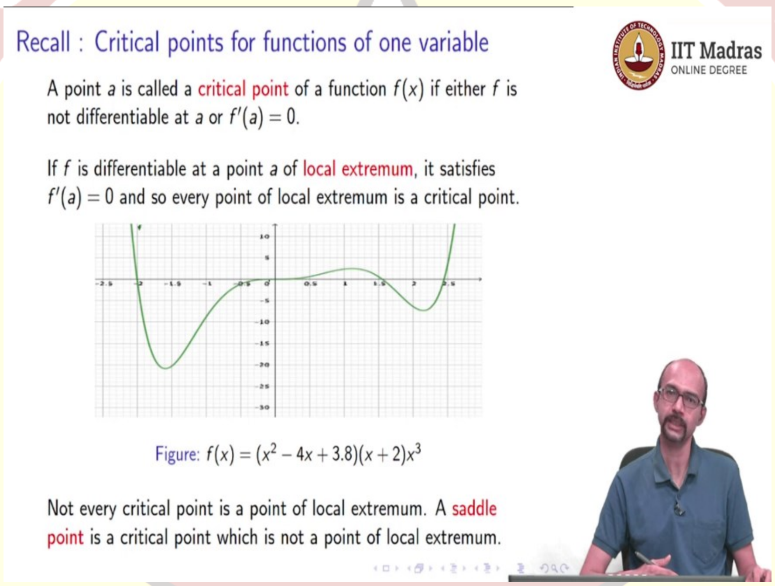
IIT Madras

ONLINE DEGREE

Mathematics for Data Science - 2
Professor. Sarang S. Sane
Department of Mathematics
Indian Institute of Technology, Madras
Critical points for multivariable functions

Hello, and welcome to the Maths 2 component of the online B.Sc. program on data science and programming. This video is about critical points for multivariable functions. So, we have studied the notion of critical points for functions of single variable.

(Refer Slide Time: 0:27)



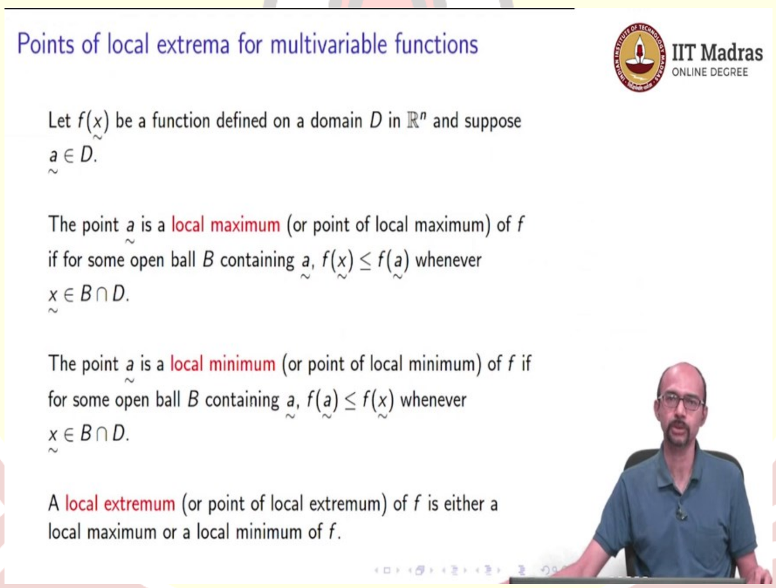
Let us recall what that was. So, for functions of one variable, a point a is called a critical point of a function fx, if either f is not differentiable at that point or f prime at a is 0. So, if f is differentiable at a point of local extremum, it satisfies that f prime a is 0 and so every point of local extremum is a critical point. This is exactly what we used in order to find the points of local extremum. So, we set f prime of the function to 0, calculated what are all the points, what are the critical points, and in amongst them, we did something further in order to determine what kinds of points they were.

So, here is an example. We had seen the same picture when we did that as well. So, we can see here that in this picture, we have several points of local extremum. For example, there is a point over here, which is a point of local extremum, close to - 1.5. There is a point close to 2. So, these are both local minimums, minima. And then there is a point between 1 and 1.5, which is a local maximum.

And unfortunately, we can also see that there is another point, which is point 0, which is what we had called saddle point. So, not every critical point is a point of local extremum and this is why we needed refined tests or more other ways of handling critical points in order to determine which of these are actually local maxima or local minima, because some of them can also be saddle points. So, what is the saddle point? It is a critical point which is not a point of local extremum.

So, here, as an example, for 0, what happens is, f prime is 0 because you can see that the tangent line is actually the x axis. But it is clearly not a point of local extremum, because on the left, you have points for which the values are, f of those values are less than 0 and on the right, they are bigger than 0, and at 0 it is equal to 0. So, this was a summary of what we did for functions of one variable. Of course, we had refined tests. For example, we went beyond this and after critical points we also looked at the second derivative and based on that we could classify them in some cases.

(Refer Slide Time: 3:00)



So, for multivariable functions, we want to do something similar. So, let us start reviewing what are points of local extrema for multivariable functions? So, let f(x) be a function defined on a domain D in Rn, and suppose you have a point a belonging to this domain D. The point a is a local maximum or a point of local maximum of f if for some open ball B containing this point a for all points in this open ball B which belong to the domain f evaluated at those points is less than or equal to f evaluated at a.

So, you want to construct a ball so that within this ball for all those points on which f is defined, f of a is the largest. So, this need not be a ball centered at a, it could be a ball which is containing a. So, but of course, if you can create such a ball you can always create one centered at a. Similarly, the point a is a local minimum. If the same thing happens except that now f of a is the smallest value within that set.

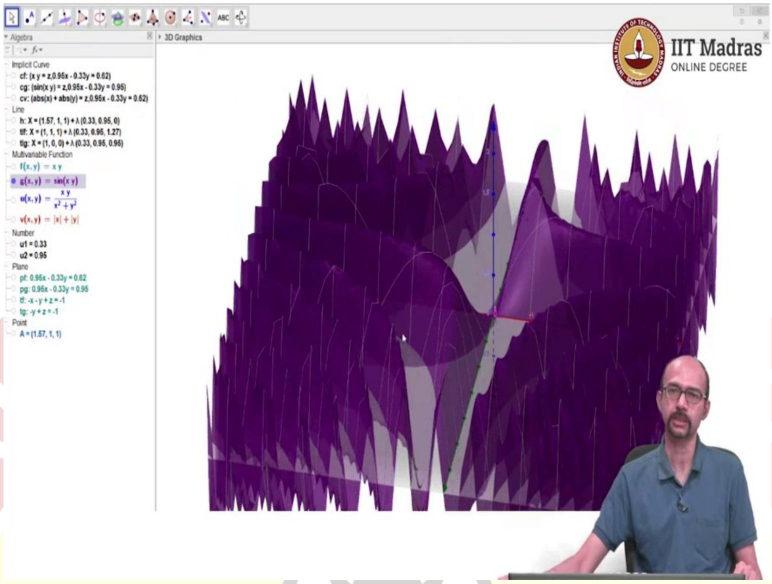
So, remember that for one variable we had open intervals. We said, if on a small open neighborhood, the point has the least or the largest value. So, here instead of intervals we have the open ball. So, in R2 you will have a disc. In R3 you will have a, you will actually have a ball, meaning a sphere. And in Rn, of course, it is the definition, meaning all those points which are at a fixed distance from or distance less than a particular number around one point.

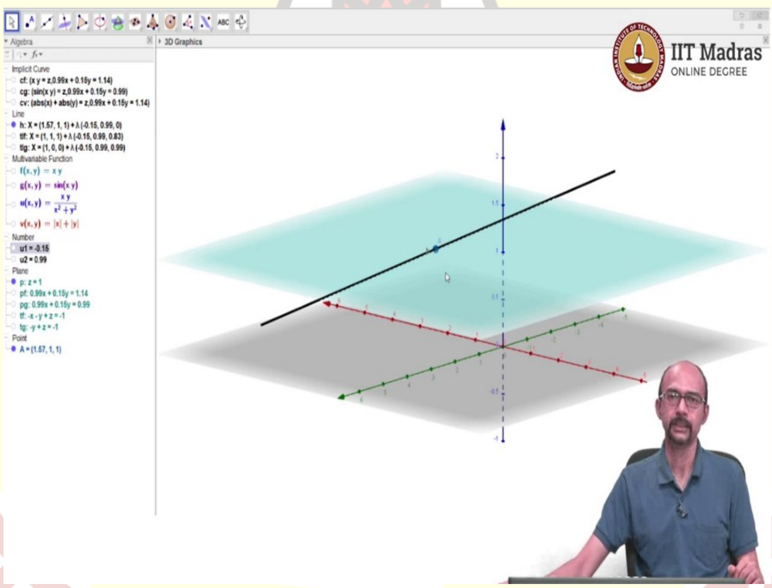
So, such a thing so, the local maxima are indeed, locally maximum values. The local minima are indeed locally minimum values. And a local extremum is either a local maximum or a local minimum. So, these are the definitions as far as multivariable functions are concerned, exactly mirroring what happens for one variable functions.

So, for one variable functions, we just saw that the derivative is always 0 at a local extremum if it exists. And the idea there was that if you are at local extremum, then the tangent is parallel to the x-axis. So, for the local extrema in the multivariable situation, you would expect something similar. So, the notion of tangent plane or tangent hyperplane is as we saw what replaces tangent line.

And as we also observed, the tangent hyperplane or the tangent plane is the equation of that is governed by the gradient function. And we have also seen, when we did the behavior of gradients, that gradients mirror, the properties of gradients mirror that of derivatives for one variable functions. So, it is natural to think that the role of f prime will be played by the gradient. So, let us look at an example first to observe what happens to the tangent plane and what happens to the gradient.

(Refer Slide Time: 6:21)





So, let us look at the function sine of x, y. So, for this function, so here is the graph of that function. We have seen this graph several time before. And the reason I have chosen this function is because for this function, it is very clear what are the local extrema. So, the sine value takes maximum value 1 and minimum value - 1. And you can see that when you are at 1, so there is this ridge over here. So, all those values are the values 1, so those are going to be local maxima.

And in fact, they will be global maxima, meaning they are maximum amongst all values and the function takes. So, in particular, their local maxima, and let us see what happens to the tangent plane at these points. Let us choose a point at which we know this local maximum occurs. For example, if you look at the function sine of x, y, we know that when x, y isπby 2, then it takes value 1.

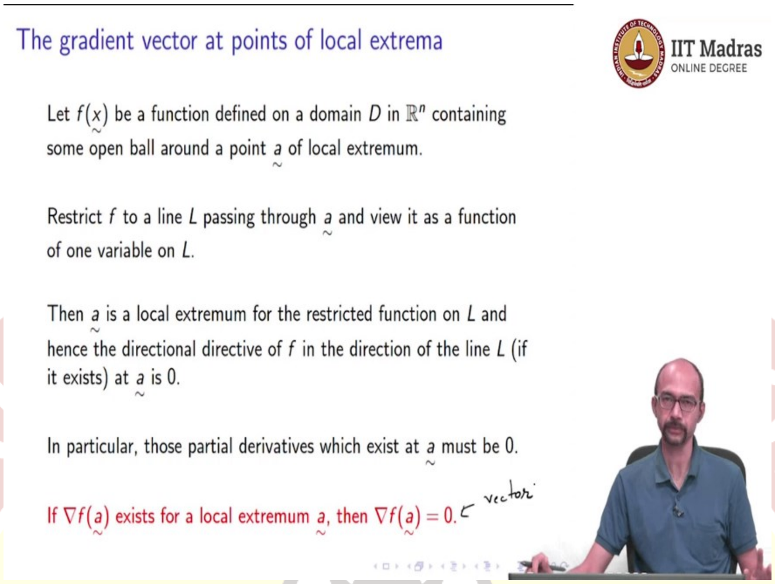
So, if I take x to beπby 2 and y to be 1, then that is going to be a point where you will have a local maximum, and the function value there is 1. So here is that point that is the point a,πby 2,1,1. And let us ask what happens to the tangent plane at that point. So, for that, we will see how all the tangent lines behave. So, if you take a tangent line at that point, let us see how that behaves.

So, here is a tangent line in the direction of the unit vector u1,u2. I hope you can see that it is indeed a tangent line. And, as I vary, u1 and u2 this tangent line is going to vary. So, here is how it varies. So, it is indeed tangent to the graph. I hope that is apparent from the picture. And if I remove this graph, so if I remove, let us see it like this, you can see it is tangent to the graph. It is touching, it is satisfying the idea that it touches the function at that point.

And if I remove the graph now, you can see it varies in the plane. And what is that plane, that plane is exactly parallel to the x, y plane. So, that plane is, in fact, in this case, it is a plane z is equal to 1. So, we can play it like this. Now, they, from this perspective, they are all lines, and you can see that this line is varying on that plane. So, it is on this plane.

So, what this picture tells us is that as in the one dimensional case where we had the situation that the local extrema, the tangent lines, if they exist were parallel to the x axis, so for the two variable case, they will be parallel to the x, y plane, and in general, you expect them to be parallel to the x1, x2, xn plane. So, let us study this in terms of the gradient, because after all the gradient determines the tangent plane or the tangent hyperplane.

(Refer Slide Time: 9:21)

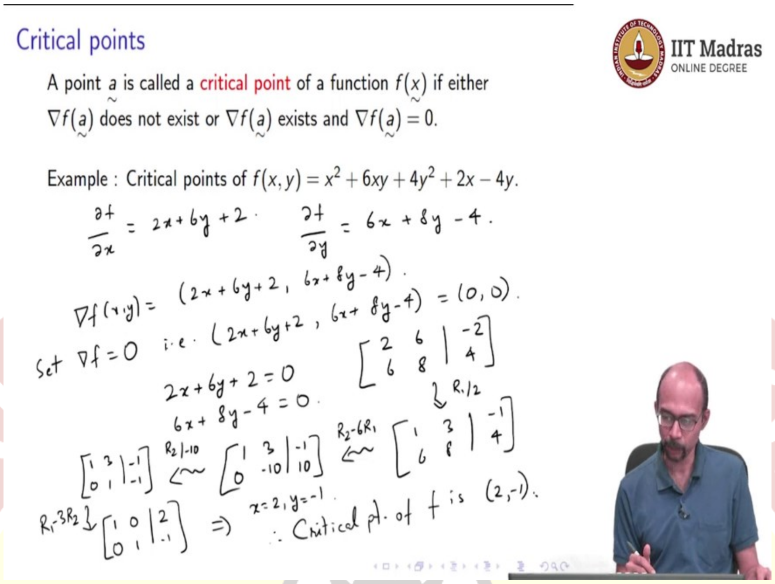


So, let us consider the gradient vector at points of local extrema. So, let f(x) be a function defined on a domain D in Rn containing some open ball around a point a of local extremum. Restrict F to a line L passing through a and view it as a function of one variable on L. Then a is a local extremum for the restricted function on L and hence the directional derivative of f in the direction of the line L at a is 0.

This is exactly what we saw in that example for sine x, y when we saw the directional, the tangents along each line in various directions. So, they were all, the directional derivative was 0. And as a result, they had no inclination with respect to that plane. So, they were all flat. So, they were parallel to the x, y plane. So, in particular, those partial derivatives which exist at a must be 0. So, if all the directional derivatives which exist must be 0, in particular, the partial derivatives, assuming they exist, must be 0.

And so the conclusion is, if gradient f exists for a local extremum a, then gradient f of a is 0. I will repeat that, if the gradient exists at that point, then the gradient must be 0. And when we say 0 here, we mean the 0 vector. So, this 0 here is the 0 vector. So, this is the vector. So, this gives us one way of understanding how to obtain local extrema. So, what we can try and do is set the gradient to 0 and then look at all those points which we obtain. And then out of them, some of them may be local extremum. This is exactly the strategy followed for the one variable case and such points are called critical points.

(Refer Slide Time: 11:10)



So, point a is called a critical point of a function f of several variables, if either the gradient does not exist or the gradient exists and equals 0. Again, 0 here means the 0 vector. So, let us look at an example. So, what are the critical points of this function f(x), y is x square + 6xy + 4y square - + 2x - 4y. Well, so this is a very nice. It is a polynomial function. So, here, there is no issue about gradient not existing. So, let us compute what are the partial derivatives.

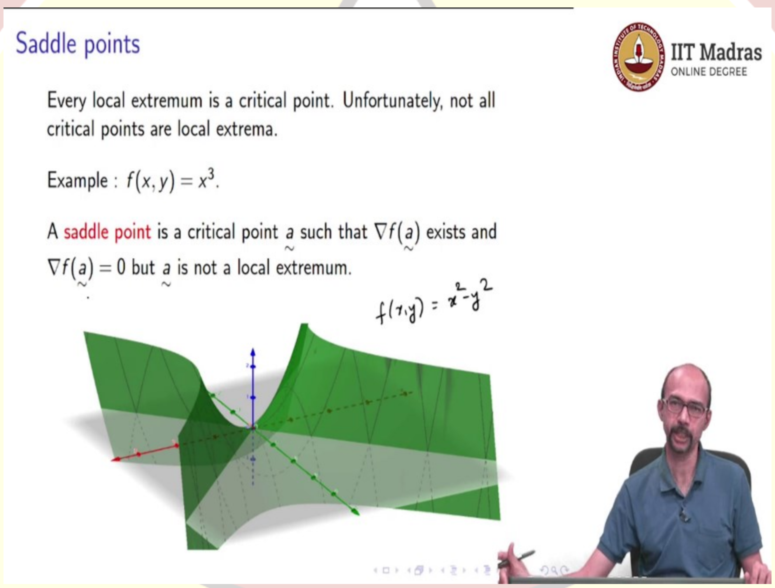
So, del f del x is 2x + 6y + 2, del f del y is 6x + 8y - 4. So, for the critical points, I set the gradient function to 0. So, the gradient function is the, at x, y is the tuple, 2x + 6y + 2,6x + 8y - 4, set it to 0, so set gradient of f to 0 and by 0 we mean the 0 vector. So, that is 2x + 6y + 2,6x + 8y - 4 is 0,0. And now we have to solve these. So, the solutions for these will give us all the critical points.

So, how do we solve this? Well, this is a very nice system of linear equations. And we actually know how to solve this. So, let us use Gaussian elimination to solve these system, this system of equations. So, if we do that, we get 2, 6, - 2 and then 6, 8, 4. Let us divide the first row by 2. So, we get 1, 3, - 1. Let us use the 1 in the 11 place knock out the 6. So, that gives us 1, 3, - 1, 0 and then 8, - 6×3, so 8 - 18 which is - 10 and then 4 - 6×- 1, so 4 + is 6 is 10.

So, if I divide the second row by 10 or - 10, I get 1, 3, - 1, 0, 1, - and now finally I can, final step I can do R1 - 3R2, so that gives me - 1, - 3×- 1, so - 1 + 3 which is 2. So, this means x is 2, y is - 1 and that is exactly the critical point. So, there is only one critical point and it is 2,- 1. So, I hope it is clear how I am computing these. So, you set your gradient vector to 0 and then you will get a bunch of equations to solve. Need not be linear this time we were lucky. And then you have to solve them and the solutions will give you the critical points.

Now, there could be lots of critical points. There is no reason for them to be for only one or two critical points to be there. For example, we saw the function sine of x, y. And if you compute the critical points there, we know what the gradient is, I will suggest you set it to 0 and see what happens. So, you will see that you get a bunch of equations. So, one solution is going to be 0, 0, but then there will be tons of solutions corresponding to cosine of x, y is 0, which we know is an infinite set. So, this is how we get critical points.

(Refer Slide Time: 15:51)



Now, sometimes it may happen that the local extrema are not the critical points, meaning there are more critical points than just the local extrema. So, every local extremum is a critical point that we observed already. Unfortunately, not all of them are, not all critical points are local extrema. Otherwise, we could be happy and not bother to go ahead. So, as an example, here is f(x), y is x^3.

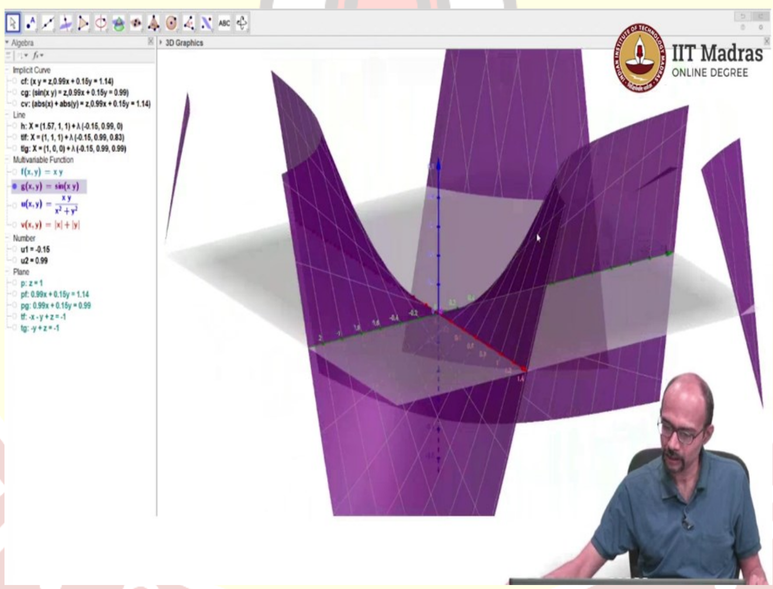
So, we have seen the x^3 actually, that is an example of a saddle point in one, the one variable situation. And the same thing happens here also, because if you recall how $f(x)^3$, $f(x) = f^3$ was, it was a function like this, and there was a problem at 0. And now you take f(x), y^3, you are taking the same function but stretching it along the y axis, because y has no role to play as such.

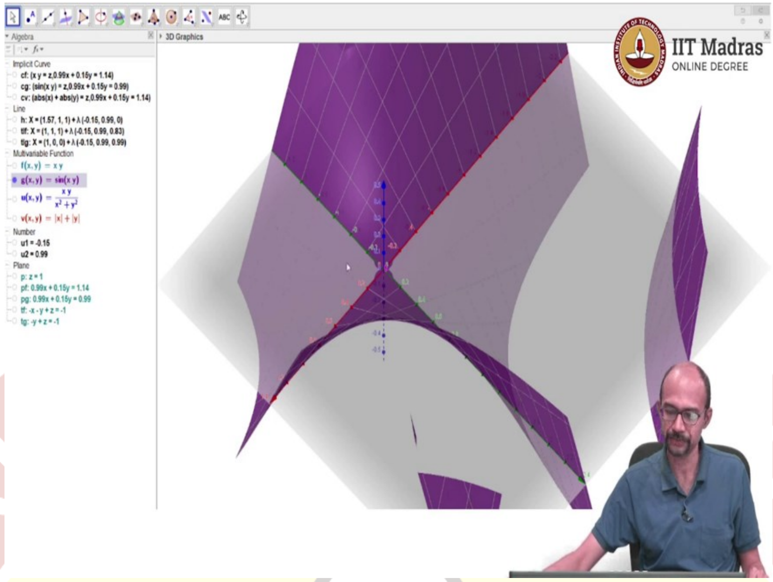
So, the entire y axis will consist of points for which the gradient is 0, but they are not local extremum. I will suggest that you check this. So, a saddle point is a critical point a such that the gradient of f at a exists and equals 0, but a is not a local extremum. So, all these points that I just said, along the y-axis, those are all going to be saddle points for the function f(x), y is x ^{3}.

So, here is an example of a saddle point, which is what gives it its name. So, this is the function f(x), y is x square - y square. And if you look at this function, what is happening is on one side, it is like this. And on the other side, it is like this. So, because you have plenty of directions, you can have this weird situation where it is a local minimum along some lines and it is a local maximum along some other lines, and that is exactly what is happening here.

So, it is, gradient is 0. So, in fact, all the directional derivatives are 0. And what is happening is that in each direction, it is either a local maximum or a local minimum. But globally, it is neither, neither a local maximum, nor a local minimum. And this function or this picture actually gives us why it is called a saddle point, because it looks like a saddle of a horse.

(Refer Slide Time: 18:24)



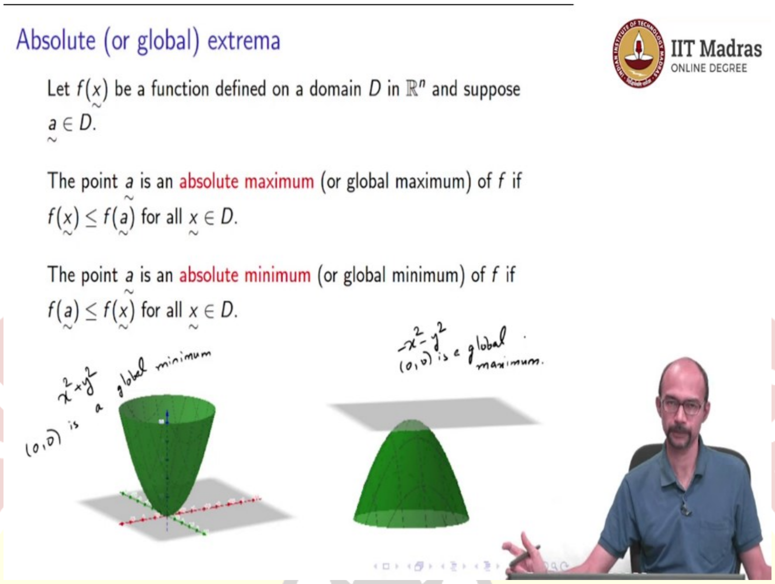


Let us look at a couple of other examples of saddle points, one such is going to be at when x is 0 and y is 0 for the function sine of x, y. So, I will zoom in and then maybe you can see what is happening. If I go in here, now this is very close to 0, 0. And let us see if I can raise this up. So, you can see what is happening. So, over here in one, on one direction, this is a local minimum, and you can see that very clearly. But on the other side, the function is actually decreasing. So, over here, it is a local maximum. And so this function is a, this function as a saddle point at the point 0, 0.

So, if you take the line y is equal to x, then the function is like sine of x square. So, x square is always positive as we know. So, on that line, the sine is going to increase from 0 onwards. So, on both sides it will increase. On the other hand, if you take the line y is equal to - x, then you are going to get sine of - x square and which is - sine x square, which will always be negative when between, when x is between 0 and $\pi/2$. So, it will, on the other side, it will decrease. And so this is a saddle point as a result. So, it is a point at which the gradient is 0, but it is not a local extremum.

So, now that we have identified the critical points as local minimum, local maximum or saddle points, we can ask how do we identify them and so for that we will, like the function of one variable where we add a second derivative test, there will be something similar for the function of many variables and we will develop theory for that in the next video.

(Refer Slide Time: 20:26)

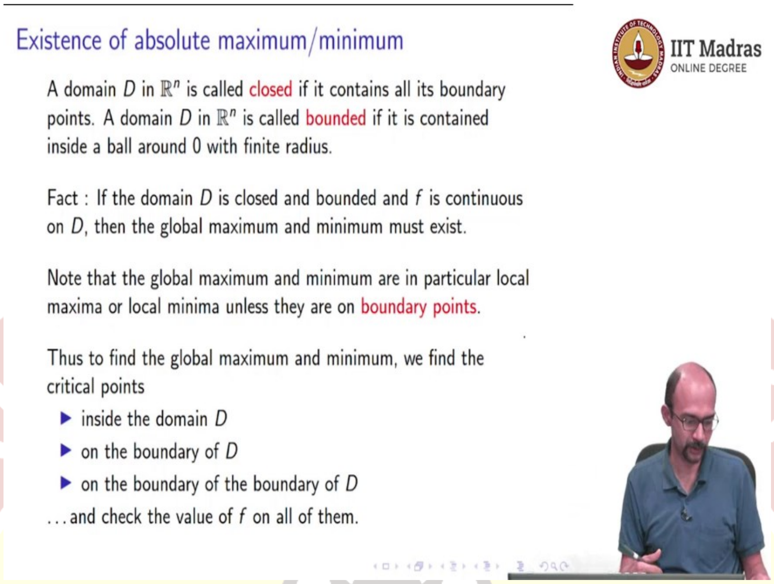


So, the final topic that we will do in this video is the notion of absolute or global extremum. So, suppose we have a function defined on a domain D in Rn and suppose we have a point in D. So, this point is an absolute maximum or global maximum if the value of f assumed on that point that is f of a is bigger than f(x) for all x within that domain. So, notice here that we are not seeing local here. So, for local maximum we said on some small ball it has a largest value that is not what we are seeing here. We are saying across the entire domain it has the largest value.

Similarly, for absolute minimum or global minimum, it has the smallest value amongst all points in the domain. So, here are two examples. These are the prototypes of absolute maximum and absolute minimum. So, this is a function of $x^2 + y^2$ and the point 0, 0 is a global minimum or the function attains a global minimum at the point 0, 0. And this is the function $- x^2 - y^2$ and 0, 0 is a global maximum.

This is clear because we know that squares are always positive. So, from there it is clear that these are global minimum and maximum respectively for these functions. So, these are prototypes of global maxima.

(Refer Slide Time: 22:06)



So, then how do we find, how do we know they exist and how do we find them. So, for the existence similar to the one variable case, where we restricted ourselves to what we called a closed and bounded interval, here we have a notion called closed and bounded domain. So, a domain D in Rn is called closed if it contains all its boundary points. So, what do we mean by boundary points.

So, a boundary point is one which is on the boundary of that domain. So, this is visually very clear what we mean by the boundary of the domain. There is a technical definition let us not get into that. So, domain D is Rn is called bounded if it is contained inside a ball around 0 with finite radius. So, what this means is it is bounded means it can vary around but it should be like this.

So, there is some huge ball. It could be very, very, very large. And the domain lies entirely inside that. It could have radius as large as you want maybe 1,000 or 20 million whatever, but there is some domain, there is some ball of as large radius as you want inside which this domain lies. So, it does not go out.

So, for example, the x, y plane is not bounded, because it shoots off. There is no ball in which, there is no disc in which you can contain that or if you take just the let us say a as a cone inside that, say the first quadrant in the x, y plane that is not bounded. So, you want something which is inside a large circle or a large disc rather.

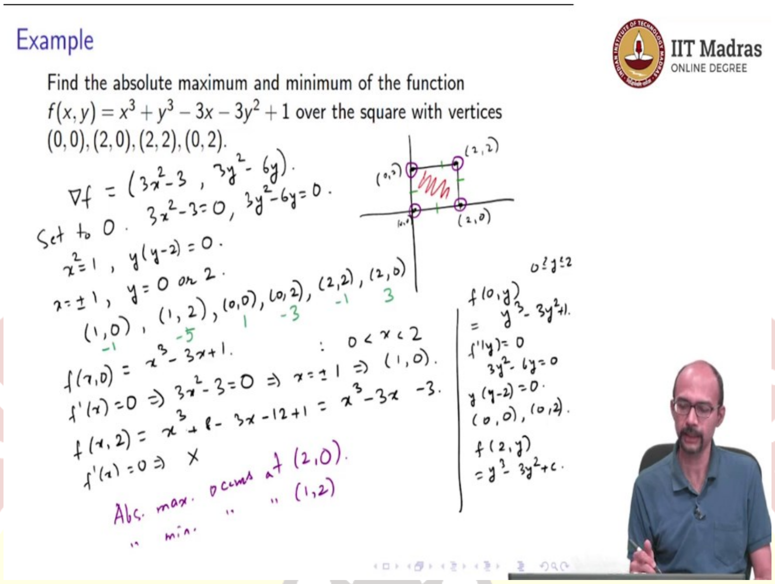
So, here is the main fact, if the domain D is closed and bounded and f is continuous on D, continuity is important, then the global maximum and minimum must exist on this domain. So, closed and bounded guarantees the existence of the global maximum and minimum. There could be more than one point. So, there could be several points in which the same thing happens. For example, we saw the function sine of x, y, several points where it takes the value 1, several points where it is the value - 1.

So, note that the global maximum and minimum are in particular local maxima or local minima, unless they are on boundary points. And this statement is what gives us the key to finding them. Thus, to find the global maximum and minimum we find the critical points inside the domain D. So, remember, critical points are inside a domain. We have to have, the critical points are defined inside where there is a ball around this point lying inside D.

So, then that leaves us with boundary points. So, you restrict your function now to the boundary. It is in one, boundary means it is one dimension less. So, you found find the critical points on the boundary. Then the boundary itself may have a boundary. So, you find the critical points on that, and then you check, so you continue this process. So, if you are in n dimensions, you have to drop dimension all the way till you reach a point.

And then you check the value of f on all the critical points that you obtain. And amongst them you choose whatever is maximum or whatever is minimum those will be of global minimum or maximum. We have done this exactly this process when we did the one variable functions, when we check for critical points within the open interval, and then we checked on the boundary points.

(Refer Slide Time: 25:39)



Let us do an example to set ideas. So, find the absolute maximum and minimum of the function, f(x), y is $x^3 + y^3 - 3x - 3y^2 + 1$ over the square with vertices 0, 0, 2, 0, 2, 2, 0, 2. Let me first draw this square. So, here 0, 0, here is 2, 0, here is 0, 2, here is 2, 2, and here is 2,0. Let us what happens for this. So, now, here, there is an open part inside. So, this is the open part of the square.

So, first, we have to check what happens, what are the critical points inside this for this function, then we have to restrict this function on each of the edges. So, then we will restrict this function on this edge, this edge, this edge and this edge, see what happens. And then finally, we have to check what happens to the functions, to the function on these four points. That is the strategy. That is what we are going to do.

Let us start with what happens inside the square. So, for that, we will find gradient. Now, first of all, this is a polynomial function. So, all the issues with, whether partial derivatives exist, etcetera, they do not play any role. So, we do not have to bother about points where gradient does not exist or they are not differentiable when you view them on the sides, etcetera. So, that is not an issue. So, let us write down what is the gradient. So, this is 3x square - 3,$3y^2$ - 3, - 6y.

Let us evaluate, let us set this to 0 and see what are the critical points. So, $3x^2 - 3$ is 0, and $3y^2 - 6y$ is 0, let us try to solve this. So, this is saying x squared - 1 is 0. So, x square is 1. The other equation is $y^2 - 2y$. So, y - y×y - 2 is 0. So, we know the solutions for these. So, the possible solutions are x is + - 1 and y is 0 or 2. These other possible solutions. Now, let us see which of these lie inside our square. So, when x is - 1 that does not lie inside our square, so we can discard that.

So, the critical points coming from, coming inside here are the points where the x coordinate is 1, the y coordinate is 0 or the x coordinate is 1 and the y coordinate is 2. So, this point 1,2 actually lies on the boundary, but we can consider it anyway just to check, although in principle this is not a critical point when we restrict to the open part, but we can keep it anyway. So, these are the critical points for the square. So, this is one set of critical points.

Now, let us restrict it to the boundary. So, first, let us do it on the bottom boundary. So, that is where y is 0. So, now we want to look at the function f(x),0, where x is between 0 and 2, actually, strictly between, but again, we can be considerate and take the end points as well because anyway we are going to take them. So, f(x),0 is $x^3 - 3x + 1$. So, the derivative let us set that to 0. That means $3x^2 - 3$ is 0.

Well, we have solved this equation above. So, that means we get x is $\pm$ 1. So, which means we should take the points, 1,0, - 1,0 does not lie in the, in this interval. So, 1,0 is already there. So, these two points, actually both these points are on the boundary. So, we will take them anyway. So, this is already taken care of.

Similarly, we can take the top edge. So, for that the function is f(x),2. So, the function is x^3 -, $x^3 + 8 - 3x - 3\times4$, which is - 12 + 1, that is $x^3 - 3x$ and + some constant, does not really matter, so - 3. So, I set it to 0. And again, I get the same thing. So, but these are not on the boundary. So, I do not have to bother about this. So, I do not get anything here.

Let us do the two ages, two other ages. So, we have f(0) ,y between 0 and 2. So, if we do that, we get $y^3 - 3y^2 + 1$. But actually, we have solved this as well. So, this gives us $3y^2$ is - 6y is 0, so $y\times y - 2$ is 0 which means y is 0 or y is 2. So, the points that we get from here are 0,0 and 0,2, which anyway we are going to take, so let us add them in 0,0, 0,2.

And for the other part, where we have f of 2,y, we will get the same function but with a different constant, so + c, so we will get the same points. So, we have added them in already. So, that means we have taken all the points which are critical points on the inside on each edge and now we will add the boundary points, the corner points as well. So, that is 2,2, 2,0, and the others are already taken. So, these are the six points at which we need to check for maximum or minimum. And if we do that, we get, so if I evaluate that, let us see what we get.

So, I had 1,0, I get 1 - 3 + 1, so that is 2 - 3, so I get - 1. For 1,2 I get 1 + 8 - 3, so 9 - 3 is 6 - 3×2 square, so - 12 so 6 - 12 so - 6 + 1 so - 5. For 0,0 I get 1. For 0,2 I get 8 - 12 + 1, so - 4 + 1 so - 3. And for 2,2, so I get 8 + 8 16 - 6 is 10 - 12 so - 2 + 1, so - 1. And for 2,0 I get 8 - 6 so 2 + 1 so 3.

So, the absolute maximum is 2,0, so absolute maximum occurs at 2,0 and absolute minimum occurs at 1,2 and the values are 3 and - 5, respectively. So, I hope this example explains what you have to do in order to get the absolute maximum and minimum over a closed and bounded domain. Thank you.