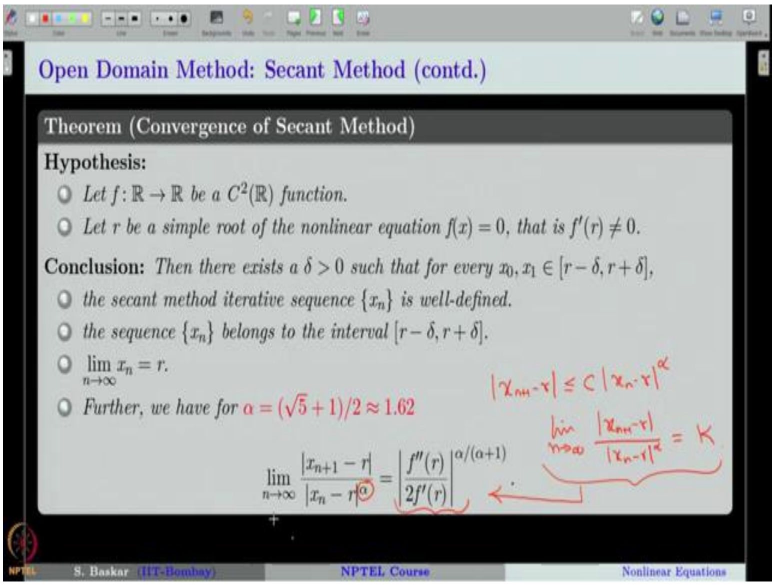
Numerical Analysis
Prof. S. Baskar
Department of Mathematics
Indian Institute of Technology – Bombay

Lecture – 31
Nonlinear Equations: Secant Method (Convergence Theorem)

Hi, we are discussing Secant Method for capturing isolated roots of a non-linear equation.

(Refer Slide Time: 00:33)



In the last class we have introduced the method and in this class we will study the convergence theorem of the secant method. We have two assumptions on the convergence of the secant method. One is that the function f which defines our equation, is a C^2 function. What is mean by a C^2 function? Well, f should be a continuous function, f' should exist and f' should also be a continuous function.

And f'' should exist and f'' should also be a continuous function. This is what we meant by saying f is a C^2 function and also we have to assume that the root that we are going to capture denoted by r is a simple root of the non-linear equation $f(x) = 0$, what it means? It means $f'(r) \neq 0$. That is what we mean by saying that the root r is a simple root.

If these two assumptions are satisfied by the function f, then we can say that there exists δ, such that for every x_0 and x_1 in the neighbourhood of r, that is the δ neighbourhood of r. We can say that the secant method iterative sequence is well defined and the second thing is that we can say that the sequence belongs to the interval $[r-\delta, r+\delta]$. That is, the δ neighbourhood from where we have taken our initial guesses.

And the third one which is what we are mostly interested in, is that the sequence x_n will converge to r as n tends to infinity. And finally, the conclusion is also that the order of convergence of the secant method is something approximately 1.62. That is what we meant by saying that $\lim\limits_{n\to\infty} \frac{|x_{n+1}-r|}{|x_n-r|^{\alpha}}$ is equal to some constant. Remember, $f'(r) \neq 0$.

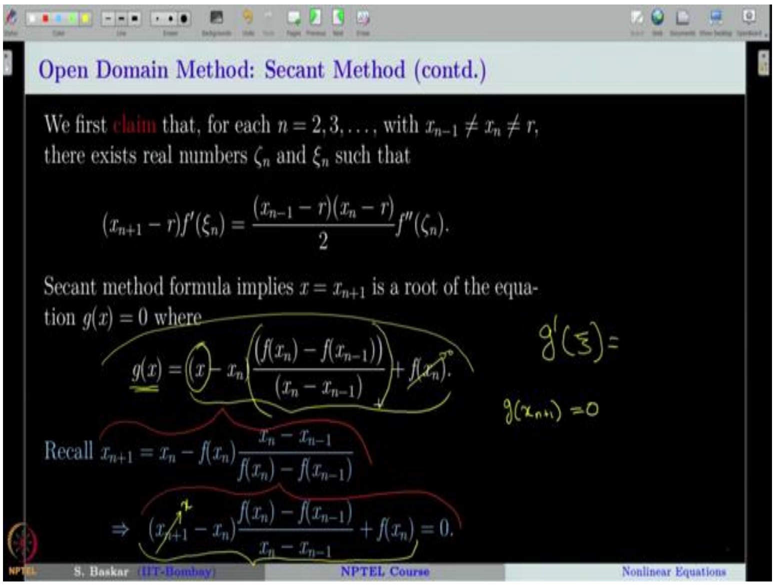
Therefore, this is some constant, non-negative constant. So, if you recall what is mean by the order of convergence, we say that a sequence converges with order at least α, if there exist a constant such that $|x_{n+1}-r|$ is less than or equal to the constant time $|x_n-r|^{\alpha}$. Another way of defining order is $\lim\limits_{n\to\infty} \frac{|x_{n+1}-r|}{|x_n-r|^{\alpha}}$ is equal to some constant.

So, this is what the condition that we are putting here and that says that the secant method is going to converge with order α and in fact, in our derivation. We will also get the value of α as $(\sqrt{5}+1)/2$ which is approximately equal to 1.62. That shows that the secant method has a super linear convergence. In fact, bisection method has a linear convergence. In that way, you can see that secant method is slightly faster than bisection method.

Of course, bisection method has its own disadvantage, also that it is a bracketing method, whereas secant method is an open domain method. Therefore, the initial guesses can be chosen arbitrarily in secant method. This is a very good advantage of secant method. In addition to that it is also nice to see that secant method has super linear convergence. Let us try to prove this theorem if you are not very good at calculus, maybe in the first go you can omit this proof.

However, I will strongly recommend you to go through the proof of this theorem very carefully. Because the techniques that are used in proving this theorem is very important for you to do any convergence analysis on an iterative method. Especially iterative method for non-linear equations. Therefore, it is very important for you to understand the proof of this theorem, at least once.

(Refer Slide Time: 05:38)



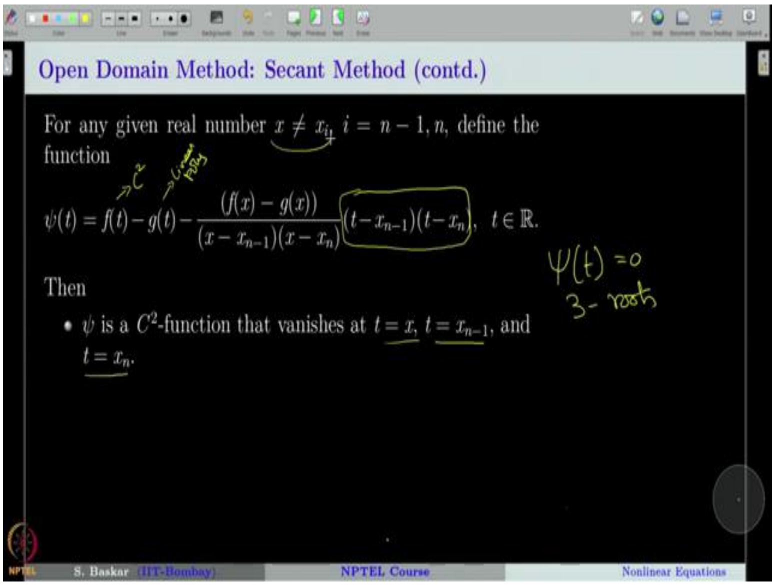
With this note, let us go into the proof of this theorem. As a first step, let us claim that for every n = 2, 3 and so on, we can find a ζ_n and a ξ_n such that this expression holds. That is $(x_{n+1} - r)f'(\xi_n) = \frac{(x_{n-1}-r)(x_n-r)}{2} f''(\zeta_n)$. This expression will be used in arriving at all the conclusions that we have listed in the statement of the theorem.

Therefore, first deriving this expression is very important for us. Let us see how to derive this, note that to derive this, we have to also assume that x_n is not equal to x_{n-1} and both are not equal to r. Now, let us start our derivation with the formula of the secant method, recall that the secant method is given by this expression. And this expression can be rewritten like this you can easily see this.

Now, what I will do is, in this expression I will replace x_{n+1} by x and then I will define this expression that is the left-hand side expression as a function denoted by $g(x)$. Now, you can see that if you plug in $x = x_{n+1}$ then clearly this is equal to 0 because that is the way we have defined the function g. Therefore, g has a root which is x_{n+1}. So, let us keep this in mind and go ahead.

Also see that g is a linear polynomial so, it depends on x with degree 1. So, these are the two observation that is g is a linear polynomial and g has the root $x = x_{n+1}$.

(Refer Slide Time: 08:04)

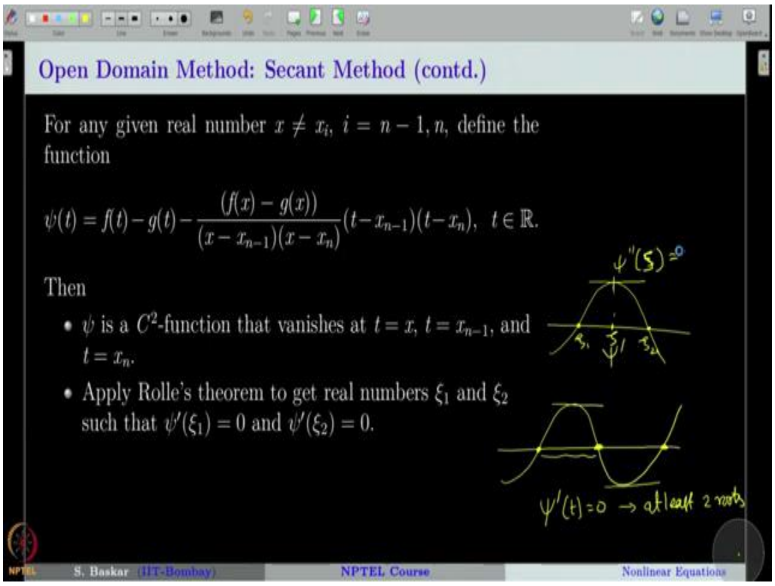


With these two properties in mind, let us go to define a function $\psi = f(t) - g(t)$. Remember ψ is a function of t where we have fixed x_{n-1} and that is not equal to x_n, Of course, both of these are not equal to r and now we are also fixing x in such a way that x is not equal to any of these two numbers. And then defining ψ as a function of t, therefore, x is fixed as far as this expression is concerned that you keep in mind.

And how $\psi(t)$ is defined? $\psi(t)$ is defined with this expression. And now let us see some important properties of the function ψ. First thing is ψ is a C^2 function. Why it is so? Because f is a C^2 function this is what we have assumed in our theorem and g is a linear polynomial. And also you can see that this third term involves a quadratic term. Therefore, ψ is a C^2 function very clearly.

Also you can check that $\psi(t) = 0$. That is this equation $\psi(t) = 0$ has three distinct roots. What are they? One is $t = x$ is a root of this equation, $t = x_{n-1}$ is the root of this equation and $t = x_n$ is a root of this equation. They all are distinct because we have chosen $x_0 = x_{n-1}$ and that is not equal to x_n.

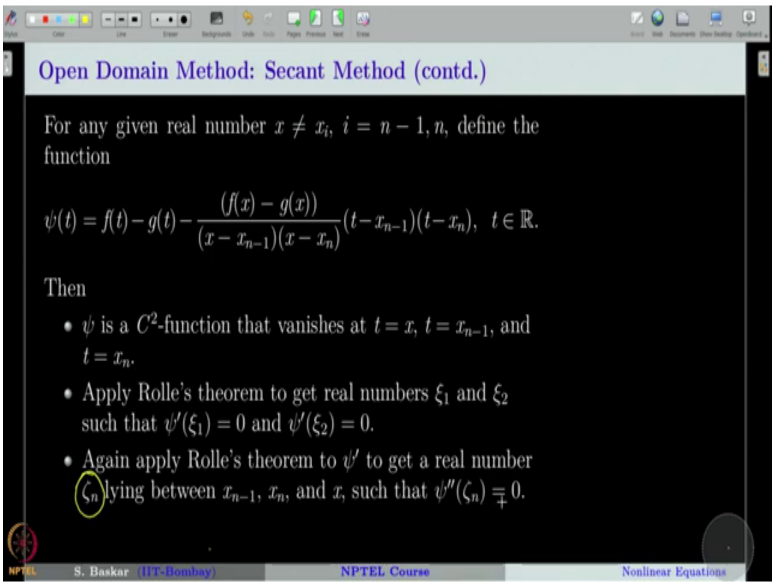
(Refer Slide Time: 09:59)



Now, apply Rolle's Theorem. What Rolle's Theorem says? If you have a function which has three roots something like this then it is derivative that is $\psi'(t) = 0$ will have at least two roots, distinct roots that is what Rolle's theorem says. Because you have three distinct roots for ψ. Therefore, in between these two if you apply Rolle's theorem, you will see that there exist at least one point at which $\psi'(t) = 0$.

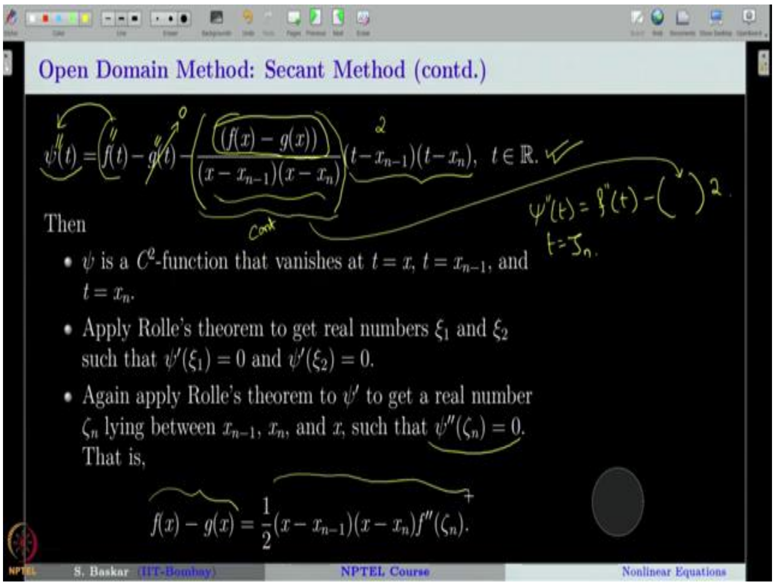
And similarly, in between these two points, you can find at least one point at which ψ' vanishes. So, Rolle's Theorem says that there exists at least two points at which ψ' vanishes. Now, ψ' is something like this at least two points ξ_1 and ξ_2 at which ψ' vanishes. Now, again apply Rolle's theorem on ψ' to see that there exists at least one point say η at which ψ''vanishes. That is what the rules theorem, when applied to ψ', will give us.

(Refer Slide Time: 11:23)



So, again applying Rolle's Theorem on ψ'. We can get a ζ_n such that $\psi''(\zeta_n) = 0$. That is what the Rolle's Theorem says.

(Refer Slide Time: 11:48)

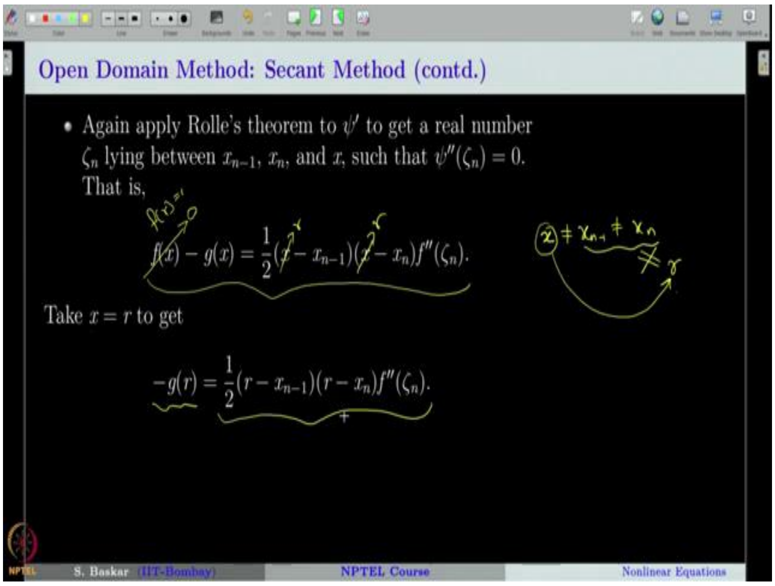


Now, we got ζ_n now, what are we going to do with this? Let us see well, we got the ζ_n let us compute ψ'' from it is expression and plug in ζ_n and see what is happening. You can see that ψ'' is equal to f'', and g'' that is equal to 0 because g is a linear polynomial. Therefore, second derivative of g is 0.

And remember this is a constant because *x* is not a variable for us as far as this expression is concerned. Therefore, this is a constant, so, do not differentiate it because we are differentiating with respect to *t* not with respect to *x*. And what will happen to this term? This term will simply give you 2, if you differentiate it twice. Therefore, we have $\psi''(t)$ is equal to $f''(t)$ minus this expression times 2.

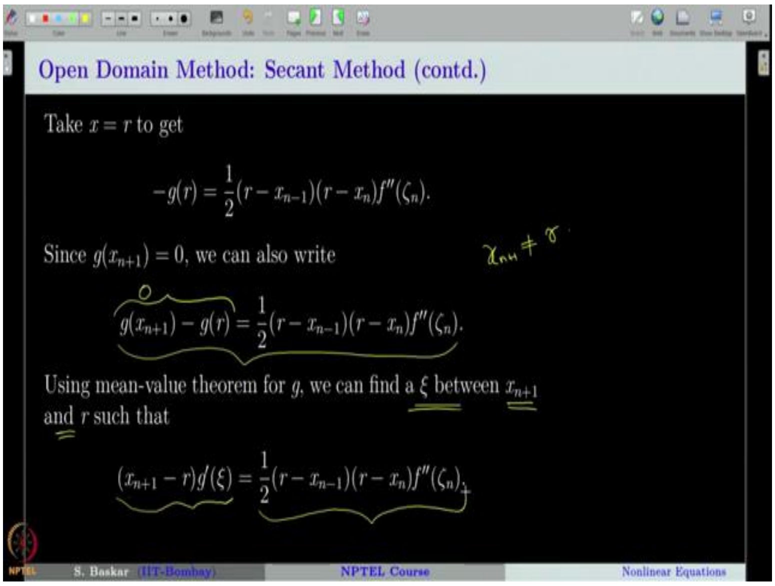
Now from there what you can write is $f(x) - g(x)$ that is, this term, is equal to 1/2. 2 is going to the other side into this term is also going to the other side after shifting this, to the left-hand side. Therefore, we will get $f(x) - g(x) = \frac{1}{2}(x - x_{n-1})(x - x_n)f''(\zeta_n)$. What I am doing? I am just plugging in $t = \zeta_n$ and then doing this simple manipulation to get this expression.

(Refer Slide Time: 13:52)



Now, let us see how to go ahead. I will put $x = r$ in this expression. Remember I have chosen $x \neq x_{n-1} \neq x_n$. Already, these two are chosen such that this is not equal to r. Therefore, I may choose my $x = r$ that is not a problem for me. Now, when I choose $x = r$ in this expression, you can see that this will go to 0 because $f(r) = 0$. So, the first term will go off and you will have – $g(r)$ = 1/2 into instead of x, I am putting r here that is what I got here.

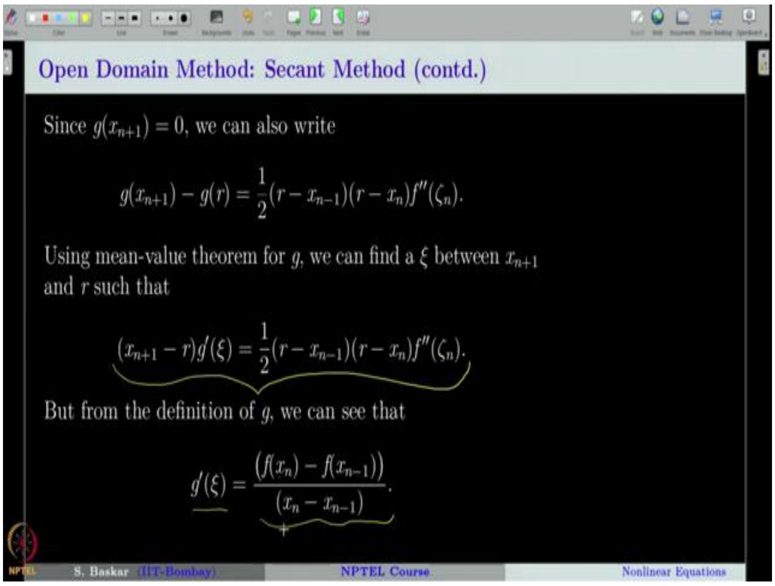
(Refer Slide Time: 14:52)



And now, let us see how it goes. If you recall, we have defined our g such that $g(x_{n+1})$ = 0. This is how we have defined g in our first step. Therefore, I can as well, write this expression as $g(x_{n+1})$– this equal to this because this is just going to contribute 0 here. Therefore, no problem, I can write like this. And now what happens? I got this expression where g is 0 keep in mind.

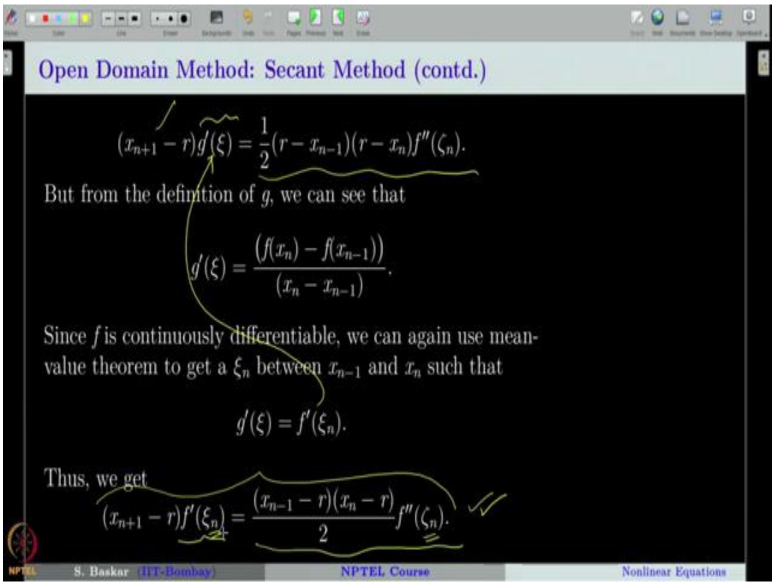
Now, what I will do is, I will put mean value theorem for this expression because x_{n+1} is not equal to r. That is our assumption as well. Therefore, you can put the mean value theorem and get $(x_{n+1} - r)g'(\xi)$. There exists ξ between x_{n+1} and r such that the left-hand side can be written like this. That is what the mean value theorem and that is equal to I am just keeping the right-hand side as it is.

(Refer Slide Time: 16:06)



Well, we got this expression now. Now, we will also see that $g'(\xi)$ is equal to this. How will you get that? Will you directly differentiate the expression g that we have taken in the first slide? That is this one, differentiate with respect to x and put $x = \xi$ here. You will see that this is 0 because this is a constant. So, when you differentiate with respect to x this will be 0 and you will be left out with this term only. That is what I am writing $g'(\xi) = \frac{(f(x_n) - f(x_{n-1}))}{(x_n - x_{n-1})}$.

(Refer Slide Time: 17:02)

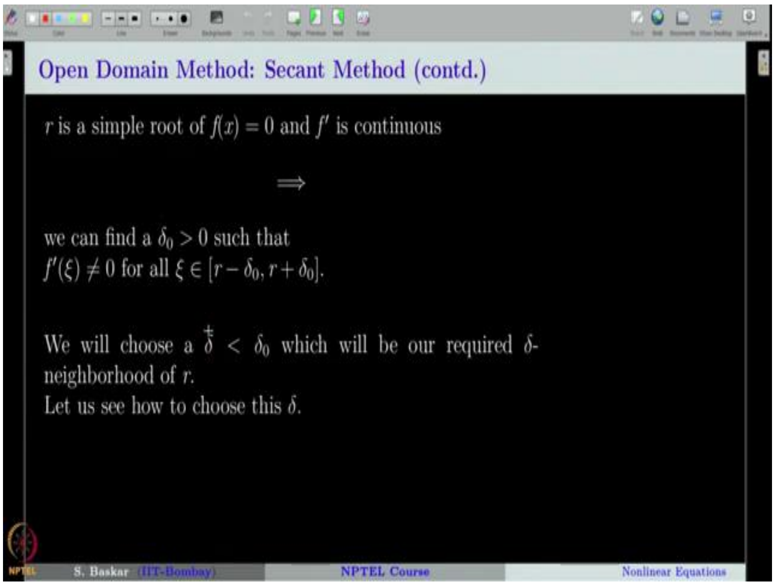


Now, you apply the mean value theorem. For this term you can see that there exist a ξ_n between x_n and x_{n-1}. Such that this is equal to $f'(\xi_n)(x_n - x_{n-1})$. This numerator equal to this therefore, $x_n - x_{n-1}$ will get cancelled and you will have $g'(\xi) = f'(\xi_n)$. Now, you put this term into the expression that you got here. You will see that x_{n+1} that is this term.

Now, instead of $g(\xi)$, you are putting $f'(\xi_n)$ and the right-hand side is kept as it is. So, this is what we wanted to derive, so, our first part of the proof is over. We have derived an expression remember. This expression is basically coming from the secant method formula, how? We have taken the secant method formula from there we have defined g and through that we have derived this expression.

Therefore, the secant method formula gives us a sequence. That sequence will surely satisfy this expression for some ζ_n and ξ_n that is what we have seen.

(Refer Slide Time: 18:42)



Now remember r is a simple root of our equation and f' is continuous. It means what you have a function f say its graph, is like this and r is the root of the equation $f(x) = 0$. This is the graph of the function $f(x)$ and we know that $f'(r) \neq 0$. Since, f' is continuous by intermediate value theorem you can see that in a small neighbourhood of r, f' will remain non-zero.

That is what the intermediate value theorem tells us that is what we are writing here. You can find a small neighbourhood of r say $[r - \delta_0, r + \delta_0]$ in which the function f' will remain non-zero. Remember in the statement of our theorem, we have to find a δ neighbourhood of r.

You remember we have to find a δ neighbourhood of r in which if you start your iteration, that is, if you choose your x_0 and x_1 in that δ neighbourhood then all these conclusions will hold. Therefore, our aim is to find such a δ. That is what we are trying to do now, so, the intermediate value theorem and the assumption that r is a simple root, will give us a δ. But remember this is not the δ that we want. We want something which is lesser than this.

So, we have to choose some other δ suitably which is less than this δ, such that all our conclusion should hold. Now, the question is how to choose this δ. That is the question let us see how are we going to get this δ?

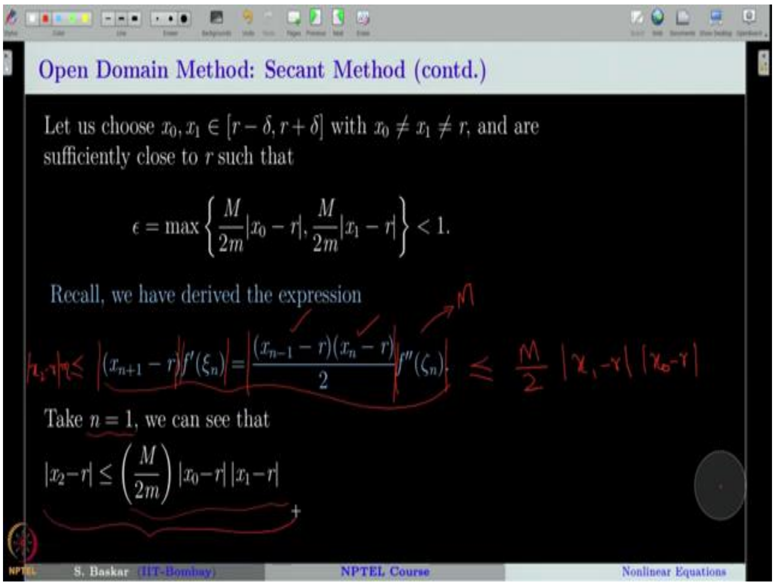
(Refer Slide Time: 21:04)

Let us have some notation. Here. I am just assuming that I got a δ and then I am just defining the notation that is all. This is not something which I am deriving or something it is just a notation I am fixing m means the minimum of f' over the δ neighbourhood that I am supposed to get now. So, δ is yet to be found but whenever there is a m it means I am just taking the minimum of f' in that neighbourhood.

And similarly, M is the maximum of f''. Now, let us see how are we going to choose our δ? This is where precisely I am going to choose my δ. I will choose my δ in such a way that whenever I choose x_0 and x_1 in that neighbourhood such that x_0 and x_1 are not equal to r then the maximum of this term should be less than 1. So, this is how I am going to choose my δ.

Now, if I choose my δ in such a way then what happens? Let us see. Recall we have already derived this expression just as the first part of our proof.

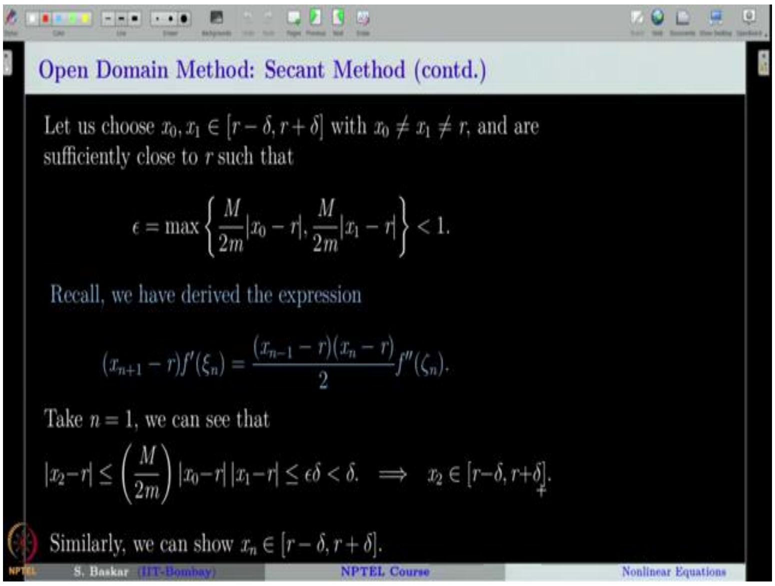
(Refer Slide Time: 22:36)



Now I am going to use this expression. How am I going to use this expression? Well, I will take n = 1 in this expression. Then you can see that $|x_2|$ because I took n = 1 and I took the modulus. And that is actually, going to be less than or equal to this term. Why it is so? You can see that this is less than or equal to $||x_2 - r||$ into minimum of f'. What is minimum of f'? Minimum of f' is denoted by m here and.

Similarly, I will take the modulus on the right hand side and that is going to be less than or equal to I will replace this by the maximum. Therefore, I will have $\frac{M}{2}|x_1 - r|$ that is, this term and $|x_0 - r|$ that is this term because I have taken n = 1. Therefore, we have this inequality by taking n = 1.

(Refer Slide Time: 23:56)

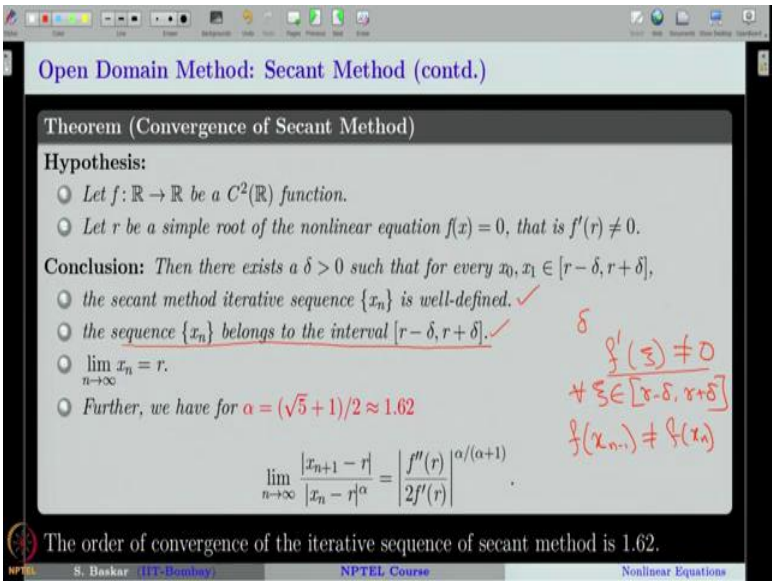


And now what happens? Let us see this term is less than ϵ because ϵ is taken as the maximum of these two. Therefore, this term is something less than ϵ and the other term is surely less than or equal to δ why? Because x_1 is chosen from the δ neighbourhood. Therefore, $|x_1 - r| \leq \delta$. Therefore, this whole quantity is less than or equal to $\epsilon\,\delta$ and remember, ϵ is less than 1.

That is the way I have chosen my δ. So, therefore, with that δ, my ϵ is less than 1. Therefore, this is less than δ and now what does this mean? This means that x_2 belongs to the δ neighbourhood of r. That is what it means by saying that $|x_2 - r| \leq \delta$ or may be less than δ means it is an open, interval. That is not important here x_2 lies in this δ neighbourhood that is more important.

And similarly, from here you can show that x_n also belongs to this δ neighbourhood for each n.

(Refer Slide Time: 25:20)

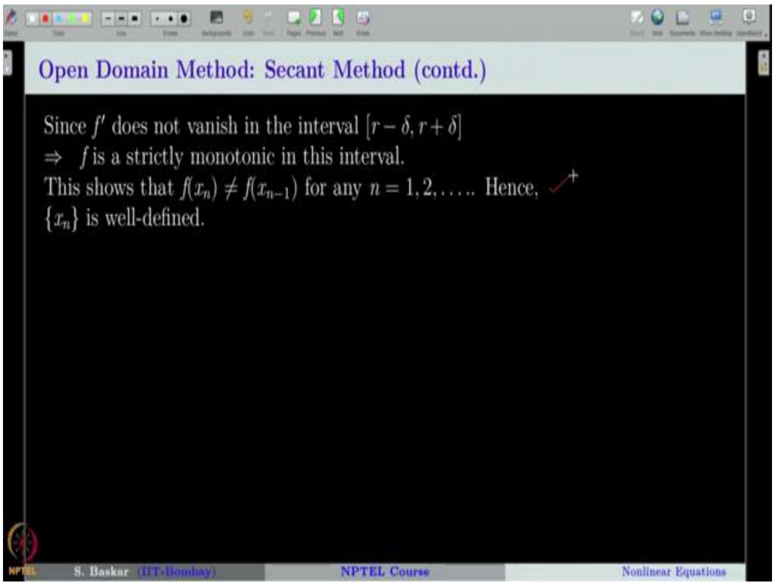


So, if you go back, you can see that this is what we wanted to so, as one of the conclusions that the whole sequence belongs to the δ neighbourhood remember, this δ is chosen such that $f'(\xi) \neq 0 \; \forall \, \xi \in [r-\delta, r+\delta]$. This is how we have chosen δ, so, you keep that in mind. You can see that f' is not equal to 0 in this interval.

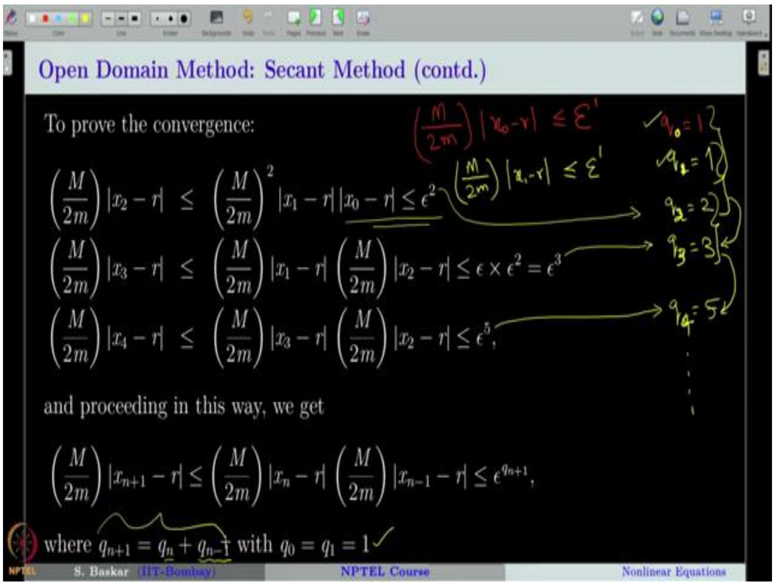
And therefore, f' is either positive or negative in this neighbourhood it means f is a strictly monotonic function. If f is strictly monotonic then it has this situation where $f(x_{n-1})$ is not equal to $f(x_n)$. Because the second conclusion says that all x_n's belongs to this neighbourhood and since all these x_n's are belonging to this neighbourhood, f' will not vanish at x_n, it means $f(x_{n-1})$ will not be equal to $f(x_n)$.

Because f will be a monotonic function. So, therefore, this conclusion is also done. Now, we have to prove the convergence and this order of convergence also.

(Refer Slide Time: 26:46)

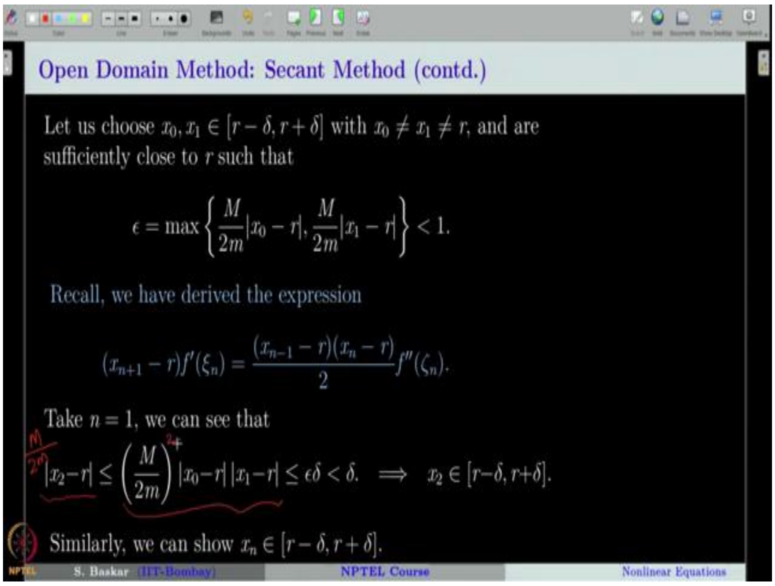


(Refer Slide Time: 26:50)



Let us see how to prove this, so, we have already discussed this. Let us go to prove the convergence, to prove the convergence first, you observe that $|x_2 - r|$ is less than or equal to.

(Refer Slide Time: 27:03)



We have already shown here, $|x_2 - r|$ is less than or equal to this quantity. Now, I am multiplying both sides by $M/2m$. Therefore, you will have a square here. That is what I am writing here $\frac{M}{2m}|x_2 - r|$ is less than or equal to $\left(\frac{M}{2m}\right)^2$ into this. And now, if you recall, $\frac{M}{2m}|x_1 - r|$ is less than or equal to ϵ and another $\frac{M}{2m}|x_0 - r|$ is also less than or equal to ϵ.

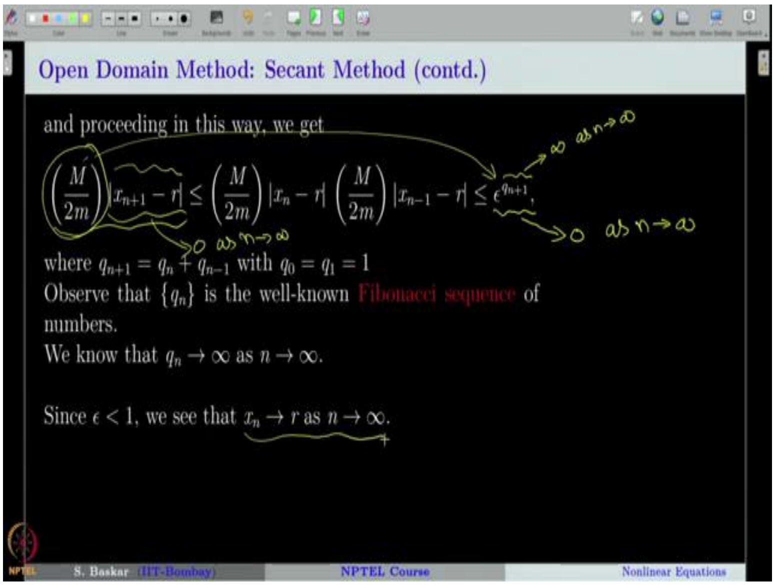
Therefore, this term is less than equal to ϵ^2. Now, take x_3, $|x_3 - r|$ is nothing but $\frac{M}{2m}|x_1 - r||x_2 - r|$ and then you multiply both sides by $M/2m$. And that will give you ϵ this is less than equal to ϵ and this is less than or equal to ϵ^2. How we have just now proved here, therefore, this entire term is less than or equal to ϵ^3.

Similarly, you can see that $\frac{M}{2m}|x_4 - r| \leq \epsilon^5$ and so on. In general, you can write $\frac{M}{2m}|x_{n+1} - r| \leq \epsilon^{q_{n+1}}$. What is q_{n+1}? Well, q_{n+1} is a sequence such that $q_{n+1} = q_n + q_{n-1}$. How can you see this? You can see that $\frac{m}{2m}|x_0 - r| \leq \epsilon$.

That is my ϵ to the power of 1 therefore, $q_1 = 1$. Now, you take $\frac{m}{2m}|x_1 - r|$ which is less than or equal to ϵ^1. Therefore, q_2 is also 1 now q_3 from here, you can see that $q_3 = 2$ which can be seen as the sum of these two terms. Now, what is q_4? q_4 is coming from here $q_4 = 3$ that can be seen as the sum of these two terms here. And what is q_5? $q_5 = 5$ that can be seen as the sum of these two terms.

In general, q_{n+1} is seen as the sum of it is immediate previous term that is, q_n + the previous to previous term that is q_{n-1} with understanding that q_0 = 1. Maybe I should start with 0 here, 1, 2, 3, 4 and so on. Sorry, therefore, q_0 is 1, q_1 is 1 and q_2 onwards we are defining like this.

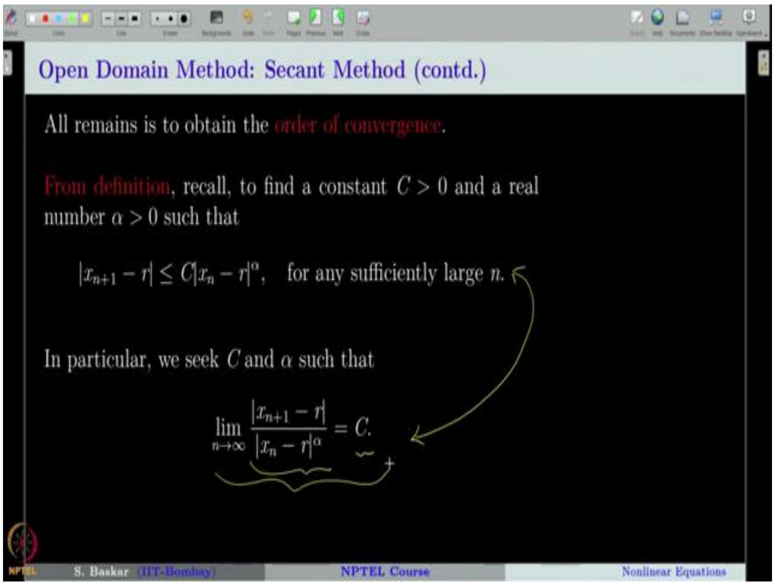
(Refer Slide Time: 30:48)



What is this sequence? This is the well-known Fibonacci sequence and once you realize that you have a Fibonacci sequence as the power of ϵ. Now, you recall that Fibonacci sequence tends to infinity, as n tends to infinity. Also recall that we have chosen our δ in such a way that ϵ is less than 1. Therefore, you have ϵ to the power of something and that is going to infinity, as n tends to infinity.

That implies that the right-hand side goes to 0, as n tends to infinity. And what is on the left-hand side? You have constant times this term this is fixed. Therefore, you can in fact push it to the right-hand side. And therefore, you can see that this term goes to 0, as n tends to infinity. That is precisely what we want, as the convergence.

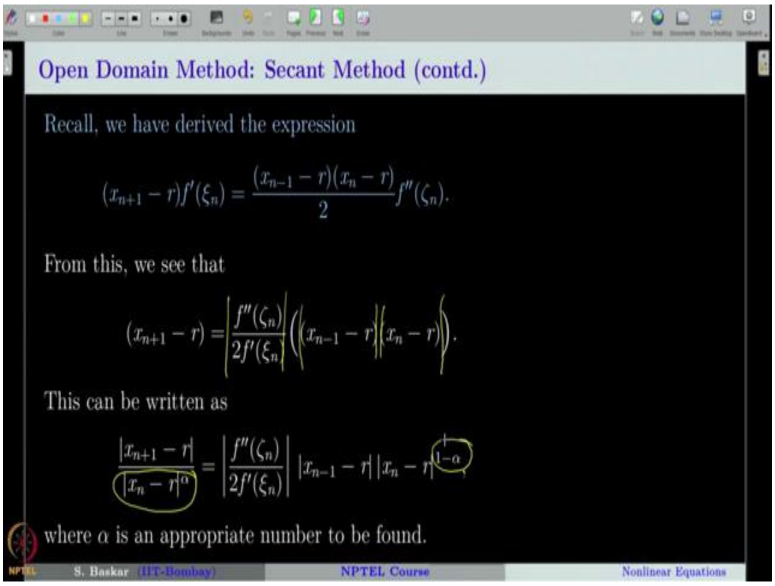
(Refer Slide Time: 31:48)



Now all remains is to get the order of convergence. How are we going to get the order of convergence? Let us see again, I will recall what is the definition of order of convergence? I have already told this at the beginning of the lecture. The definition of the order of convergence is that there should exists a constant C such that $|x_{n+1} - r| \le C|x_n - r|^\alpha$.

Then we say that the sequence converges with order at least α. That is what we have seen and that should happen, as n tends to infinity. Another way to define order of convergence is to use the following definition called asymptotic order of convergence. This says that $\lim_{n\to\infty}$ this term should be constant. This both are often used for defining order of convergence. Let us try to derive or find the constant C such that this happens with an appropriate α.

(Refer Slide Time: 33:10)

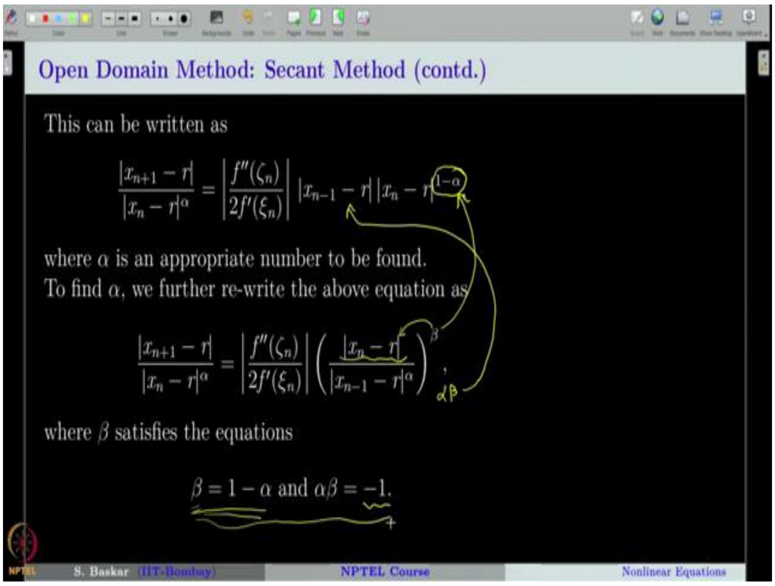


Let us see both of this. Again, you start with the expression that we have derived at the beginning of the proof and then you just take this f' on the other side with the understanding that f' never vanishes in the neighbourhood that we are working with. Remember we are working with $[r - \delta, r + \delta]$. This is the neighbourhood in which we are working.

All our x_n's belongs to this neighbourhood. Therefore, ξ_n which lies between x_n and x_{n-1} and also between r in whatever it is. It belongs to this interval therefore, $f'(\xi_n)$ will not vanish. Therefore, you can just divide both sides by $f'(\xi_n)$ and that gives you this expression. Now, take modulus on both sides and divide by $|x_n - r|^{\alpha}$. We have to find what is α but just take α and divide both sides.

You can see when you take modulus, you are taking modulus here and also here, since you are dividing both sides by $|x_n - r|^{\alpha}$, you will have $|x_n - r|^{1-\alpha}$. Where alpha has to be chosen appropriately.

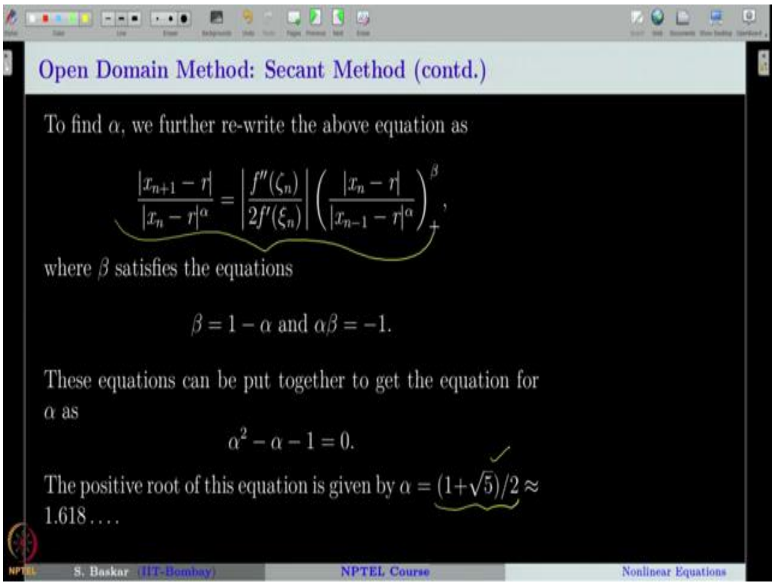
(Refer Slide Time: 34:42)



So, we have this expression. Now, what we will do is we will write this term in this form. I am just rewriting it with an appropriate β. That is all if that is so then what should be our β? You can see that β should satisfy that β should be equal to 1 – α. Why? Because $|x_n - r|^{\beta}$ will come and that should be equal to $1 - \alpha$ from this expression.

Therefore, we want β to be 1 – α and in the denominator we have α into β that when you compare with this term, you want α into β to be – 1. Now, you just have to find α such that these two things happen simultaneously.

(Refer Slide Time: 35:38)

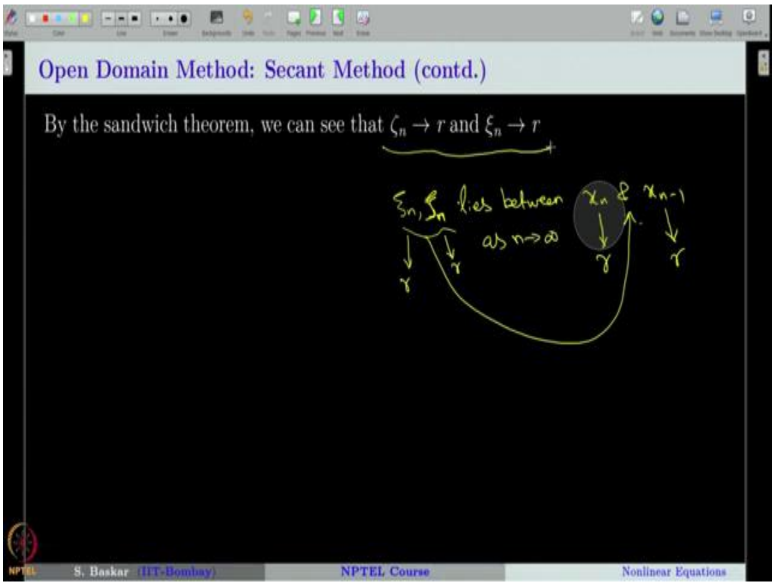


That implies that α should come as the root of this equation. Why? Because α should be equal to – 1/β or β = – 1/ α that value of β, you can put here to get this equation. Therefore, the α that we want should come as the root of this quadratic equation. Remember this quadratic equation has two roots and one root is given like this and the other root of this quadratic equation is a negative number.

We are concerned about the order of convergence. Therefore, we will not choose the negative value of the α. So, we will choose this α and this number will surely satisfy this expression which is precisely satisfied by the secant method. **(Video Starts: 36:35)** Because the expression that we have derived if you recall, we started with this expression.

And this expression is derived from the secant method and from this expression only we have landed up with this and further rewritten that expression in this form and that is further written like this. Therefore, everything is coming from the secant method. And the α is therefore, chosen very naturally, from the secant method and that shows that the secant method, if it converges, will converge with order α. And what is this value? This value is approximately 1.618. **(Video Ends: 37:15)**

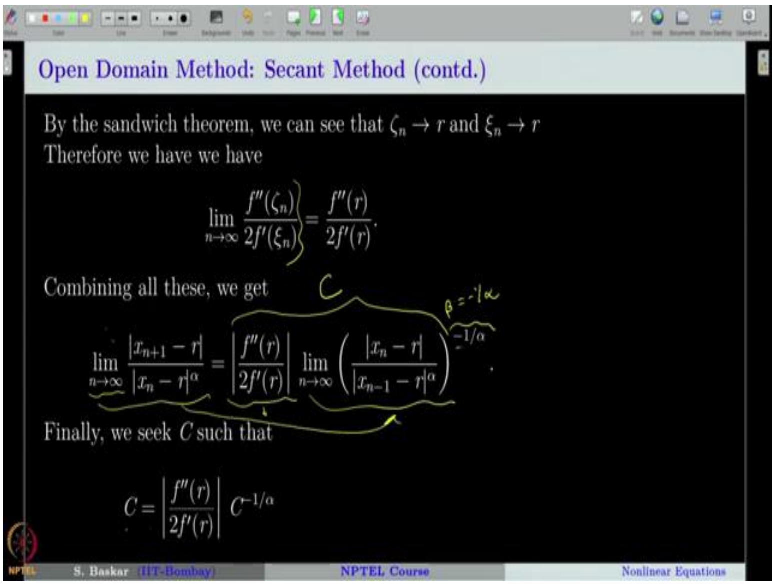
(Refer Slide Time: 37:15)



All remains is to now, find the constant *C*. How will you find? Recall that ξ_n will converge to *r* and also ζ_n will converge to *r*. How can you see? You can see it from the sandwich theorem because how we have chosen ξ_n and ζ_n? Well, ξ_n and ζ_n both lies between $[x_n, x_{n-1}]$. But we know that x_n converges to *r* that we have already shown and also x_{n-1} converges to *r*.

And these two numbers always lie between these two numbers. Therefore, they will also converge to *r* as *n* tends to infinity. That is what is very clear from the sandwich theorem.

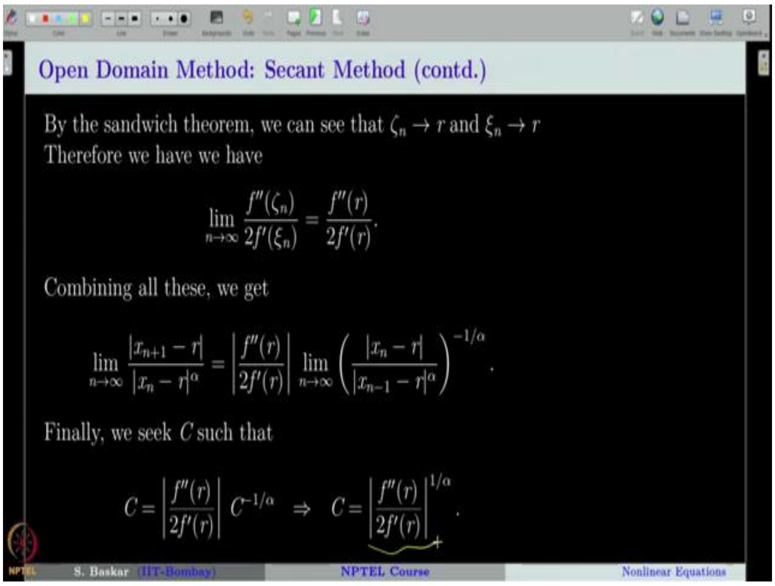
(Refer Slide Time: 38:18)



Now, you take the limit *n* tends to infinity in this expression. You can see that this is equal to $\frac{f''(r)}{2f'(r)}$. Now, putting that into your expression because you are taking limit, as *n* tends to infinity, it will become like this into this term. Remember we had β here but β = – 1/α that is what I am putting here. And therefore, we want the constant *C* such that, remember you want the whole thing to be equal to some constant.

And that constant should be such that you have *C* which is this one this should be equal to *C* therefore, *C* is equal to this term and this is nothing but this. This to the power of – 1/α.

(Refer Slide Time: 39:28)



Therefore, we want the constant *C* such that *C* equal to this term into *C* to the power of – 1/α. That will immediately give us what is *C*? *C* is given by this. This is precisely what we wanted to show in the statement of our theorem. Well, this is little long but very interesting and important for us to understand. With this, I will end this lecture. Thank you for your attention.