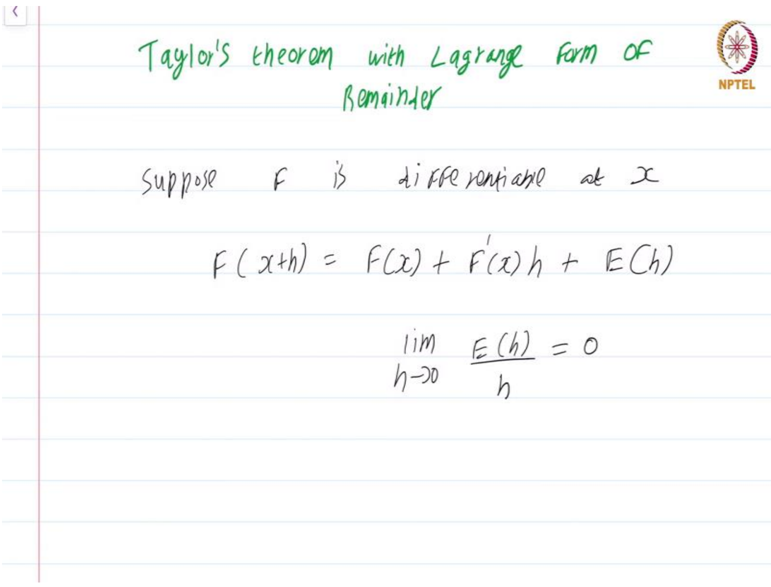
Real Analysis - I
Dr. Jaikrishnan J
Department of Mathematics
Indian Institute of Technology, Palakkad

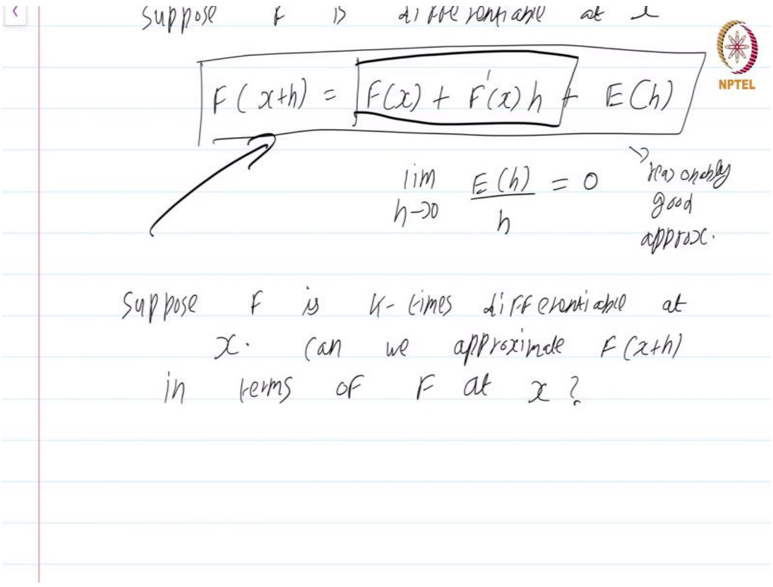
Lecture - 24.1
Taylors Theorem

(Refer Slide Time: 00:13)



In this module, we are going to study the very important Taylor's Theorem. Recall our definition of derivatives interpreted in terms of linear approximations. Suppose, f is differentiable at the point x, then $f(x+h) = f(x) + f'(x)h + E(h)$ and this error term has the property that $\lim_{h \to 0} \frac{E(h)}{h} = 0$

(Refer Slide Time: 01:09)



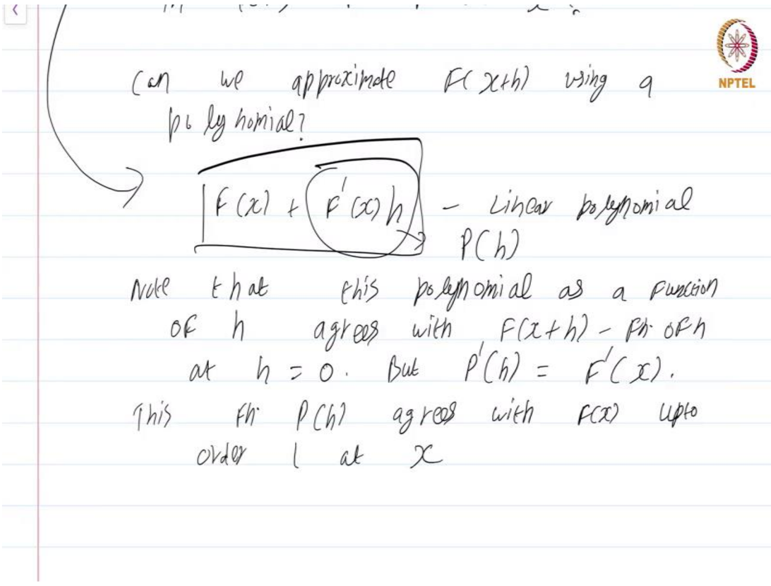
So, this approximation is a reasonable one, a reasonably good approximation. The question naturally arises can we make a better approximation if we know not only that f is differentiable at the point x, but f is k-times differentiable at the point x. So, let us see what happens.

Suppose f is k-times differentiable at x. Now, one would like to have an approximation of $f(x+h)$ in terms of the data that you have about f at the point x. So, can we approximate $f(x+h)$ in terms of f at x.

So, we can do an approximation because k-times differentiable automatically means, it is once differentiable and we have this approximation. Of course, the goal is to do much better than this by using the higher derivatives.

Now, what will be our guess? We have approximated $f(x+h)$ using something that is linear. So, the natural thing to do would be to try to approximate $f(x+h)$ using a polynomial. Let us try to do that.

(Refer Slide Time: 02:54)



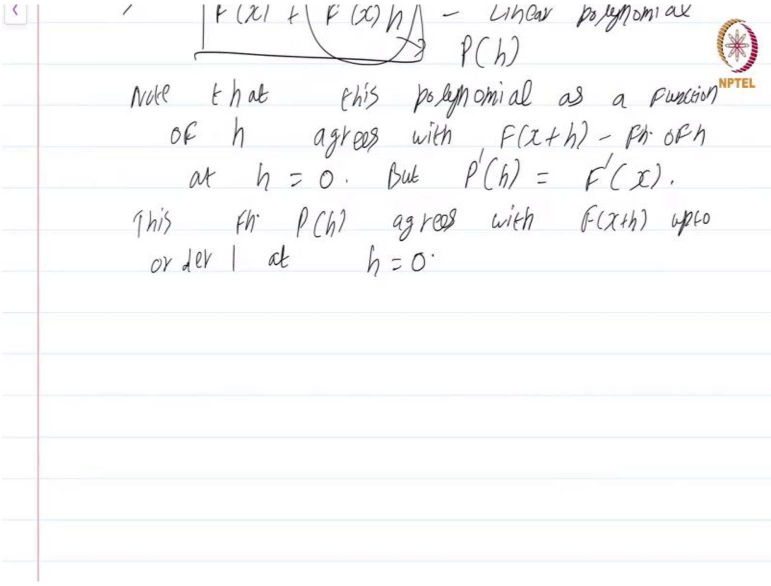
Can we approximate $f(x+h)$ using a polynomial? Now, what polynomial would you choose? Note that the polynomial that we have considered before is $f(x) + f'(x)h$; this is a linear polynomial, a polynomial of degree 1.

Not only is it a polynomial of degree 1, it is a special polynomial. So, note this is treated as a polynomial in h of course, it is not a polynomial in x, x has been fixed.

Note that this polynomial as a function of h agrees with $f(x+h)$ treated again as a function of h at h =0. Well yes, when you substitute $h = 0$, this term vanishes you get $f(x) = f(x)$.

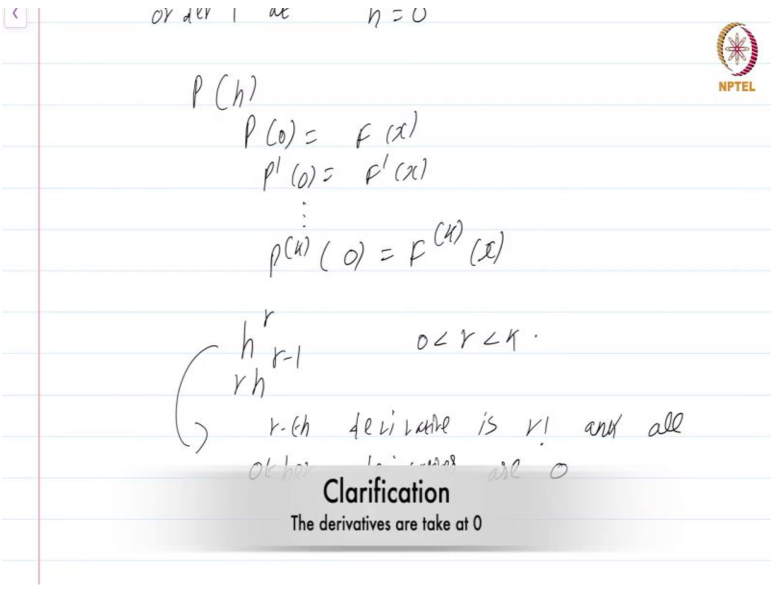
But $p'(h) = f'(x)$, right because when you differentiate this with respect to h, this $f(x)$ term vanishes and derivative is just $f'(x)$. So, this function $p(h)$ agrees with $f(x)$ up to order 1 at x.

(Refer Slide Time: 05:10)



Or rather to write it more precisely agrees with $f(x+h)$ up to order 1 at $h = 0$.

(Refer Slide Time: 05:48)



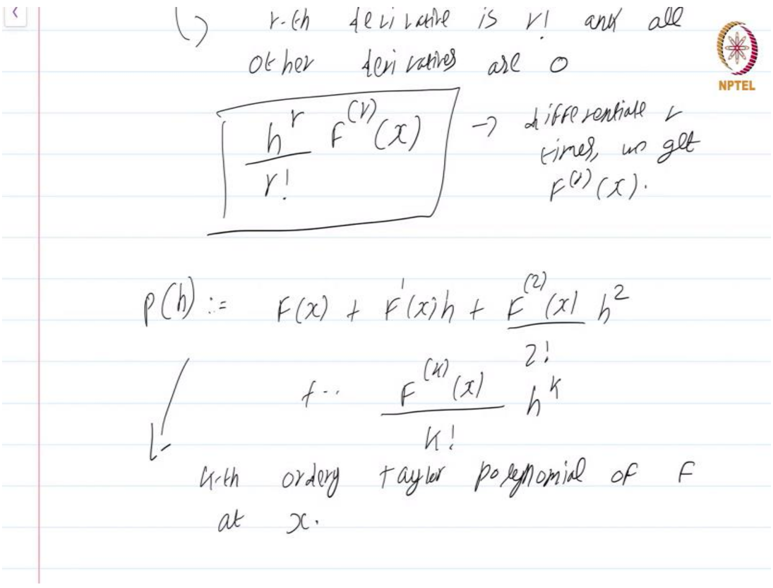
Now, we want to produce a polynomial that agrees with f at the point $h = 0$, not just till order 1, but till order k or $k+1$ or what not. Here, we are assuming f is k times differentiable so, up till order k ok.

So, how would you do that? What we want is a polynomial $p(h)$ with the property that $p(0) = f(x), p'(0) = f'(x), \ldots\, p^{(k)}(0) = f^{(k)}(x)$.

Now, if you think about this for a moment, you will soon understand that it might be a good idea to use these special monomials of the form h^r, $0 \leq r \leq k$, right. Look at these special monomials h^r, look at the derivative. The derivative of this is rh^{r-1}.

So, this h^r, have the special feature that rth derivative is $r!$ and all other derivatives are 0; all other derivatives are 0. Of course, I am treating it as a function of h and differentiating with respect to h.

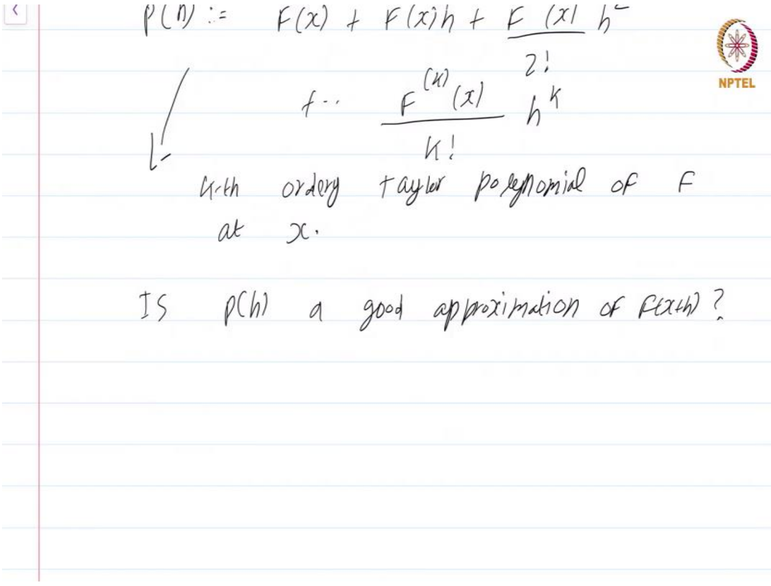
(Refer Slide Time: 07:05)



So, if you consider $\frac{h^r}{r!} f^{(r)}(x)$, if you look at this; if you look at this term, differentiate r times; differentiate r times, we get $f^{(r)}(x)$, well right because the rth derivative of h^r would be nothing but $r!$ factorial which will get cancelled with the $r!$ factorial in the denominator and you are left with $f^{(r)}(x)$.

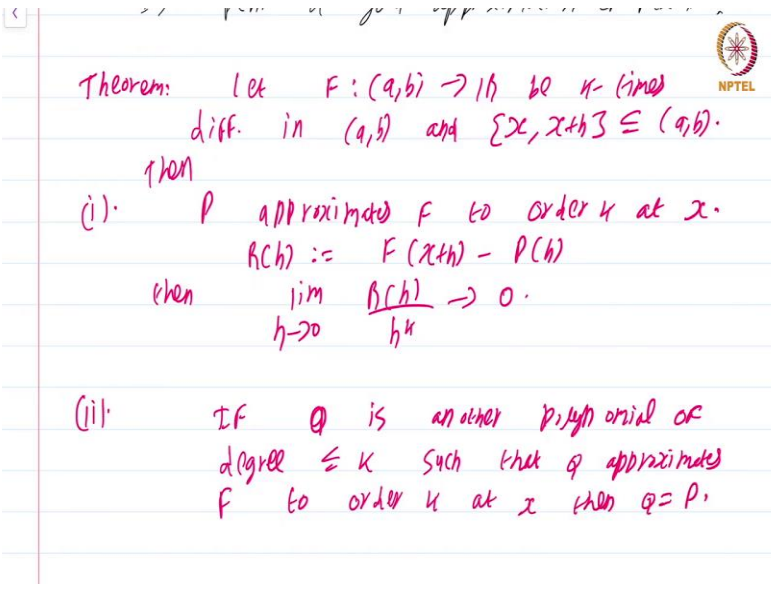
So, this prompts us to consider this polynomial $p(h){:} = f(x) + f'(x)h + \frac{f^{(2)}(x)}{2!} h^2 + \cdots + \frac{f^{(k)}(x)}{k!} h^k$ and this is called the kth order Taylor polynomial of f at x ok.

(Refer Slide Time: 08:33)



Now, is $p(h)$ a good approximation of $f(x+h)$? The following theorem answers this question.

(Refer Slide Time: 08:57)



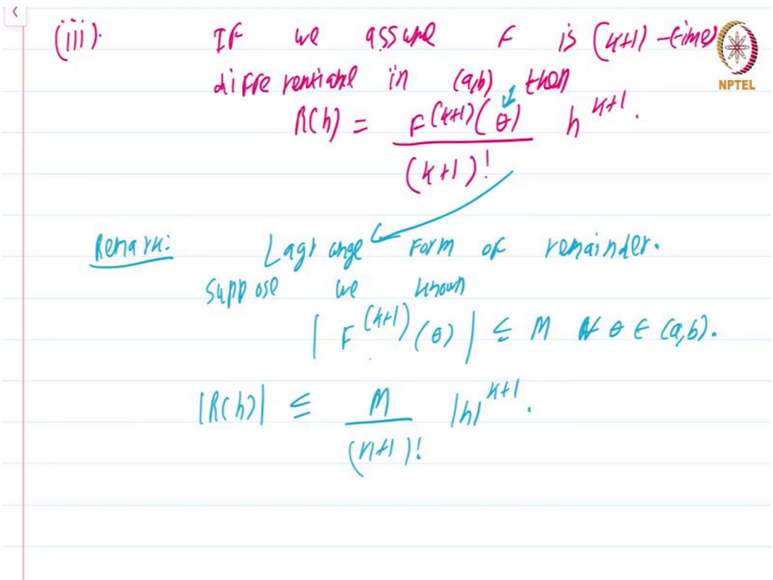
Theorem: Let $f:(a,b) \longrightarrow R$ be k-times differentiable in (a,b) and $[x, x+h]$ closed interval let it be a subset of $[a,b]$, then

(1) p approximates f to order k at x. What does this mean? This means if you define $R(h)$ $= f(x+h) - p(h)$, then $\lim_{h \to 0} \frac{R(h)}{h^k} = 0$, ok.

So, this just says that the remainder term $f(x+h) - p(h)$ gets really small as h goes to 0 and this is quantified by saying it approximates to order k, $\frac{R(h)}{h^k}$ itself goes to 0, not just $\frac{R(h)}{h}$ as what happened in the definition of the derivative.

(2), If Q is another polynomial of degree less than or equal to k such that Q approximates f to order k at x, then $Q = p$ right; that means, there is only a unique polynomial which has the feature that f, that polynomial approximates f till order k.

(Refer Slide Time: 11:28)



(3), If we assume f is $k+1$ times differentiable in (a,b), then $R(h)$ you can write it as $\frac{f^{(k+1)}(\theta)}{(k+1)!} h^{k+1}$. You can write an explicit expression for $R(h)$. Let me immediately make a remark. This is known as Lagrange form of remainder and it is very useful in the following sense.

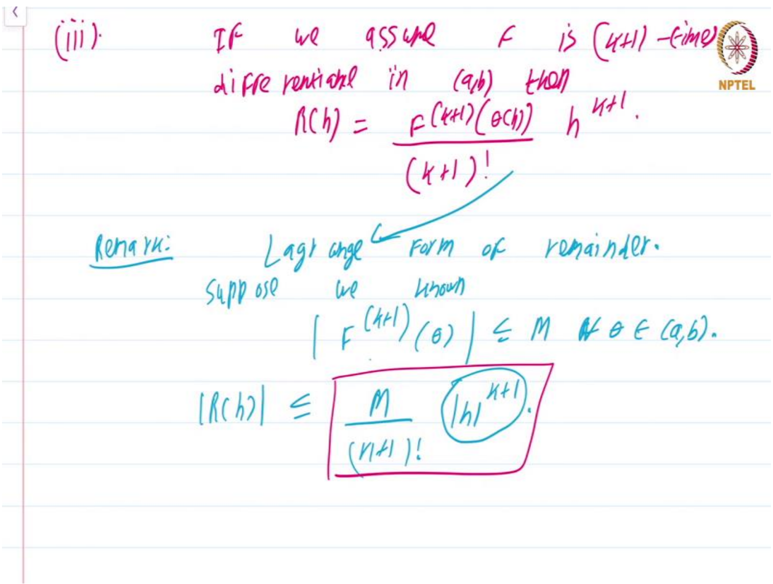
Suppose we know that $|f^{(k+1)}(\theta)| \leq M$ for all $\theta \in (a,b)$. for instance, we know this for sine and cosine. We know that the derivative of sine is cosine, and the derivative of cosine is - sine and these just get repeated, the higher order derivative just get repeated.

So, we know that if you are considering the sine function, no matter what derivative you take, it is always going to be modulus less than or equal to 1 right. So, in many functions we do know such data about any derivative.

So, we have this $|f^{(k+1)}(\theta)| \leq M$ when θ comes from (a, b). So, $|R(h)| \leq \frac{M}{(k+1)!} h^{k+1}$. Now, observe that in the Lagrange form of the remainder, there is a quantity θ that is unknown to us.

This θ will in general change when you perturb the point h. So, even if you have fixed x, depending on the h, this θ will change. So, this remainder form, this remainder term even though I have written $R(h)$, this expression on the right there is an implicit dependence of h even in θ.

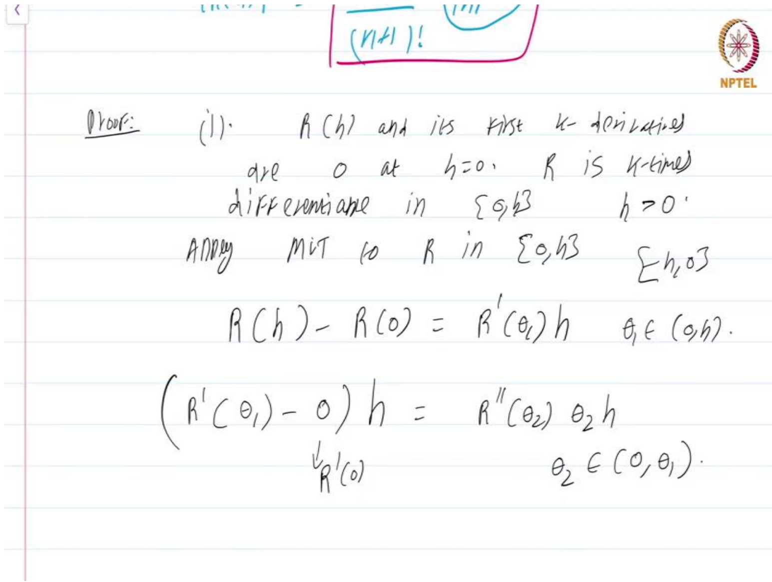
(Refer Slide Time: 14:17)



So, a better way to write it is to write this as $\theta(h)$ and this $\theta(h)$ function is sort of unknown to us, how it behaves, whereas, in this expression that additional dependence on h has disappeared, the dependence on h is solely coming from this; solely coming from this. A function that we understand really well $|h^{k+1}|$ is a very simple function ok.

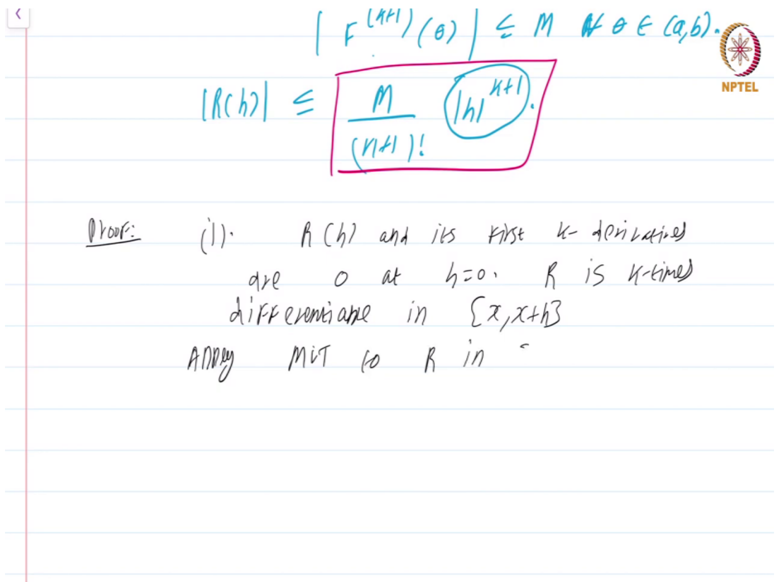
So, the additional dependence on this function θ which is in general unknown is gone. So, this Lagrange form of the remainder when you know you have some estimate on the derivatives is very very useful and this will be illustrated throughout in the coming modules when we talk about the elementary function sine, cosine and so on defined in terms of power series ok.

(Refer Slide Time: 15:13)



On to the proof. Well, let us first start with the first part. We already know that $R(h)$ and its first k derivatives are 0 at $h = 0$ and R is certainly R is k times differentiable in $[a, b]$ in the closed interval $[a, b]$.

(Refer Slide Time: 15:58).



Not in the closed interval $[a, b]$ sorry in the closed interval $[x, x + h]$. So, we can apply mean value theorem to R in not $[x, x + h]$, in $[0, h]$. We can apply mean value theorem to R in $[0, h]$. What do we get? Well, $R(h) – R(0) = R’(\theta_1)h$.

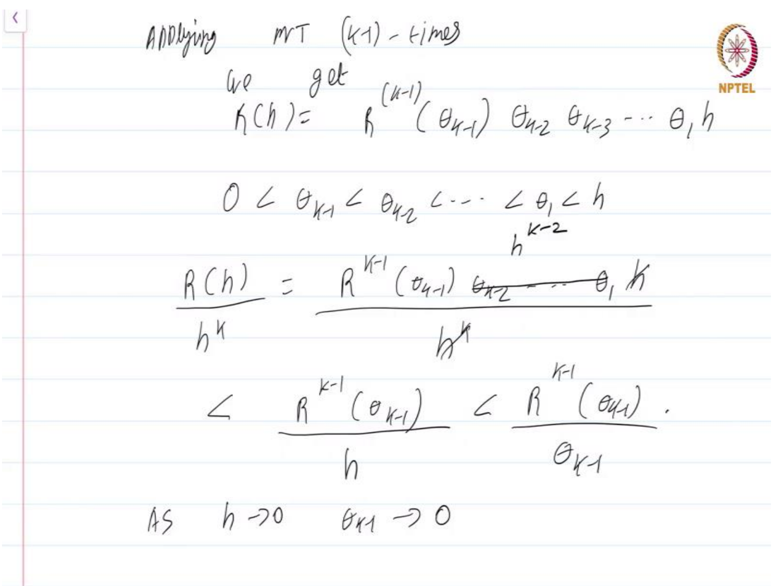
for convenience, I am assuming $h > 0$ because I am writing $[0, h]$. Whatever I am about to write if $h < 0$, you just have to apply it to $[-h, 0]$. You will just have to apply an analogous argument a $[-h, 0]$; for convenience I am assuming $h > 0$, ok. So, $R(h) – R(0) = R'(\theta_1)h$ by mean value theorem, where $\theta_1 \in [0, h]$ or rather the open interval $(0, h)$; open interval $(0, h)$.

Now, we can of course, apply the mean value theorem again to the function $R(\theta_1) - 0$. Look at $R'(\theta_1) - 0$. Well, this is just $R'(0)$, right. So, you can apply the mean value theorem again now to the function R' to get $R''(\theta_2)h$ where $\theta_2 \in [0, \theta_1]$ right. So, this is just applying the mean value theorem again.

Now, note we know that the function R is k times differentiable in the closed interval $[0, h]$, but we cannot apply the mean value theorem k-times on this closed interval simply because we do not know whether kth derivative of R is continuous in closed interval $[0, h]$ that is unknown to us.

Because we are just assuming kth order differentiability of the function f in the open interval (a, b). So, we cannot apply the mean value theorem k times, but we can apply it (k-1)times ok.

(Refer Slide Time: 18:34)



So, applying mean value theorem (k-1)times, we get $R(h) = R^{(k-1)}(\theta_{k-1})\theta_{k-2}\theta_{k-3} \ldots \theta_1 h$, where we have $0 < \theta_{k-1} < \theta_{k-2} < \theta_{k-3} \ldots < \theta_1 < h$ ok. So, this is by repeatedly applying the mean value theorem.

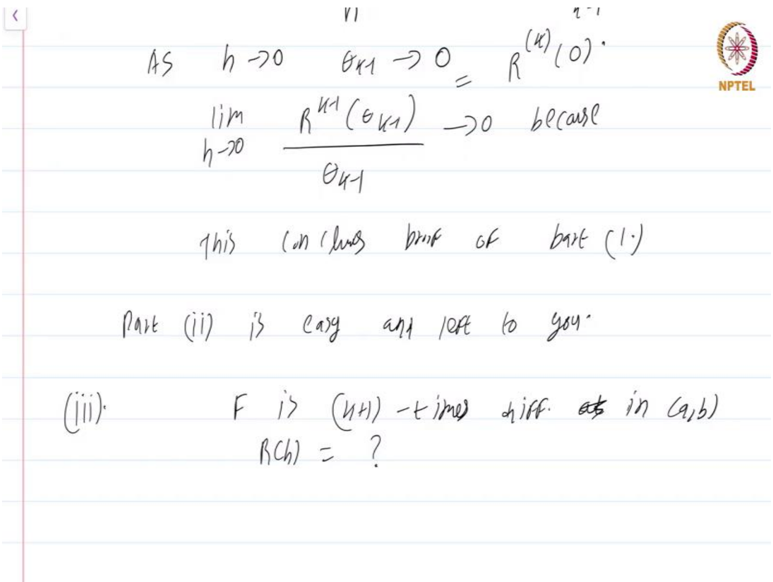
Now, the quantity we are interested in is $\frac{R(h)}{h^k}$ and this just turns out to be $R(h) = R^{(k-1)}(\theta_{k-1})\theta_{k-2}\theta_{k-3} \ldots \theta_1 \frac{h}{h^k}$ Now, how is this going to be, how are we going to simplify this?

Well, observe that this $\theta_{k-1}, \theta_{k-2}, \theta_{k-3}, \ldots, \theta_1$ are all less than h. So, this whole thing I can write it as $R^{(k-1)}(\theta_{k-1})$ and cancel off h and this and cancel off all of these also. So, what I am doing is in the numerator, have a quantity which is less than the denominator.

So, I am going to replace these essentially by h^{k-2}, right and cancel it off. So, this is strictly less than $\frac{R^{k-1}(\theta k-1)}{h}$. Now, what I am going to do is I am going to be clever.

The denominator I am going to replace by a smaller quantity and write this is less than $\frac{R^{k-1}(\theta k-1)}{\theta_{k-1}}$ which I can do. Now, as h goes to 0, well θ_{k-1} also goes to 0. We do not know how it goes to 0, but it certainly is going to go to 0, right.

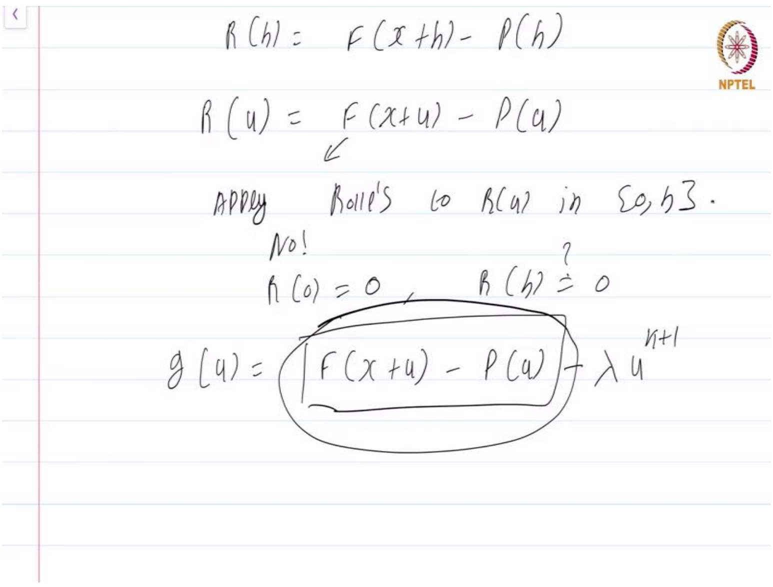
(Refer Slide Time: 21:14)



That means, $\lim_{h\to 0} \frac{R^{k-1}(\theta k-1)}{\theta_{k-1}} = 0$ because this is just $R^k(0)$, right. So, what happens is the quotient goes to $R^k + 0$ which is 0 we know that because that is how the function R was defined, we had defined the function p to agree with function f till order k at the point $h = 0$ ok.

So, this concludes the proof of part 1 ok. Part 2 is easy and left to you. Now, let us go to part 3. In this part, let me recall we are assuming that f is k+1 times differentiable at or in (a, b) and we want to get an explicit form for $R(h)$; $R(h)$ explicit form we want ok.

(Refer Slide Time: 22:40)



So, now, we already know that $R(h) = f(x+h) - p(h)$. Now, what I am going to do is I am going to treat R as a function of u for a moment because at the end of the day, even though I write $R(h)$ sort of h is fixed right so, I am going to treat it as $f(x+u) - p(u)$.

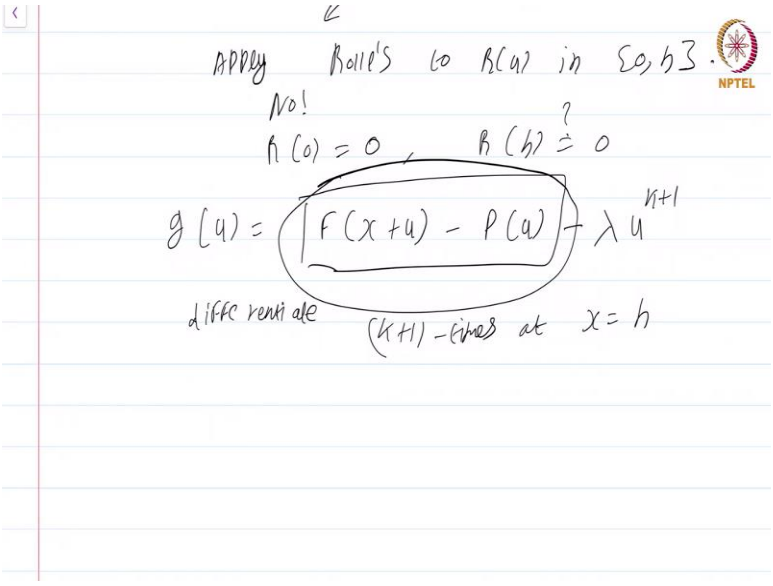
Now, what I plan to do is look at $f(x+u) - p(u)$, I want to so, I want to apply Rolle's theorem to $R(u)$ in $[0, h]$. I want to apply Rolle's theorem to $R(u)$ in $[0, h]$, but can I apply Rolle's theorem to $R(u)$ in $[0, h]$? No. So, even though R(0) = 0, we do not know whether R(h) = 0. Is this 0? Well, that is not clear; that is not clear.

So, what we do is we consider this new function $g(u) = f(x+u) - p(u)$ something to make R, to make the value at h, 0, but if I make the value at h, 0, I will be modifying these first few terms at the point $u = 0$ also. If I want to make this $g(u) = 0$ at $u = h$, I do not want to touch the value of g at the point 0.

So, what I do is I subtract a quantity λ; λ is going to be a constant which I am going to determine times u^{k+1}. Why did I do u^{k+1} because of the convenient property that at $u = 0$, the first k derivatives vanish, the first k derivatives vanish at $u = 0$.

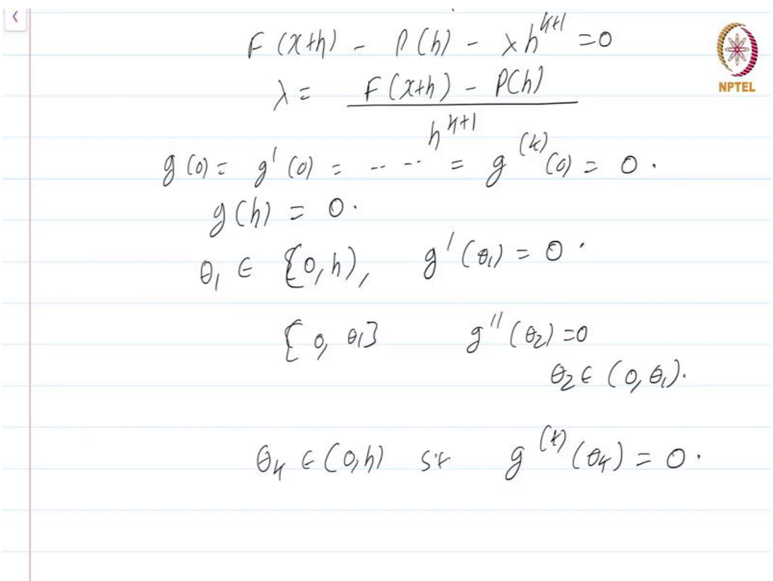
So, this modification to the function R that I am doing by subtracting λu^{k+1} does not affect the behaviour of the first k derivatives at the origin that is the logic behind subtracting λu^{k+1}, ok.

(Refer Slide Time: 25:16)



Now, how do we determine what this λ is well simple. Differentiate $(k+1)$ times at $x = h$; at $x = h$, rather you do not have to do all that sorry about that, you do not have to do that at least now.

(Refer Slide Time: 25:45)



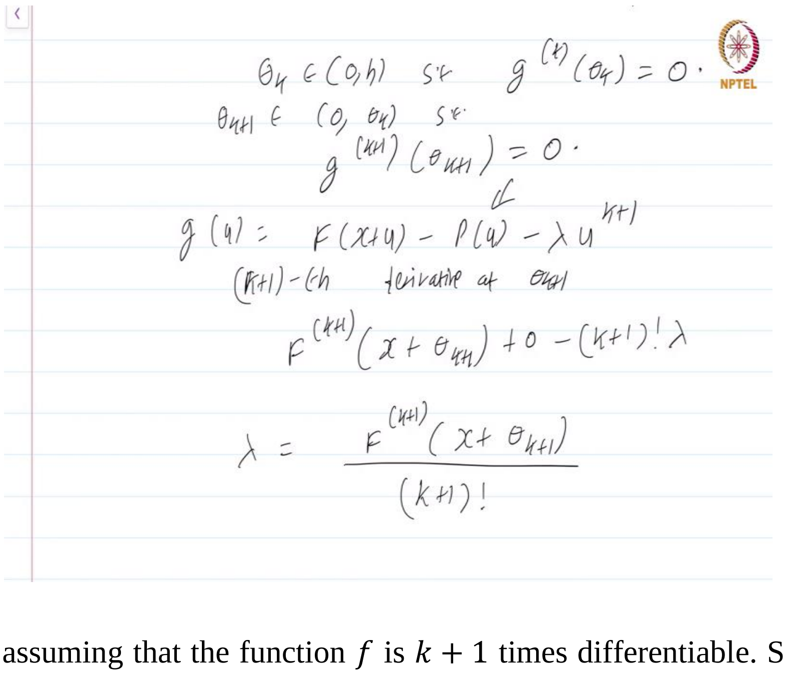
You just want to determine what λ is. Well, we want $f(x+h) - p(h) - \lambda h^{k+1} = 0$ right. So, that just gives λ is going to be $\frac{f(x+h)-p(h)}{h^{k+1}}$. Now, it really does not matter what exactly λ is I am just telling you that there is a λ that will make Rolle's theorem applicable to the function g.

Now, observe the way we have done $g(0) = g'(0) = \cdots = g^k(0) = 0$, we know this and we also know that $g(h) = 0$ that is all we know ok.

Now, applying Rolle's theorem, we get a point θ_1. Again, I am going to assume that $h > 0$ for convenience analogous arguments will hold when $h < 0$, there is a $\theta_1 \in (0, h)$, in open interval $(0, h)$ such that $g'(\theta_1) = 0$.

Similarly, applying it again, applying Rolle's theorem again to $[0,\ \theta_1]$, we will be able to conclude that $g''(\theta_2) = 0$, where θ_2 comes from $[0,\ \theta_1]$. When you apply this repeatedly, what will happen is we will get all the way till we will get a point θ_k such that this will be somewhere in $[0, h]$ such that $g^k(\theta_k) = 0$, right.

(Refer Slide Time: 28:03)

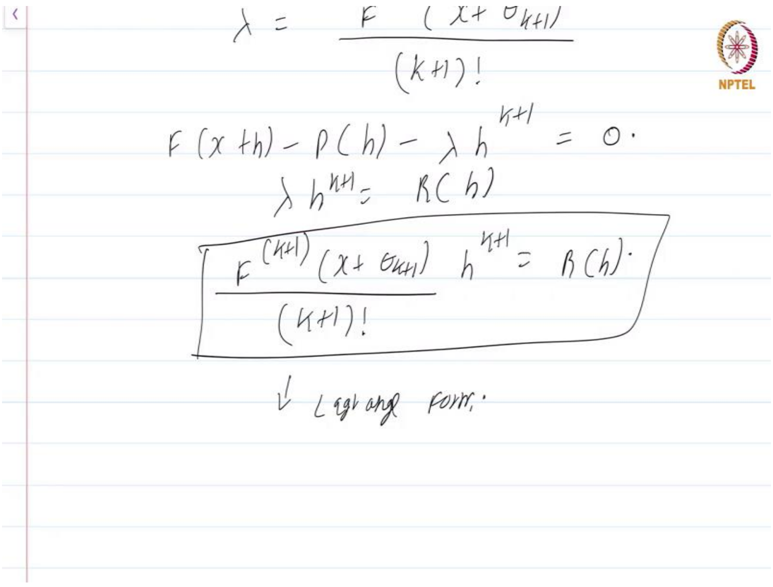


But we are now assuming that the function f is $k+1$ times differentiable. So, we can apply Rolle's theorem once more. We can apply Rolle's theorem once more to get a point $\theta_{k+1} \in [0, \theta_k]$ such that $g^{(k+1)}(\theta_{k+1}) = 0$, ok.

Now, this is what I want. Well, how is this useful? So, we have $f(x+h)-p(h)$ sorry we do not have that sorry about that we have $g(u)$; $g(u) = f(x+u) - p(u) - \lambda u^{k+1}$.

What will be the $(k+1)$th derivative; what will be the $(k+1)$th derivative of this at θ_{k+1}, well that is just going to be $f^{(k+1)}(x+\theta_{k+1})$. This polynomial is a of degree k, so, that vanishes and this will just give us $-(k+1)!\,\lambda$. In other words, $\lambda = \frac{f^{(k+1)}(x+\theta_{k+1})}{(k+1)!}$ this is what λ is going to be.

(Refer Slide Time: 29:48)



But wait a second, how was λ chosen? λ was chosen so that $f(x+h) - p(h) - \lambda h^{k+1} = 0$, right. In other words, λh^{k+1} is what we have been calling $R(h)$ all along. It was just a different name that we have given right.

But $\lambda = \frac{f^{(k+1)}(x+\theta_{k+1})}{(k+1)!}$ right that is what we have concluded by repeated application of Rolle's theorem the; that means, λh^{k+1} is $R(h)$ and if you look carefully, this is what we wanted, this is the Lagrange form.

So, I hope the proof is clear and what is going on and the logic behind the proof. So, it is easy to understand this proof, it is not that hard, but the logic behind the proof is what I want you to appreciate. So, please go through this proof once more and try to make sure that you digest it completely.

We will anyway revisit Taylor's theorem once again after we study integration and give a different form of the remainder term. This is a course on real analysis, and you have just watched the module on Taylor's theorem with Lagrange form of the remainder.