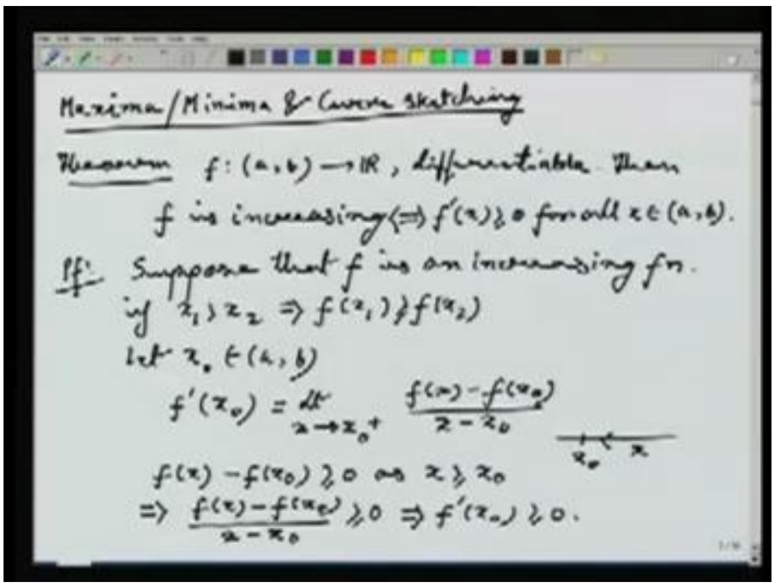
Mathematics-I

Prof. S.K. Ray
Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur

Lecture – 10
Maxima, Minima

Today we are going to start with maxima and minima and convex sketching.

(Refer Slide Time: 00:27)



First, we start with the behavior of the derivative and the behavior of the function. How can we join these two things together? So the first theorem we are going to deal with is as follows: Suppose f is a function on the open interval [a, b] to R. We assume that f is differentiable. So f is a function from [a, b] to R and we assume that f is a differentiable function. Then the following thing is true, that f is an increasing if and only if f prime x is bigger than or equal to 0 for all x in the open interval [a, b].

We will see that one part of the result just follows from the usual definition of derivative of the function and the other part uses mean value theorem which we have proved in the last lecture. That is, you can think of this theorem as an application of the mean value theorem. So let us start with the proof of the function, of this result.

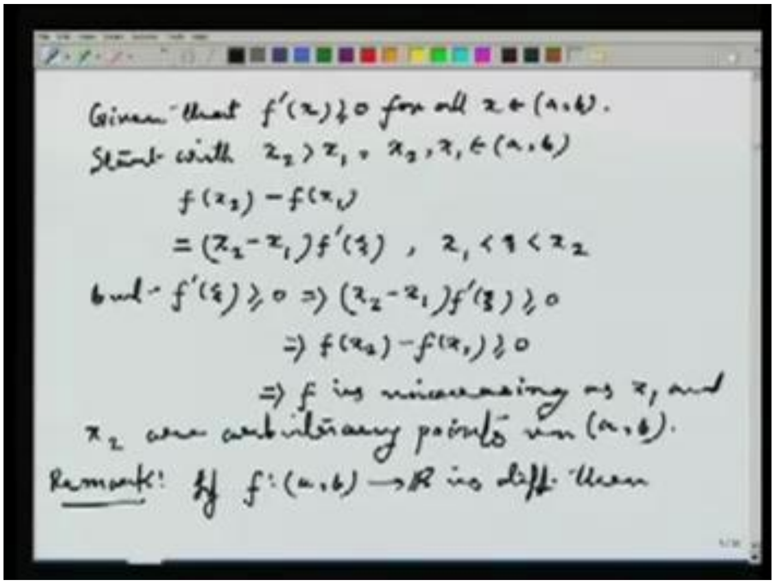
Suppose that f is an increasing function. That is, if x 1 is bigger than x 2, then I know that f x 1 is strictly bigger than f x 2. Since I just said increasing and I did not say strictly increasing, I should say f(x) 1 is bigger than or equal to f(x) 2. Now that is given. Let x naught be an element in the open interval (a, b). Then I look at the definition of f prime at x naught. This, by definition, is limit x going to x naught plus f(x) minus f(x) naught divided by x minus x naught.

Now notice that in the definition, actually we have limit x going to x naught plus f(x) minus f(x) naught divided by x minus x naught but since we are assuming that f is differentiable at x naught, one sided limit is good enough to look at. So I have looked only at the right hand side. The meaning is that x approaches x naught from the right hand side of (x) naught. So in picture, it will look like this that I have x naught here, x is here. It is going towards x naught. Now since I have that f is increasing, this would imply that f(x) minus f(x) naught bigger than or equal to 0 as x is bigger than or equal to x naught.

This would then imply f(x) minus f(x) naught divided by x minus x naught. Notice that x is bigger than or equal to x naught. So x minus x naught is also non-negative. This is bigger than or equal to 0. Since all the terms are bigger than or equal to 0, that would certainly mean that the limit which f prime x naught is bigger than or equal to 0. That is one part of the result. That is we have assumed that f is increasing and from that it follows that the derivative at any point x naught in that interval is also bigger than or equal to 0. Now I want to go about the other part.

That is, it will be given to me that the derivative at each and every point is non negative from that I want to infer that f increasing.

(Refer Slide Time: 05:03)



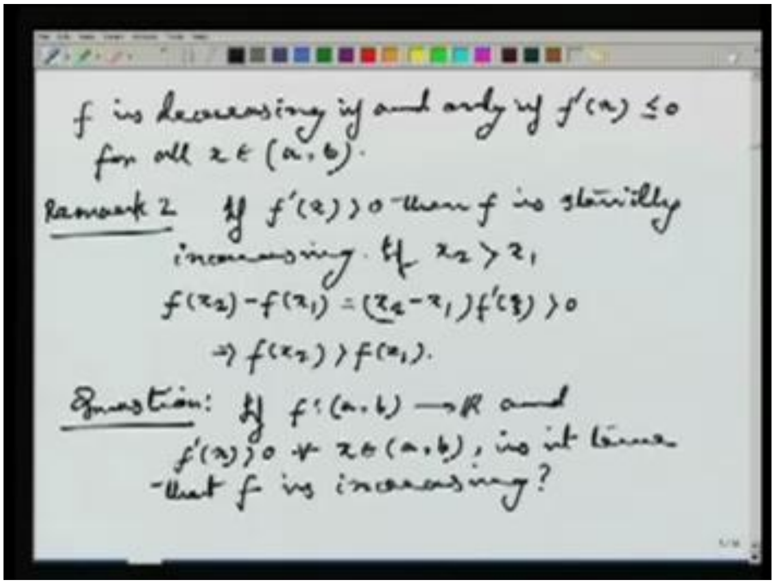
Given that f prime at x is bigger than or equal to 0 for all x in the open interval [a, b]. So if I want to show that f is increasing what I have to do is, start with some x 2 bigger than x 1 such that x 2 and x 1 both are in the open interval ab. That is, they are in the domain of the function, so that applying f on these two points do makes sense. Now if I have to do this, what I have to do is I have to look at f(x) 2 minus f(x) 1. Somehow I have to show that this is non-negative. If I do that, it will prove that f is increasing but at this point I can apply mean value theorem which I have already proved.

The statement of the mean value theorem says that f of(x) 2 minus f of(x) 1 is equal to x 2 minus x 1 f prime at a point xi where xi lies between x 1 and x 2 but by the given condition I anyway know f prime xi is bigger than or equal to 0. This, then it would imply that x 2 minus x 1 which is anyway nonnegative because x 2 is bigger than x 1 times f prime of(x)I, this is bigger than or equal to 0 but that would imply f of(x) 2 minus f of(x)1 is bigger than or equal to 0. This implies f is increasing as x 1 and x 2 are arbitrary points in the interval [a, b].

Just for two points, x 1 and x 2 if I check this inequality, that does not suffice that the function is increasing. All we have to do is choose arbitrarily two points, which I called x

1 and x2 and then you try to prove that f(x) 2 minus f(x) 1 is bigger than or equal to 0 which I could do. That means f is increasing. That means, to understand given a differentiable function whether the function increasing or not, all I have to do is I have to check the derivative of the function at each and every point. If the derivative turns out to be non-negative then the function is increasing and it is equally easy to show which I put as remark.

(Refer Slide Time: 08:24)

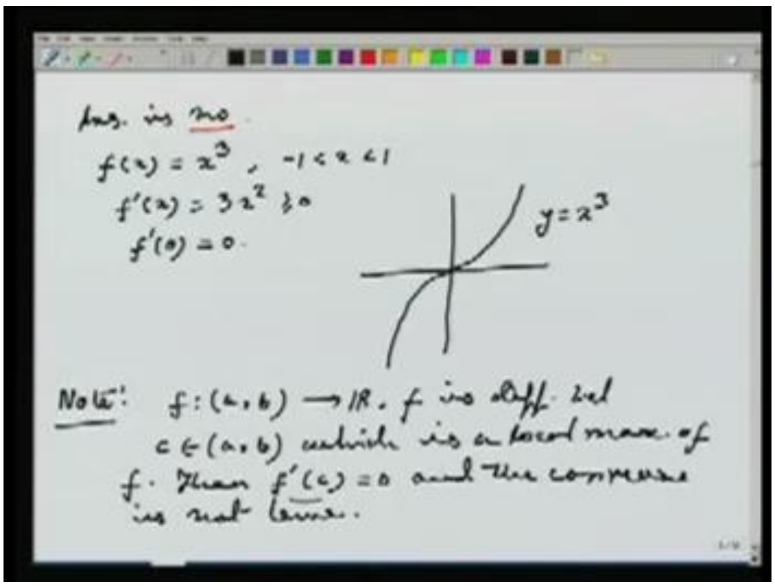


If f from open interval (a, b) to R is differentiable, then f is decreasing if and only if f prime x is less than or equal to 0 for all x in the open interval (a, b). The proof is exactly similar as in preceding cases or what one can do is given a decreasing function f, construct a new function g which I will call f, then that function will be increasing and then the previous theorem certainly apply to it. What about strict inequality? Again, I will put it as a second remark.

If f prime at x is strictly bigger than 0, then f is strictly increasing follows again from the mean value theorem. Actually it is if and only if. This is not actually if and only if. If f prime at x is bigger than 0 then f is strictly increasing. It is very easy to prove, if you look at f(x) 2 minus f(x) 1. Write it as x 2 minus x 1 times f prime of(x) i. Then I know anyway this is strictly bigger than 0 by the assumed condition. That would imply f of(x) 2 is strictly bigger than f of(x) 1. But what about the converse? Is it true that if f is strictly increasing then the derivative every where is strictly bigger than 0?

So the question is, if f is from [a, b] to R and f prime x strictly bigger than 0 for all x in the open interval (a, b) is it true that f is increasing? The answer actually is no and it can be shown by an example which we have dealt already with. The answer is no.

(Refer Slide Time: 11:18)



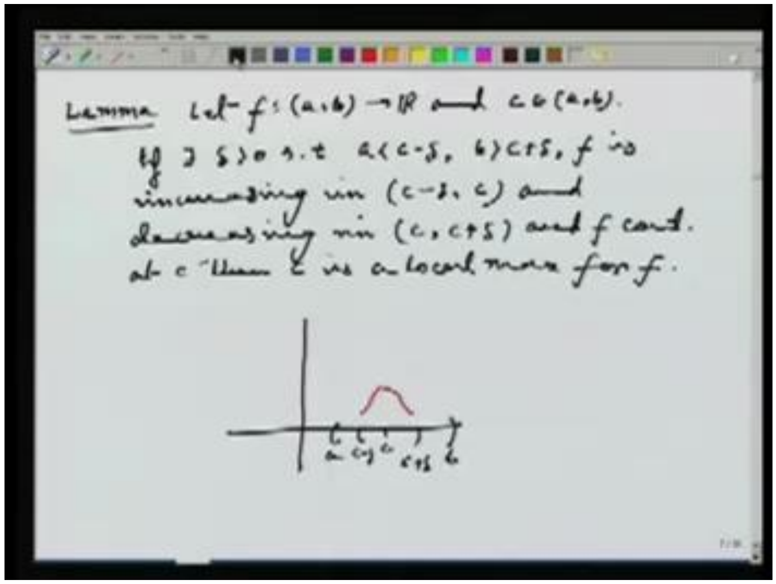
If you look at the function f(x) equal to x cube where minus 1 less x less 1, then I look at the derivative f prime x that is 3 x squared; it is bigger than or equal to 0. Actually f prime at 0 is 0 but what is the function if I try to draw the graph of the function. It looks exactly like this. We will see after few minutes how do I draw the curve of this, but just assume for the time being that this is curve of y equal to x cube which can be shown. Then you can see the function is strictly increasing.

Nevertheless at the point 0, the derivative is not positive; it is equal to 0. That means, it can happen even if you have function which is strictly positive, differentiable but at certain points its derivatives can vanish. It is not necessary that derivative will be strictly bigger than 0. The converse of this, that is, if you have a differentiable function whose derivative is always strictly bigger than 0, then the function is certainly strictly increasing. Let us observe something more which we have already seen. So I will call it a note.

Suppose I have a function f defined on an interval (a, b) and assume f is differentiable. Let c is a point in [a, b] which is a local maximum of f. Then we have proved in the previous lecture that then the derivative of the function at that point has to be equal to 0 and the converse is not true. It is again given by the previous example y equal to x cube. I can look at the function y equal to x cube. Its derivative at 0 is 0 but if you look at the graph of the function y equal to x cube as drawn here, the point is neither a local maximum nor a local minimum but the derivative is 0 there.

That means the derivative of the function at a point being 0 does not guarantee that the point is either local maximum or local minimum. So the question we are asking now is what other condition we can put on the derivative which will guarantee the point is either a local maximum or local minimum. Well, it turns out that staying just with in the derivative is not that neat. One has to go to the double derivative. That is the thing we are going to try now.

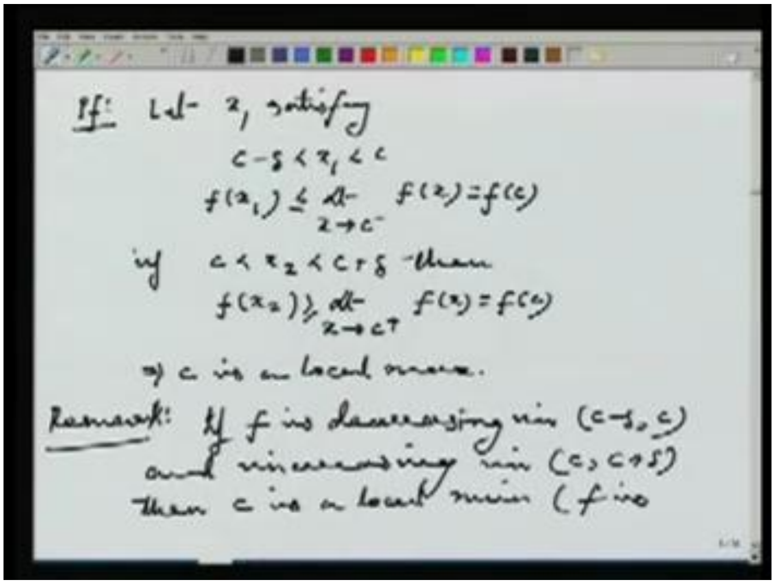
(Refer Slide Time: 14:47)



Let us first observe the following thing which I will call it as a lemma. Let f be a function defined on the open interval [a, b] to R and c belonging to [a b]. Now suppose if there exist delta bigger than 0, such that a is less than c minus delta, b is bigger than c plus delta, f is increasing in c minus delta c and decreasing in c, c plus delta and f is continuous at c, then c is a local maximum for f.

So let us draw a picture which will help us understanding the statement clearly. This is a this is b now this is c, this is c minus delta, this is c plus delta and on c minus delta, the behavior of the function is that is increases and on c, c plus delta the behavior of the function is, it decreases. Then it is believable at the point c the function is having the local maximum but this need to be proved now. So this is how we go about it.

(Refer Slide Time: 16:54)

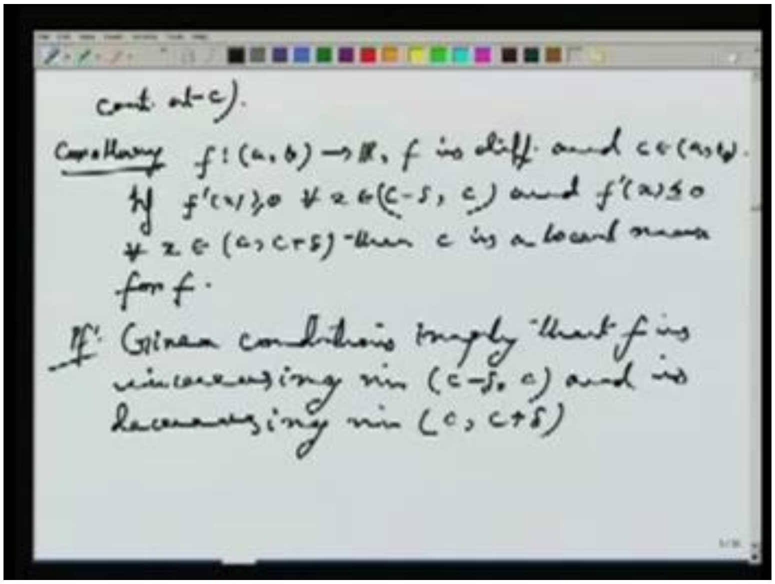


Let x 1 satisfy c minus delta less x 1 less c. Then I can say that f(x) 1 is certainly less than or equal to limit x going to c minus f(x). Why so? Because when I look at x going to c minus it means that I am approaching c from the left hand side. Now if you think in terms of the sequences, the corresponding values for all these f(x)s are increasing because I said that on c minus delta c, the function f is increasing.

Now then the sequence will converge to the supremum which has to be anyway bigger than or equal to f of(x) 1. That is what is written here but now I can use the fact f at c is continuous. That means this is equal to f c. That means if I choose any x 1 in c minus delta c f of(x) 1 is less than or equal to fc. Similarly x 2 satisfies, then I know that f of(x) 2 has to be bigger than or equal to limit x going to c plus f (x) which again by continuity is equal to f c. Now this implies c is local max.

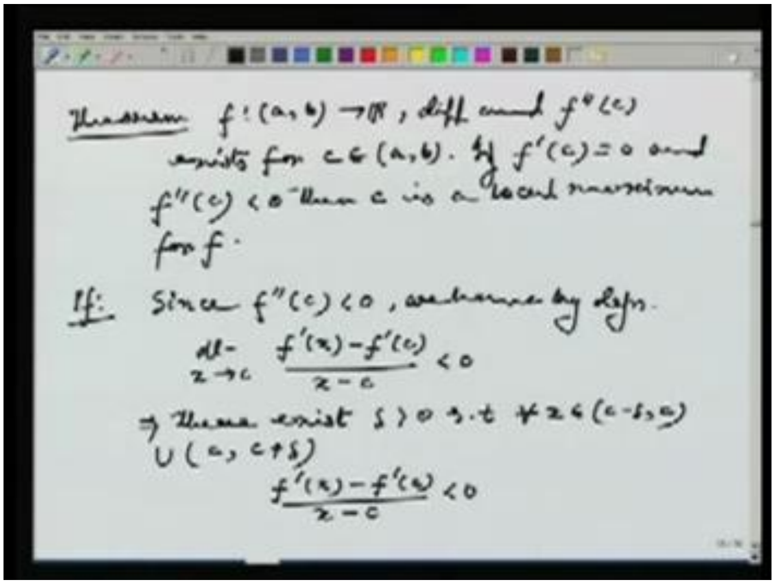
Now again as a remark I will tell you what happens in the case of minimum, that is, if f is decreasing in c minus delta c and increasing in c, c plus delta then c is a local minimum. Of course, our assumption is f is continuous at c.

(Refer Slide Time: 19:28)



Now this result has an interesting corollary. Now the corollary is that suppose f is from a, b to r, f is differentiable and c belonging to a, b. If f prime x is bigger than or equal to 0 for all x in c minus delta c, and f prime x is less than or equal to 0 for all x in c, c plus delta then c is a local max for f. It actually follows from the previous result which we have done. Proof is very simple that the given conditions imply that f is increasing in c minus delta c and is decreasing in c, c plus delta then I can use the previous lemma which would imply that c is a local maximum for the function. Now these can be used in the next theorem which is very useful and time and again we apply that. So this I will mention as a theorem.

(Refer Slide Time: 21:30)

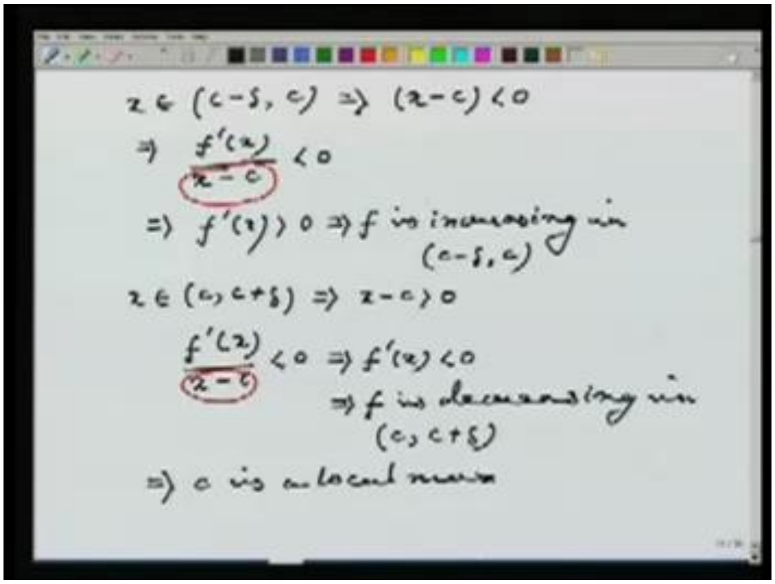


This is usually called the double derivative test of maximum and minimum. Let f from a,b to r again, differentiable and f double prime c exist. If f prime c is equal to 0 and f double prime is strictly less than 0, then c is a local maximum for f. So the point is that I have defined the function f on the open interval (a, b). Suppose there exist a point c where f prime c is equal to 0, if f is twice differentiable at c and the double derivative of f at c is less than 0, strictly less than 0, that is important then c turns out to be local maximum for the function f.

Let us go for the proof of it. Most of the times, while talking about maximum and minimum we only talk about this result but notice here one thing that the point c where we are testing for the maximum and minimum that point c is the interior point of the interval (a, b). That is, c not an end point of the interval. That means, if you have function whose maximum and minimum occur at the end points for those, this theorem will not apply. This theorem applies only in the case searching for maximum and minimum in the interior of the interval where f is defined.

Well, now since f double prime c is strictly less than 0, we have, by definition, limit x going to c f prime x minus f prime c divided by x minus c to be less than 0. This is the definition. This would imply, there exist delta bigger than 0 such that for all x in c minus delta c union c, c plus delta, we have f prime x minus f prime c divided by x minus c is less than 0.

(Refer Slide Time: 24:58)



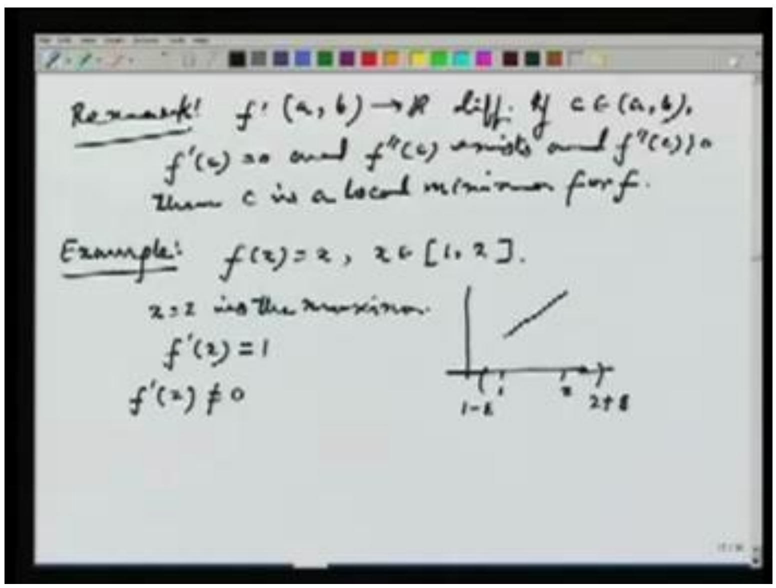
Now let us look at the cases when x belongs c minus delta c. What does it mean? Now x belonging to c minus delta c, let us look at this case first. This implies, f prime x, I can forget f prime c because f prime c is 0, one of my assumption, divided by x minus c. This is less than 0. Now what does this mean, because x is in c minus delta c, this implies anyway that x minus c is less than 0. That means, the denominator here is less than 0. The total quantity is anyway less than 0. That means the numerator has to be positive or at least non negative. This would then imply that f prime x is strictly bigger than 0. So f is in c minus delta c. That means, f prime x is strictly bigger than 0, but this then implies that f is increasing in c minus delta c.

On the other hand, if I choose x in c, c plus delta, notice that this would imply that x minus c is strictly bigger than 0. Then again, from the condition, on the double derivative I have that f prime x divided x minus c is strictly less than 0. That anyway is given to me.

Now I have the numerator and the denominator and the denominator is positive but the whole quantity is negative.

That would then imply that the numerator f prime x, this must be strictly less than 0 which then implies that f is decreasing in c, c plus delta. So the situation is, I have the point c on c minus delta c, f is increasing. On c, c plus delta, that is on the right side of c, f is decreasing but then, in the previous case we proved that if f is continuous at c. that means, c is a local maximum. But f is differentiable that c, so it is certainly continuous. This would then imply by the previous lemma, c is a local maximum.

(Refer Slide Time: 27:48)



Analogously again, I can deal the case of minimum that is given as remark that f from a b to r differentiable, if c belongs to a b with the condition f prime c is equal to 0 and f double prime c exists and positive, then c is a local minimum for f. Now let us look at the following example. f(x) is equal to x, if x belongs to closed interval 1, 2. You see that f is a nice, differentiable function. I can easily draw the graph of the function also. This is the point 1, this is the point 2 and at the point 1, the value of the function is 1. At 2, the value of the function is 2.

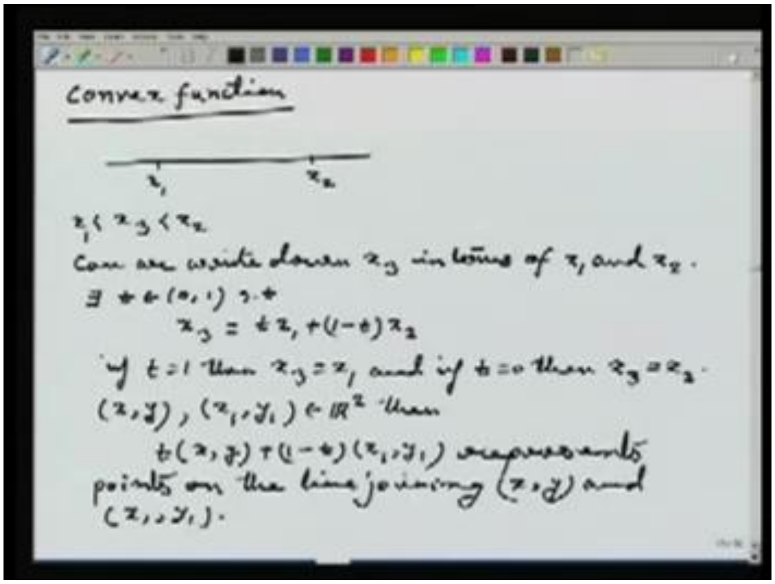
So it look like you can see the point x is equal to 2 is the maximum. But if I look at f prime x, that is continuously equal to 1 for all x, it is the constant function 1. So f prime 2 is not equal to 0. But I said that if I have a local maximum, then at that function has to be equal to 0. x is equal to 2 is the local maximum for the function but the derivative is not 0. Why it is happening?

Well, if look back at the proof of the theorem where we have proved that at the local extremum, the derivative of the function is 0. It assumes that the point which we are looking at, that is the extremum the maximum or the minimum, whatever you say, that point has to be in the interior of the domain of definition of f. That means, it cannot be a boundary point. But look at the situation here, I have my function f(x) is equal to x, the maximum is happening at the boundary point x is equal to 2 and that is why the derivative is not equal to 0.

You notice the same thing is happening at the minimum also. x is equal to 1 is the minimum of the function and it attain at the end point, one which is not in the interior of the function and hence the derivative at 1 is also is not equal to 0. Now you might say well, I might increase my intervals little bit, that is, I look at 1 minus epsilon and increase this to 2 plus epsilon. Then the point 2 is an interior point, but the derivative is not 0 there, which is obvious, because if x is equal to 2, then f(x) is neither maximum nor the minimum. It ceases to be the local maximum and hence the derivative is not 0.

So, no contradiction in the previous theorem but this is just to remind you that have to be careful about the situation of the local maximum or the local minimum. If it is not in the interior, none of the theorems which we have proved so far works. At the boundary points, the theorems do fail. Now we want to talk about something called convex function.

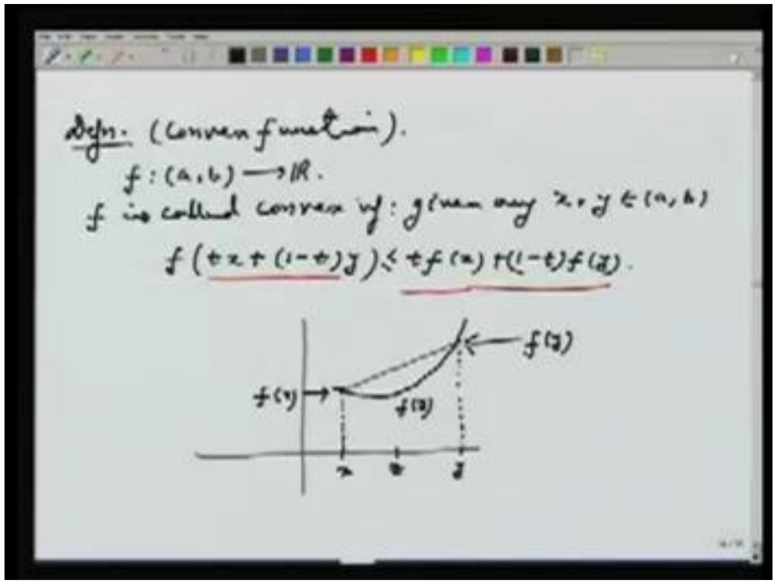
(Refer Slide Time: 31:33)



We will first start with convex combination of points. The question is this that suppose I have the real line. This is a point x 1 and this is point x 2. Now if I look at x 3, which satisfies it is bigger than x 1 and it is less than x 2, then is there any way of expressing the point x 3 in terms of x 1 and x 2? The question is, can we write down x 3 in terms of x 1 and x 2? Well, it turns out to be, the answer is, there exist then t in the open interval 0 1, such that, x 3 is equal to t x 1 plus 1 minus t x 2, which I leave as an exercise for you to verify that given any x 3 between x 1 and x 2, there exist a t in the open interval 0, 1, such that, x 3 is t x 1 plus 1 minus t x 2.

Well, if x 3 turns out to be equal to x 1, what you do is, you just take t to be equal 1. So if t is equal to 1, then x 3 is equal to x 1 and if t is equal to 0, then x 3 is equal to x 2. The same thing works in the Cartesian plane also that if I have 2 points x, y and x 1, y 1 in r 2, then the points of the form t times x y plus 1 minus t x 1 y 1 represents points on the line joining x, y and x 1, y 1. Now with this preparation, we come to the definition of convex functions. This is the definition of convex function.

(Refer Slide Time: 35:00)

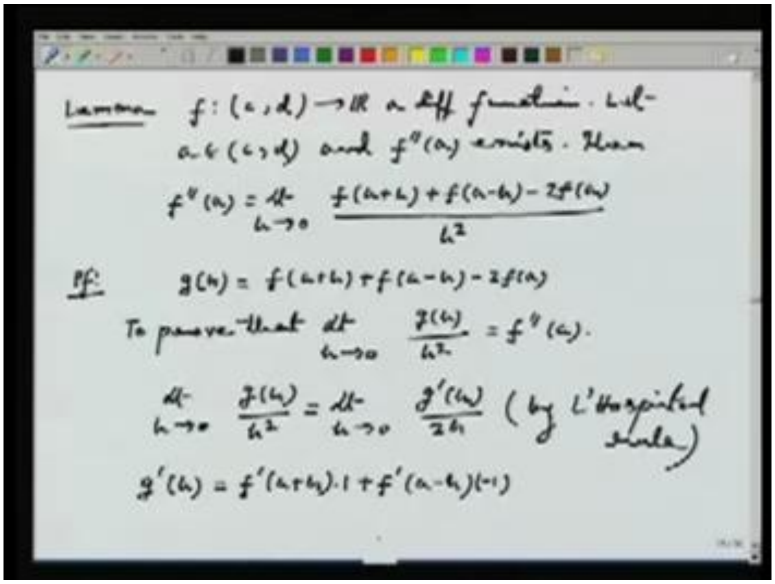


Suppose, f is a function on the open interval (a, b) to R. It might be closed interval also, but open interval suffices for us. Then f is called convex, if given any x and y in the interval (a, b). f of t x plus 1 minus t, y is less than or equal to t times f(x) plus 1 minus t times f y. Well, let us try to understand the meaning of this. First, look at the quantity t x plus 1 minus t y. What does this mean? It represents the points on the line joining x and y. That means it is all the points between x and y. If I look at the right hand side, that is here, it represents all points on the line which joins f(x) and f (y). So in the picture then, what should it mean? I draw the axis. Let us say this is x and this is y, then t x plus 1 minus t y represents points between x and y.

Now let us say this is this height is f(x) and let us say this is f(y). Now then, if I look at the right hand side of the above equation, the right hand actually represents the line joining f(x) and f(y). Now the definition of the convex function says, if I take in between points, for example, if I take z here, then f(z) should lie below the above line. f(z) should be somewhere, that is, it is saying, the given definition is saying that the graph of the function then must look like this. That is, given any 2 points, you look at the graph of the function, then at the points (x, f(x)) and (y, f(y)), look at those two pair of points, join them by a line, then the graph of the function lies below that line. That is what the definition of convexity says.

Now just in these form it might become difficult to understand which functions are actually convex. So to understand this criteria, to make it more applicable, what we do is, we assume the differentiability of the function and we try to device criteria involving differentiability which will produce the convexity of the function, that is our aim in the next theorem. Now let us first observe the following thing about the derivatives of functions. So I will call it as a little lemma, the proof is very easy.

(Refer Slide Time: 39:09)



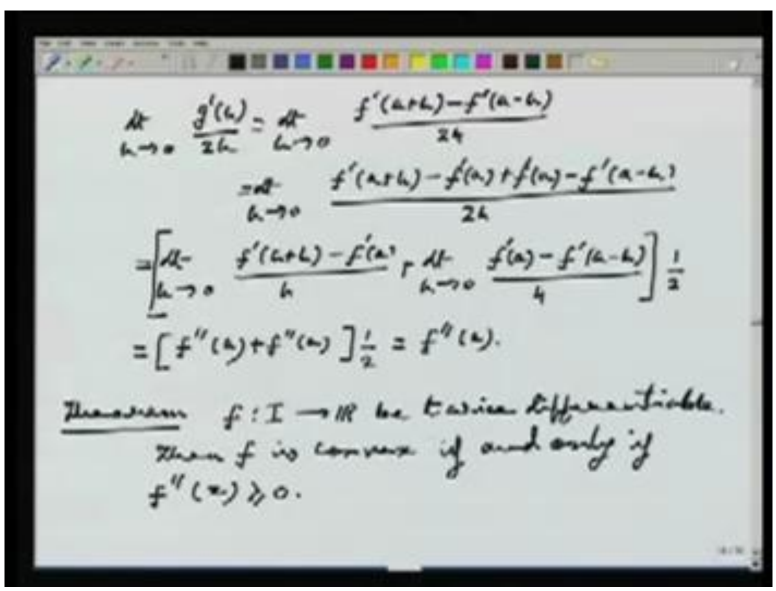
Suppose f from c to d to R is a differentiable function and let a belongs to (c, d) and also suppose that f double prime a exists. We are saying that f is differentiable function in (c, d) and a is a point in (c, d). The double derivative of f at a exists. Then the following is true. f double prime a is equal to limit h going to 0 f of a plus h plus f of a minus h minus twice f a divided by h squared. Now this is a lemma which we are going to use in our device of convexity of function using derivatives. So the proof goes as follows. What I do is, I look at the function g h is equal to f of a plus h plus f of a minus h minus twice f a.

Then we need to prove that the limit h going to 0 g of h divided by s square is equal to f double prime a.

Now, if you just think naively by putting h equal to 0 in the numerator and in the denominator, you quickly observe that this is the famous 0 by 0 form and as we have done in the last lecture to take care of the 0 by 0 form, you use the L' Hospital rule. The condition is for the L' Hospital rule for the both functions has to be differentiable. Certainly, the second one is the denominator, if h is not equal to 0, should be not equal to 0. Well, that is the case. h squared is always not equal to 0, if h is not equal to 0. So what I can do is, I can write the limit h going to 0 g of h divided by h squared is equal to limit h going to 0 g prime h divided by twice h. This is by L' Hospital.

Now let us calculate the g prime h. How to do that? Well, we will just use the chain rule here that g prime of h is first, f prime at a plus h and then the derivative of f plus h as a function h which is 1. So this is into 1 plus f prime at a minus h into derivative of a minus h as a function of h which is minus 1. The rest is constant since twice f a does not involve h. So the derivative is anyway 0.

(Refer Slide Time: 43:30)

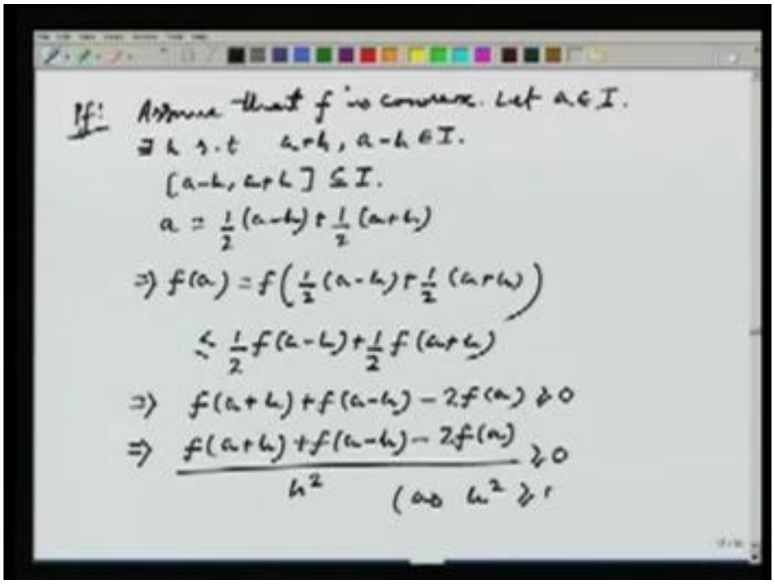


So what we get is that limit h going to 0 g prime h by twice h is same as limit h going to 0 f prime at a plus h minus f prime at a minus h divided by twice h, which now we write as limit h going to 0 f prime at a plus h minus f of a plus f of a minus f prime at a minus h whole divided by twice h, which is same as limit h going to 0 f prime at a plus h minus f of a divided by h plus limit h going to 0 f of a minus f prime of a minus h. The whole thing, I will put a half outside.

Now look at the two individual limits inside. The first limit is certainly f double prime a. So is the second one. This is also f double prime a and then, I have half outside, that is, f double prime at a. This proves our lemma. Now using this, we are going to prove the next result, which is the most fundamental result for applications as far as convex functions are concerned. So the theorem is this.

Suppose f is defined on the interval i to r. Suppose this is twice differentiable. Then f is convex if and only if f double prime at every point of the interval i is bigger than or equal to 0. Well, right now we are in a position to prove only one part of the theorem. For the other, we will need some thing called Taylor's theorem which we are going to do next. But let us only one part of the theorem and then apply it to draw some curves of functions. So the proof of this goes as follows.

(Refer Slide Time: 47:07)

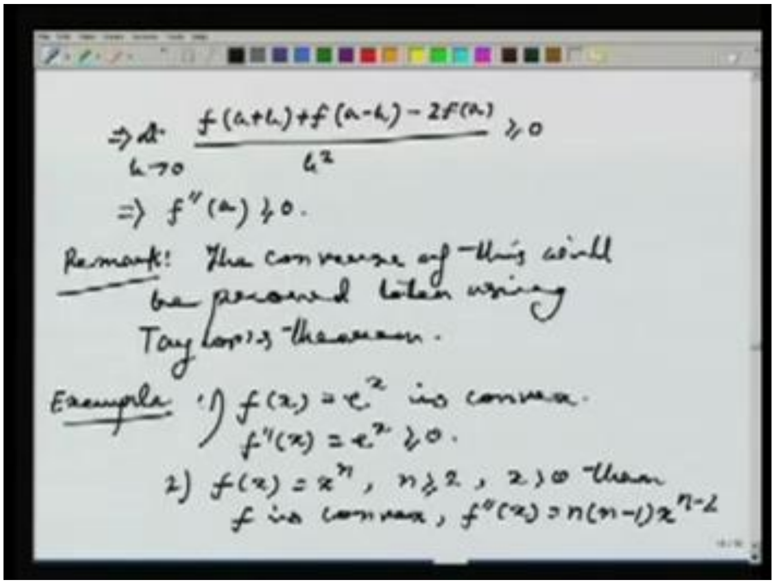


Assume that f is convex. We are going to prove that the double derivative of f is non-negative. Well, let a be an arbitrary point in the interval I. All I know is, f double prime at a, I have to prove that f double prime at a is bigger than or equal to 0. I have just assumed that f is convex. I have taken a point a in I. I need to prove f double prime at a is bigger than or equal to 0.

Now since I is an open interval and a is in I, this implies there exist h, such that, these two points a plus h and a minus h both are in I. That means, this whole interval a minus h and a plus h, both are contained in I. Now I can certainly write a is equal to half of a minus h plus half of a plus h because a is the mid point of a minus h and a plus h. Notice that then this is a convex combination. Since f is convex, this would imply that f of a which is equal to f of half of a minus h plus half of a plus h is certainly less than or equal to by definition of convexity, it is half of f a minus h plus half of f a plus h. If I multiply this inequation by 2 on both sides, this would then imply that f of a plus h plus f of a minus h minus twice f of a is bigger than or equal to 0.

Now this would then imply, if I divide by h squared which is anyway non-negative that f of a plus h plus f of a minus h minus twice f a divided by h square is also bigger than or equal to 0, as h squared bigger than or equal to 0.

(Refer Slide Time: 50:23)



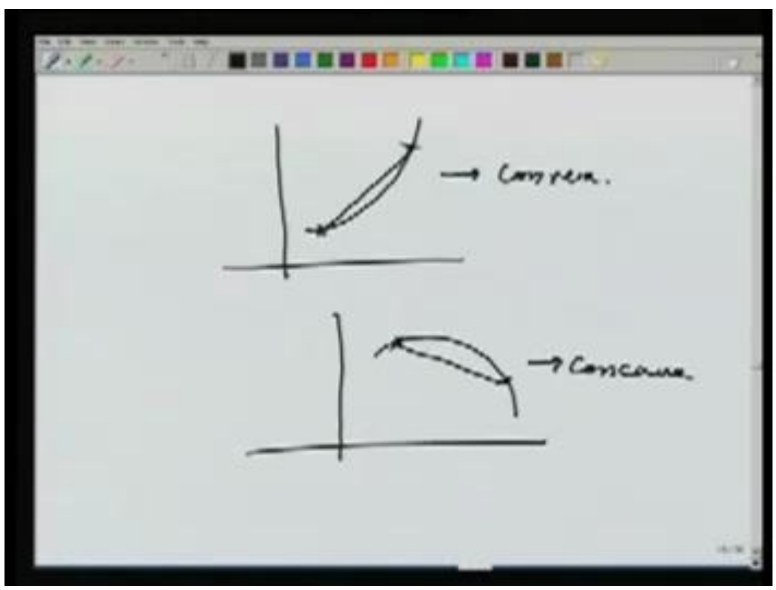
Now since this is true for all small h, this would certainly imply that limit h going to 0 f of a plus h plus f of a minus h minus twice f a divided by h squared is bigger than or equal to 0. Now we recall the previous lemma, where we have proved that this limit actually is f double prime a. This would then imply that f double prime a is bigger than or equal to 0. So I will remark here so that you can remember.

The convex of this will be proved later using something called Taylor's theorem. That means assuming that using Taylor's theorem, we can prove this. It means that if I have a function defined on an interval which is differentiable twice and the double derivative is everywhere non-negative, that means the function is actually a convex function. Now then, it because very easy to give examples of convex function that you can see very easily. I can say f(x) that is equal to e to the power x. f(x) is equal to e to the power x is convex because double derivative of f of(x) is e to the power x which is bigger than or equal to 0.

For example, if I look at f(x) is equal to x to the power n where n is bigger than or equal to 2 and x is bigger than 0, then f is convex because the double derivative of f is n into n minus 1 into x to the power of n minus 2. Since n is bigger than or equal to 0, this quantity is always non-negative because it is power of(x) and x is positive. So the function is actually convex.

Using this theorem, we can very easily detect if the function is twice differentiable, whether it is convex or not. Now similarly now, I can talk about concave function also. We will just define it by saying that f is concave if minus f is convex, which you can check easily that the inequality which defines convexity, it will go on the reverse side for concave function.

(Refer Slide Time: 53: 38)



So a typical way of convex function would be something like this to remember that this is a convex function and if I want to talk about concave functions, it would look something like this because if you take the points here and here. Look at the line joining these two. The function lies above the line. This means concave and here the situation is, if I take this point and this point, join the line. The graph of the function lies below the line. That means this is convex.