

Ordinary Differential Equations (noc 24 ma 78)

Dr Kaushik Bal

Department of Mathematics and Statistics

Indian Institute of Technology, Kanpur

Week-01

Lecture-04

Calculus in Several Variables

students in this video we are going to talk about the differentiation the concept of derivative for a real valued function so essentially we are assuming so I am assuming that

The function $f: \Omega \rightarrow \mathbb{R}^m$ and Ω is an open subset of \mathbb{R}^n . Let us say that I am assuming that here let me write Ω it is an open subset of \mathbb{R}^n .

of \mathbb{R}^n , okay?

And this is hold for any $n \geq$.

And so, we are assuming that the function in this video, at least, whenever I am not specifying anything, is from \mathbb{R}^n to \mathbb{R}^m , let us say, okay?

And where both n and m are natural numbers, right?

So, f is a function from \mathbb{R}^n to \mathbb{R}^m . And generally speaking, first of all, we will introduce the concept of addition.

So, first of all, we will introduce the concept of the concept of

concept of derivative, right?

Derivative for a real valued function for $f: \Omega \rightarrow \mathbb{R}$

And once we do that, then we will see that we can do the exact same thing for \mathbb{R}^n .

So, let me first of all introduce this concept and then we will go to the other part.

Now, please bear in mind that all of these I am assuming that you have already known the concepts, but I am just this is kind of a revision in case that you have some gaps or in misunderstanding this will fill it up.

And also, if you want to study more on these subjects, of course, you have the books by Rudin Math Analysis, Principles of Mathematical Analysis.

I am not writing the whole book name, but you guys know Rudin.

And you can, of course, look at the Apostol book, Tom Apostol book.

This is also Math Analysis, Mathematical Analysis.

So, these are the two books which you can look at for these things.

So, let me introduce you with the concept.

So, essentially, let us say $f: (a,b) \rightarrow \mathbb{R}$ So, this is motivation.

And this is differentiable,

Is differentiable at

a point $x \in (a, b)$.

What does that mean?

This will imply that $\lim_{h \rightarrow 0} \frac{f(x+h) - f(x) - f'(x)h}{h} = 0$;

Right, so this is what it means now of course the equivalent form of this so basically you can actually see that $\lim_{h \rightarrow 0} \frac{f(x+h) - f(x) - f'(x)h}{h} = 0$ is equivalent to this expression that if I put a modulus here nothing changes okay $\lim_{h \rightarrow 0} \frac{|f(x+h) - f(x) - f'(x)h|}{|h|} = 0$.

Okay?

If I do that, nothing changes.

So, essentially, I am going to use this idea to motivate my definition of derivative for real valued function, okay?

Preliminaries (Differentiation of real valued functions)

Assuming $f: \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ (Ω is an open subset of \mathbb{R}^n , $n \geq 1$)

Introduce the concept of derivative for $f: \Omega \rightarrow \mathbb{R}$.

Rudin: Math. Anal.
Apostol: Math. Anal.

Motivation
 $f: (a,b) \rightarrow \mathbb{R}$ is differentiable at $x_0 \in (a,b) \Rightarrow$

$$\lim_{h \rightarrow 0} \frac{f(x_0+h) - f(x_0) - f'(x_0)h}{h} = 0 =$$

equivalent, $\lim_{h \rightarrow 0} \frac{|f(x_0+h) - f(x_0) - f'(x_0)h|}{|h|} = 0 =$

So, definition, definition.

So, please remember this, of course, we can, the definition which I am putting it for a function from \mathbb{R}^n to \mathbb{R}^m , okay?

So, let me do it this way.

So, let,

$$f: \Omega \rightarrow \mathbb{R}^m.$$

So, again whenever I am writing Ω , please remember that it is in subset of \mathbb{R}^m , okay, is any given function, any given function, function, okay.

Then, then f is said to be differentiable, okay, to be differentiable

at the point x , clear and where x is a any element in Ω , see here omega open is important because we are basically looking at a neighbourhood of x right and that neighbourhood exists because Ω is open okay so we are assuming that okay at x in Ω if right if

There exists, there exists, so I will write it like this, if there exists, this is the notation which I am using, okay, there exists a linear transformation.

So, we have talked about linear transformation in the earlier lectures $A \in (\mathbb{R}^n, \mathbb{R}^m)$, right,

So, this is Ω subset of \mathbb{R}^m , right.

So, the linear transformation will be from \mathbb{R}^n to \mathbb{R}^m , okay.

So, wherever Ω lies, not necessarily from Ω to \mathbb{R}^m , \mathbb{R}^m to \mathbb{R}^n , clear, okay, such that

such that the following holds okay so this definition is motivated by our earlier
 $\lim_{h \rightarrow 0} \frac{|f(x+h) - f(x) - f'(x)h|}{|h|} = 0$ definition okay $\lim_{h \rightarrow 0} \frac{|f(x+h) - f(x) - f'(x)h|}{|h|} = 0$ definition using this
 definition we are writing this so what is it $\lim_{h \rightarrow 0} \frac{|f(x+h) - f(x) - Ah|_m}{|h|_n} = 0$.

So please remember the, so this has to be , sorry, this has to be , okay.

And so basically the $A_x = Df(x)$

that A is nothing but the derivative of f at the point x .

This is what we, this is the notation which we use.

This is how we put it together, okay.

In many books, you can also see it will be written like $A = Df(x) = f'(x)$, but anyways, since it is does not really matter, we will write that it is capital Df at the point x .

Yeah.

Okay.

Now, if it is differentiable at every point x in Ω , then of course, it is differentiable in Ω .

So, that is always there.

You guys already know that I am not doing that.

Please remember this.

The norm, the norm here is in \mathbb{R}^m .

Okay.

And the norm in the denominator, this is in \mathbb{R}^n .

Okay.

And okay, now the thing is this, see why I did not use this $\lim_{h \rightarrow 0} \frac{f(x+h) - f(x) - f'(x)h}{h} =$ particular definition, but I used $\lim_{h \rightarrow 0} \frac{|f(x+h) - f(x) - f'(x)h|}{|h|} =$ definition because if I am using this definition without the normage, it will not make sense because in that case you are basically dividing out with a vector.

So, we have to use this particular definition.

I hope this is clear why we did this.

So, basically what I am saying is the difference between the $f(x+h) - f(x)$ can be actually approximated with a linear approximation.

This is what it means.

Now, the thing is this, with this definition, of course, we know that if the derivative exists, this is going to be, so small remarks, small remarks.

So, please check this, check, check that.

The derivative is derivative is unique, is unique.

So, basically if it exists, this is unique.

If it exists, of course, if it exists.

And please remember that this A depends on the point x .

So, let me write it this way.

This depends on the point x .

So, basically if you change x , A is also going to be changed.

So, this is not a linear transformation for all points in Ω .

Every point, for every point you have one-one unique linear transformation.

So, let me write it this way.

Please note

Please note that for every point, every point, A depends, A changes.

Let me put it this way.

So, you do realize what I am trying to say.

So, that is there.

Now, the thing is this.

See, and we call it this.

This we say as this is the total derivative.

This is called a total derivative of f at the point x .

This is called a total real value.

Yes.

Okay.

So, now the thing is this.

See.

Since derivative of the function, you know, function from \mathbb{R}^n to \mathbb{R}^m , okay?

So, that is basically a linear transformation to \mathbb{R}^n to \mathbb{R}^m .

So, basically, let me put as an example what is the derivative of a linear transformation, okay?

So, basically, if you are starting out with a linear transformation, what should be its derivative?

So, as an example, let us put as an example, okay?

So, let f is a function, f is a function from, for now let us just assume that it is from \mathbb{R}^n to \mathbb{R}^m , $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$, and is linear, okay, is linear, is linear and now the question is this, I want to put down that and let us say and x is in \mathbb{R}^n , okay, then I want to see that then

I want to see what is $A'(x)$. And essentially, you can actually see that this is going to be A . So, how is it possible?

You see, what happens is $A(x+h)-A(x)=Ah$, because of linearity. You understand?

So, you do realize that $A'(x)$, this is definitely going to be A .

okay.

So, and essentially here the remainder term is going to be .

So, essentially this is what we are we have.

So, basically and you also expect it right because you see since it is linear what is the best approximation of the function?

The function itself right.

So, that is why this is the case.

Okay, now b, another small thing is this, that let us say $f: \mathbb{R}^n \rightarrow \mathbb{R}$, right. Now the question which we need to know, and this is important, okay, $\mathbb{R}^n \rightarrow \mathbb{R}$, given by $f(x_1, \dots, x_n) = k$, let us say, any constant k .

k and k is a constant, k is a constant, okay, any constant k . Now, the question is this, what is the derivative?

Please check, please check that the derivative, please check that the derivative at any point, okay.

So, $Df(x) \equiv$, linear map.

So, basically, this is not zero in \mathbb{R}^n , please remember, this is a and belongs to the linear

from map from \mathbb{R}^n to \mathbb{R}^m .

So, basically the map which takes every element of \mathbb{R}^n to that is the map yes please you have to check this part.

So, please do that ok.

Definition :- Let $f: \Omega \rightarrow \mathbb{R}^m$ is any given function. Then f is said to be differentiable at $x_0 \in \Omega$, if $\exists A_{x_0} \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ such that

$$\lim_{h \rightarrow 0} \frac{\|f(x_0+h) - f(x_0) - A_{x_0}h\|_m}{\|h\|_n} = 0$$

We say $A_{x_0} = Df(x_0) = f'(x_0)$ (Total Derivative)

Remark :- Check that the derivative is unique (if it exists)

(b) Please note that for every point, A changes.

Example :- (a) $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is linear and $x \in \mathbb{R}^n$ then $A'(x) = A$
($A(x+h) - A(x) = Ah$)

(b) $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $f(x_1, \dots, x_n) = k$ (k is constant). Please check that $Df(x) = 0$.
[$0 \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$]

Now, the question is this what are some of the properties that we know one of the most important properties which we sorry which we use

we are talking about derivative, is the so-called chain rule, right?

What is chain rule?

So, essentially, let me just write down the chain rule.

Of course, you know how to prove it.

I mean, you can do the exact same thing what you did in \mathbb{R} .

So, same sort of proof works, okay?

So, let us say, let Ω subset of \mathbb{R}^n

Okay.

This is of course I am assuming all the time.

And $f: \Omega \rightarrow \mathbb{R}^m$.

Okay.

is differentiable at the point x .

Is differentiable let us say.

Differentiable at x in \mathbb{R}^m .

At x .

x in Ω .

Okay.

So let us say it is differentiable at the point x .

Now you see we are looking for another map g . So let us say g maps.

An open set.

An open set.

open set containing, containing $f(\Omega)$, okay, into \mathbb{R}^k , okay.

So, basically, what is happening is f is from Ω to \mathbb{R}^m and g is from $f(\Omega)$ to \mathbb{R}^k , okay.

And g , we are also assuming that g is differentiable, j is differentiable, differentiable

that $f(x)$.

So, we are assuming this.

Now, if this is the case, then if you define $F(x) = g(f(x))$, then we can say that F is differentiable

is differentiable at the point x differentiable and what happens is this is very very important please remember this $F'(x) = g'(f(x)) * f'(x)$

times $F'(x)$

So, in this case what is happening is this since F' is nothing but a linear transformation which again actually is nothing but a matrix.

So, basically it is a m cross n matrix ok.

So, another small remark before I move on to the other important parts let us do this small remark which I forgot actually to put it here.

You see if f is different.

So, $f: \Omega \rightarrow \mathbb{R}^m$

is differentiable at the point x , right, is differentiable, differentiable at x .

Now, you see what is happening is this, if you

So, then what happens is this, you have this Df , right?

So, $Df(x)$, this is a linear map from \mathbb{R}^n to \mathbb{R}^m , this is given, right?

Then we know that any linear map from \mathbb{R}^n to \mathbb{R}^m is basically nothing but a matrix, right?

Then $Df(x)$ is, you can actually write it as a m cross n matrix, right?

m cross n matrix clear ok now the question is this how does this m cross n matrix looks like that is that idea and we need to put this idea together that how does it look like then we are done more or less that is the that is the

Okay, question which we need to answer.

Okay, so to do that what we are going to do is we are going to start with something called a concept of partial derivatives.

You understand, this is a m cross n matrix, that is what we know.

But we can actually, you know, for generally calculating the derivative, we actually need to know what exactly is the matrix, how do we compute something like this.

And to do that, we need the concept of partial derivatives, partial derivatives.

derivatives.

And please, again, please remember this, I am skipping a lot of ideas here, since it is assumed that you already know this, okay.

So, what about partial derivative see,

Essentially, what we are doing is, we are basically looking at derivative in a particular direction, right, not in any arbitrary direction, but the unit direction, okay.

So, what we have is this.

Let us say that f , again, given $f: \Omega$ is a subset of $\mathbb{R}^n, \mathbb{R}^m$.

\mathbb{R}^n , okay, is a given function.

Nothing is given here.

We are not assuming in this case that it is differentiable.

So, for this function, let us say this is any given function, any given function, okay, and I am going to define the concept of partial derivative for any given function.

Okay, so any given function, right?

And now what we are doing is I am going to, as I told you, I am going to look at the concept of derivative in a particular direction, right?

So what I am going to do is let, I am going to choose a particular direction means unit directions, okay, the basis directions.

Okay, so let us look at the particular direction.

So let $\{e_1, \dots, e_n\}$ and $\{u_1, \dots, u_m\}$.

be the standard basis, be the standard basis, standard basis of \mathbb{R}^n and \mathbb{R}^m respectively, and okay.

Now, what happens is this, then, then, you see,

Then $f(x)$, if for any point x in Ω , f of x will look like this, no?

It will look like $\sum_{i=1}^m f_i(x)u_i$, because you see $f(x)$ is a element of \mathbb{R}^n .

So, there are, we have m components, right?

So, it is, it will look like f_i of x u_i , right?

u_i . And this x is in Ω

okay so for any x in Ω $f(x)$ will look like this so it can be written as $\{u_1, \dots, u_m\}$, the linear combination of those and those coefficients I am writing of course this coefficient depends on x I am writing it as $f_i(x)$ okay so so therefore what do I have what I wrote is $f_i(x) = f(x) \cdot u_i$

That is how you calculate what $f_i(x)$ is.

And this holds for $\leq i \leq m$. Now the thing is this.

Chain Rule: Let $\Omega \subseteq \mathbb{R}^n$ and $f: \Omega \rightarrow \mathbb{R}^m$ is differentiable at $x_0 \in \Omega$, g maps an open set containing $f(x)$ into \mathbb{R}^k , g is differentiable at $f(x_0)$. Then $F(x) = g(f(x))$. Then F is differentiable at x_0 and $F'(x_0) = g'(f(x_0)) f'(x_0)$.

Remark: $f: \Omega \rightarrow \mathbb{R}^m$ is differentiable at x_0 . $DF(x_0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ then $DF(x_0)$ is a $(m \times n)$ matrix.

Partial Derivative: $f: \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ is any given function. Let $\{e_1, \dots, e_n\}$ and $\{u_1, \dots, u_m\}$ be the standard basis of \mathbb{R}^n and \mathbb{R}^m respectively.

$$\text{then } f(x) = \sum_{i=1}^m f_i(x) u_i \quad (x \in \Omega)$$

$$\Rightarrow f_i(x) = f(x) \cdot u_i \quad ; \quad 1 \leq i \leq m.$$

With this idea, I am going to define the concept of partial derivative.

So what is the partial derivative?

So for x in Ω and i between 1 and m .

j between 1 and n , okay.

We define the partial derivative, we define, we define.

So, I will define the partial derivative like this, $D_j f_i(x)$, clear.

So, basically what it means is I am looking at the j th direction, okay, e_j th direction, okay, of the component function f_i , okay.

So, that is defined by limit

t - .

So, basically $\{f_1, \dots, f_n\}$, these are the functions, right?

You see, if you write it like this, what are, see f_i , f_i is a function.

So, if you think of it, so see, note, note, f_i is a function from what to where?

x is in \mathbb{R}^n , right?

So, $f_i: \Omega \rightarrow \mathbb{R}$, right?

So, basically f_i is a real valued function, clear.

So, I am going to define for a real valued function, I am going to define the derivative in the j th direction, okay.

So, you see I am looking at the how f is changing in the j th direction and then $D_j f_i(x) = \lim_{t \rightarrow 0} \frac{f_i(x+te_j) - f_i(x)}{t}$.

So, how f is provided?

The limit exists, of course, provided.

the limit exist.

So, you do understand what I am trying to say?

I am not looking what I am doing is this first of all for a given f you see given f_i can I break that f of into components and what happens is each each components are nothing, but real valued function and for real valued function I am defining.

So, this is my definition I am defining the derivative of a real valued function in the j th

So, this is think of it like derivative is just the first order derivative in a particular direction in a one dimensional case.

So, derivative think of it like derivative of f_i in one dimension in particular dimension in particular the j th direction.

So, this is the idea.

okay and so that's your partial derivative and we write it like this clear so once that is there you see we have that this theorem okay so theorem i hope the idea is clear for a real valued function we define the partial derivative like this so basically it is just the derivative but in that particular direction whatever direction we are choosing okay so let's say suppose

$f: \Omega \rightarrow \mathbb{R}^m$, is differentiable.

So, what is the theorem?

This theorem will actually guide us to answer this question that $Df(x)$ at some point x wherever is differentiable is a m cross n matrix.

Now, the question is what exactly is that matrix?

So, to do that we define the concept of partial derivatives and here I am writing this theorem.

So, $f: \Omega \rightarrow \mathbb{R}^m$, let us say.

and at some point, at some point, x in Ω , right, that is given.

Then, what happens is, then the partial derivatives, the partial derivatives, so this is the relation between derivative and partial derivative, the partial derivatives, sorry, not derivatives, derivatives, yeah, sorry, derivative $D_j f_i(x)$, okay.

So, you see, for each f_i , we have j th \mathbb{R}

j th direction j directions right so m directions basically because j varies between 1 and n so basically m directions are there i for each f_i and i is varying between 1 and m j is varying between 1 and n right so i have n j since j varies between 1 and n i have n derivatives here partial derivatives okay so uh they this partial derivative exists

Exists, okay?

And what happens is, and this is very, very important, okay?

What happens is $Df(x) e_j$, clear?

Sorry, see $DF(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, right?

So, this is a linear map from \mathbb{R}^n to \mathbb{R}^m .

So, basically and I we already discussed talking about linear maps that you can actually determine the linear map based on how it acts on elements of the basis element right.

So, you see since e_i is an element of \mathbb{R}^n the basis elements ok.

So, we just want to see what happens to $Df(x)$ at the point e_j and that to answer that question this becomes nothing but $\sum_{i=1}^m (D_j f_i)(x) u_i$ ($\leq j \leq n$).

clear?

So, what I meant by this is you see if you write down this DF.

So, basically this is acting on e_j and here the end result will acting will be on u_i .

So, essentially if you want to write down what is $DF(x) = \begin{pmatrix} D_1 f(x) & \dots & D_n f(x) \\ D_1 f_m(x) & \dots & D_n f_m(x) \end{pmatrix}_{m \times n}$

okay, the partial derivative with respect to x_n .

So, basically nth direction of f with respect to x , okay.

So, basically on the first, all the derivatives of the first component and then the derivative of the second component and in the nth row, you will have the derivative of the mth component, sorry, mth row, okay.

Yeah.

And you do realize that this matrix is a m cross n matrix.

It has m row and n columns.

Okay.

So, we have this.

So, essentially what is happening is this, for as a small remark, as a small remark, if f is just a real valued function, if $f: \Omega$ is a subset of $\mathbb{R}^n \rightarrow \mathbb{R}$, is differentiable, is differentiable.

Okay.

Differentiable, let us say at the point x .

Then, $DF(x)$, okay, is nothing but a, you do realize m will be then, m will be then, so this is nothing but a m cross n matrix, is a m cross n matrix, matrix given by, let me write it here only, okay,
 $DF(x) = (D_1 f(x) \dots D_n f(x))_{*n}$

And these we will call it, so essentially we will actually, so this is a m cross n matrix, sorry, this is,

This is a m cross n matrix.

Let me put it this way, m cross n matrix, okay.

For $x \in \Omega$, $1 \leq i \leq m$, $1 \leq j \leq n$ we define

$$D_j f_i(x) := \lim_{t \rightarrow 0} \frac{f_i(x + te_j) - f_i(x)}{t} \text{ provided the limit exists}$$

[Derivative of f_i in 1-dimension (j^{th} direction)]

Theorem: $f: \Omega \rightarrow \mathbb{R}^m$ is differentiable at $x_0 \in \Omega$. Then the partial derivatives $D_j f_i(x)$ exist and $Df(x) e_j = \sum_{i=1}^m (D_j f_i(x)) u_i$ ($1 \leq j \leq n$)

$$Df(x) = \begin{pmatrix} D_1 f_1(x) & \dots & D_n f_1(x) \\ \dots & \dots & \dots \\ D_1 f_m(x) & \dots & D_n f_m(x) \end{pmatrix}_{m \times n}$$

Remark: If $f: \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ is differentiable at x_0 , then $Df(x_0)$ is a $(1 \times n)$ matrix given by $Df(x_0) = (D_1 f(x_0) \dots D_n f(x_0))_{1 \times n}$

And what happens is we will, we can define, so one defines, defines the gradient, okay, gradient of f , of f . In this case, f is from, sorry, gradient is defined as this.

$\nabla f(x) \cong Df(x)$ and the gradient will be a function $\nabla f(x) = (D_1 f(x) \dots D_n f(x))$ and so basically this will be a vector,

vector in \mathbb{R}^n .

And why this identification is there?

Let me just put this part together.

So, you see this is there because of this theorem that let us say let $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ to \mathbb{R}^m is a linear transformation is a linear transformation transformation

Then, what happens is, it is actually nothing, you can actually represent them, then there exists a vector, okay, let me put it this way, $\mathbb{R}^n \rightarrow \mathbb{R}$, is a linear transformation, then there exists a unique vector, unique, unique vector, vector in \mathbb{R}^n , okay.

such that one can identify that $T(x)$, unit vector let us say x in \mathbb{R}^n , let me put it this way, unit vector x in \mathbb{R}^n such that $T(x) = x$.

This holds for all x in \mathbb{R}^n .

What is the proof?

Let us quickly look at the proof.

proof why this works.

See for any x in \mathbb{R}^n what happens is x can be written as, $x = \sum_{i=1}^n x_i e_i$ and then that will imply that $t_x = \sum_{i=1}^n x_i t(e_i)$, and then what happens is you can actually write it as $(x_1, \dots, x_n) \cdot (T(e_1), \dots, T(e_n)) = x \cdot x$.

where we write x to be this vector, $T(e_1), \dots, T(e_n)$.

I hope this is clear.

So, this is the idea.

So, essentially what I am trying to say is this.

You see, for a real valued function, whenever we talk about the

derivative map, okay, that is nothing but a linear transformation from \mathbb{R}^n to \mathbb{R} , which is nothing but a matrix, cross n matrix, okay, which will look like this, right, okay.

Now, the thing is, you see,

we define the gradient of f to be the vector.

The vector is given like this, exactly the same sort of thing, but this is a vector in \mathbb{R}^n .

And how do you actually, why can we classify this thing?

How can you identify this?

Because we know that if you have a linear transformation from \mathbb{R}^n to \mathbb{R}^m , then corresponding to this t , corresponding to this t , you have a unique x in \mathbb{R}^n .

So, basically, which actually gives you

identification.

For every element here in $\mathcal{L}(\mathbb{R}^n \rightarrow \mathbb{R}^m) \cong \mathbb{R}^{nm}$.

So, this identification is there and this is what we are identifying it.

So, basically, let us say as an example, let us just put the term here as an example.

If you have a function, let us say $f(x, x) = -(x^2 + x^2)$

Then what happens is, then we of course know $DF(x, x) = (X, X)$ is nothing but, you see, we write it as the derivative.

The first derivative is X here and the second is X .

So, this is how we write it.

See, this, please remember, (x, x) is nothing but an element of (\mathbb{R}, \mathbb{R})

and (X, X) is the element of \mathbb{R} ok, but we identify this, this equals to is basically means this is identification, identification clear ok.

So, with this, let me go back to this original theorem here.

So, what is our theorem?

It says that $f: \Omega \rightarrow \mathbb{R}^m$.

Let us say it is differentiable at the point x , okay?

Now, we want to classify what exactly is the derivative map.

How does the derivative map look like?

So, basically, you see, this is a linear transformation.

How do you classify, characterize a linear transformation?

How do you know what exactly is that?

You just look at how it acts on the basis element.

So, what does it do?

It actually gives you this matrix, okay?

Okay.

One defines the gradient of f , $\nabla f(x_0) \simeq Df(x_0)$

And $\nabla f(x_0) = (D_1 f(x_0), \dots, D_n f(x_0)) \leftarrow$ vector in \mathbb{R}^n .

$\simeq (\mathbb{R}^n, \mathbb{R}) \simeq \mathbb{R}^n$

[Let $T: \mathbb{R}^n \rightarrow \mathbb{R}$ is a linear transformation; then \exists a unique vector in $x_0 \in \mathbb{R}^n$ s.t.

$T(x) = x \cdot x_0 \quad \forall x \in \mathbb{R}^n$.

Proof: $x \in \mathbb{R}^n, x = \sum_{i=1}^n x_i e_i \Rightarrow Tx = \sum_{i=1}^n x_i T(e_i) = (x_1, \dots, x_n) \cdot (T(e_1), \dots, T(e_n))$

$= x \cdot x_0$ where $x_0 := (T(e_1), \dots, T(e_n))$]

Example: $f(x_1, x_2) = (x_1^2 + x_2^2)$

then, $DF(x_1, x_2) = (x_1, x_2)$

\uparrow \mathbb{R}^2 \uparrow \mathbb{R}^2

$\simeq (\mathbb{R}^2, \mathbb{R})$ Identification

So I want to let us look at the proof of this thing.

This is very very important.

Okay.

Right.

So let us look at the proof of this.

The proof is considerably simple actually.

Okay.

So you first of all you fix a J . You fix a j .

And since f is differentiable at the point x , let us say, yeah, and since f is differentiable at the point, differentiable at x , one has, I can write it like this, right, $f(x + te_j)$,

Okay.

See, f is differentiable here.

See, what is our definition of differentiability?

Sorry.

$$\lim_{h \rightarrow 0} \frac{|f(x+h) - f(x) - f'(x)h|}{|h|} = \text{definition, right?}$$

This is the definition of differentiability.

So, this definition and the definition which I am writing, these are exactly the same definition.

It is just broken down parts.

So, please check that those two are equivalent essentially.

But what I am meaning to say is this.

You see, this h is any direction h in \mathbb{R}^n , right?

So, I will choose a particular direction which is e_j .

So, we can write that since f is differentiable at the point x , we can write it as $f(x + te_j) - f(x) = Df(x)(te_j) + r(te_j)$.

So, this is h in that direction.

So, basically I am using a particular direction.

Which depends on $p = j$ clear if you if you break down that definition you can just write it down so if you are not convinced you please check that those these two definitions are actually equivalent clear so and where

So, what is the remainder term?

The remainder term is $r(te_j)$.

So, you remember the norm of this by t that should go to 0, right?

So, $\frac{|r(te_j)|}{|t|}$, as $t \rightarrow 0$.

So, this is exactly the same definition, I am just writing it in this way.

Now, you see since if $DF(x)$ is linear,

So, essentially I can write it as $\lim_{t \rightarrow 0} \frac{f(x + te_j) - f(x)}{t} = DF(x) e_j$

here now you see you just represent f in terms of components okay and what happens is then ah this you can write it so since f is in terms of components so basically f is $\{f_1, \dots, f_m\}$ right okay so you can write it as $\lim_{t \rightarrow 0} \sum_{i=1}^m \frac{f_i(x + te_j) - f_i(x)}{t} \cdot u_j = DF(x) e_j$ (*Breaking into components*).

One gets this.

And which is nothing but.

Our assumption.

So you do realize.

What I am trying to say is this.

Although it is a derivative.

It is nothing but a matrix.

Okay.

So that is there.

Proof - Fix j and since f is differentiable at x_0 ,
 $f(x_0 + te_j) - f(x_0) = DF(x_0)(te_j) + r(te_j)$ (check)

where, $\frac{|r(te_j)|}{|t|} \rightarrow 0$ as $t \rightarrow 0$.

$\therefore DF(x_0)$ is linear,

$$\lim_{t \rightarrow 0} \frac{f(x_0 + te_j) - f(x_0)}{t} = DF(x_0)e_j.$$

\Downarrow Breaking into component.

$$\lim_{t \rightarrow 0} \sum_{i=1}^m \frac{f_i(x_0 + te_j) - f_i(x_0)}{t} u_j = DF(x_0)e_j$$

And just before I finish.

You see.

I need to know the concept of.

We need to know the concept of.

Continuous differentiability.

Continuous differentiability.

Differentiability.

So, what is the continuous differentiability?

It is nothing but when the function is, differentiable function, we are looking at a differentiable function.

So, basically, a differentiable map, okay, differentiable map $f: \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, okay, this is said to be continuously differentiable, is said to be

continuously differentiable continuously differentiable in Ω let us, if the map DF right DF is a continuous map is a continuous map continuous map ok.

from Ω to $\mathcal{L}(\mathbb{R}^n \rightarrow \mathbb{R})$. So, what I mean by this is, see DF at the point x , let us say x is any point in Ω , what I mean by this is nothing but a linear map from $\mathbb{R}^n \rightarrow \mathbb{R}$. This is for any x in Ω .

Okay, now I am doing is what I am doing is this, this x varies in Ω , right?

So you see if you are changing, so this DF is nothing but a function from Ω to $\mathcal{L}(\mathbb{R}^n \rightarrow \mathbb{R})$ You can think of it like this, right?

Now if this function is continuous.

If this function is continuous, okay, then we say it is continuously differentiable.

I hope this is clear, okay.

What are some examples?

So, most polynomial functions, you know, into variables that is, sorry, n variable or, you know, most exponential function, if you define it like that, using the composition function, you can, of course, say that they are going to be continuously differentiable.

It is not a very difficult thing to prove, okay.

So, with this, I am going to end this video.

Continuous Differentiability \Leftrightarrow A differentiable map $f: \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ is said to be continuously differentiable in Ω if DF is a continuous map from Ω to $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$

$$\left[\begin{array}{l} DF(x_0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \\ \Downarrow \\ DF: \Omega \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \\ \text{continuous} \end{array} \right]$$