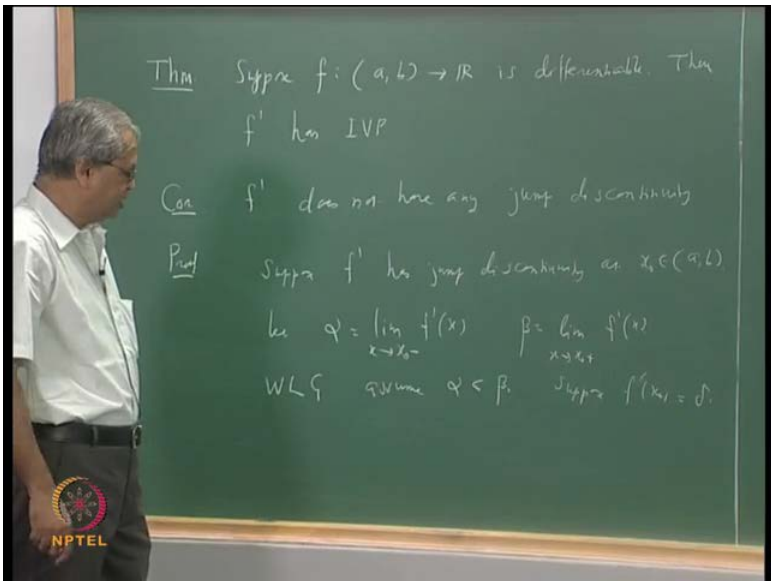
Real Analysis
Prof. S. H. Kulkarni
Department of Mathematics
Indian Institute of Technology, Madras

Lecture - 36
Taylor's Theorem

Well, we recall the last theorem that we proved yesterday, timely that when a function is differentiable, its derivative has intermediate value property, let us just recall that.

(Refer Slide Time: 00:24)



So, this is the theorem suppose f from a b to r is differentiable, then f prime has intermediate value property. Now, let us just take all intermediate value property, means if the derivative takes any two values, then it takes all the values lying in between those two. Those values are taken in between those two values and those values are taken actually in between those two points where the other values are taken. Anyways, we have explained in detail in the last class, one of the corollaries, this is the one which we wanted.

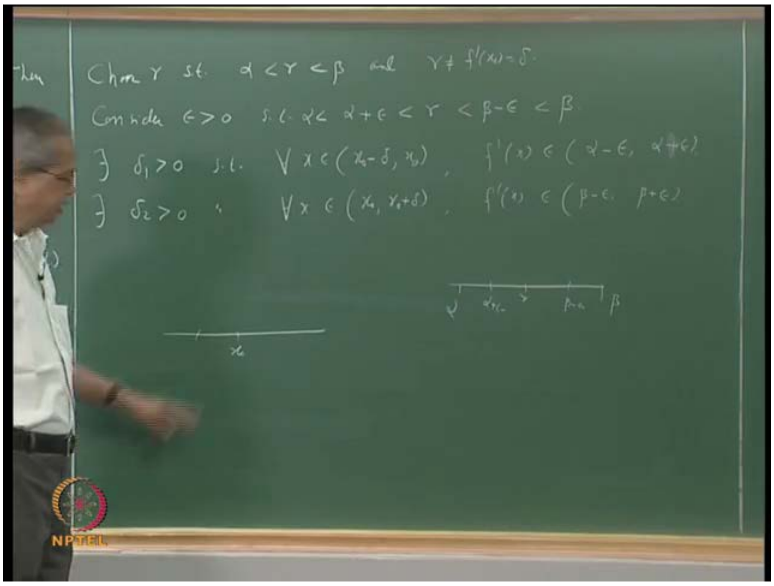
Again, we have the same hypothesis, then f prime, then let us say f prime does not have any discontinuity of the first, let us say let us say any jumped discontinuity. It can have discontinuity of the second kind and we have seen an example of that type of function. Now, how does one do this, of course there is there is basically one way of proving that you assumed that it has a jumped discontinuity and arrive at a contradiction.

So, that is how we proceed, suppose f prime has jumped discontinuity at some point x naught in it, what is that mean? It means that the limit of f prime as f goes to x naught exist, that is left hand limit exist, right hand limit exist, but those are different, that is the meaning of jumped discontinuity. So, let us say suppose I call this left hand limit as alpha, let alpha be equal to limit f prime x as x goes to x naught from negative side from minus from left side.

So, x goes to x naught minus and let us say beta is limit as x goes to x naught from the right hand side f prime x, then jumped discontinuity means alpha is not equal to beta means what either alpha is less than beta or alpha is bigger than beta. Let us take any one of the cases and give the proof in that, the proof in the other case will be similar. So, this is standard terminology for this kind of thing one say that without loss of generality assume that alpha is less than beta that is the standard way of saying.

Without loss of generality, assume alpha is less than beta also we have to say something about the value of this f prime at the point x naught also value of the point f prime at x naught. It may be alpha, it may be beta or some totally different number, anyway it does not matter, so suppose f prime at x naught is equal to let us say delta.

(Refer Slide Time: 05:05)



Now, we will choose some number which lies between alpha and beta which lies between alpha and beta, but is different from delta. So, suppose this we can always find because there are infinite limit of real numbers lying between alpha and beta, so we can easily choose one which is different from delta. Choose gamma such that alpha less than gamma less than beta and gamma is also different from this f prime x naught which we had delta. It this does not matter really whether we call it delta or not, now if we knew that at some point f prime x is alpha f and at some point f prime x is beta. We could have said that some point it has to take value of gamma, but that is something that we do not know, all that we know is that the limit of f prime x is x tends to x naught minus alpha, but does not matter.

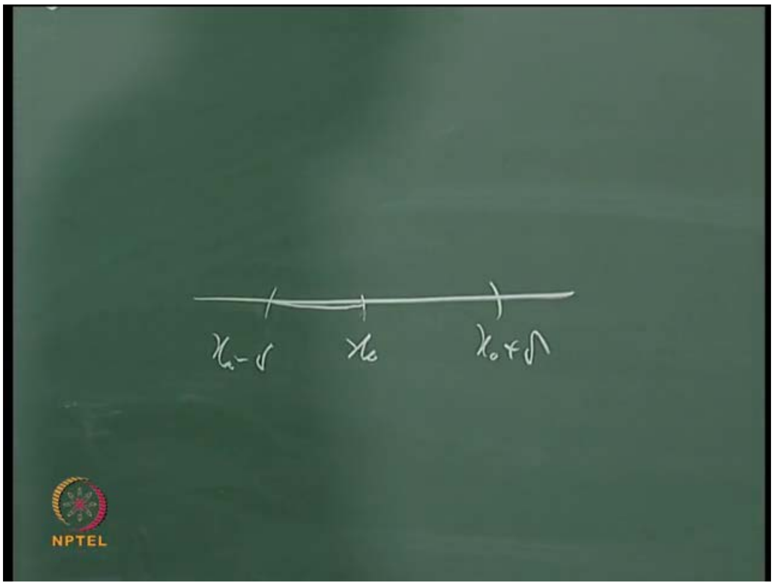
We can use that, this is the limit, so we can say that suppose we consider any see since alpha is less than gamma, gamma is less than beta, I can always think of epsilon such that alpha plus epsilon is less than gamma. Similarly, beta minus epsilon is bigger than gamma, so we can say that consider epsilon bigger than 0 such that alpha plus epsilon is less than gamma, of course gamma will be less than beta and gamma is less than beta minus epsilon. This will be bigger than alpha this will be bigger than alpha and this will be less than beta and we shall use the definition of this left hand limit.

For this epsilon, there exists let me say delta on bigger than 0 such that when u take any x between x naught minus delta to x, if you take any x between x naught minus delta to x such that foe all x in x naught minus delta to x. What should happen f prime x should lie between alpha plus epsilon and alpha minus epsilon every x and x naught that is the meaning of left hand limit, that is meaning of left hand limit. So, for every x and x naught minus delta to x f prime x lies between alpha minus epsilon to alpha minus epsilon to alpha plus epsilon.

Similarly, we use this fact also since right hand limit is beta, they will exist for the same epsilon, there will exist some delta to there exist some delta to bigger than 0 such that for every x is, now this time this should be x naught x naught to x. It is a right hand limit for every x in x naught 2 x naught plus delta for every x in x naught 2 x naught plus delta, what should happen f prime x should belong to beta minus epsilon to beta plus epsilon.

You can see that this contradicts this intermediate value property, see what is happening here you have this x naught plus delta f prime x takes the value between let me draw this alpha and beta. So, this is alpha this is beta this is gamma and this is something like let us say this may be plus epsilon and this may be beta minus epsilon in the interval x naught minus delta, let us suppose this is x naught minus delta and this is x naught plus delta.

(Refer Slide Time: 09:33)

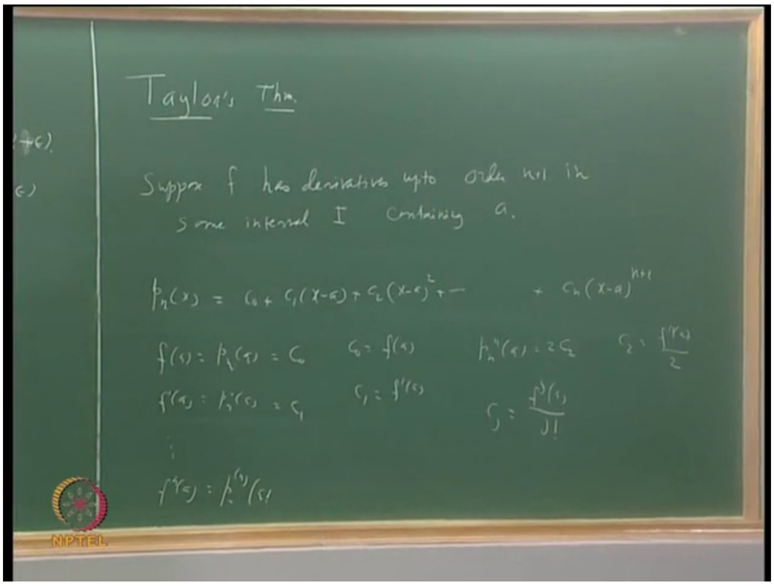


In this interval, x naught minus delta to x naught all values of f prime x lie between this alpha minus epsilon to alpha plus epsilon f prime x naught is some value delta which is not same as gamma. For all values in the x naught to x naught plus delta, the value lies in beta minus epsilon to beta plus epsilon, that means in this interval alpha plus epsilon to beta minus epsilon no value is taken, that is clear.

So, that clearly contradicts this intermediate value property, so that means the derivative cannot have a jumped discontinuity, you can compare this with the properties of monotonic functions. In case of monotonic functions, we have shown that it has only jumped if all these studies are jumped discontinuities. Now, let us move forward to some other natural extensions of mean value theorem, see all this we have seen as consequences of mean value theorem. There is a most general mean value theorem which you would have, I am sure you would have heard of what is called Taylor's theorem.

Now, what is involved in Taylor's theorem, in Taylor's theorem we basically are interested in approximating a given function by a polynomial and a polynomial of as high degree as possible. Now, what is the highest possible degree that can go to, that depends on how many derivatives the given function f has. Let us begin with saying that suppose f has derivatives up to order n, then let us say that suppose f has derivatives up to order n, let us say in some interval.

(Refer Slide Time: 10:48)



Suppose, I call this interval i, in some interval i containing the given point a, now suppose we want to find a polynomial in such a way that we want to instead of taking the order n. Usually, the thing is if you want to find the polynomial of degree n, then in order to understand the difference between that polynomial and f that difference is given in terms for the derivative of order n plus 1. So, let us assume that it is up to order n plus 1 and suppose I take the polynomial p and x, suppose p and x is the polynomial and since we are interested in what is happening near a, we shall write this polynomial in terms of x minus a.

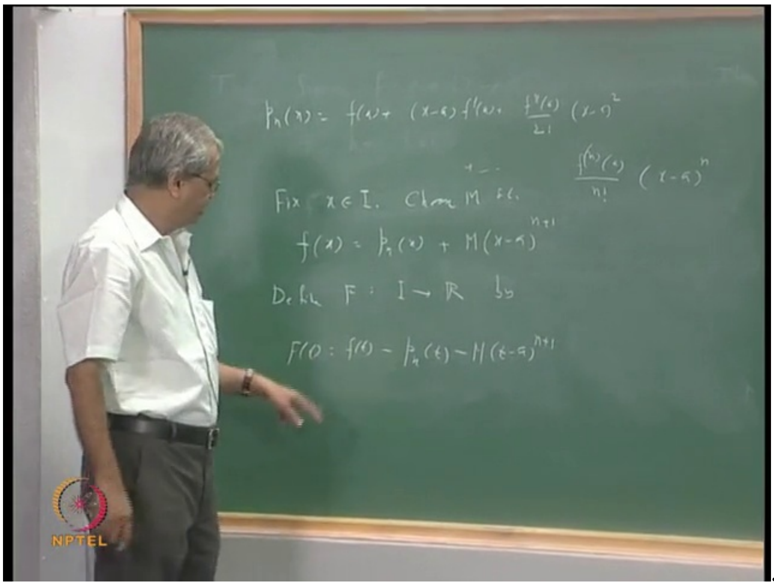
So, suppose the coefficients or let us say c naught plus c 1 into x minus a etcetera plus c 2 into x minus a square etcetera going in this way, the last will be c n into x minus a to the power n plus 1. So, how do we want to choose this polynomial, we want to choose it in such a way that the values of the function and its derivatives should coincide with the values of this polynomial at each points. So, that means first thing that is need is this f of a is equal to p n of a f of a is equal to p n of a. Similarly, I want let us say f prime a equal p n prime a etcetera up to n th derivative of f at a that should be same as n th derivative p n at the point a.

Suppose, I want to that is I want to choose a polynomial that satisfies this conditions, choose a polynomial means what choose the coefficients, choose the polynomial means choose the coefficients. Now, what we are saying is that this conditions immediately fix this coefficients c naught, c 1 c 2, there are n plus 1 coefficients, there are n plus 1 conditions here, how does that fix. Let us look at this, what is p n of a p n of a is c naught, p n of a is c naught, so that immediately says c naught must be same as f of a and what about what about p n prime at a.

You have to differentiate this p n prime at a, so it would be this this c 1 plus c 2 type 2 c 2 into x minus a etcetera, so when you put x equal to a all the other terms will vanish. So, p n prime a that is equal to c 1 which is same as say that c 1 is same as f prime at a let us do for one more if you look at p and double prime at a.

Suppose, I want this, then when you differentiate two times, this will be 2 c 2 into x minus a 2 c 2 into x minus a and then all the other terms will involve x minus a. So, p n double prime will just be 2 c 2, it will be just 2 c 2, so p n prime a will be just 2 c 2 or which is same as saying that c 2 is f prime a divided by 2, sorry f double prime a divided by 2. In the same way, you will be able to put that c 3 is f double prime a divided by 2 into 3 which is factorial 3 etcetera. In general, you will get this relations c is a is equal to z th derivative of a divided by factorial j.

(Refer Slide Time: 16:32)



So, using all this what can say that is polynomial P n x becomes f a plus x minus a into f prime a etcetera plus f double prime a divided by factorial 2 into x minus a is whole square etcetera. The n th term will be nth derivative of a divided by n factorial multiplied by x minus a to the power n. This is just a basically elementary calculation just using this conditions and calculating the values of c naught, but Taylor's theorem is not about that what Taylor's theorem says is that how to express the difference between P f and P n x Taylor's theorem say.

That difference is expressed in terms of n plus once derivatives and in order to express that difference we are going to use either mean value theorem or Rolle's theorem repeatedly. Now, how does one do that, so suppose I want to write the difference let us say write a difference in this way f x is equal to or f x is equal P n x plus some term here P n x because we do not expect that f x will be equal to P n x for all values of x unless if itself is a polynomial unless the thing like this will be some term.

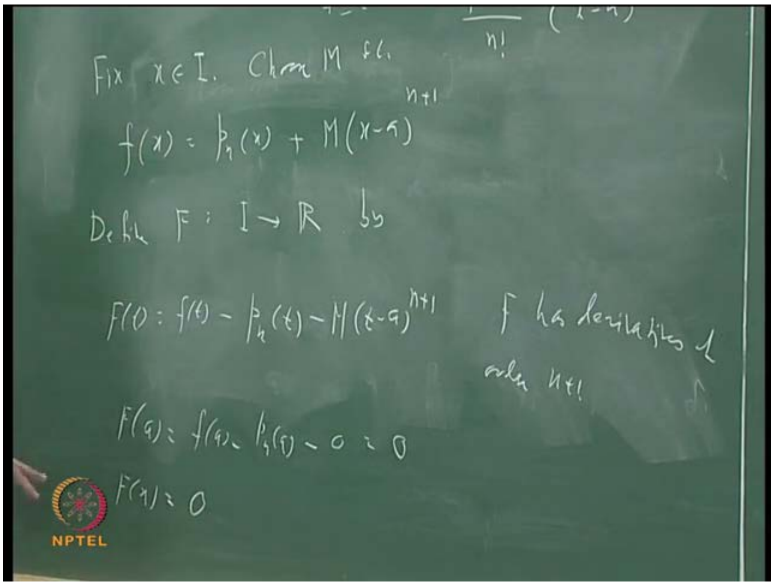
So, to express the difference in fact that will also depend on what is x the difference will also depend on what is x. So, what will do is that let us to begin with let us just fix some x fix some x in i what is i, i is an interval containing the given point a then fix some x. Then, we will choose numbers such that this difference can be expressed as P n x plus something like m into x minus a to the power n plus 1 x minus a to the power n plus 1. Then, we shall say how we can find m, so let us say fix x and I choose m such that f of x is equal P n x plus m into x minus a to the power n, of course you cannot that the same m will work for all x.

This m will also depend on x, but now we have fix x, this m works for that, so we can say that m is nothing but f x minus P n x divided by x minus a to the power n plus 1. Now, we will do something similar that we have done in case of proof, the mean value theorem what we did there was that we constructed an auxiliary function and applied Rolle's theorem or mean value theorem to that function. Now, will define auxiliary function, suppose I call that auxiliary function is F, so define F, define b, F from i to r as follows, F since we already we fix x for variable.

So, I shall use some other letter so let us suppose I call F at t suppose that is same as P n at t plus m into x minus not x, I will say t minus a t minus a to the power n plus 1 F of t. If I simply say that it will have no relation with the function small f it has to depend on small f of t minus this. Now, what can we say about the properties of this function F, after all what does this F depend on? It depends on f P n and this t minus a to the power in fact this is also polynomial.

So, these two are polynomial, so they will have derivatives of as many orders as you want and this small f has derivatives of n plus 1. So, F also have derivatives of n plus 1 in the interval i, so let us just take that purpose f has derivatives of order n plus 1 that means all derivatives up to n plus 1.

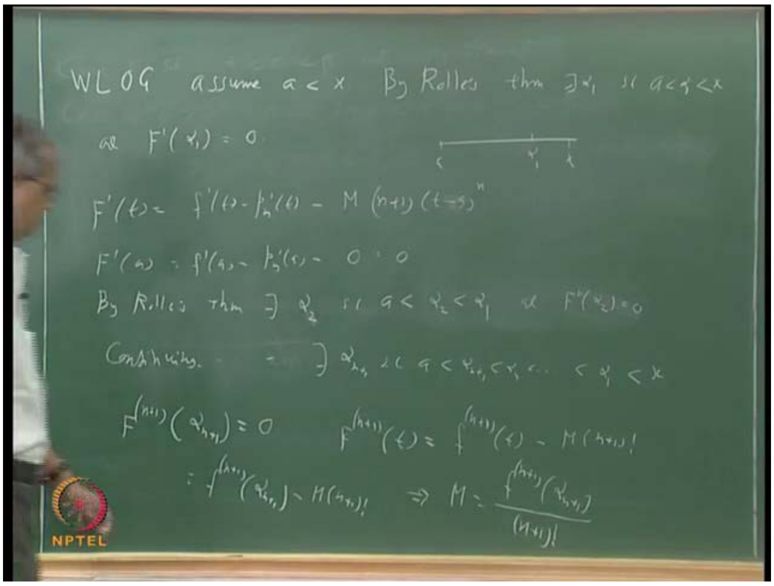
(Refer Slide Time: 21:56)



One more thing what is F of a, what is F at a it is small f at a minus this P n at a, put t equal to a here, so that is 0 so minus 0, so what is F of a minus P n of a that. In fact, that is our starting point both of them, so this is 0, so f of a is 0, what is F of x. You look at this so F of x big that F at x is f of x minus P n x minus m into x minus a to the power n plus 1 that is 0 that is for choosing m in this manner this choice of m use F at x is equal to 0.

Now, what is the obvious thing to say after this, you have a function F which have derivatives of sufficient order and it is 0 at a and x, so what follows from Rolle's theorem that there is some point lying between a and x. Let us assume that x is bigger than a, there exist some point lying between a and x such that F prime at that point should be 0. So, suppose we all again this big x may be either bigger than a or it may be smaller than a, but it does not matter, so again I will say that as we did earlier.

(Refer Slide Time: 23:43)



Let us say without loss of generality we will use this argument again without loss of generality assume a less than x a less than x then there exist a point. Suppose, I call that point alpha 1 such that a is less than alpha 1 less than x such that at that point F prime at alpha 1 is 0 at that F prime at alpha 1 is 0. So, you can say that by Rolle's theorem, there exist alpha 1 such that a less than alpha 1 less than x and F prime at alpha 1 is equal to 0.

Let us go once is it clear, so we let us say this is the interval a to x, we have find some points alpha 1 such that at this point f prime at alpha 1 is 0. You look at the function F prime, so what is f prime f of t is given there, so differentiate that F prime at t that is equal to f prime at t minus P n prime at t and minus m into n plus 1 into t minus a t minus a to the power n. What can we say about F prime at a F prime at a is same as f prime at a c minus P n prime at a. This is 0, but what is F prime at a minus P n prime a because that is how we choose that is, we choose the polynomials to start with, so this is 0.

Now, you look at the function F prime at a F prime, it has the next because anyway F has derivatives up to order n plus 1, so the second derivative of f is this, so that this function is differentiable it is 0 at a and at alpha 1. So, again what follows from Rolle's theorem that there exist some point between a and alpha 1, suppose I call that point alpha 2 such that at that point, the second derivative will be 0. So, we can say that again by Rolle's theorem there exist alpha 2 such that a less than alpha 2 less than alpha 1 and F double prime at alpha 2 is equal to 0.

Now, do you get the idea, we will proceed in this way differentiate this once again you will get the second derivative of t will get the second derivative of t that will be F double prime t minus P n double prime t. This will n plus 1 into n into t minus a to the power a plus 2 and again by the similar conditions you will get the second derivatives of f at a is also 0. You will find some alpha 3 lying between a and alpha 2 such that f double triple prime at alpha 3 is 0, continue with this way, so continuing how many steps we can go continuing in this way?

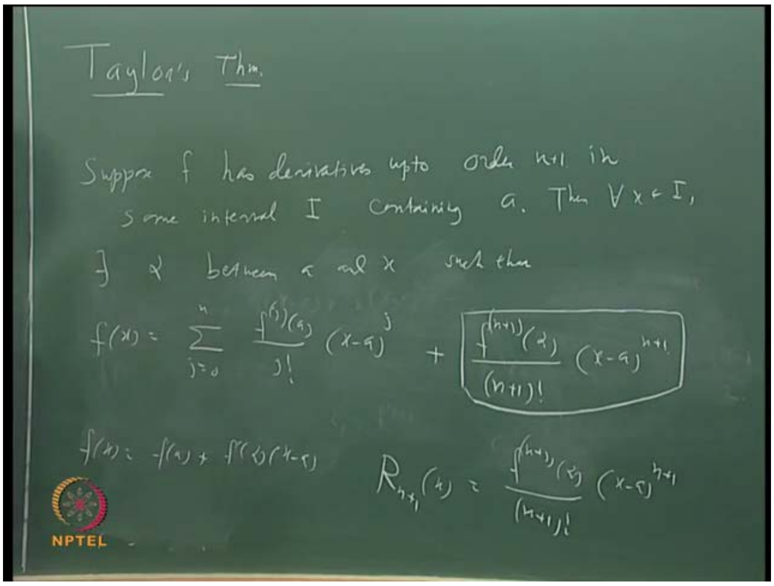
We can go as long you can differentiate it and this how many times you can do, you can do n plus 1 times, this you can do n plus 1 times because it has n plus 1 derivatives. If we do that, it has some more derivatives we could have preceded further, but we know only up to n plus 1, so continuing what we will get. There should exist some point proceeding in this way, we will call that point alpha n plus 1 alpha n plus 1 continuing will say that there exist alpha n such that no alpha n plus 1 a less than alpha n plus 1 less than alpha n etcetera less than alpha 1 and that alpha 1 is less than x.

So, all this point alpha 1 alpha n alpha n plus 1, they are all lying between a and x, so is that what should happen in n plus 1 th derivative n plus 1 th derivative at this alpha n plus 1 should be 0, but we should have some idea what is this n plus 1 derivative. We have differentiated once, twice here, suppose we do it n times n plus 1 times differentiate this n plus 1 times. So, what will happen so what is n plus one derivative of f at t n plus 1 derivative of f at t n plus 1 derivative of f at t what is n plus 1 derivative of this polynomial P n t. It is clear that that will be 0 it is a polynomial degree n, so its n plus 1 derivative will be 0, what about this it will be m into n plus 1 factorial.

So, minus m into n plus 1 factorial put it back here, so f f n plus 1 alpha at is 0 that is same as saying that this is equal to small f n plus 1, f at this point alpha n plus 1 minus m into n plus 1 factorial. That is same as saying that this m must be n plus 1 derivative at this point divided by n plus 1 factorial that is same as saying that this means that value of m is equal to m n plus 1 at alpha n plus 1 divided by n plus 1 factorial.

So, what did we say that let us come back what does, suppose f has derivatives of to order n plus 1 in some interval i containing a, then what did we prove that given x any x in that interval i. We can find a point in this case, the point is n plus 1 which lies between a and x such that at that point at that point x f x is given by this polynomial plus m into x minus a to the power m is this m is given by this f n plus 1 etcetera n plus 1. So, let us say we can we can now write the statement precisely.

(Refer Slide Time: 32:21)



Suppose, f has derivatives of up to order n plus 1 in some interval i containing a, then for every x in i there exist instead of calling that point alpha n plus 1 I will simply that point some other number or we can simply call it call that point alpha.

There exists alpha between a and x such that f of x is equal to P n of x let us write this P n of x given by this and we can use the summation notation to write in the short form so f of x is sigma. Let us say j going from 0 to n j is derivative at a divided by factorial j multiplied by x by a to the power j that is P n x that is P n x plus n plus 1 th derivative of f at alpha divided by n plus 1 factorial n plus 1 factorial multiplied by x minus a to the power n plus 1. This is the statement of the theorem and proof we already consider this time; we discuss the proof first, then finally given the statement does not matter by the way

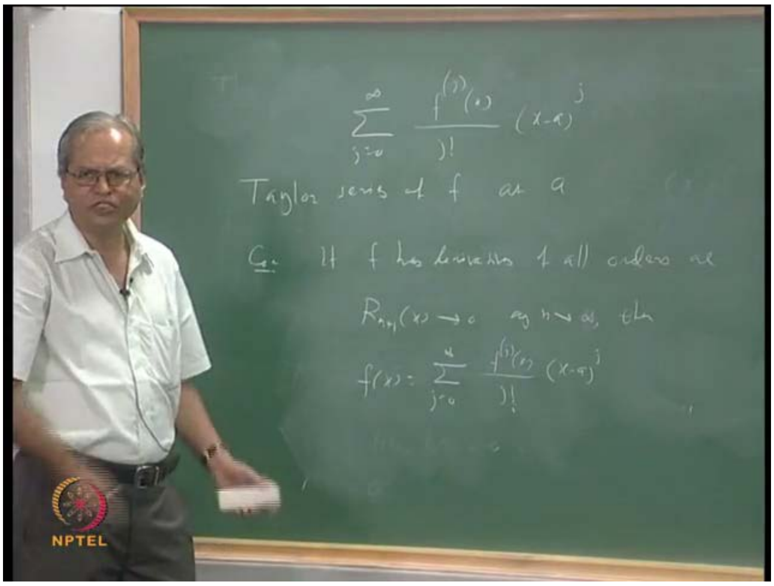
By the way though here are a few comments, first comment is if you can see that if n equal to 1, this is basically same as mean value theorem. Suppose, n equal to 1 what will happen to the right hand side f x is equal to 3 have only two term say j j equal to 0 and j equal to 1. So, what are those terms f x is equal to f a plus f a plus f a plus you can say if prime at alpha into x minus a or which is same as saying that f x minus f a divided by x minus a is equal to f prime at alpha for some alpha lying between x and a. That is just mean value theorem that is just mean value theorem, so that is why Taylor's theorem can be considered as extension of mean value theorem.

One more thing this term is called reminder after n terms, this called reminder after n terms, so usually denoted by this is R n plus 1 of x. So, R n plus 1 that is reminder after n terms R n plus 1 x and add it is it is f n plus 1 alpha divided by n plus 1 factorial multiplied by x minus a to the power n plus 1, of course. This is just a notation some books may be calling it R r and x also it does not matter it is a it is just a matter of notation. Now, you can consider the situation, suppose f has a we have this is what Taylor's theorem says if f has derivatives of order n plus up to order n plus one now suppose f has derivative each n.

You can write the expression like this and then you can say that suppose you consider an infinite series suppose an infinite series of f z a by z factorial multiplied by x minus a to the power j. Then, this pol what you what you have called polynomial P and x is nothing but what you will normally call partial sum of that series partial sum of that series. This reminder term will express the difference between the partial sum and the value of the function between the partial sum of the series and the value of the function.

Now, if you are able to show that for some value of x this see this for a for a fix value of x this R n plus 1 x will be a sequence. For each, value so this will be a sequence. So, if this sequence goes to 0, if this sequence goes to 0 what at what does it mean, it mean that for that particular x, this infinite series will converse what is let us see what is that infinite series.

(Refer Slide Time: 37:13)



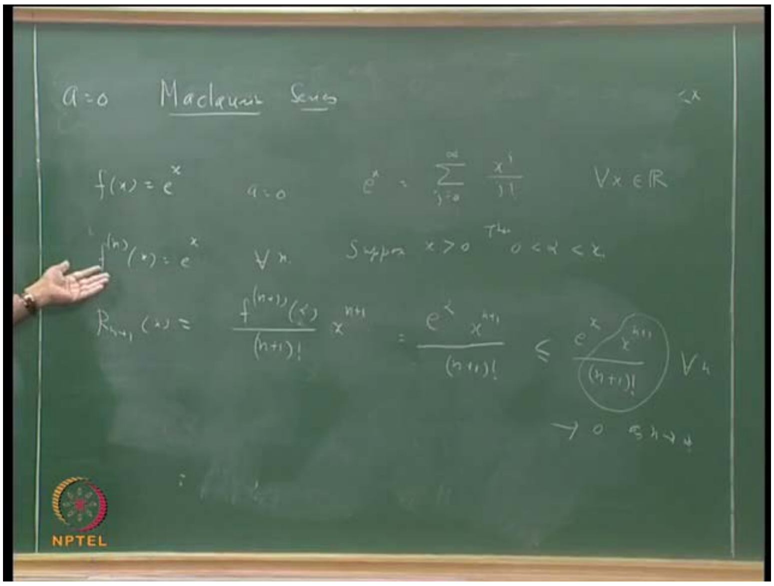
It is this sigma j going from 0 to infinity f j x divided by factorial j and multiplied by x minus a to the power j. This is what is call Taylor series of f at a Taylor series of f at some books also call around a or about a and things like that, but those are basically just matter of preferences of different authors. So, what we observe just now is that when does this series converge when does this series converse it converges.

If you can show that that sequence of that sequence of reminder terms sequence of reminder terms goes to 0 if the sequence of reminder term goes to 0. It means f x minus p n x goes to 0 which is same as saying that f n x that p n x converges to or which is same as saying that this is equal to f x.

So, that is the meaning, so let us just make that observation or we can write as a corollary if f has derivatives of all orders sometimes this is express by saying that f in infinitely differentiable. What we say that f is a derivative of all orders some people have the habit of calling that f is infinitely differentiable so f has derivatives of all orders and this R n plus 1 x tends to 0 as n tends to 0. Then, f x is equal to the sigma j going from 0 to infinity f j x by factorial j multiplied by x minus a to the power j. So, this that means the Taylor series converges to the values of the function f, remember this may or may not happen this may or may not happen for all function there have various possibilities that is the Taylor series may or may not converge.

It may converge, but it may not converge to the value of the function f all those possibilities are there, but none of those things happens if R n plus 1 x tends to 0 as n tends to sorry n tends to infinity. Now, in various functions we can check this in case of various functions we can check this, let us just take one or two examples one more thing. It is just a matter of notation here we have taken this a this a is called the center of the power series.

(Refer Slide Time: 40:37)



Instead, If we take a is equal to 0, then the corresponding series is called as Maclaurin's series if a is equal to 0, this is the corresponding series is called Maclaurin's series and this is what we discuss more often.

So, in order to see whether this happens or not for a particular function, we should take some function which have derivatives of all orders, of course there are many such functions. If you can of, course takes polynomial, but in that case the whole question is trivial in case of polynomials. What is going to happen if the polynomial is of degree n then this R n plus 1 x will be 0 for all x after n plus 1. Obviously, the Taylor’s series will coincide with that given polynomial.

So, let us take something else so the simplest such function is exponential function take f x is equal to e power x, then what is what is so simple about this function first of all it is differentiable any number of times.

Not only that all its derivatives are also coincide with the given function, so we know that f and x is also same as e power x for all x for all n, so for any x for any x. If you look this R n plus 1 x what is that it is it is suppose we are considering Maclaurin's series suppose we take a is equal to 0. So, instead of this, this will be R n plus 1, x will be let us it will be f n plus 1, alpha f n plus 1 alpha divided by n plus 1 factorial and multiplied by at since a. We have taken at a to be 0 will say x to the power n plus one x to the power n plus 1 now.

If f x is e power x this f n plus one derivative of f at alpha is going to be e to the power f, so this is nothing but e power alpha multiplied by x to the power n plus 1 divided by n plus 1 factorial. Now, out of these this e power alpha is now independent of n e power alpha is it is not clearly speaking independent of n because it will depend on n because of because what is that alpha it is the one which is given by that by Taylor’s theorem. It is not independent of n, but we can say we know that alpha lies between a and x and in our case for example, suppose x is bigger than 0. Then, suppose x is bigger than 0 suppose x is bigger than 0, then then we will have 0 less than alpha then 0 less than alpha less than x.

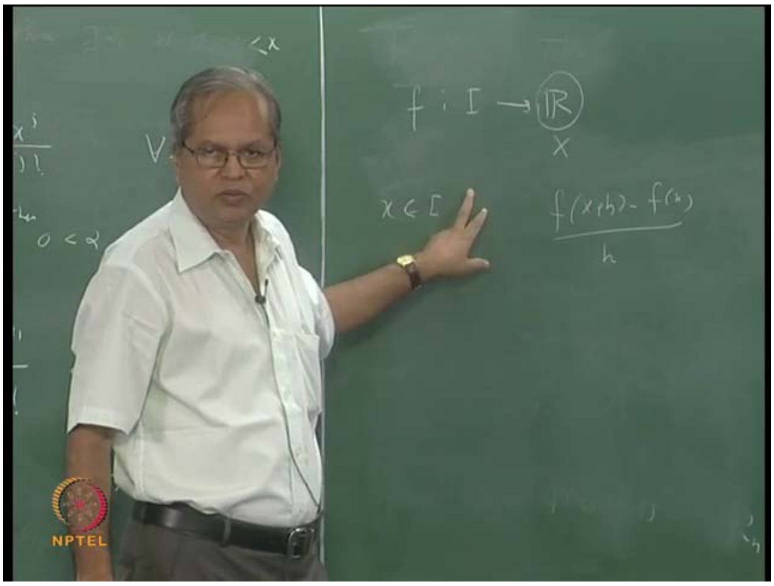
Then, alpha has to lie between 0 and x and let us assume that we know that e power x is monotonically increasing. So, e power alpha will be less than e power x e power alpha will be less than e power x, so we can say that this s less than or less not equal to e power x multiplied by x to the power n plus 1 divided by n plus 1 factorial. Now, we can say and this is true for all n, it does not depend on n this term no does not depend on n this term. So, whether this tends to 0 or not depends on whether this part x to the power n plus 1 divided by n plus 1 factorial whether that tends to 0 or not as n goes to infinity and what is the answer for that it is well known.

In fact one can say that this is a this is a n th term of a power series which converges everywhere which we can show by a ratio test etcetera, but I mean that is making a thing more complicated. In a much simpler way, you can show that this sequence goes to 0 as n goes to infinity. So, one can say that power series Maclaurin's series for e power x converges and because this remember what is the series then series is given by this f to the power f j th derivative of x divided by factorial j multiplied by x. Since, in this case j th derivative is again e power x, so we can say that what is what is says that because we will get that f x is equal to sigma j going from that is e.

In this case, e power x is equal to sigma j going from 0 to infinity x to the power j divided by factorial j and this is true for all x for all x in r. So, whatever we have done in case of this exponential function in a similar way you can you can do for many other well known functions which has derivatives. For example, sin x cos x etcetera, what is the property of sin x or cos x functions their derivatives are again in the same forms sin x and cos x.

So, suppose you come up to this step this will either cos alpha or sine alpha, then we know that that has to be less not equal to 1 that has to be less not equal to one and all other considerations will remain same. Now, so far we have discuss the derivatives of functions which are defined on some interval in the real line and take the values also in the real line.

(Refer Slide Time: 47:19)



So far, we have been talking of functions of the type f from i to r f from i to r, now suppose I want to change this co domain. Suppose, I want to instead of this R suppose I want to take the functions which takes values somewhere, now what is what is the kind of structure that we can think, of course we can we can think of the functions for example, instead of real value function. I can think of complex valued functions or one can take the function because which takes the value R 2 R 3 or it can be set of all metrics this can be some set of functions.

So, in general one could think of replacing this R by some other metric space, there is slightly problem there you cannot take any arbitrary metric space because suppose you want to take side. Suppose, we take some x in i and suppose I want to talk of derivative of the function at this point x, then you have to think of limits of elements like this f of x plus h minus f of x divided by h, of course diving by h is no problem.

So, the metric space should have this property that there should be something like x plus h, suppose you have x and h both R in x. This will be something like x plus h and there should be something 1 by h etcetera F of in short there should these operations of addition of two elements in that and multiplication by scalar and that elementary is that. I think by now you would have understood that what we need there is this must be at least a vector space. This must be a and in addition to that that vector space we have to talk convergence, so it should also be a metric space and it also cannot happen just arbitrary vector space.

You are given some arbitrary metric ton that that that vector space and that metric should have some relationship with each other and perhaps you would have heard of what is called non linear space. So, that is what we can say that this R can be replaced by any non linear space x, so in particular we can think of functions which takes values in complex numbers are 2 R k etcetera. We will not go into very detailed discussions of functions of that kind, but we shall see some few elementary things about this and what are the common things between the derivatives of the usual functions at those kinds of functions. What are the main differences, but that something, we shall discuss in the next class.