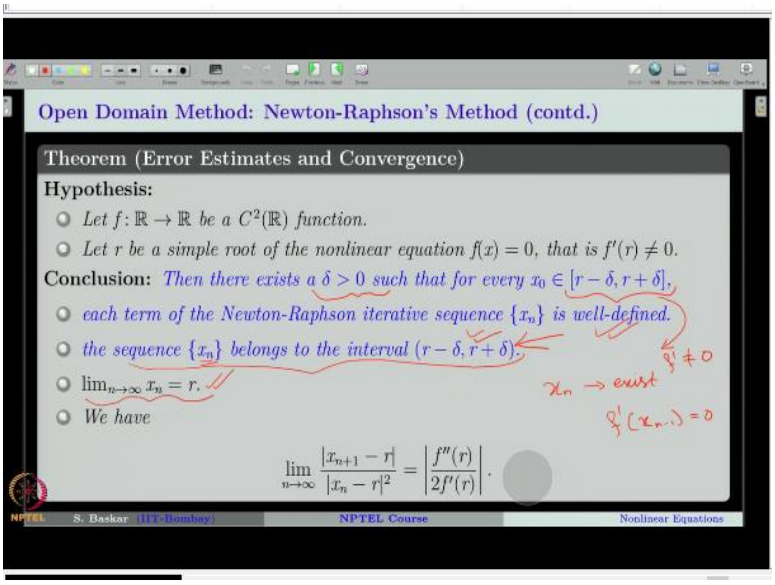
It means that you see $|x_1 - r| < \delta$ means x_1 belongs to $[r - \delta, r + \delta]$. So, it means if you start your x_0 from this neighbourhood, we have shown that x_1 will also belong to that neighbourhood. So, that is what we have shown.

(Refer Slide Time: 18:09)



Now you can use an induction argument to show that all x_n's belongs to this neighbourhood. That is very easy for you to show.

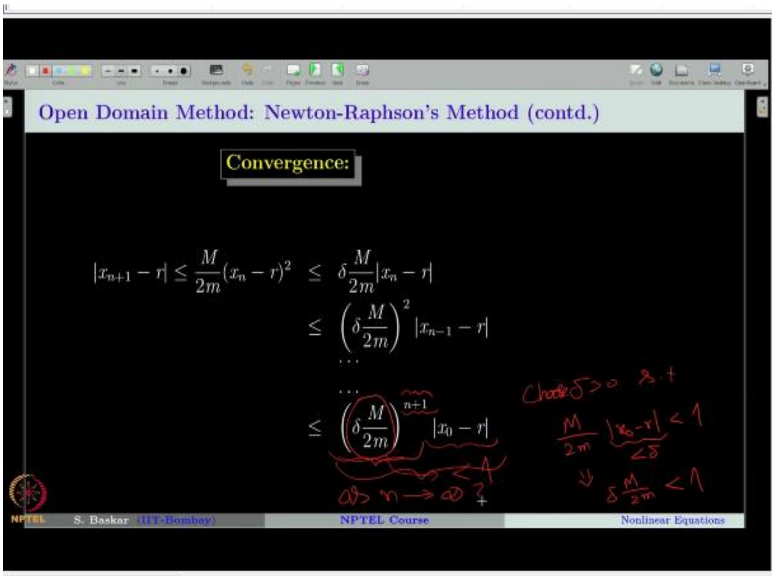
(Refer Slide Time: 18:21)



So, what we have shown? We have obtained a δ neighbourhood. Remember how we obtain? We obtained this δ such that this inequality holds. That is always the demand. So, once you choose such a δ, we have seen that all the x_n's belong to this neighbourhood. That is what we have seen and also you can see that as long as x_n belongs to this neighbourhood, all x_n will exist why because the only way they do not exist is when $f'(x_{n-1}) = 0$.

But that can never happen because in this interval, in this neighbourhood basically, we have chosen this δ such that f' is not equal to 0. Therefore all the x_n's will be well defined. Therefore, we have also shown this property. Let us now prove the convergence.

(Refer Slide Time: 19:28)

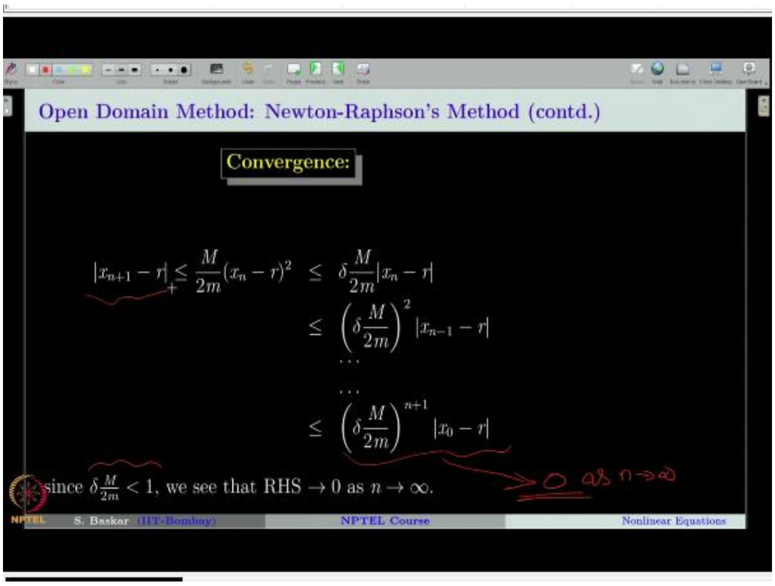


It is not very difficult. Let us start with the inequality that we have already derived in our previous slide. First thing is you can see that $|x_n - r| < \delta$, because x_n belongs to the δ neighbourhood. Therefore, one term in this square I will take and put it in fact this can be put as strictly less than δ and remaining $\frac{M}{2m}|x_n - r|$ is kept as it is and now let us go to apply the same inequality for $|x_n - r|$.

See this kind of recursive idea is used in many places. So, once you use this inequality for $|x_n - r|$ you will get $\left(\delta\frac{M}{2m}\right)^2 |x_{n-1} - r|$. Again, you use this inequality for $|x_{n-1} - r|$ that gives you less than or equal to it is strictly less than $\left(\delta\frac{M}{2m}\right)^3 |x_{n-2} - r|$, like that you can keep on going till what, you will go till you hit $|x_0 - r|$. At that level you will have this constant as $\left(\delta\frac{M}{2m}\right)^{n+1}$.

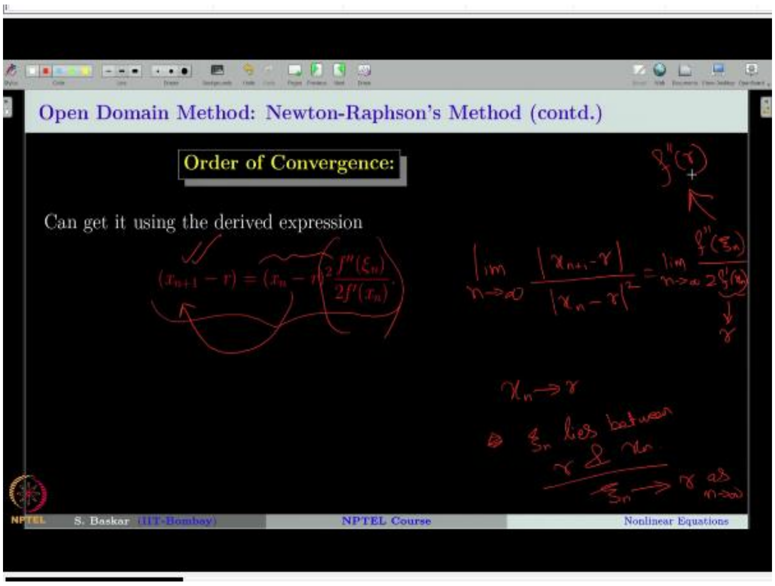
Now you see what is this constant? You can show that this constant is less than 1, why it is so? We have chosen our δ such that $\frac{M}{2m}|x_0 - r| < 1$. What is $|x_0 - r|$ that is less than δ? Therefore, this implies $\delta\frac{M}{2m}$ is actually less than 1. Now you see you have this constant which is less than 1 and that constant is given with power *n* + 1. Now what can you say about this right hand side as *n* tends to infinity that is the question.

(Refer Slide Time: 21:57)



As *n* tends to infinity this term goes to 0, because this constant is less than 1. That is precisely what we want to show because once this side goes to 0 then left hand side will also tends to 0 and that is precisely the convergence of the sequence x_n. So, we have proved the convergence also. Now, all remains is to prove this order of convergence. That is not very difficult.

(Refer Slide Time: 22:35)



Again, you start with this expression which we have derived as the first step of our proof. Now you can see you want to show $lim_{n\to\infty}|x_{n+1} - r|$, you see that is there already with you divided by $|x_n - r|^2$. That is already there so you bring it to the left hand side and that will be equal to $lim_{n\to\infty}\frac{f''(\xi_n)}{2f'(\xi_n)}$.

That is what is given here. Now you see $x_n \to r$, we have just now proved the convergence. Therefore, this is converging to r. Similarly, ξ_n lies between r and x_n. Therefore, you use the Sandwich theorem to conclude that ξ_n also converges to r as n tends to infinity because x_n is converging to r and ξ_n is lying between r and x_n. Therefore, ξ_n will also converge.

Therefore, your numerator will converge to $f''(r)$. This is what precisely we wanted to show in the order of convergence and that is also achieved. By this we have proved the convergence theorem of Newton-Raphson's method. The Newton-Raphson's method converges if you start your initial guess pretty close to your proof. That is what the theorem says and the important takeaway of this theorem is that if the sequence converges then it converges quadratically. With this we will end this lecture. Thank you for your attention.