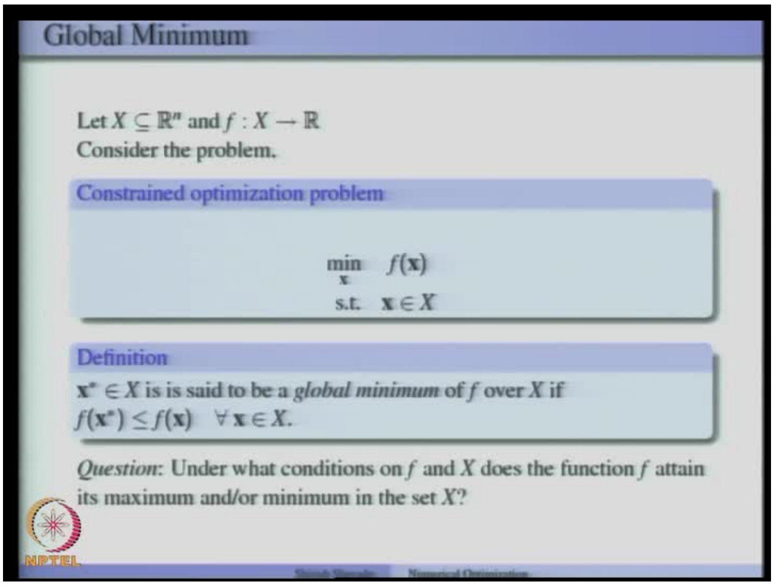
Numerical Optimization

Prof. Shirish K. Shevade
Department of Computer Science and Automation
Indian Institute of Technology, Bangalore

Lecture - 4
One Dimensional Optimization Optimality Conditions

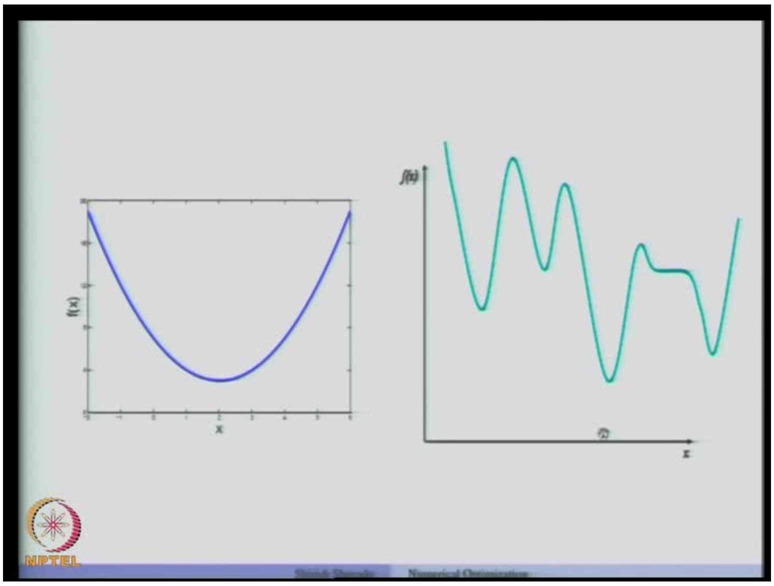
Welcome to this 4th lecture, in this series of lectures on Numerical Optimization. So far, in the last two lectures, we saw some mathematical background needed for this course; and some of those concepts will be useful for understandings on the theory that we are going to see today especially, the concepts on differential calculus and some analysis. So, in today's lecture, we will, we will look at unconstrained optimization problems, how to solve those problems, what are the necessary and sufficient conditions for the solutions of those problems.

(Refer Slide Time: 01:10)



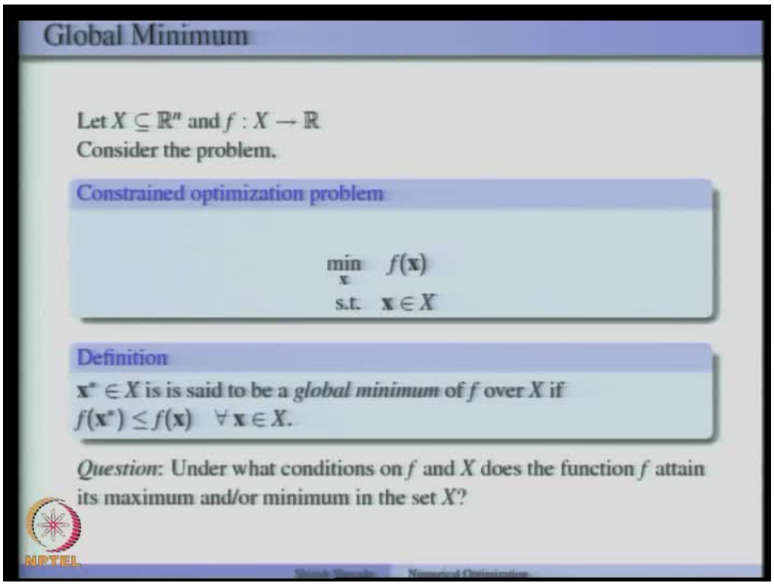
So, to begin with, let us look at a constraint optimization problem, which is given here so, you have a function f from the domain X to R and X is a subset of R n. And our aim is to minimize the function f, where x belongs to the set X. So, this is the constraint optimization problem but, any unconstrainted optimization problem is a special case of this constrainted optimization problem; in the sense that, if I replace this X by R n then, it becomes a unconstrainted optimization problem. So, our aim is to solve this problem and for that, we need the definition of minimum. So, x star belonging to the set X is said to be a global minimum of f over X, if the value of x star, the value of the function at x star is less than or equal to the value of the function at any other point in the set X. So, such a point x star is said to be a global minimum.

(Refer Slide Time: 02:18)



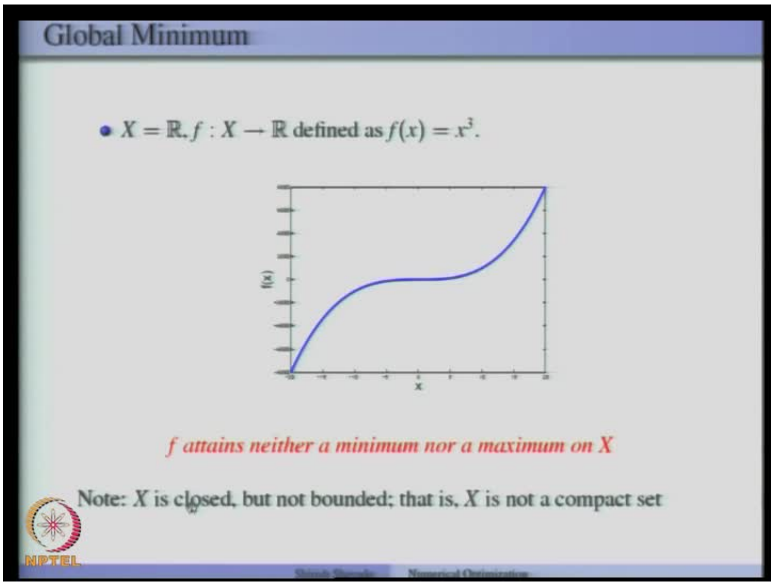
For example, let us consider a problem, where you have a function on the left panel, which is shown in the black line and let us assume that the constrainted set is a only this interval. Then, you will see that the global minimum of this problem is, it occurs at this x and the corresponding minimum value of the objective function is shown here. Now, on the right panel, you will see another function, which is not as nice as the function on the left panel so, you will see that there are lots of peaks and lots of valleys. And suppose, we restrict ourselves to this interval and assuming that the function goes to plus infinity beyond these points although, it is not of interest to us, we have only interested in this small interval. And you will see that in this interval, this is the point where the function achieves the least value and the corresponding x star turns out to be this. So, among all possible points in the domain of that function, on which we want to minimize the function, this is the point where the function achieves this the least value so, this is called a global minimum.

(Refer Slide Time: 03:40)



Similarly, one can define what is called a global maximum so, in the definition of global maximum, this inequality will be reversal, f of x star greater than or equal to f of x, for all x belongs to X, for x star to be your global maximum. Now, in this course many times I will be talking about minimization problems of this type but, the ideas that we discussed here can be easily extended by writing to the maximization problem by writing a maximization problem as a minimization problem, which we saw how to do, that was seen in the first lecture. Now, the question is that, are there any conditions on f or on x or both so that, the function f does attain its maximum or minimum in the set X. So, can we have some conditions on this f as well as on x so that, the the minimum or maximum is guaranteed in the set X.

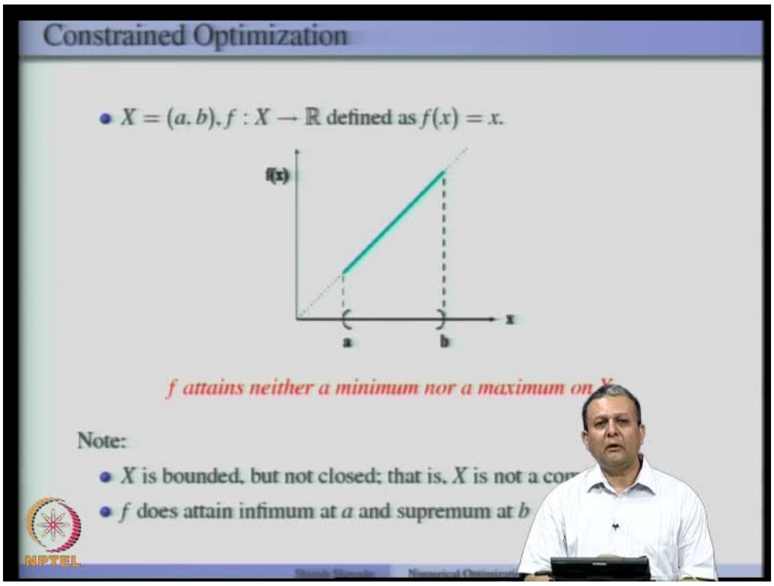
(Refer Slide Time: 04:51)



Now, for that let us look at some examples so, let us take a function f x equal to x cube where, the domain of the function is a set of real numbers and the function is defined from R to R. Now, the function is shown here in this plot so, on the right side the function goes to plus infinity, as x goes to plus infinity. And on the left side the function goes to, when x goes to minus infinity the function goes to minus infinity. Now, you will see that, over the domain X equal to R, this function has neither a minimum nor a maximum.

Because, the function extents to plus infinity and minus infinity so, it has no minimum or no maximum. Now, in this case, if you look at the domain set, which is X that is a close set, the set of real numbers R, that is a close set. But, it is not a bounded set. So, that means, naturally x is not a compact set now, do we want X to be bounded and not necessarily closed, we do not know as of now, let us see another example.

(Refer Slide Time: 06:06)

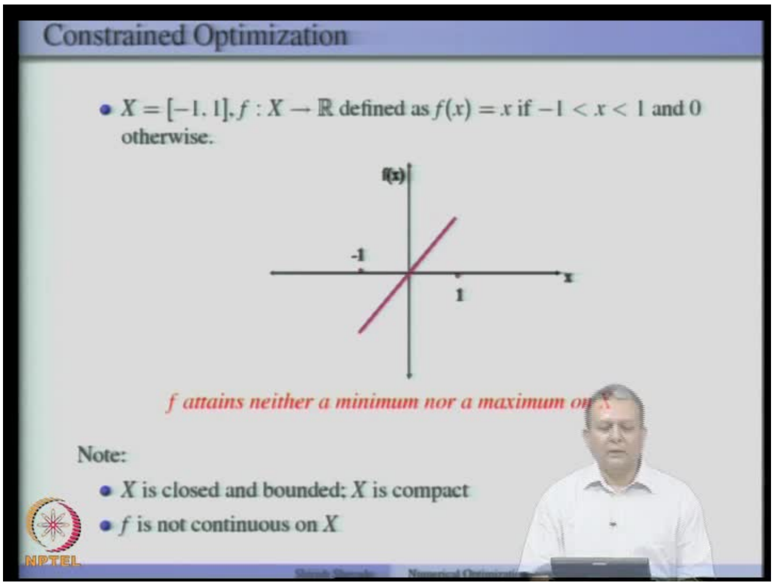


Now, here is an example where, the domain X is a open interval from a to b and the function f is from X to R, and defined as f of x equal to x. So, the function is shown in this figure and so, you will see that, these end points are not included in the domain. So, that is why, they are shown using some special symbols now, if you look at this domain, this domain is now bounded. So, unlike the previous case where, the domain was not bounded here, the domain is bounded but, it is not closed so, again x is not a compact set.

Now, you will see that, f attains neither a minimum nor a maximum on X because, f of a and f of b are not defined so, we cannot have a minimum value of the function in this open interval a to b. But, in this case, you can get a bound on this minimum and maximum value, and those are called infimum and supremum. So, the infimum of this function is f of a, which is attained at x is equal to a and the supremum of this function is f of b, which is attained at b.

But remember that a and b do not belong to the domain so, we cannot say that, minimum or maximum is attained but, in this case, the infimum and supremum are possible. So so far, we have considered two examples where, the domain X in one case was not bounded but, closed in the other case, it was bounded but, not closed, so that that essentially means that the domain was not a compact set.

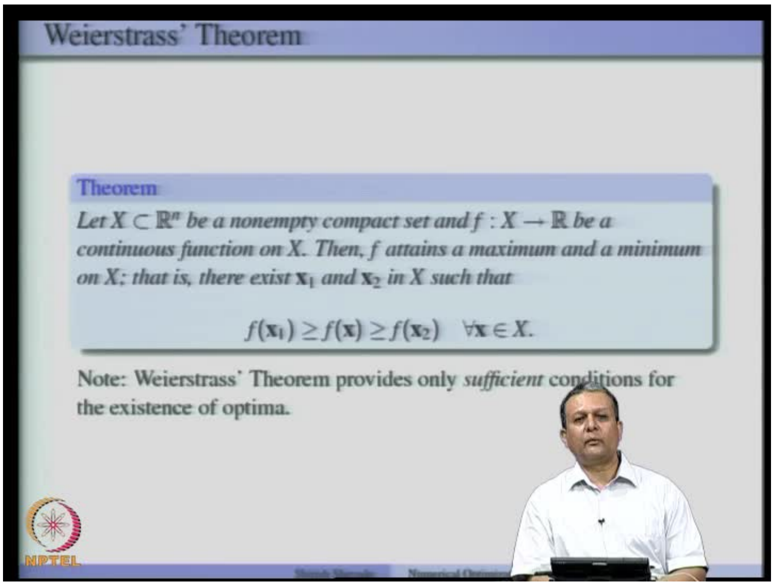
(Refer Slide Time: 08:33)



Now, so, does the, does that mean that, do we need X to be compact to attain a minimum so, let us look atone example. So, let us consider a domain X to be the closed interval minus 1 to 1 now, it is a close set as well as a bounded set so, hence it is a compact set. And let us define a function f from X to R, as f of x equal to x, if x is in the open interval minus 1 to 1 and 0 otherwise. So, in the open interval minus 1 to 1, the function is a straight line and at the end points, the function has value 0.

So, clearly you will see that, the function is a discontinuous function on the set X now, again we can say that, f does not attain a minimum or a maximum in this set X. So, remember that, here X is a compact set and f is not continuous. So, compactness of X alone was not enough probably, we need something some more conditions on f, and that condition is the continuity of function f on x. So, suppose, if we have X to be a compact set and f to be a continuous function on the set X, then can we say something.

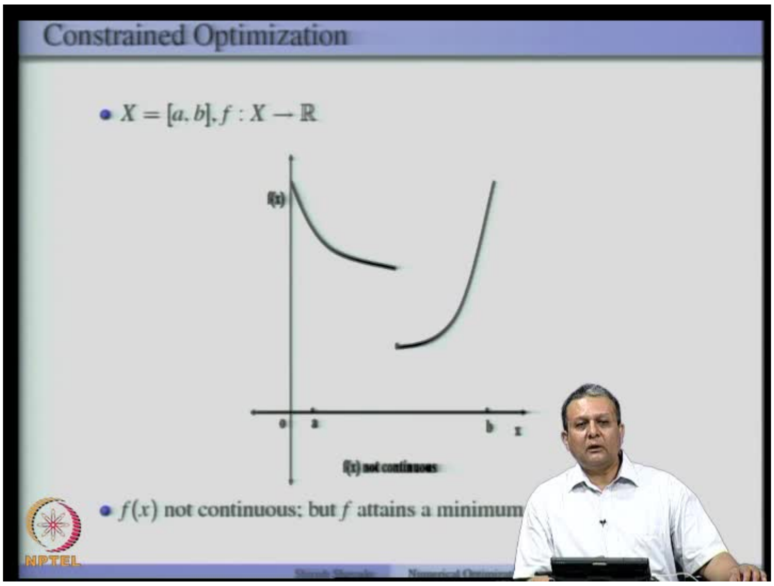
(Refer Slide Time: 10:00)



So, this leads us to the Weierstrass, the Weierstrass theorem, which states that, if X which is a subset of R n is a nonempty compact set and f is from X to R is a continuous function on X. Then, f does attain a maximum and a minimum on X that is, there exist some x 1 and some x 2 belonging to X. So, as that f of x is less than or equal to f of x 1, for all x belongs to X and f of x greater than or equal to f of x 2, for all x belongs to X.

So, So this theorem was proposed by Weierstrass, we will not go into the proof of this theorem but, it gives us sufficient conditions for the existence of optima. That is, the if f is a function different from X to R then we need X to be compact and f to be continuous function on x. Then, it is guaranteed that the, there exist two points x 1 and x 2 such that f of x 1 greater than or equal to f of x, greater than or equal to f of x 2, for all x belongs to X. So, as I said that, this Weierstrass theorem just provides us sufficient conditions for existence of optima but does not give you the necessary conditions for the existence of optima.

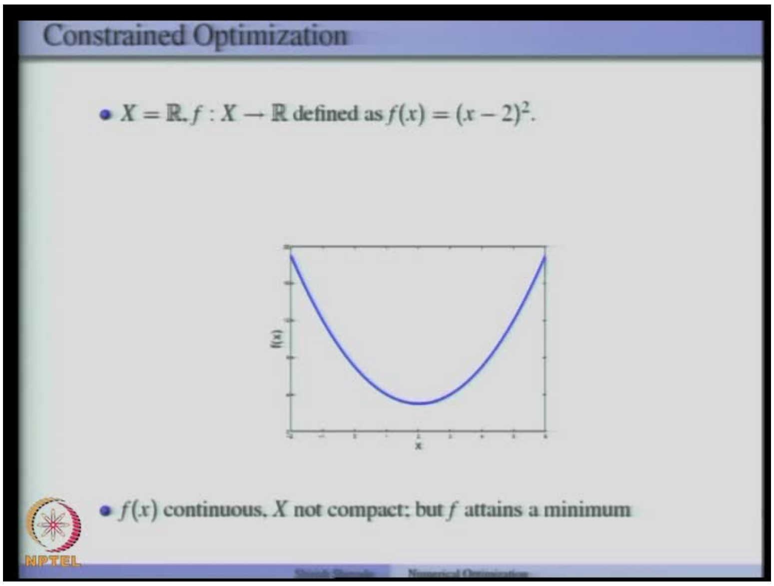
(Refer Slide Time: 11:32)



For example, suppose we take a function f, which is a, I have not written the explicit form of this function here but, assume that the function has a form like this. Now, the function is defined on a closed interval a to b, the closed interval a to b is a close and bounded set and hence, it is compact. But the function f is not continuous here; you will see that, there is a break in the function at this, at this point.

Now, although f is not continuous, you will see that, the function does attain a minimum at this point and the minimum value of this function is somewhere here. So, although f x is not continuous, f does attain a minimum in the closed interval a b, for this case. So, this example illustrates that, the conditions of a Weierstrass theorem are not made but the function does have a minimum. So, Weierstrass theorem does not talk about necessary conditions or existence of optima, it only talks about the sufficient conditions for optima.

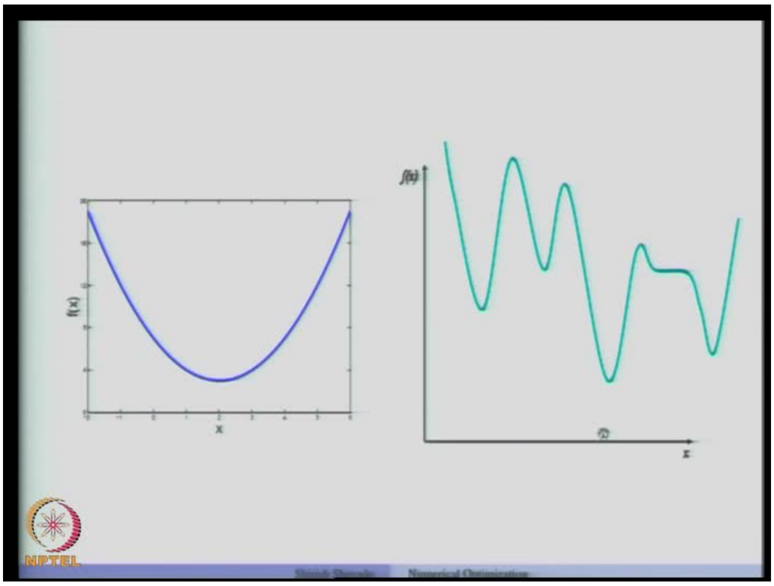
(Refer Slide Time: 12:54)



Now, here is another example, let us define a function f x as x minus 2 square and it is defined from R to R, the function is shown here. Now, you will see that the function is continuous, but X is not compact, the set R of real numbers is not a compact set. So, again the conditions required by Weierstrass theorem are not met because, x is not compact. But despite this fact, the function does attain a minimum at this point and the corresponding x star is somewhere here.

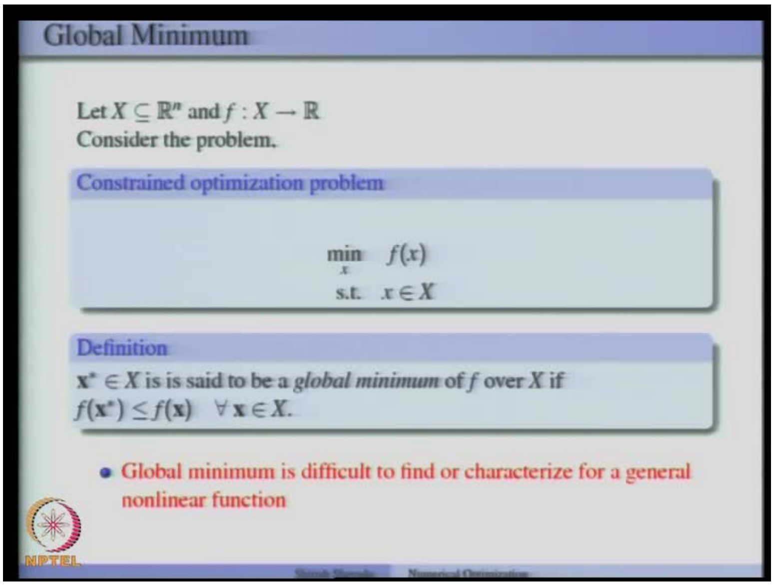
So, we will so, you will see that Weierstrass theorem just gives sufficient conditions for existences of optima or many optimum points but, does not give necessary conditions for the existences of optima. But any way it is a useful theorem in the sense that, under if those conditions are made like, if f is a function from X to R and X is compact and f is continuous on X then, the minimum and the maximum R of the function f are guaranteed in the set X.

(Refer Slide Time: 14:13)



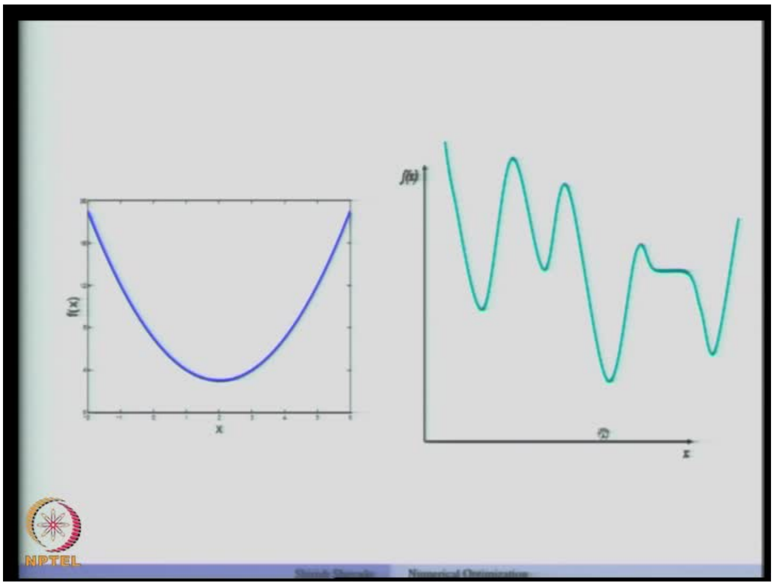
So now, let us look at general problem so, you will see that the left panel u will see that figure, which we saw earlier in this lecture. That, the function is very nice and attains minimum at a single point and on the right panel you will see a function where, there are lots of peaks and valleys. And if you look at this point for example so, in the neighborhood, you will see that the function in increasing and the function has a least value in this, if you consider on neighborhood around this point. But this value is still higher than the value, which is called a global minimum now, it turns out that finding this global minima is difficult.

(Refer Slide Time: 15:10)



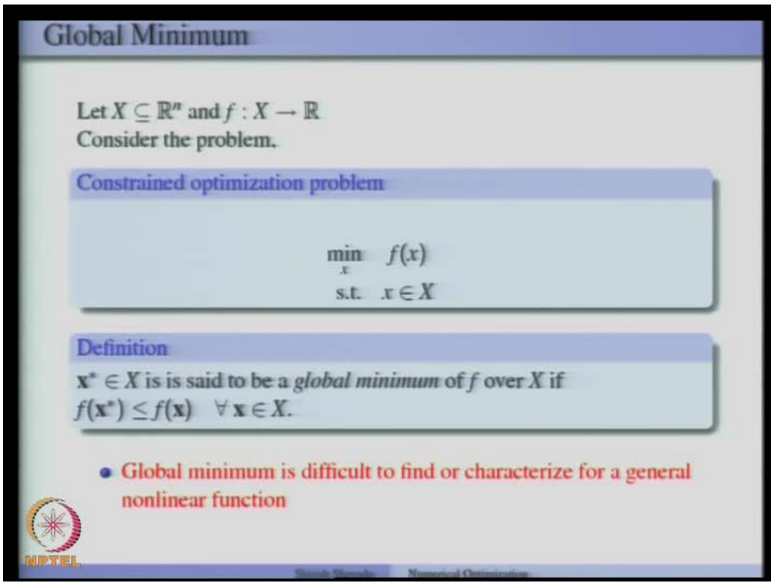
For example, it is very difficult to characterize or find the global minimum for a general non-linear function.

(Refer Slide Time: 15:22)



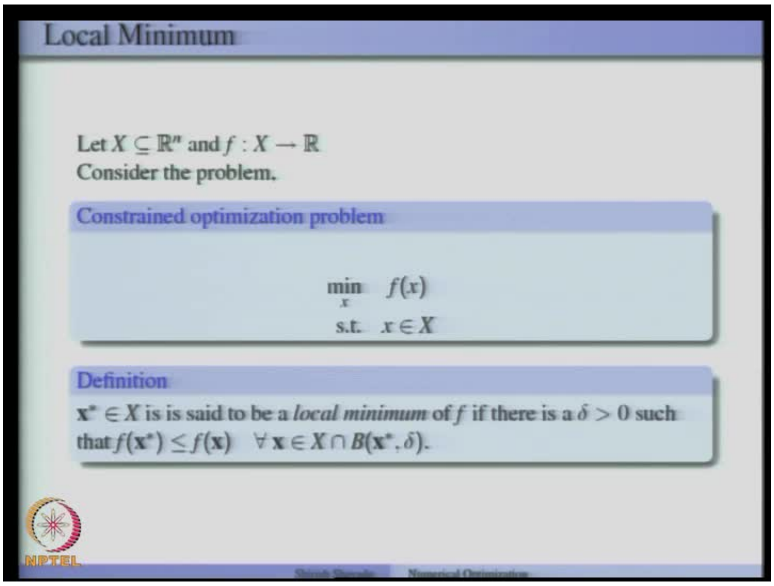
So, if you have a function, which is nice as on the left panel, it is easy to find a global minimum. While on the, on the right panel, you will see a function, whose global minimum is very difficult to find. And suppose, we find this point, which is a global minimum then, can we say that, it is a global minimum. So, for that purpose, we need to look at the value of the function at all the points and to ascertain that, f of x star is less than or equal to f of x, for all x belong to X and that is going to be a very difficult task.

(Refer Slide Time: 16:04)



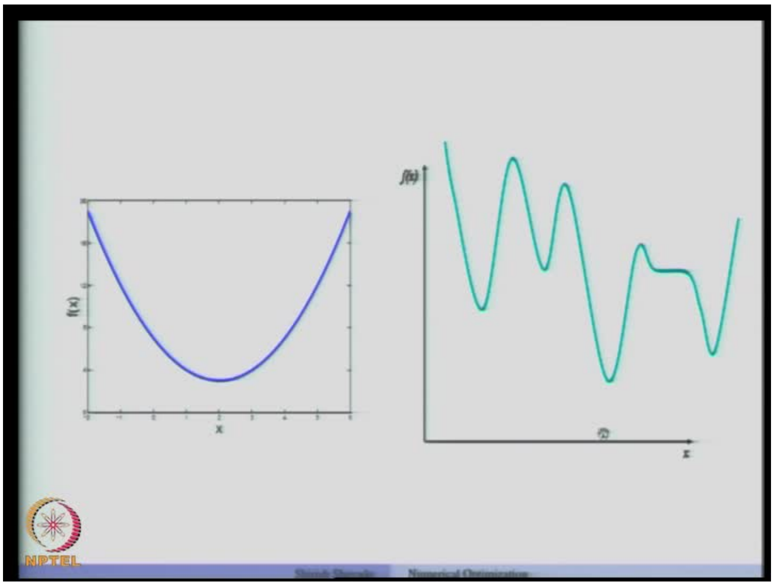
So, identifying a global minimum or characterizing it, is very difficult for a general non liner function, only in some special cases, one can characterize them very easily. We will see those cases some time later in this course but as of now, we know that, it is difficult to characterize this, this global minima.

(Refer Slide Time: 16:27)



So, it resulted in a concept of local minimum so, again let us consider x to be a subset of R n and f to be a function from X to R. And let us consider constraint optimization problem and we are trying to minimize f of x subject to x belongs to X. Now, x star belongs to X is said to be a local minimum of f, if there is a delta greater than 0 such that, in the delta neighborhood of x star, the value of the function is less than R equal to the value of the function, at a all points in the neighborhood.

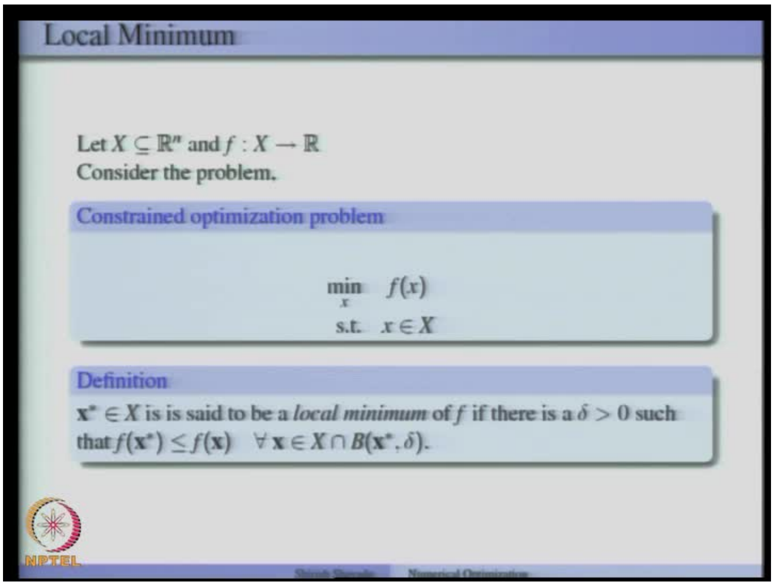
(Refer Slide Time: 17:16)



So, if you look at this, this point in the neighborhood, you will see that the value of the function is less than or equal to all the points so, if you take a small neighborhood remember, the delta in the definition of local minimum can be very small. So, if we take a small neighborhood around this point and find out the function values at all the points, you will see that, the value of the function is at least the value of all, the value of the function at all the points in the neighborhood.

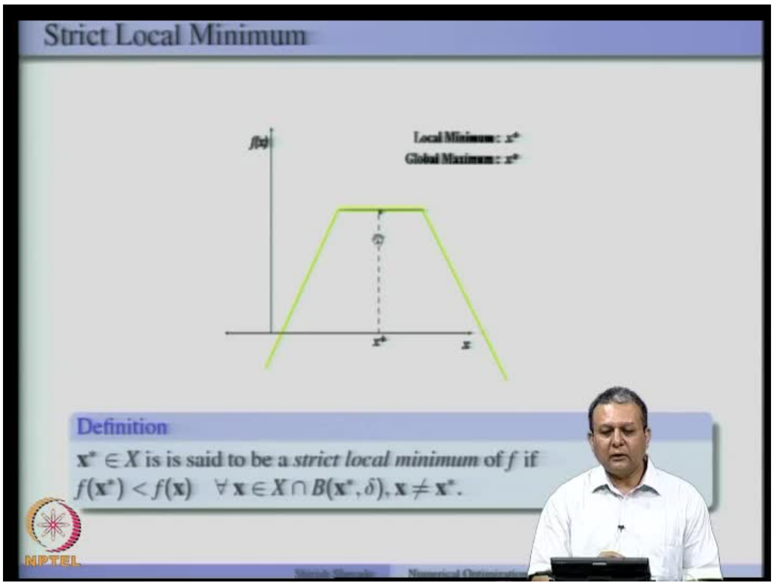
So, similar similarly, in this case it is true that, the value of the function is at least the value of the function, that neighborhood however, small neighborhood that you take and same is true in this case also. But here interestingly the function is almost flat in some range and then, in one direction it starts increasing and other direction it starts decreasing. So, such points in which, the function increases in one direction and decreases in the other direction, they are, they are called saddle points. So and similarly, in this case of course, this is a global minimum so it also is a local minimum because in the neighborhood the, the function value increases.

(Refer Slide Time: 18:43)



So, from this definition, one can extend this definition to the definition of local maximum so again as I said earlier that, the inequality reverses in this definition. So, again one has to take a local neighborhood, can be sufficiently small and find out the value of the function in the local neighborhood and you will see that, these points are local maximum. So the definitions of local minimum and local maximum, they are center around the neighborhood of a given point.

(Refer Slide Time: 19:35)

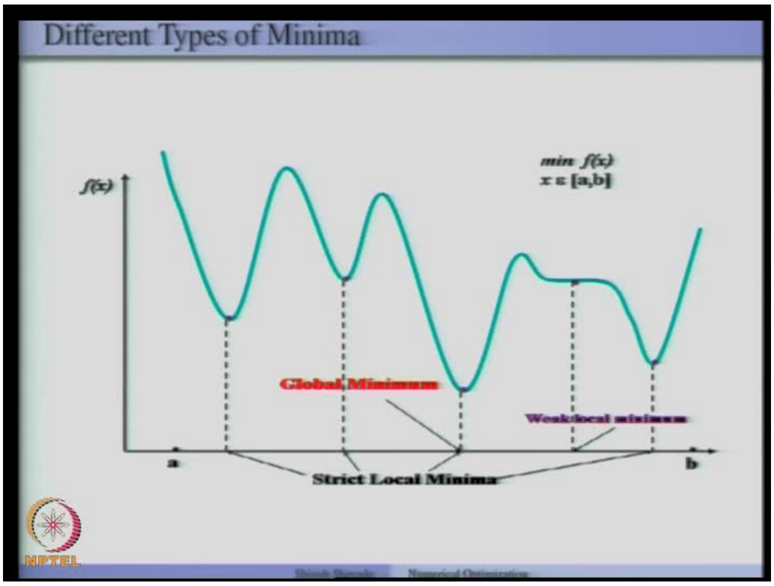


Now, let us consider one example so, here the function is shown here now, let us take a point x star now, x star by the definition of local minimum that, their exist neighborhood around x star. So that, the value of the function is f of x, is greater than equal to f of x star so, because of that definition, there is a problem in this case. Mainly that, since we are allowing f of x star greater than equal to f of x, for all x in the neighborhood of x star so, you will see that, x star terms of to be local minimum, because of that definition.

But, by looking at the function assuming that, the function goes to minus infinity in both the directions, when x goes to plus infinity and minus infinity. So, you will see that, x star indeed is a global maximum so, you have x star to be a local min and as well as a global maximum. Now, to avoid such cases, one can think of definition of a strict local minimum now, x star belong to X is said to be a strict local minimum of f, if f of x star is less than, strictly less than f of x, for all x in the neighborhood of x, delta neighborhood of x star and x naught equal to x star.

So, you will see that then, if you use this definition then, x star is a, is a local min but, not a strict local min. And also, it is a global max because, the function value is highest at x star of course, at the other points in the, at the other neighboring points also, the function values highest. So, remember that the global minimum or global maximum, they did not be unique. So, this is one example, where you have multiple global maxima and x star is not a strict local minimum. So, all these points, which are in the neighborhood, they are not strict local minima.

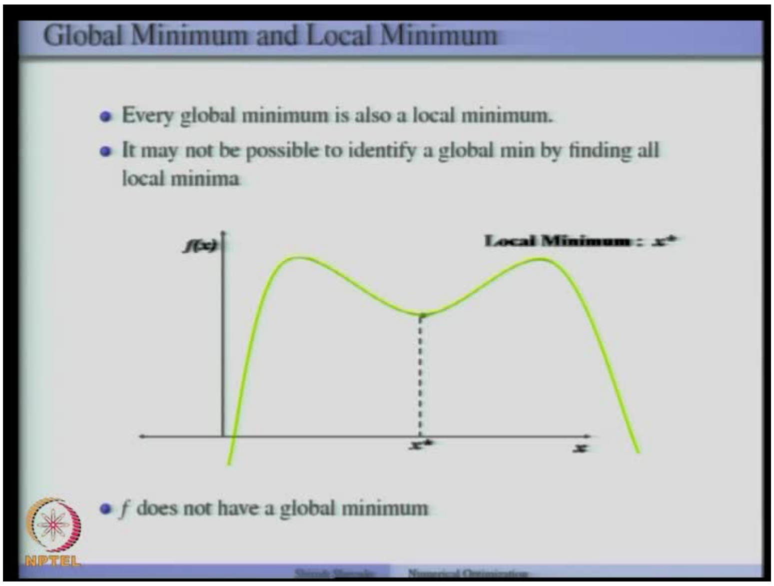
(Refer Slide Time: 21:53)



Now, let us look the function again, the function we saw earlier so, these are the points, which are local minima and not only local minima, they are strict local minima. Now, this is a point, which is a weak local minima so, you will see that, f of x star is less than or equal to f of x in the neighborhood of x star. So, such points are weak local minimum and this point is a global minima.

Now, earlier I said that, this global minima are difficult to characterize so, one way wonder that, how about that finding different local minima, strict local minima or weak local minima, collecting them together. And then, trying to find out the least among those among those so, this idea may work in some cases but may not work always.

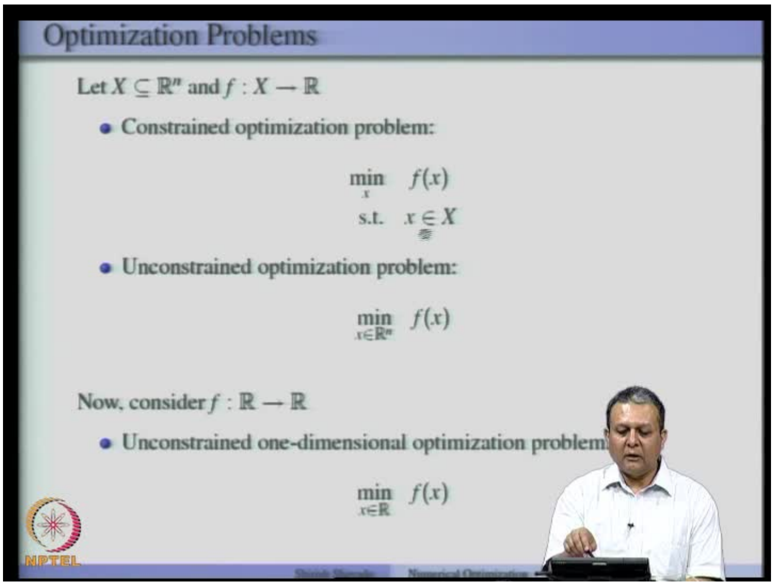
(Refer Slide Time: 22:59)



For example, suppose I have function, which is shown here and it goes to minus infinity so, the domain of the function is the set of real numbers and you will see that, this x star is the local minimum. In fact, it is a strict local minimum because the value of the function in the neighborhood of x star does not reached f of x star at any point, other than x star so, it always it always exceeds f of x star.

So now, if you look at this function and if you collect all the local minima, which is one in this case, that will not give you the global minimum, because a function does not have a global minimum, the function goes to minus infinity in a both positive x as well as negative x direction. So, by collecting all local minima, it is not always granted that, we can find a global minimum of a function. So, one has to check, whether the function has a proper form so that, collecting all local minima can give us global minima.

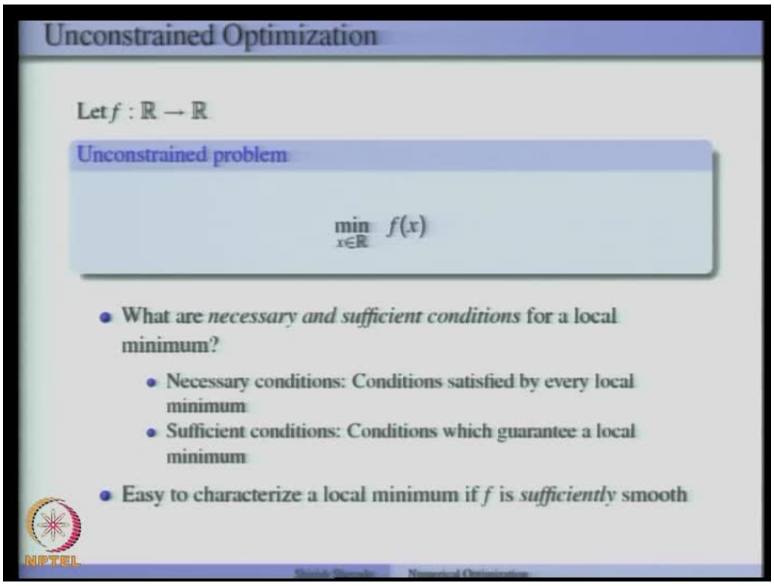
(Refer Slide Time: 24:14)



Now, let us look at unconstrained optimization problems, as I said earlier that, we are trying to solve a problem from, to minimize a function f from where, x belongs to capital X and f is a function from X to R and x is a subside of R n. So, this is a general concerned optimization problem and unconstrained optimization problem is a special case of this. Now, to solve a concerned optimization problem, many times we need to solve an unconstrained optimization problems, first we will show see those techniques later but, solving an unconstrained optimization problem is important to solve a concerned optimization problem.

Now, to solve a unconstrained optimization problem, one has to solve any one dimensional unconstrained optimization problem, which are of this type. So, in today's talk will focus on solving one dimensional unconstrained of optimization problems, they are very important to solve a unconstrained multidimensional optimization problem. And this unconstrained multidimensional optimization problems are important to solve general concerned optimization problem. So, this important problem will spend time studding about the solutions of these types of problems.

(Refer Slide Time: 25:48)



So, here is a unconstrained one dimensional optimization problems where, f is the function from R to R. So, remember that now, we are not worried about the set X in this case, the domain of the function is the entire set of real numbers, the ranges is also the entire set of real numbers. Now, we want to solve this problem and we saw earlier that, it is very difficult to identify the global minima of function. So, we will restrict ourselves to local minimum now, the ideas that we see here can be easily extended to finding local maxima also.

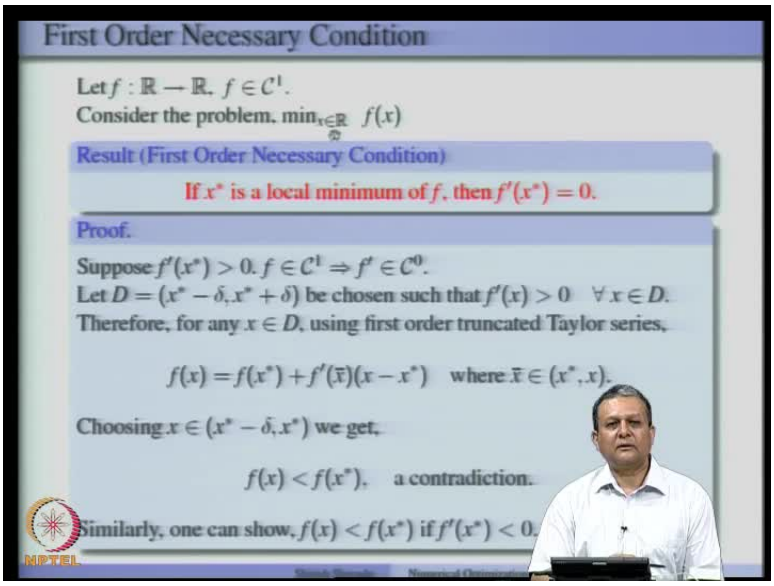
So, the natural question that one would like to ask still that, what are the necessary and sufficient conditions for a local minimum? Now what, what do you mean by the necessary and sufficient conditions so, necessary conditions are the conditions, which are satisfy by every local minimum. For example, if suppose x star is a local minimum of this problem then, we can say that, if x star is a local min then, certain conditions are satisfied and those conditions are called necessary conditions.

And by sufficient conditions, we mean those conditions, which guarantee a local minimum for example, if add x star belong to R certain conditions are satisfied then, we can say with grantee that, x star is a local minimum. Now, as we saw earlier that, there exist lots of their might exists lot of local minima or a general non linear functions then, how do we characterize such local minima, that is a next question that, we would like to answer.

Because again we do not want to a find the local neighborhood along around x star and check whether the value of the function exceeds the value of the function at x star. Because, there could be infinitely many points again in the neighborhood and we do not want to ensure we do not want to have any conditions, which will require finding different points in the neighborhood and taken the function values.

So, instead, is there any algebraic approach, to check whether a particular point x star is a local minimum now, that ideas of differential calculates become very useful in this case. So, if the function is sufficiently smooth that means, is first order derivative second order derivative and so on. They exist and they are continuous then, one can arrive at different conditions for characterization of local minima so, we are going to see those conditions.

(Refer Slide Time: 28:54)



Now, now let us consider a function from R to R and f belongs to C 1 so, this C 1 denotes the class of function, whose first derivatives are continuous. So, now we are restricting ourselves to functions, whose first derivatives are continuous and let us consider the problem to minimize f of x where, f belongs to R. Now, here is a first set of conditions called first order necessary conditions, the result is very simple, if it says that if x star is a local minimum then, f dash x star is equal to 0.

So, which means that, the derivative of the function vanishes so, if x star is a local minimum, the derivative of the function at x star vanishes. So, these conditions are called the first order necessary conditions, they are necessary conditions because, it extends a local minimum of f then, this conditions are satisfied. So, that is why they are necessary conditions and they use the first order derivative information so, that is why, we call them as first order necessary conditions.

Now, let us look at the proof of this result, the proof is best on the fact that, let us so, we have statement, which is of the type a implies b. So, what we do is that, we assume that b is not true and then, say that a is not true so we saw this method of proof in one of our earlier lectures. So, let us assume that f dash x star is greater than 0 so, the result says that, if x star is a local minimum then, f dash x star is equal to 0. So, on the contrary let us assume that f dash x star is greater than 0 and then, we will prove that, in such a case x star cannot be a local minimum.

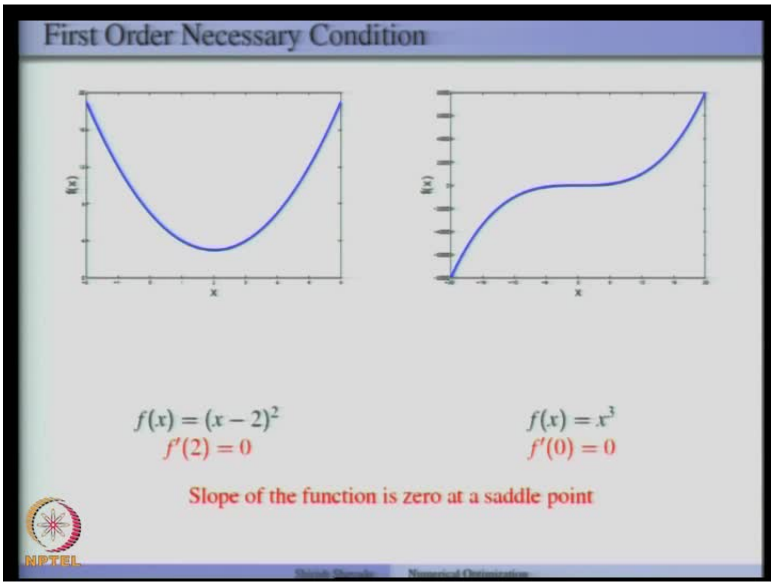
Now, since f belongs to C 1 so, it is derivate, it is a continuous functions so that means, f dash belongs to C naught. So, let us take a interval D, which is of size 2 delta around x star so, x star minus delta to x star, x star delta is open interval. So, let this interval be chosen so that, derivative of the function at any point in this interval, that derivative of the function at any point in this interval is greater than 0. Now, that is possible it is possible to get such a interval because, we are assuming that, f belongs to C 1.

So, atleast in the neighborhood, f dash is continuous and we are assuming that, f dash x dash greater than 0 so, at least in the neighborhood f dash x is greater than 0. Now, if you write the truncated Taylor series expansion of f of x around x star and truncated to the first order. So, what we get is, f of x is nothing but f of x star plus f dash x bar into x minus x star where, x bar is a point on the line segment joining x star and x. Now, remember that, f dash x bar is greater than 0, because x bar is a point in the interval D.

So, f dash x bar is greater than 0 now, we are free to choose any x in the interval b so, suppose, we choose x to be in the open interval x star minus delta to x star. So, x minus x bar x minus x star will become negative because, x lies in this interval so, x minus x star becomes negative. So, x minus x star becomes negative, f dash x bar is greater than 0 so, therefore, f of x will become less than f of x star and that contradicts the fact that, x star is a local minimum.

Now, when can use similar ideas by assuming that, f dash x star less than 0 and then the rest of the things follow, as they are except that, you can choose point in the open interval x star to x star minus x star plus delta. So, in that case, f dash x bar will be less than 0 and this quantity is greater than 0 so, that product will be less than 0 and therefore, f of x should be less than f of x star. So, in either case, we come off with the contradiction that, x star is not a local minimum so, if x star is local minimum then the first derivate of the function should vanish.

(Refer Slide Time: 34:11)

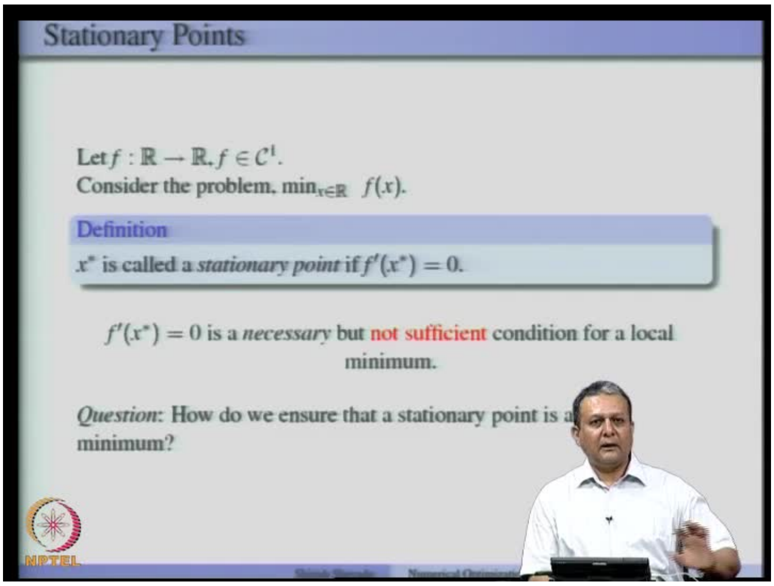


Now, here are some examples so, on the left panel, you will see a function f of x, which is x minus 2 square and the derivate of the function at two vanishes. So, you will see that the function has a zero slope at x is equal to 0 and then, on either side the function is increasing. Now, on the right panel, you will see one function, which is f of x equal to a minus x square and at x equal to 0, you will see that slope of the function is 0.

But in this case, x is equal to 0 turns out to be a global maximum and in this case, it turns out to be a global minimum assuming that, the function x tends to plus infinity here and function goes to minus infinity in this case. So, the slope of the function is 0 at a local minimum as well as at the local maximum so, looking at the slope really does not tell us much about the minimum or maximum of a function at that point.

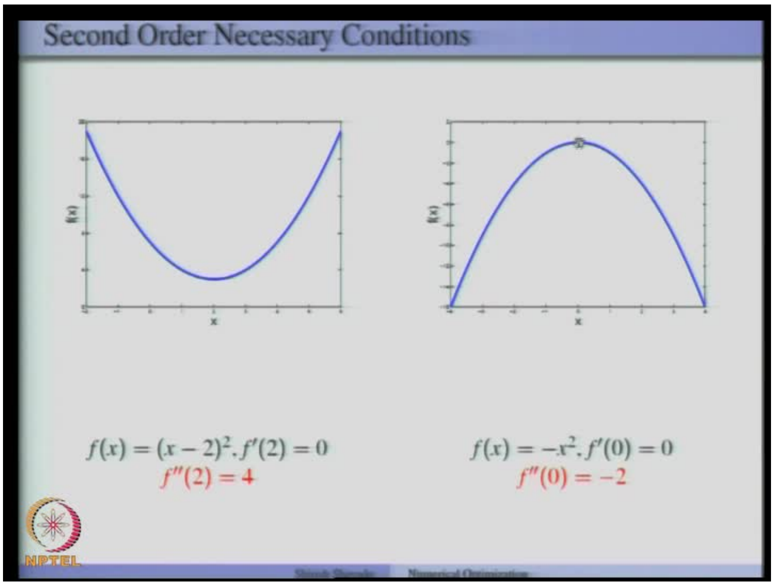
Now, let us consider one more case so, on the right panel, you will see a function f of x is equal to x cube so, f dash x at x equal to 0, if 0. But you will see that, the function increases in one direction and decreasing in the other direction. So, if you move x to minus infinity, the function goes to minus infinity so, slope of the function really does not tell us anything about the existences of local minimum or local maximum at that point. So, the points at which, the slope of the function is 0 such points are called saddle points sometimes people also call it stationary points. So, this saddle points really do not give us any idea about the local minima or maxima so, we need some extra information.

(Refer Slide Time: 36:23)



And as I said that f dash x is equal to 0, is just a necessary condition for a local minimum but, not a sufficient condition. So, if you collect all points for which, f dash star equal to 0, such points are called stationary points. Now, since we are interested in solving a problem of the type minimize f of x where, x belongs to R. The natural question that you would like to ask is that, how do we ensure that stationary point is indeed local minimum. And we have seen in the case that, checking the gradient of the function or checking the derivate of the function in one dimensional case does not really guarantee a local minimum.

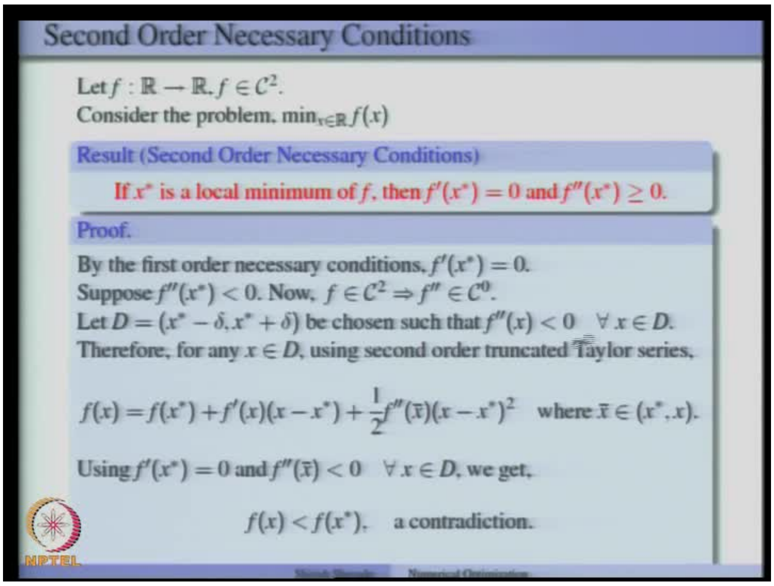
(Refer Slide Time: 37:14)



So, let us look at those second order information now, on the left panel you will see the same function f of x equal to x minus 2 square. Now, the derivative of the function vanishes at this point, the second derivate of this function at x equal to 2 is 4 and that is greater than 0. Now, on the right panel, you will see the function f of x equal to minus x square, the first derivative of the function at x equal to 0 is 0, and the second derivate of the function at x is equal to 0 is minus 2.

So, you will see that, for this function the second derivative of the function at this point is positive and for this function this second derivate of the function at this point is negative. So, the second derivate in some sense, talks about the curvature of the function at that point so, this function has a positive curvature at this point and this function has a negative curvature at this point. So, you will see that, although the first derivatives of these two functions at these points are 0, it is a second derivative which tells us about the curvature and that tells that this point indeed is a local minimum and this point indeed is a local maximum.

(Refer Slide Time: 38:37)



So, let us look at the second order necessary conditions now, for the second order necessary conditions, we need the second derivate of the function. So, let us assume that f belongs to C 2; that is a second derivative of the function is continuous and the function is from R to R and we are interested in solving in one dimensional unconstraint optimization problems. Now, here is a result about the second order necessary conditions now, if x star is a local minimum of f then, the first derivative at x star vanishes and second derivative is non-negative.

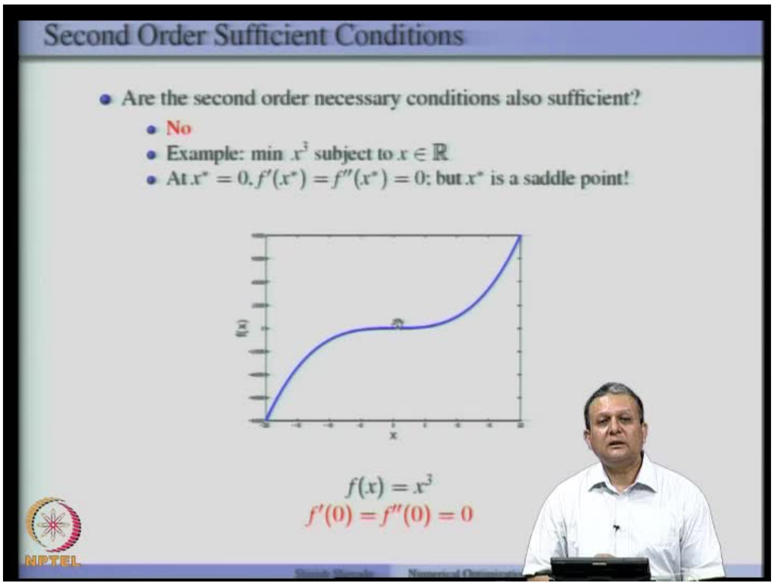
So, this result gives us so, the necessary conditions, which use the second order information of the function. Now, how do you prove this So, the proof is very easy now, the by the first order necessary conditions we know that, if x star is a local minimum then f dash x star equal to 0, we already proved that result. Now, let us assume that, the second derivative is less than 0 at x star and then, will come up with the contradiction.

Now, since f belongs to C 2, the second derivative of the function is a continuous function so, again as in the previous case let us choose an interval around x star of size 2 delta. So, this a interval D, x star minus delta to x star plus delta now, since f 2 dash is continuous, we can always choose this interval such that, for any point in this interval D, the second derivative is less than 0. Because, we assume that, f 2 dash x star is less than 0 so, now, let us write down the second order truncated Taylor series.

So, f of x is nothing but, f of x star, plus f dash x into x minus x star plus half f 2 x bar, into x minus x star square where, x bar is a point on the line segment joining x star and x. Now, we know that, from the first around the necessary conditions that f dash x star equal to 0 now, we are so far chosen this interval so that, f 2 dash x is less than 0, for all x belongs to D.

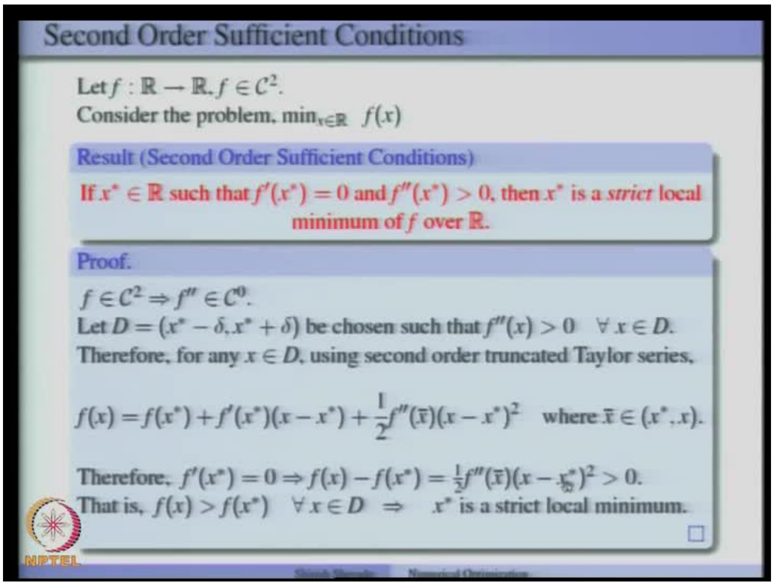
And x bar belongs to this interval so, f 2 dash x bar is also less then 0 now, for this quantity this should be x star. So, this quantity is 0 and f 2 dash x bar is less than 0, x minus x star square is greater than 0, x minus x star square is 0. So, you will see that this quantity vanishes and this quantity is less than 0. So, which means that, f of x is less then f of x star and which is a contradiction. So, we will see the we see that, if x star is a local minimum then, f dash x star equal to 0 and f 2 dash x star greater than equal to 0. So, remember that, this is f dash x star and by first order necessary conditions, f dash x star equal to 0.

(Refer Slide Time: 42:05)



Now, the question is, are this second order sufficient conditions, are the second order necessary conditions also sufficient, and the answer is no.

(Refer Slide Time: 42:46)



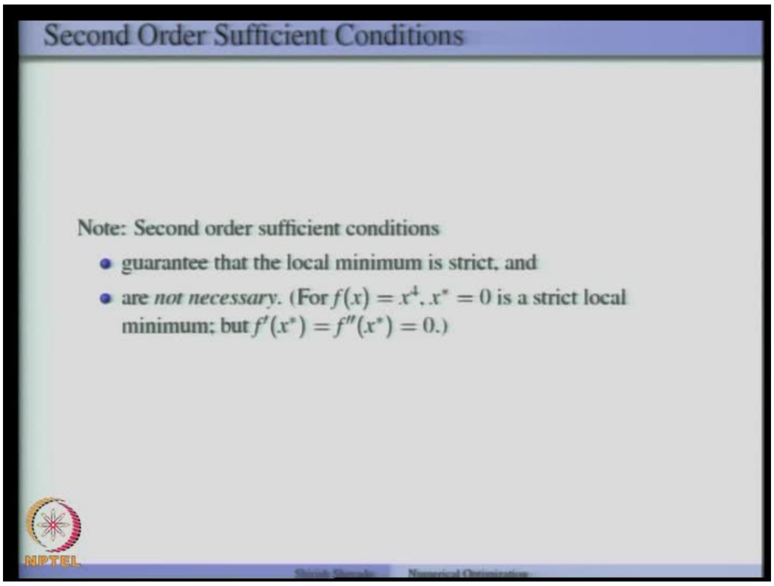
Because, if we consider problem, which is to minimize x cube subject to x belongs to R, the function is drawn here. Now, at this point, the first derivative and the second derivative both are 0 but then, x R turn out to be saddle point. So, these conditions are necessary for existence of a local minimum, but not sufficient. So, what are the sufficient conditions for the existence of the local minimum so, let f be a function from R to R, f belongs to C 2.

The second order sufficient conditions results states that, if f star is (( )) from the set R says that, the first derivative of the function vanishes at that x star and the second derivative is greater then 0 then, x star is a strict local minimum of f over R. So, remember that, this is result, which guarantee strict local minimum of f over r now, the proof of this is again very easy that, since f belongs to C 2, f 2 dash is continuous. So, will choose the interval D so that, in the, in that interval for any x, in that interval f 2 dash x is greater then 0.

And that is possible because f 2 dash x star is greater than 0 so, let us write down the truncated Taylor series expansion of f around x star. So, f of x is nothing but, f of star plus f dash x star into x minus x star plus half, f two dash x bar in to x minus f dash square where, x bar is a point on the line segment x star to x. Now, if f dash x star is equal to 0 then, f x minus f of x star is nothing but half f 2 dash x bar in to x minus x star square.

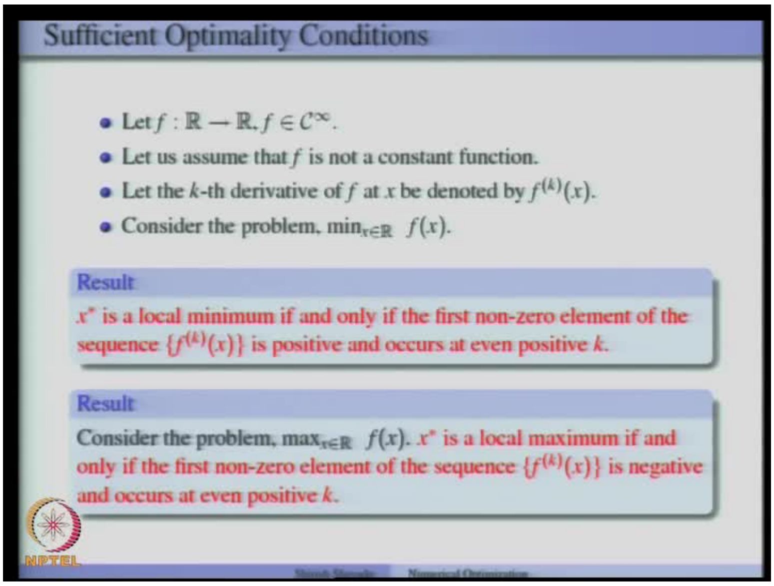
Now, x bar is in the point interval D so, f 2 dash x bar is greater than 0, this quantity is also greater than 0 because, it is a square. So, the whole quantity is greater than 0 so, which means that, f of x is greater than f of x star, for all x belongs to D. So, at least in the, the existing delta greater than 0 such that, in the neighborhood delta, delta neighborhood of x star, the value of the function does not attain f of x star in fact, it is, it exist f of x star. So, which implies that x star is a strict local minimum so, remember that, so, this sufficient conditions guarantee a strict local minimum of f over R.

(Refer Slide Time: 45:11)



Now, are these conditions necessary, though they guarantee that the local minimum is strict, they turn out to be not necessary conditions, because if we consider case where, f of x is equal to x to the power 4 then, x star is equal to 0, is strict local minimum. But then, f dash x star and f 2 dash x star or both 0 so, this second order sufficient conditions, they are not necessary in this case.

(Refer Slide Time: 45:49)

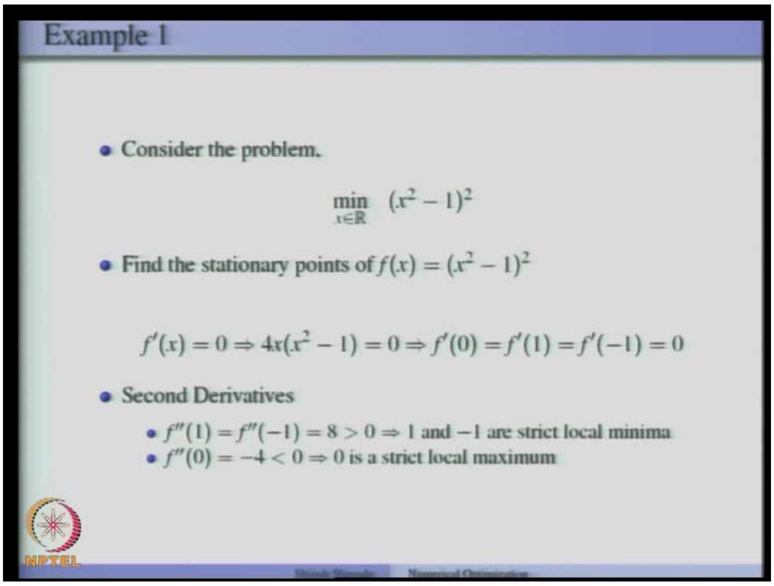


So, now, we have important result, which will state without the proof so, let us assume that f is a function from R to r and f belongs to C infinity that is, all the derivatives of the function exists and they are continuous. Let us also assume that, f is not a constant function so, let us rule out that possibility also and let us denote the k eth derivative of f at x, f k x. And consider, the problem to minimize f of x, x belongs to R. Now, here is the result, which says that, x star is a local minimum of f, if and only if, the first non-zero element of the sequence f k x is positive and occurs at even positive k.

So, this is the very important result we say that, if you start to take a derivative of the function from the first derivative second derivative and so on. Now, if you consider that sequences of derivatives of function f at x so, the first non-zero element of the sequence f k x star is positive and occurs at positive even and positive so, that is very important. So, if you consider the example of f of x equal to x to the power 4 so, you will see that, x star equal to 0 is a local minimum of that function and the first three derivatives at x star are 0 and the fourth derivative is positive.

So, the fourth derivative that means, the k is equal to 4, which is even and positive number and fourth derivative of the function x, x to the power 4 is positive. So, that is why, x star equal to 0 is a local minimum of f x equal to x to the power 4. Now, one can have a similar result for the local maximum so, the only change that one has to make is that, the sequence of k x star is negative. The first element of the first non-zero element of the sequence f k x star is negative and it occurs at even positive.

(Refer Slide Time: 48:21)

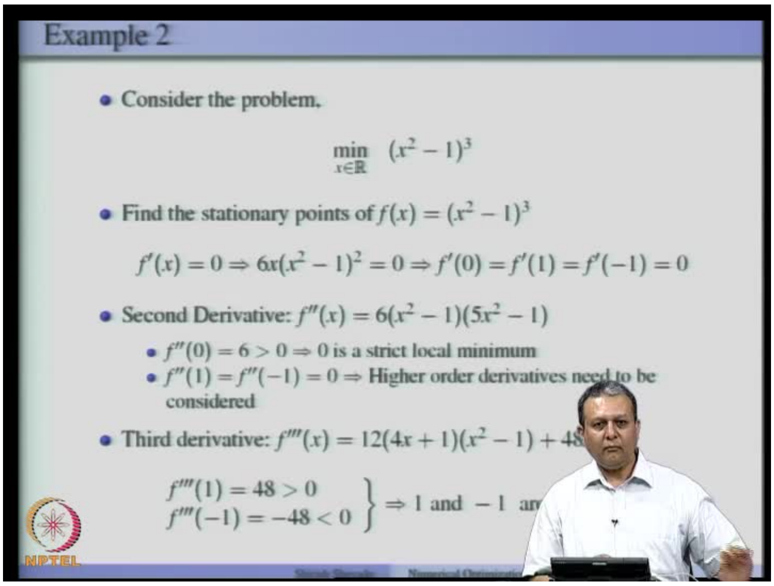


Now, let us see some examples, to find out the minima or maxima, minima of a problem now, let us consider the problem minima is x square minus 1 square. Now, the first step to solve such problems is always to indentify the stationary points and the stationary points are found by equating the derivative to 0. So, if you take the first derivative of this function and that transfer to be 4 x n to x square minus 1 and that, equal to 0 implies that, either x is equal to 0 or x equal to plus 1 or x equal to minus 1 so, it has three stationary points.

Now, we need to check, which of these stationary points are local minima or local maxima so, then, we need to go for the second derivative information. Now, f 2 dash 1, in this case transfer to be yet and f 2 dash minus 1 also transfer to be yet and that is also greater then 0. So, you see that the first derivative is 0 at 1 and minus 1 and the second derivative is positive. So, the first non-zero element of the sequence f 2 dash k, of the sequence f k x star is positive and occurs at u and k, which is 2 in this case.

So, 1 and minus 1 are strict local minima and along similar line one can say that, f 2 dash zero, which is minus 4 and less then 0 implies that 0 is a strict local maximum, because this the second derivatives negative, the first derivatives is 0 and 2 is a even positive integer so, 0 is a strict local maximum.

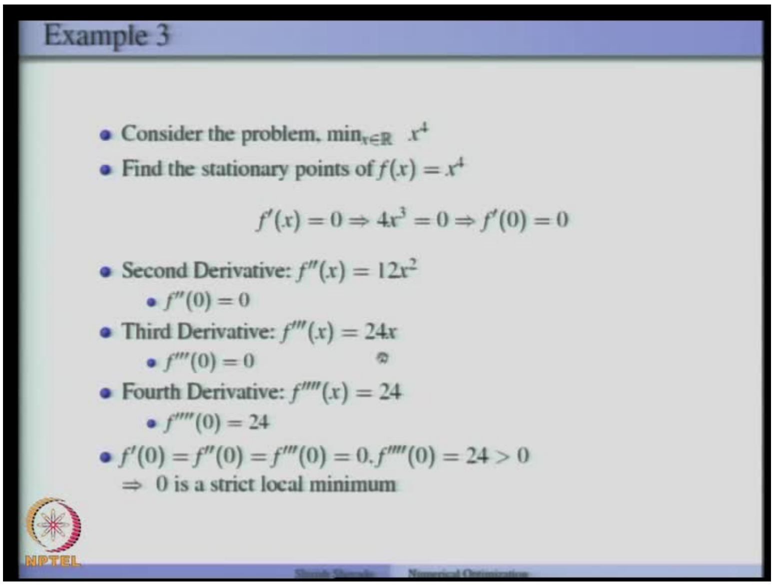
(Refer Slide Time: 50:17)



Now, let us consider another case where minimize f of x where, f of x equal x square minus 1 cube. Now, if look at the stationary points, again we need three stationary points 0, 1 and minus 1, at which points the derivative of the function vanishes. So, we look at the second derivative now, second derivative information tells us that, the second derivative of the function at 0 is 6, which is positive which means that, 0 is a strict local minimum.

But, at the stationary points 1 and minus 1, the second derivative also vanishes, so we have to look for the higher order derivatives. So, let us look at the third derivative now, f 3 dash x is a function like this and you will see that, f 3 dash x f 3 dash 1 is positive, f 3 dash minus 1 is negative. So, we really cannot conclude any thing from our result about sufficient optimality conditions because, that this derivatives are non-zero at k, which is R. So, we can just conclude that, 1 and minus 1 are saddle points of this problem and 0 is a strict local minimum.

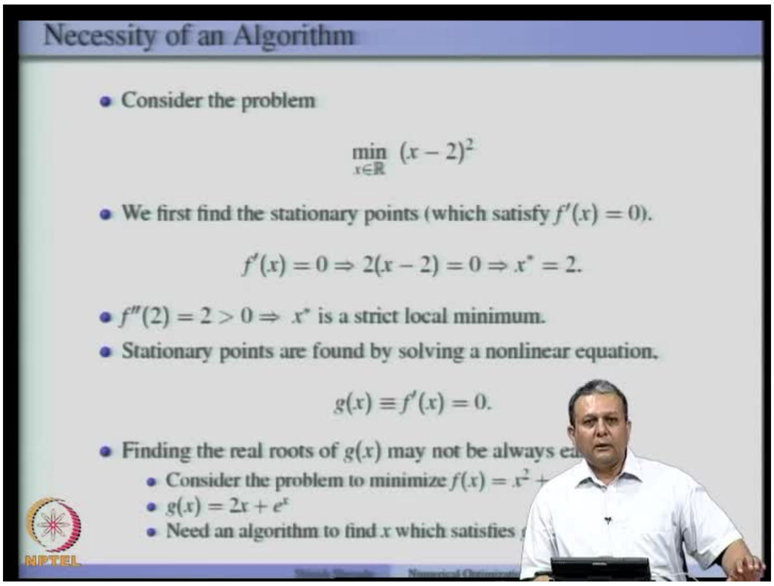
(Refer Slide Time: 51:35)



Now, if you look at example x to the power 4 so, you will see that f 3, f 4 dash x is 24, which is positive and all the earlier derivatives, right from the first, second and third, all these derivatives are 0. So, the first non-zero element of the sequence f k x star occurs at k, which is even and positive and the element is positive so, which means that, 0 is a strict local minimum.

So, so you will see that, to solve such problems, we need to find the derivative of the function and identify the stationary points by equating the derivative to 0 and then go to the higher other derivatives from the second order derivative onwards, to see whether we can conclude anything about the stationary points, whether they are local minima or local maxima.

(Refer Slide Time: 52:42)



Now, if the problem is such a simple problem then, why do we need any algorithm to solve the problem. So, here is the simple example, minimize x minus 2 square and f dash x equal to 0 implies, x star equal to 2. So, 2 is a stationary point and at that stationary point, you will see that the f two dash 2 is greater than 0, which implies that, x star is the strict local minimum.

So, as I said that, if it is such a easy thing to solve, to, to solve f dash x equal to 0 and identify stationary points then, where is the need of an algorithm. So, here is an example so, let us consider the example to minimize f of x equal to x square plus e to the power x where, f is the function from R to R. Now, the derivative of the function, which will denote in this course mainly using the function g of x, that turns out to be 2 x plus e to the power x.

And if we equate it to 0, we really cannot find a solution of this problem in the straight forward way like what we did here. So, we need some algorithm to find out x, which satisfies 2 x plus e to the power x equal to 0 and that is where we will need a numerical procedure or algorithm to solve to find a stationary points of these kinds of functions. So, in general, it is difficult to find the stationary points using a simple way like this and one has to go for some algorithm to solve this.

Now, there exist different kinds of algorithms to solve this problem so, some algorithms they do use the derivative information, some algorithms do not use the derivative information. So but assume certain form of the function called unimodal function so we will see those things in the next class. So, in the next class, we will look at some algorithms to solve the problem of the type f dash x is equal to 0 or finding the minimum of the function f of x without resorting to any derivative information.

So, those are called the derivative free methods and the methods, which we use derivatives they are derivative based methods. So, we will see some of them, to solve the problems of the type g of x equal to 0 or the, or the problems where the, the derivative of the function vanishes or the function has the minimum. So, we will stop here and continue in the next class.

Thank you.