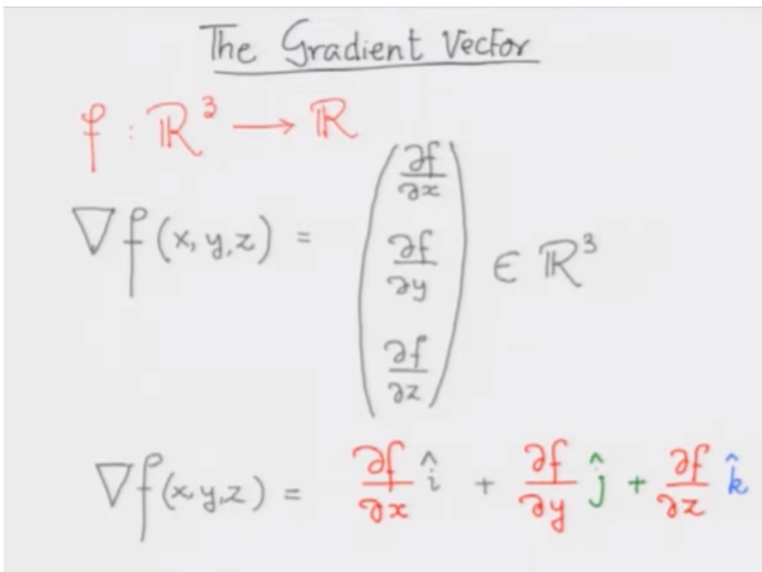
Calculus of Several Real Variables
Prof. Joydeep Dutta
Department of Economic Sciences
Indian Institute of Technology – Kanpur

Module No # 03
Lecture No # 11
The Gradient Vector and Directional Derivative

So we start of third week and as we proceed you will see the level gets slightly advanced we have spoken about a gradient vector when we spoke about the derivative of a function in high dimension f from R, N to R, R2 to R whatever we spoke about the gradient. Gradient is nothing but a column vector which consists of all the partial derivatives.

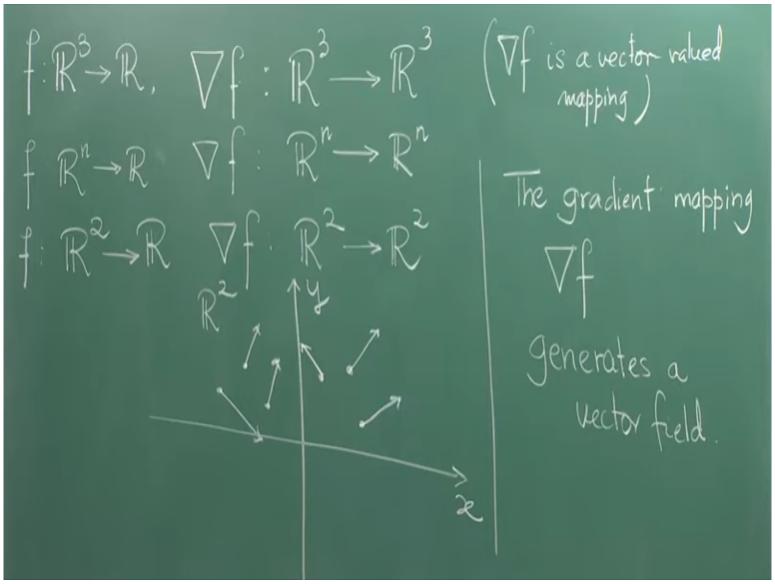
(Refer Slide Time: 00:46)



So gradient vector so if I have f say f from say from R3 to R so in this case the gradient of f at any point x, y and z is a vector whose entries are the 3 partial derivatives. So it is a vector with 3 rows and 1 column 3 cross 1 so this is the vector in R3 of course this and these are coordinates vector grad f can be f at x, y, z can be also written in the way which the physicist would like it to be written as follows that del f, del x i vector + del f del y j vector + del f del z k vector.

So physicist would possibly expect writing like this so why are we interested in talking about the gradient vector. The first thing to be note about the gradient vector is the following in fact with other matter where which space you are in R3 or R2 or R1 it does not really matter but.

(Refer Slide Time: 03:00)



So what does grad f do so if have function f from R3 to R grad f picks up an element in R3 and sends them to another element in R3 so grad f is on mapping so grad f is a vector valued mapping. So in general if f is from Rn to Rm sorry (()) (044:03) f is from Rn to R and grad of f gradient vector is also mapping from Rn to Rn. Interesting part is that this would domain and the core domain of this mapping are same.

Now let us look at the mapping f from R2 to R and grad f is the mapping from so it takes in any element in R2. Assuming that the gradient exist so of course when I am just writing you do not take it to be any arbitrary function to take a function for which the partial derivative is exist. So let us take this kind of function and observe what does it mean at any point in R2 I get another point in R2. So if this is my R2 the x axis and y axis this is my space R2 then take any point the at any point in R2 grad f take this and makes gives me another vector.

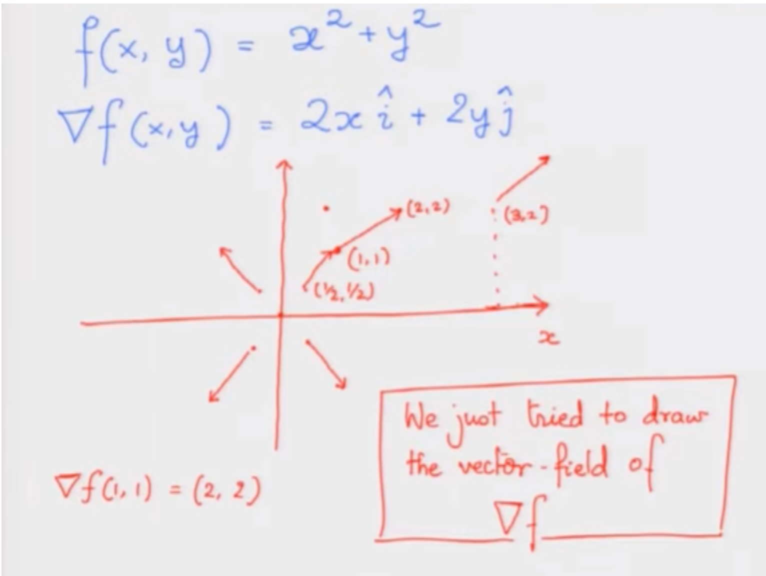
So it gives me some vector in a given direction so from here it talks me about a vector in given direction another point it can show me a different direction another point will show another direction right. So these is this full of such (()) (05:51) at every point there is a kind of direction. This is an arbitrary drawing taken off from the mind it just to tell you and at every point you can have this arrows.

So basically if you take a large number of points the feel would be like filled with small arrows so we say the gradient mapping generates of vector field the gradient mapping grad f generates a vector field it is like a kind of you can say why this name field this field name as (()) (06:47) when the time and (()) (06:50) though about electric field that every point in space there is a charge.

So as if the charge seems to flow from one point of the space to allow so it is like the that is what is called as velocity field in mechanics that if a particle is moving in space then at every point there is a vector which is tangent who which is tangent to the path of velocity vector the tangent to the path of the motion of the curve and motion of the particle along the curve. At every point when you take a tangent you generate a field of velocity.

So the velocity field is a vector field so this that is again generated so gradient mapping right or the derivative mapping it is a very simple for example let us take an example right out of this book which we are referring continuously the book again by (()) (07:48) Einstein that is the standard book which we have using from very first day.

(Refer Slide Time: 07:56)



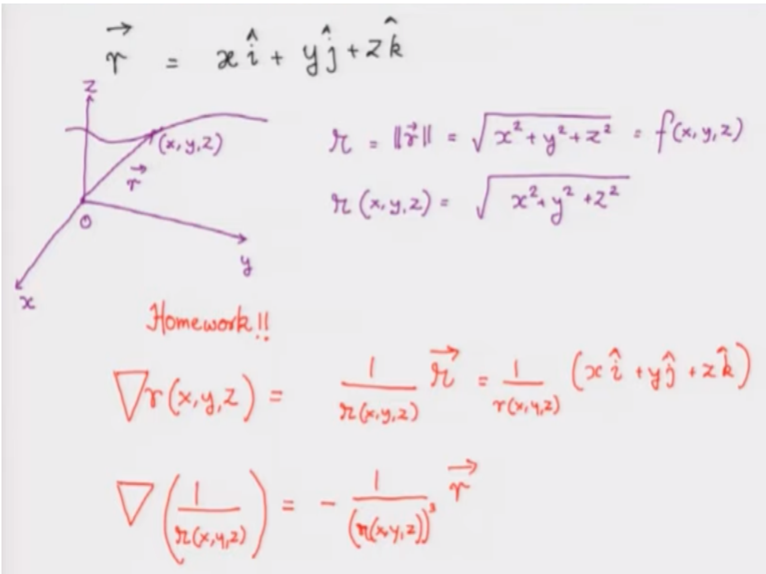
So if we want to take a function fx, y from R2 to R2 and let us function become most simplest one simplest looking quadratic one x square + y square. So grad of f at x, y I can write it in the physicist style as 2x i vector + 2y j vector. Now let me try me to draw the field okay this is xy now let we look at the point 1, 1 at the point 1, 1 the gradient vector is 2, 2 so gradient vector is 2, 2 so the gradient vector is 2, 2 so the gradient vector points in this direction because 2, 2 is here.

So gradient vector is 2, 2 so this is your field at 1, 1 so grad for example here grad f 1, 1 which is 2, 2 0, 0 it remains at 0, 0. Suppose you have half-half it is 1, 1 so if it is half-half if it is half and half then it is the field is here. So at any point you have a field so suppose it is for suppose I take it as 3, 2 right so this is 3 and 2 suppose this is a point 3, 2 at this point where is the vector field pointing would exist 3 to be 6 and then y would be 4 6, 4.

So you go 6 and some direction it is so you look pointing somewhere like this so at every point you know take a point like this it will point in this direction. So you can take a little bit of time and have some fun sketching the vector field of this particular function. So what we did was we just try to draw the vector field of course that is infinite point you do not try at every point you cannot spend lives and lives many lives are insufficient to it that is the whole notion of infinity that it is so incomprehensible but so useful we just try to draw the vector field of grad f that is it this is what we try to do.

Now I give you a homework or a very useful computation of gradient sometime computing gradients is very and these this computation is important for physics.

(Refer Slide Time: 11:22)



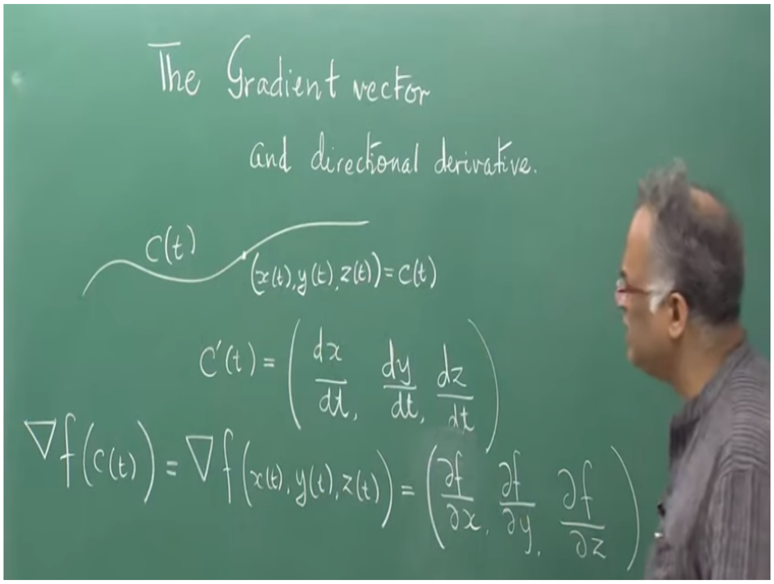
Consider the radius vector of a particle moving in 3 dimensions you already we have already studied of particles move in the very beginning and made vector presentation. So if you have forgotten for you to just remind that okay here is the kind of particle moving through space and here is the radius vector and this point has coordinates x, y and z this as a origin the x axis, y axis and z axis. Some people also call z as z it is up to you what I will always call it z.

Corresponding to the I construct of small function r which I had call as mod r the length of the vector r and that is root over xs square + y square + z square. This is my f of x, y, z. You can also call it r of x, y, z so if you want to be more you do not want to get this f business in you just call it r of x, y, z is root over x square + y square + z square and know this is simple. Now I am giving the homework.

Homework is important I simply hope that you really do it if you are sincere about the course you will try to do it. So gradient of mapping r x, y, z it is 1 / r x , y, z into vector r. That is 1/ r x, y, z scalar multiplying the vector which is x of i vector y of j vector z of k vector you can compute this you can try it out computing the partial derivatives. Also an important one is and in this case the answer should be – 1 / r x, y, z whole cube in to the vector. I am not writing the remaining part you can easily figure out what we have to do.

Now let us focus ourselves on something called a chain rule and we will see how the chain rule brings into four the idea of the directional derivative. So chain rule associated with gradients chain rule. So you might say “oh come on another chain rule yes”.

(Refer Slide Time: 14:55)



Yes another chain rule “my god you must be worried what is happening?” another chain rule is this. So we will take f to be a function of 2 or 3 variables this is just for your visualization it does not matter the number of variables I will let me tell you to have chain rules you do not need really the number of variables can be very different but let me just for simplicity for visualization is just let us be restricted to this.

Now ct let ct we should to 2 I have the chain rule okay so let us ct be a curve on the plain or a curve in 3 space 3 dimensional space. Let h of t be f of ct, ct as xt, yt and zt the curve is defined by that right. And ct has 3 coordinates so xt, yt and zt I will tell you in details let this then assuming differentiability of everything FC everything dh dt of h is equal to the gradient of f evaluated ct dot product with the derivative of this curve ct.

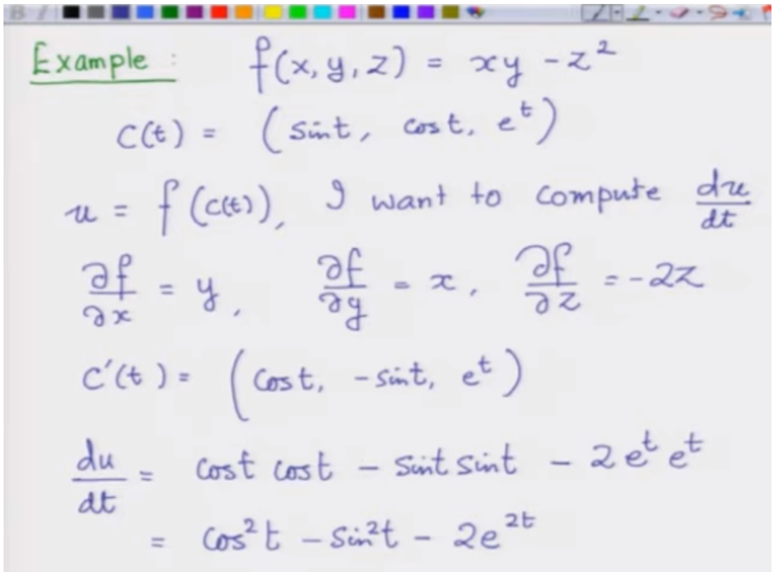
So let us see what does this mean so you are telling that there is a curve ct along which say a particle is moving. So this curve is given so at any point here is defined by xt, yt is defined by a parameter t zt and this any point is ct. ct is actually the radius nothing else the ct the position so that is how what is defined as a curve. Now you know that c dash t is nothing but the derivative of c basic is dx dt, dy dt and dz dt that is it.

See if you look into this whole thing this is basically a dot produce you have to do with the gradient of this but okay what is the gradient of this? So gradient of f at ct means gradient of f at xt, yt, zt basically this is nothing but I am just writing it as a row vector you can write it as a column vector also sometimes I write it just for the space maybe I should write this also the row because you might not be able to see because the table seems to block.

So let me just write it as this del f del x, del f del y, del f del z obviously each of them can be that is how you basically take the gradient and evaluate the gradient at the point xt, yt, zt take it with respect to x, y and t and evaluate or now you take the dot product of this 2 and that basically dot product is a kind of when a del f, del x here evaluated at xt del f del y evaluated at yt del f del x evaluated at zt and then take the dot product evaluated at t.

So this if you write it down in more details is del f, del x evaluated at ct into dx dt + del f del y evaluated at ct dy dt + del f del z evaluated at ct into dz dt that is it that as what dot product means and that is exactly what you will do. Now let us take a small example exactly from the books if somebody was a book can actually check it up. Let us just do it as if we are in a class so let us see.

(Refer Slide Time: 21:31)



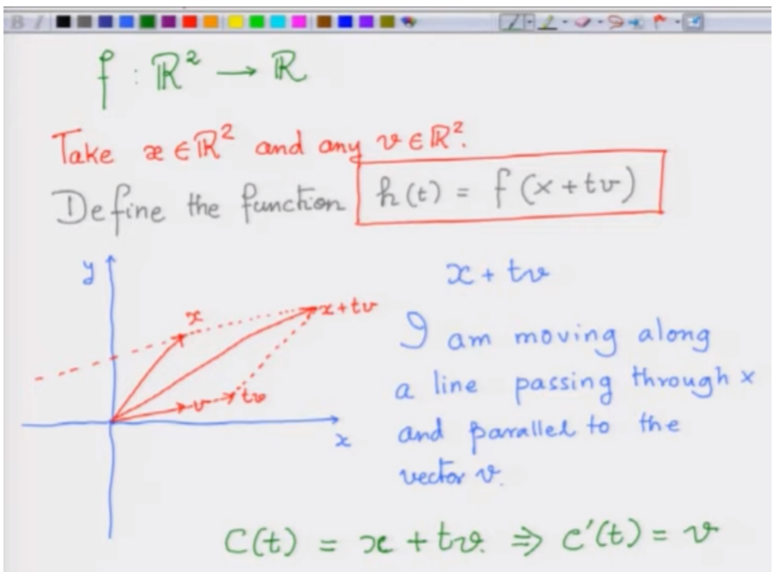
So we will consider the function f x, y, z example of using the chain rule so let us say f of x, y and z which is evaluated it is written as xy – z square where ct is given as so x so sin t x is sin txt sin t, y t is Cos t these are t the t which are not written explicitly it is et. Now what I do I evaluate now I call this so u is f of ct I want to compute du dt so I will just do whatever I did on the board you see I will just have the look at the board and I just first compute the individually del f del x, del f del y, del f del z and evaluate them at the point sin t Cos t and et.

So what is my del f del x my del f del x is y so y would be evaluated at that point so basically which means del f del y is x and del f del z is – 2z. Now c dash t which you know very well is Cos t – sin t and et just taking the derivatives now am taking the inner product so du dt now when I take the inner product what I get when I am taking the inner product I will put in place of y I will put Cos t it will become Cos t into Cos t then I will put x as sin t into – sin t (())(24:10) evaluated at the point ct and then I will put et that will become -2 et into et.

So it will become Cos square t – sin square t – 2 e to the power 2t so this is how very simply using the chain rule where been able to compute the derivative of this function when given that the x, y, z as this following pattern. In fact u can actually be viewed as a function of t itself if I put here this can be written as sin t Cos t – et square and then if you take the derivative du dt this is exactly what you would get as the derivative okay.

Now we shall come to a very interesting idea of directional derivative this idea of directional derivative works even when you are higher dimensions but we are going to talk about high dimensional issue now what we are going to talk about higher dimensional issue and we are going to talk about maxima and minima which will come very soon but let me just remain in space R2, R3 is fine but let us just be in R2 function of 2 variable I will interested to look at things.

(Refer Slide Time: 25:50)



So consider a function f from R2 to R take an x take x element of R2 means you are in the plane and any v element of R2 take these two things. Define the function we could choose v to be a unit vector that is 1 which as normal v or length 1 norm v is 1 define the function ht as so you have a t which is in r f of x so f is fixed + tv so define a function like this. So what is the f of x + t let us understand this whole terminology.

I always put things on positive of into of this side R2 + it is because of just as a matter of habit you can put it anywhere please understand it is just a matter of habit. So take a point x so here is the x vector and take another vector v like this v is vector x is the vector I am not putting the arrow on the top where is should but I am assuming that when i say v element of R2 basically talking about a coordinate and that coordinate represents a vector what is mean tv?

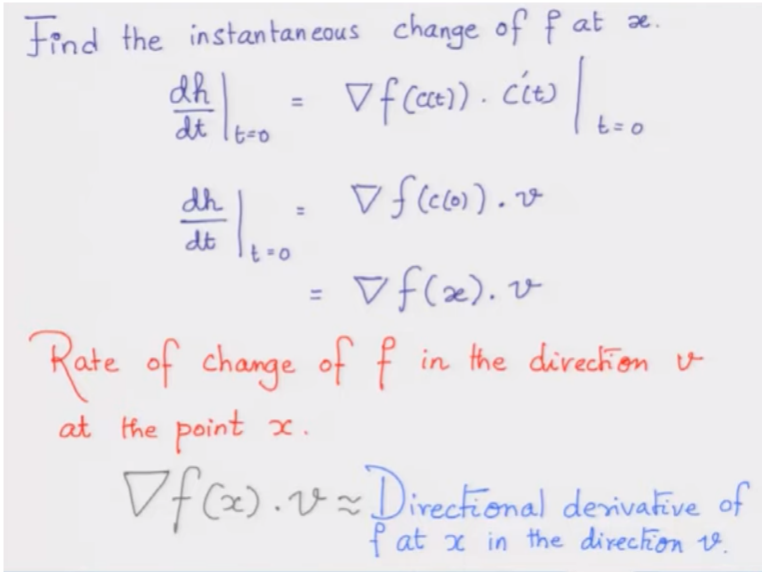
tv means suppose ti say take is greater than 1 for the means I pull this vector v to some point tv t is so now this new vector has become tv whose length is t times v. Now x + tv by the parallelogram law so I will have to draw of vector from x starting from x but parallel to the direction of v and then from the point of tv you know how to draw a vector exactly parallel to x. This new vector is x + tv so what I did I am moving from x along the vector in the direction of vector v and moved a distance tv away from x that is the meaning.

So essentially if I take t2 in negative I can come in this direction so basically when I am talking about this point x + tv what I am doing is I am moving along a line passing through x along a line in R2 of course passing through x and parallel to vv to the vector v. When t = 0 means xo I have f x I have fx when t = 0 so what is my ct here? So we are again using the chain rule so what is my ct? Think of it your ct in this case is x + tv and this implies that my c dash t is v.

So my rate of change of then function along the direction x see when I am when i am talking about the function ht is the restriction of the function f along this line which is passing through x and it is parallel to vector v. So is in the direction of the vector v so what I am trying to say is that c dash t tells you at what rate you can move along this line so it is a as if the velocity along this line would remain a constant this velocity vector would remain constant so you have basically kind of moving along the line with constant velocity even think in that terms.

So what you are doing here you are now going to ask how is the function changing at x what is the instantaneous change of f at x just we talk about instantaneous velocity of particle at time t we are talking about a instantaneous change of the function f at x because as I move from x to x + something x + tv x + as you move but as we move within x and x + tv is the function value keeps on changing.

(Refer Slide Time: 32:04)

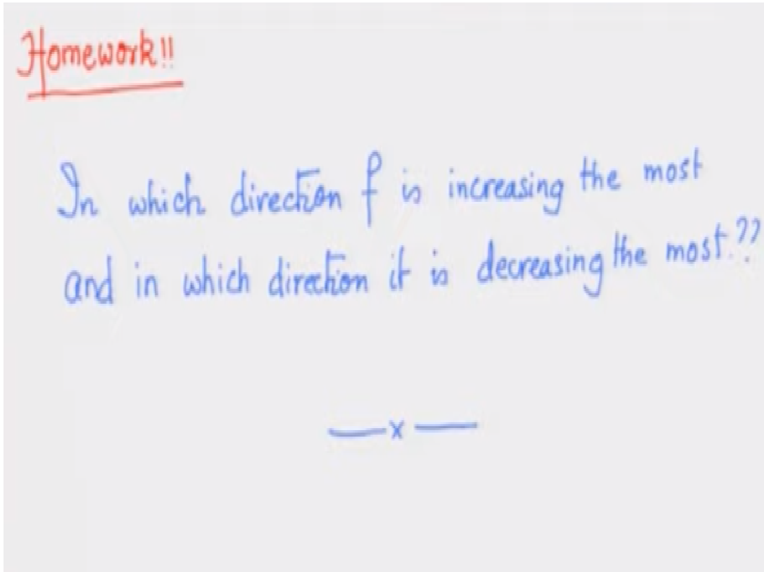


So what is the change at x so my question now is find the instantaneous change of f at x so what I do? So I take the derivative of h and if I put t = 0 I will get back I am at the point x so basically I want to now know what is the dh of dt at t = 0 my joy to my joy I will just do the chain rule the grad f of ct dot c dash t but I know it no want it at t = 0. So my instantaneous rate of change of the function f at the point x.

Because when I put t = 0 x + tv becomes x is nothing but grad f of c0 dot v which is nothing grad f of c0 is nothing but x grad f of x dot v. This rate of change is often called the directional derivative of f in the direction v so it is the rate of change of f at the point of x in the direction v so what is this so it is let us rate of change of x so rate of change of f the function value the rate of change of f in the direction v at the point x and voila this is what is called the directional derivative.

So this thing grad fx v grad f of x dot v is called directional derivative of f in the direction v directional derivative where t of f at x in the direction v or along v direction v. So here we reach our definition of direction derivative I ask you a simple question and you really need to answer me this before if tell you something I will tell you a little bit about how gradient behave in respect to level surfaces I will just ask you a very simple question please take it as the homework.

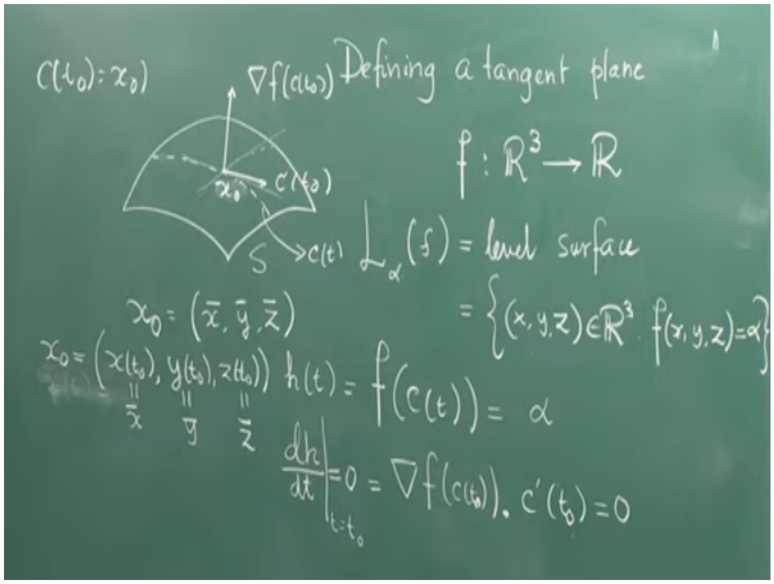
(Refer Slide Time: 35:39)



I will my TA's to look into the lectures and I will ask them to put the solution or I think the solutions are not required fine if you ask the solutions will give the solutions. The homework is simple in which direction f is increasing the fastest and in which direction vf is decreasing the fastest in which direction this very question has deep link to the subject of optimization the part of mathematics which deals with the maximization and minimization of functions in which so this is the important in which direction f is increasing the fastest the most and which direction it is decreasing the most.

Today we will not go to talk about anything more on the this electronic board but will go to the blackboard and try to explain a very very important thing about the structure of the relation how a gradient of a function behaves with the surface associated with the function.

(Refer Slide Time: 37:33)



Take any surface and again any point x0 how would you define a tangent plane to the surface s at a x0 so defining tangent plane how will you do it is not a very difficult job if you just extend your imagination a bit. So I threw x0 think of smooth curves I am just drawing you like dot but they are essentially smooth curve passing. So to every smooth curve take another smooth curve which is passing through x0.

So to this smooth curve the first one I can draw a tangent so the other smooth curve I can draw another tangent and similarly I can draw uncountably infinite such tangent lines and that tangent lines form what is called as tangent plane. So it is not a very difficult thing to define a tangent plane. So basically consider a function consider the set of all x, y, z that is the level consider a function in 3 dimension f from R 3 to R consider function f from R3 to R and consider the level surface alpha of f the level surface.

So what is the meaning of the level surface? Level surface means the set of all x, y, z which is in R3 such that alpha is my required parameter here that is the level at that we are setting it f, y, z xy is alpha. So I suppose that this that is represented by this surface this is the level surface so basically what is happening through each through x naught there are curves to parameter by t's ct is passing. Now what direction or what relation does the gradient vector at a x naught.

So this x naught maybe I should make it right at this point x naught means it is in 3 dimension maybe I should write this at some points x naught which are maybe x naught is actually a point in 3 dimension representing the vector the x bar y bar and z bar. Now what is if want to compute the gradient of f at x bar y bar and z bar but in which direction it is pointing? Is there any relationship of the direction with the surface is there any particular relationship?

This is the question that we are going to ask ourselves now let us see what we get out of it so take this curve as ct consider this curve this one this curve whose tangent line I have drawn this curve let this be given by ct then because this curve lies completely on the surface and for in that surface any point on that surface as to satisfy this I have f of ct is equal to alpha. So if I take the derivative with respect to t I will be have 0 because they said so basically so if I call this as ht I know that dh dt = 0.

And again I know by the chain rule dh dt grad f of ct right dot c dash t and that is equal to 0 at and at what point I am computing it? I am computing it at the point x naught or x bar y bar z bar which is the t bar. So x bar t right I am computing x naught means where x basically I am computing x naught actually means the point. Suppose on the curve ct x naught x bar means actually means the point xt naught, yt naught and zt naught.

So this is actually your x bar this is your y bar this is your z bar so there is a t naught with we respect to what it which corresponds to these values x bar y bar and z bar. But if there is another is on c on t then may be some other point t0 dash where x you have this kind of correspondences but the x bar, y bar, z bar is the same point because it appears in the same point x naught vector in R3. So now what happens what did you get?

So here so this is dh dt = 0 so at this particular point if you want to evaluate so at t naught so t naught you will get 0 so if you evaluated at t = t naught you will get 0 means at this point c dash t naught is tangential see if ct is the curve along which the particle is moving for example on the surface then c dash t naught is its velocity and this is 0 the dot product is 0 which means we have already shown earlier that the direction of the gradient vector.

The ct naught is nothing but a x dot your ct naught is nothing but your x naught so your grad f of x naught is just perpendicular to the c dash t naught but c dash t naught is lying on the tangent plane because it is tangent to the curve ct is lying the tangent plane. So if at a point where the tangent plane that they state that tangent plane touches the x naught the since it is perpendicular to one vector of c dash the grad fx naught is perpendicular to the tangent plane at x naught which we have taken the convention that if a vector is perpendicular to a tangent plane at a surface at x naught then we said that vector is perpendicular to surface at x naught.

Hence the relation between the gradient of f at x naught and the surface is that the gradient is always pointing perpendicular to the surface this is possible because of this simple chain rule that we had done and this is so simple that it the geometry chain rule is actually elucidating the geometry and this is something which you really have to keep in mind thank you very much I hope you are a good time I had a good time discussing this with you and in the next class we are going to talk about a very important thing implicit function theorem thank you.