Analysis II, Term I,

lectures by Davoud Cheraghi

Email: d.cheraghi@imperial.ac.uk

Q: What is the essence of analysis, as opposed to algebra, number theory, geometry?

"Analysis is the theory of infinite constructions."

"Analysis allows us to create new objects by infinite procedures."

- to build $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$, we do algebra
- to build $\mathbb{R}$, need to do analysis.

Analysis 1 is mostly on $\mathbb{R}^1$,

Analysis 2: Part I: mostly on $\mathbb{R}^n$,

Part II: more general spaces, metric spaces, topological spaces.

1.1. Euclidean spaces

1.1.1. preliminaries from analysis 1, about $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, |\cdot|$, etc.

(see typed notes)

1.1.2 Euclidean spaces of dim n.

For $n \geq 1$,

$$\mathbb{R}^n = \{ (x^1, x^2, \dots, x^n) \mid \forall i = 1, 2, \dots, n, \ x^i \in \mathbb{R} \}$$

vector space over $\mathbb{R}$, $\dim = n$.

coordinates of x.

the innerproduct, $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, is

$$\langle (x^1, x^2, \dots, x^n), (y^1, y^2, \dots, y^n) \rangle = \sum_{i=1}^n x^i y^i$$

The norm function $\| \cdot \| : \mathbb{R}^n \to \mathbb{R}$, is defined as

$$\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{\sum_{i=1}^n (x^i)^2}$$

The norm function satisfies the following properties.

(i) for all $x \in \mathbb{R}^n$, we have $\|x\| \geq 0$, and $\|x\| = 0$ iff $x = (0, \dots, 0)$.

(ii) for all $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^1$, $\|\lambda x\| = |\lambda| \cdot \|x\|$

(iii) for all $x, y, z \in \mathbb{R}^n$,

$$\|x + y\| \leq \|x\| + \|y\|.$$

↳ triangle inequality.

There is an important relation between $\|x\|$, and $|\cdot|$ of the entries of x.

$$\min_{k=1,2,\dots,n} |x^k| \leq |x^k| \leq \|x\| \leq \sqrt{n} \max_{k=1,2,\dots,n} |x^k|$$

(see problem sheets).

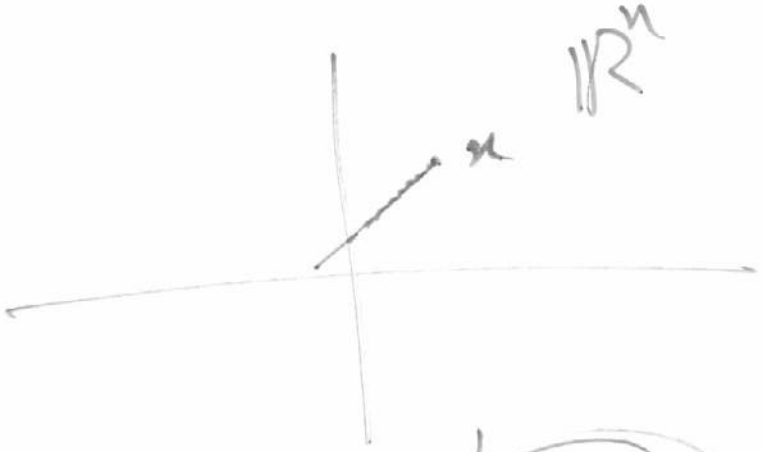
1.1.3 Convergence of sequences in $\mathbb{R}^n$.

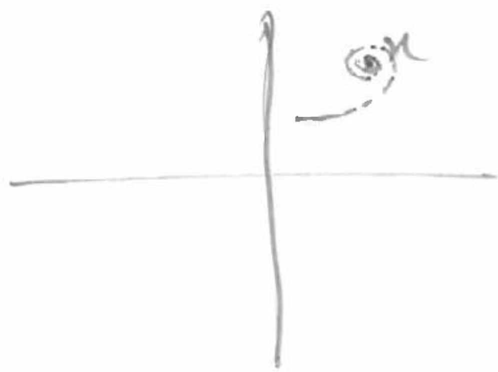
A sequence in $\mathbb{R}^n$, is any ordered list

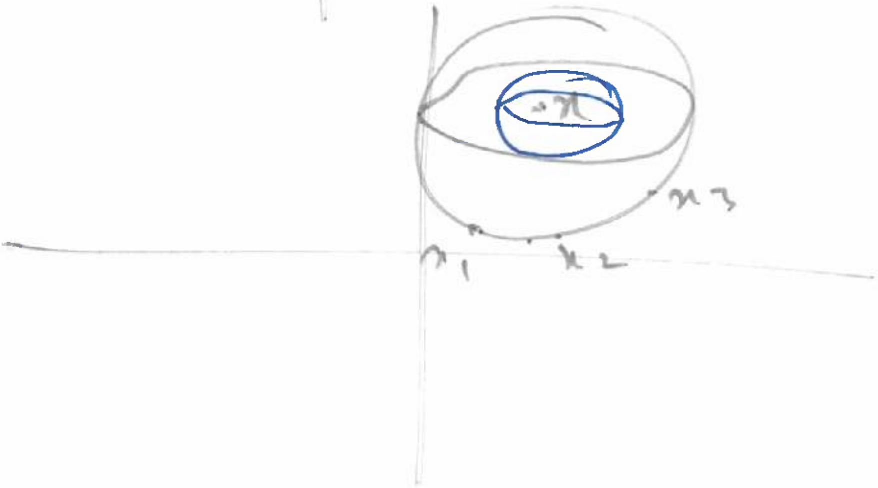
$$x_1, x_2, x_3, \dots,$$

s.t. $\forall i \geq 1$, $x_i \in \mathbb{R}^n$.

Def 1.1 A sequence $(x_i)_{i\geq 1}$ in $\mathbb{R}^n$ converges to $x \in \mathbb{R}^n$, if for any $\varepsilon > 0$, $\exists\, N \in \mathbb{N}$, s.t. $\forall\, k \geq N$, $\| x - x_k \| < \varepsilon$.







Prop 1.1. A sequence $(x_i)_{i=1}^{\infty}$ in $\mathbb{R}^n$ converges to some $x \in \mathbb{R}^n$, iff each component of $(x_i)_{i=1}^{\infty}$ converges to the corresponding component of x, that is,

if $$x_i = (x_i^1, x_i^2, \ldots, x_i^n)$$

$$x = (x^1, x^2, \ldots, x^n)$$

then
$$x_i \longrightarrow x \iff x_i^k \longrightarrow x^k, \text{ for all } k = 1, 2, \ldots, n.$$

proof: see typed notes.

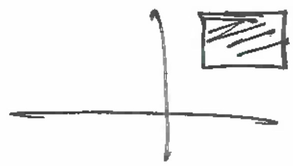
1.1.4 Open sets in Euclidean spaces

In dim 1, we consider maps on $[a, b]$, or (a, b).

We may consider sets of the form

$$[a^1, b^1] \times [a^2, b^2] \times \cdots \times [a^n, b^n]$$

$$= \{ (x^1, x^2, \ldots, x^n) \mid \forall i = 1, \ldots, n, \quad a^i \leq x^i \leq b^i \}$$

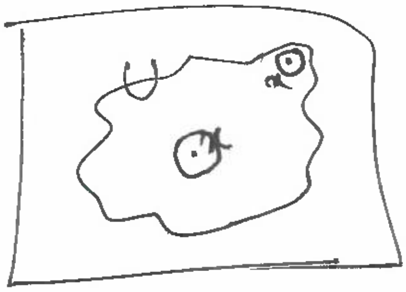


For $x \in \mathbb{R}^n$, $r > 0$, the open ball of radius r about x is

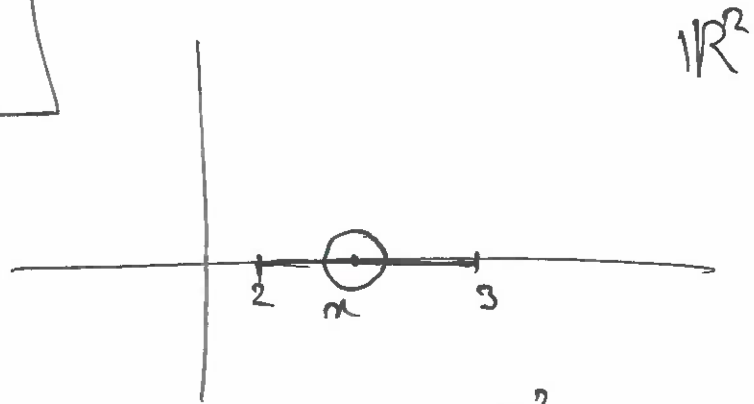
$$B_r(x) = B(x, r) = \{ y \in \mathbb{R}^n \mid \|x - y\| < r \}$$

Def 1.2 A set $U \subseteq \mathbb{R}^n$ is called open in $\mathbb{R}^n$, if for every $x \in U$, there is $r > 0$ s.t.

$$B_r(x) \subseteq U.$$



r depends on x.

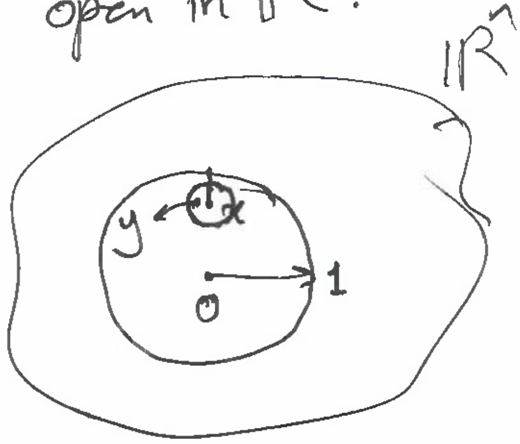


the interval $(2,3)$ is not open in $\mathbb{R}^2$

$\{(x,y) \mid 2<x<3,\ y=0\} \subseteq \mathbb{R}^2$

$(2,3)$ is open in $\mathbb{R}^1$.

Example 1.1. The ball $B_1(0)$ is open in $\mathbb{R}^n$.



proof: Fix $x \in B_1(0)$

let $r = \frac{1-\|x\|}{2} > 0$

need to prove $B_r(x) \subseteq B_1(0)$.

Fix $y \in B_r(x)$. $\|y-0\| = \|y-x+x\|$

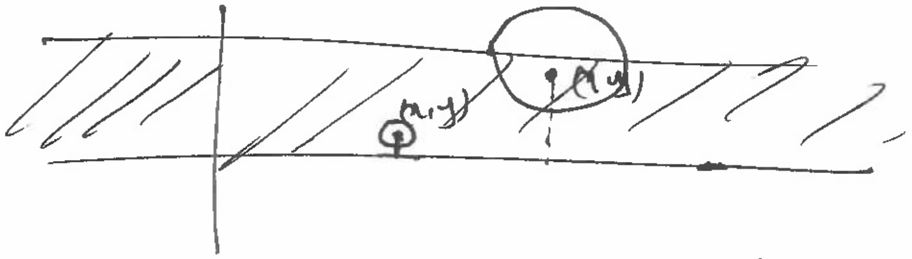
$\leq \|y-x\| + \|x\|$

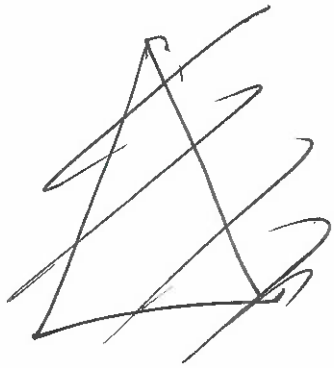
$< r + \|x\|$

$= \frac{1-\|x\|}{2} + \|x\| = \frac{1+\|x\|}{2} < 1$

Thus $y \in B_1(0)$.

$U = \{ (x,y) \in \mathbb{R}^2 \mid x \in \mathbb{R}^1, 0 < y < 1 \}$





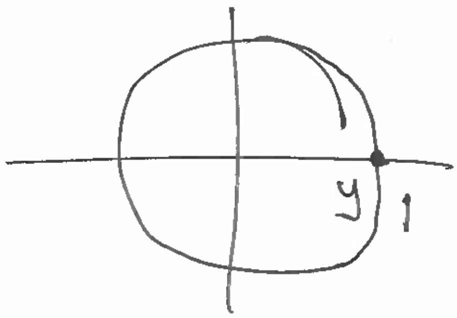
let $(x,y) \in U$. let $r \leq \frac{y}{2}$

$r < \frac{1-y}{2}$

let $r = \min\{ y/2, \frac{1-y}{2} \}$.

let $A = \{ y \in \mathbb{R}^n \mid \|y\| \leq 1 \}$

let $y = (1, 0, 0 \ldots, 0)$, $y \in A$.



there is no $r > 0$ s.t. $B_r(y) \subseteq A$.

W2, L1

1.2.9 Continuity at a point, and on an open set in $\mathbb{R}^n$.

Def 1.3 let $A \subseteq \mathbb{R}^n$ be an open set, and suppose $f: A \to \mathbb{R}^m$. The map f is continuous at some $p \in A$, if

for every $\varepsilon > 0$, there is $\delta > 0$, s.t.
for every $x \in A$ satisfying $\|x - p\| < \delta$, we have $\|f(x) - f(p)\| < \varepsilon$.

If f is continuous at every $p \in A$, we say f is continuous on A.

$\forall \varepsilon > 0$, $\exists \delta > 0$, s.t.

$$f(A \cap B_\delta(p)) \subseteq B_\varepsilon(f(p)).$$

$\forall \varepsilon > 0$, $\exists \delta > 0$ s.t.

$$f(B_\delta(p)) \subseteq B_\varepsilon(f(p)).$$

Example 1.4. let $\Lambda : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map. Then Λ is continuous on $\mathbb{R}^n$.

Λ is linear if $\forall x, y \in \mathbb{R}^n$, ~~[illegible]~~ $r \in \mathbb{R}$,

$$\Lambda(x + ry) = \Lambda(x) + r \cdot \Lambda(y).$$

$(1,2) = 1 \cdot e^1 + 2 e^2.$

$e^1 = (1, 0, \dots, 0)$

$e^2 = (0, 1, 0 \dots 0).$

$e^k = (0 \dots 0, 1, 0 \dots 0)$

Proof: fix $p \in \mathbb{R}^n$. fix $\varepsilon > 0$.

$$\begin{aligned}
\|\Lambda(x) - \Lambda(p)\| &= \\
&= \|\Lambda(x-p)\| = \|\Lambda(\sum_{k=1}^n (x-p)^k e^k)\| \\
&= \|\sum_{k=1}^n \Lambda((x-p)^k e^k)\| \\
&= \|\sum_{k=1}^n (x-p)^k \cdot \Lambda(e^k)\| \\
&\le \sum_{k=1}^n |(x-p)^k| \cdot \|\Lambda(e^k)\| \\
&\le \sum_{k=1}^n \|x-p\| \, \|\Lambda(e^k)\| \\
&\le M n \cdot \|x-p\|.
\end{aligned}$$

$\|z+w\| \le \|z\| + \|w\|$

let $M = \max\{\|\Lambda(e^k)\|, k = 1, 2, \dots, n\}$

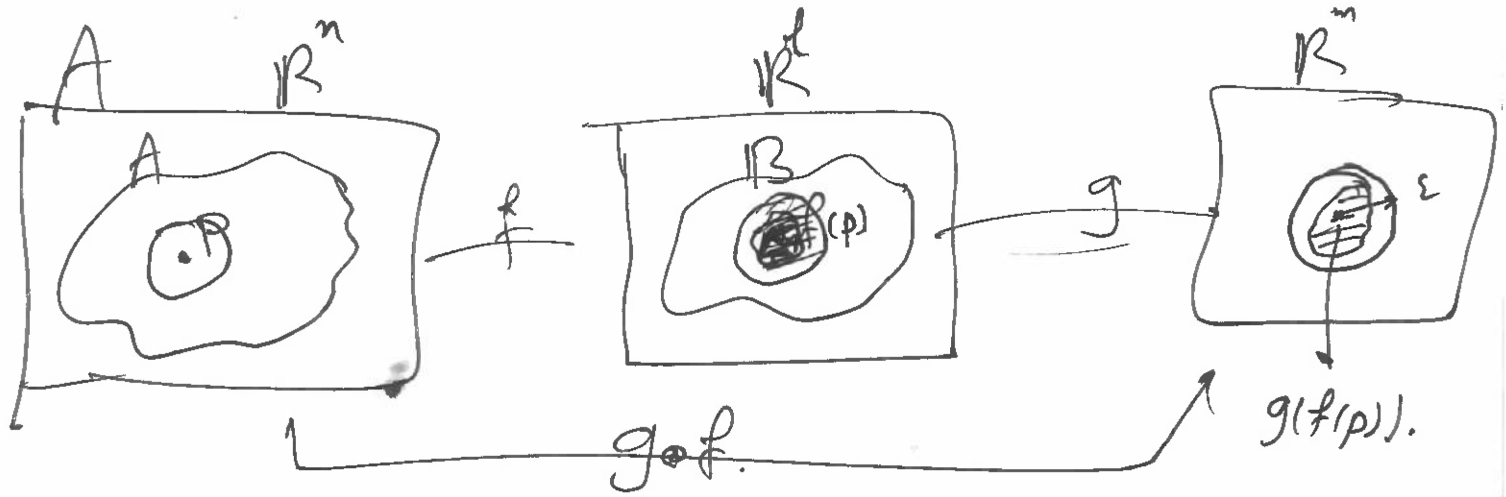
let $\delta = \dfrac{\varepsilon}{nM}$.

if $\|x-p\| < \delta$, then

$$\|\Lambda(x)-\Lambda(p)\| \le M \cdot n \cdot \|x-p\|$$

$$< \cancel{M}\cancel{n} \cdot \frac{\varepsilon}{\cancel{n \cdot M}} = \varepsilon.$$

Thm 1.2. Let $A \subseteq \mathbb{R}^n$ and $B \subseteq \mathbb{R}^{d}$ be open subsets. Suppose $f: A \to B$ is continuous at $p \in A$. and $g: B \to \mathbb{R}^m$ is continuous at $f(p)$. Then $g \circ f: A \to \mathbb{R}^m$ is continuous at p.



See ~~tg~~ pdf notes for more examples.

more properties s.t. $f+g$, $f \cdot g$.

$\lim_{x \to p} f(x)$.

1.3 Derivative of a map of Euclidean spaces.

For $f: (a,b) \to \mathbb{R}^1$, if the limit

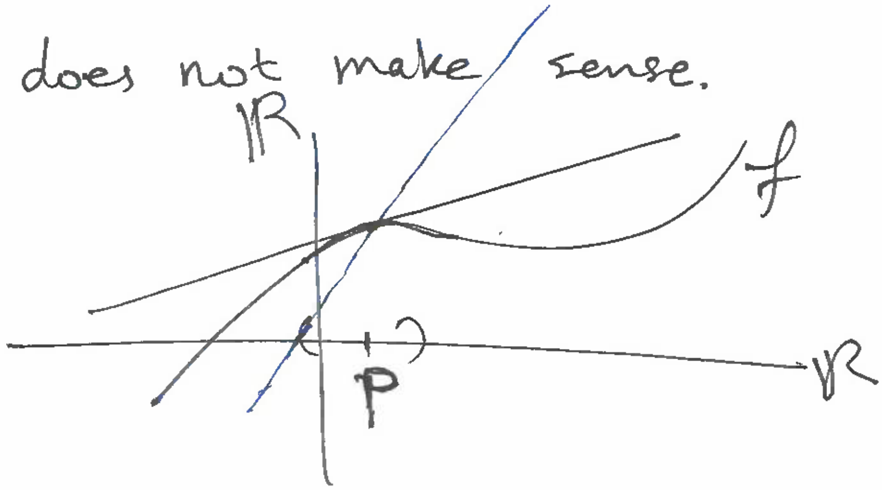
$$\boxed{\lim_{p \to x} \frac{f(x) - f(p)}{x - p}}$$

exists, f is differentiable at x, and the derivative of f at x is the value of the limit.

If $\Omega \subseteq \mathbb{R}^n$ is open, $f: \Omega \to \mathbb{R}^m$ is a "nicely behaving map"

for $x, p \in \Omega \subseteq \mathbb{R}^n$, $f(x), f(p) \in \mathbb{R}^m$.

the ratio does not make sense.



$f'(p)$ is the slope of the tangent line at p.

to the "graph of f near p."

the tangent line is of the form

$$x \longmapsto ax + b$$

Infact, the tangent line is the graph of the function $A_\lambda(x) = \lambda(x-p) + f(p)$ with $\lambda = f'(p)$.

A_λ is a "good approximation" of f near p.

$$\lim_{x\to p} \left(f(x) - A_\lambda(x)\right) = \lim_{x\to p} f(x) - \lim_{x\to p}\left(\lambda(x-p) + f(p)\right)$$

$$= f(p) - f(p)$$

$$= 0.$$

This holds for any value of $\lambda \in \mathbb{R}$.

When $\lambda = f'(p)$, the approximation is better

$$\lim_{x\to p} \left(\frac{f(x) - A_\lambda(x)}{x-p}\right) = \lim \frac{f(x) - f(p) - \lambda(x-p)}{x-p}$$

$$= \lambda - \lambda = 0.$$

$A_\lambda(x) = \lambda(x-p) + f(p)$

is a composition of a linear map & a translation, these are called affine maps.

let $L(\mathbb{R}^n ; \mathbb{R}^m)$ denote the set of all linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$.

Def. 1.5 let Ω be an open set in $\mathbb{R}^n$. and $f: \Omega \to \mathbb{R}^m$ be a map, $p \in \Omega$. The map $f: \Omega \to \mathbb{R}^m$ is called differentiable at p if there is a linear map $\Lambda \in L(\mathbb{R}^n ; \mathbb{R}^m)$

s.t. $$\lim \frac{\| f(x) - (\Lambda(x-p) + f(p)) \|}{\| x - p \|} = 0 . \quad \ast$$

The map Λ is called the derivative of f at p. and write $\Lambda = Df(p)$

replace $x-p$ with h

~~the~~ the ration ~~*~~ can be written as

$$\lim_{h\to 0} \frac{\| f(p+h) - f(p) - \Lambda[h] \|}{\|h\|} = 0 .$$

Example 1.8. The map $f : \mathbb{R}^n \to \mathbb{R}^1$,

$f(x) = \|x\|^2$ is differentiable at each

$p \in \mathbb{R}^n$, with

$$\underbrace{Df(p)}_{\text{linear}} [\overset{\in \mathbb{R}^n}{h}] = 2 \langle p, h \rangle .$$

~~the~~ For each fixed $p \in \mathbb{R}^n$, the map

$$h \longmapsto 2 \langle p, h \rangle$$

is linear.

$$f(p+h) = \|p+h\|^2 = \langle p+h, p+h\rangle$$

$$= \|p\|^2 + 2\langle p, h\rangle + \|h\|^2$$

Then

$$\lim_{h\to 0} \frac{\| f(p+h) - f(p) - 2\langle p, h\rangle \|}{\|h\|} = \lim_{h\to 0} \frac{\cancel{\|h\|^2}\ \|h\|}{\cancel{\|h\|}} = 0$$

Example 1.9. Let $m \geq 1$, and assume that $f^1, f^2, \ldots, f^m$ are maps from (a,b) to $\mathbb{R}^1$ which are differentiable at some $p \in (a,b)$. Then

$$f(x) = (f^1(x), f^2(x), \ldots, f^m(x)).$$

$$(a,b) \longrightarrow \mathbb{R}^m$$

is differentiable at p, with Jacobian

$$Df(p) = \begin{pmatrix} (f^1)'(p) \\ (f^2)'(p) \\ \vdots \\ (f^m)'(p) \end{pmatrix}.$$

$$f(p+h) - f(p) - \begin{pmatrix} (f^1)'(p) \\ \vdots \\ (f^m)'(p) \end{pmatrix} h$$

$$= \begin{pmatrix} f^1(p+h) - f^1(p) - (f^1)'(p).h \\ f^2(p+h) - f^2(p) - (f^2)'(p).h \\ \vdots \qquad \vdots \qquad \vdots \\ f^m(p+h) - f^m(p) - (f^m)'(p).h \end{pmatrix}$$

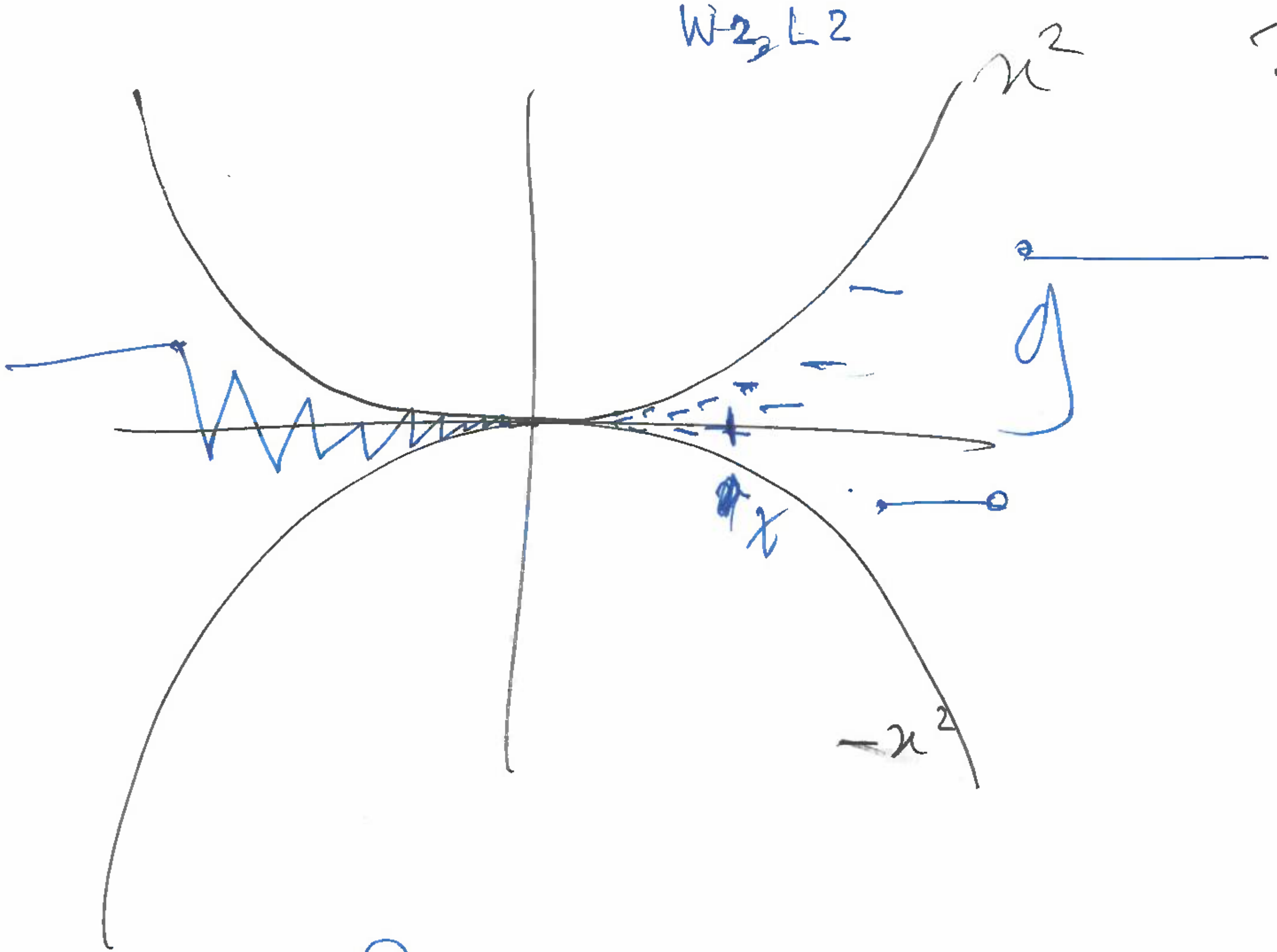
$$\lim_{h\to 0} \frac{\| f(p+h) - f(p) - Df(p)h \|}{\|h\|} =$$

using $\|w\| \le \sqrt{n} \cdot \max_{j=1,\dots,m} w^j$

there is a $j^* \in \{1, 2, \dots, m\}$

$$\le \sqrt{n} \frac{|(f^j)(p+h) - f^j(p) - (f^j)'(p)..h|}{|h|} = 0$$

W2, L2 7.



$$\lim_{x \to 0} \frac{g(x) - g(0)}{x}$$

$$0 \leftarrow \frac{-x^2}{x} \leq \lim_{x \to 0} \frac{g(x)}{x} \leq \frac{x^2}{x} \to 0$$

1.3.2 Chain Rule

Let $g:(a,b)\to(c,d)$ be differentiable at $p\in(a,b)$

and $f:(c,d)\to\mathbb{R}$ diff. at $g(p)$.

Then $f\circ g:(a,b)\to\mathbb{R}$ is differentiable at p,

with

$$D f\circ g = Df(g(p)) \cdot Dg(p)$$

multiplication

Thm 1.8 Assume $\Omega\subseteq\mathbb{R}^n$ is open, $\Omega'\subseteq\mathbb{R}^m$ is open, and $g:\Omega\to\Omega'$ is differentiable a $p\in\Omega$.

and $f:\Omega'\to\mathbb{R}^l$ is differentiable at $g(p)$.

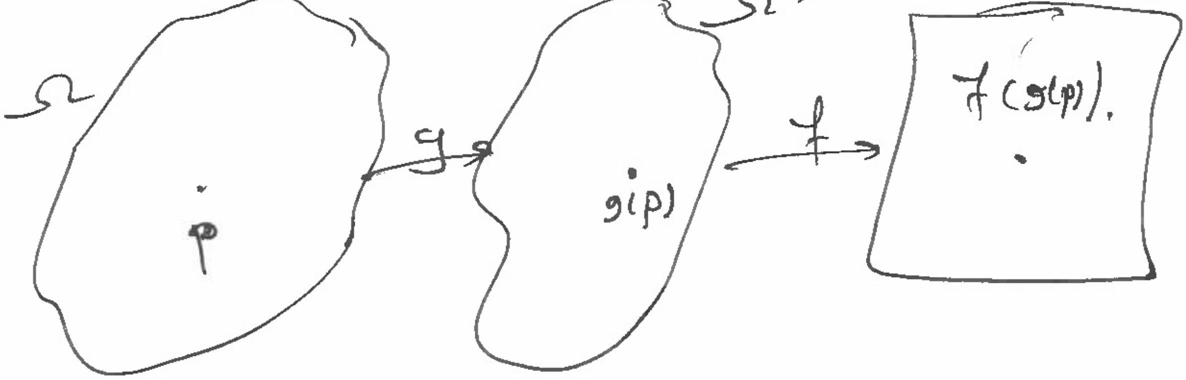
Then $h=f\circ g:\Omega\to\mathbb{R}^l$ is differentiable at p,

with derivative

$$Dh(p) = Df(g(p))\circ Dg(p).$$

composition of linear maps

proof: typed notes. optional. (*)



g near p is approximated by the affine map

$$x \longmapsto g(p) + Dg(p)(x-p)$$

and f near $g(p)$ is approximated by the affine map

$$y \longmapsto f(g(p)) + Df(g(p)).(y - g(p)).$$

Compose these affine maps, we obtain

$$x \longmapsto f(g(p)) + Df(g(p)).(\cancel{g(p)} + Dg(p)(x-p) - \cancel{g(p)})$$

$$= f(g(p)) + Df(g(p))(Dg(p)(x-p))$$

$$= f(g(p)) + \underbrace{Df(g(p)) \circ Dg(p)}[x-p]$$

Example 1.10 let $m \geq 1$, and for $i = 1, 2, \ldots, m$, the functions

$$g^i : (a,b) \to \mathbb{R} \text{ are differentiable at}$$

$p \in (a,b)$. the function $k : (a,b) \to \mathbb{R}$ defined as

$$k(x) = \| (g^1(x), g^2(x), \ldots, g^m(x)) \|^2$$

is differentiable at p and its derivative ~~has~~ is multiplication by

$$2g^1(p) \cdot (g^1)'(p) + 2g^2(p) \cdot (g^2)'(p) + \cdots + 2g^m(p) \cdot (g^m)'(p).$$

let $g(x) = (g^1(x), g^2(x), \ldots, g^m(x))$

and $f(x) = \|x\|^2$.

$g : (a,b) \to \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$,

$$f \circ g(x) = k(x)$$

By Example 1.9, $Dg(p) = \begin{pmatrix} (g^1)'(p) \\ (g^2)'(p) \\ \vdots \\ (g^m)'(p) \end{pmatrix}$

By Example 1.8

$$Df(q)[h] = 2\langle q, h\rangle.$$

By the chain Rule: $(k:(a,b)\to\mathbb{R}^1)$

$$Dk(p)[h] = Df(g(p))\circ Dg(p)\cdot[h]$$

$$= Df(g(p))\cdot[(g^1)'(p), (g^2)'(p), \dots, (g^m)'(p)]\cdot h$$

$$= Df(g(p))\cdot[(g^1)'(p).h, (g^2)'(p).h, \dots, (g^m)'(p).h]$$

$$= 2\langle (g^1(p), g^2(p), \dots, g^m(p)), ((g^1)'(p).h, \dots, (g^m)'(p).h)\rangle$$

$$= 2\, g^1(p)\cdot(g^1)'(p).h + \dots + 2\, g^m(p)(g^m)'(p).h$$

$$= \left(2\, g^1(p).(g^1)'(p) + \dots + g^m(p).(g^m)'(p)\right) h$$

1.4 Directional derivatives

~~1.4.~~ We would like to find a candidate for the derivative of a given map.

1.4.1 Rates of Change & partial derivatives

Let $\Omega \subseteq \mathbb{R}^n$, open, $f : \Omega \to \mathbb{R}^m$ be differentiable at $p \in \Omega$, and let $v \in \mathbb{R}^n$ be a unit vector.

We aim to identify $Df(p)[v] \in \mathbb{R}^m$.

Recall that by definition,

$$\lim_{h \to 0} \frac{\| f(p+h) - f(p) - Df(p).[h] \|}{\|h\|} = 0$$

In particular, if we replace $h = tv$, with $t \to 0$ in $\mathbb{R}$, we obtain

$$\lim_{t \to 0} \frac{\| f(p+tv) - f(p) - Df(p)[tv]}{\| tv \|} = 0$$

Df(p) is linear

$$\Rightarrow \lim_{t \to 0} \frac{\| f(p+tv) - f(p) - t \cdot Df(p)[v] \|}{|t|}$$

$\leftarrow \|v\| = 1$

$$\Rightarrow \lim_{t \to 0} \left\| \frac{f(p+tv) - f(p)}{t} - Df(p)[v] \right\| = 0$$

(using $\frac{1}{|t|} \cdot \|w\| = \| \frac{w}{t} \|$)

Therefore

$$\lim_{t \to 0} \frac{f(p+tv) - f(p)}{t} = Df(p)[v] \in \mathbb{R}^m$$

note that $f(p+tv) - f(p) \in \mathbb{R}^m$ & $t \in \mathbb{R}$,

so $$\frac{f(p+tv) - f(p)}{t} \in \mathbb{R}^m.$$

so the limite makes sense.

~~often~~ we write $Df(p).[v] = \frac{\partial f}{\partial v}(p)$.

This is called the partial derivative of f at p in direction v.

Directional derivatives.

$$f : \underset{\text{open}}{\Omega} \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

f is differentiable at some $p \in \Omega$, i.e.

- $\exists\, \Lambda \in L(\mathbb{R}^n ; \mathbb{R}^m)$

  s.t. a certain limit is 0.

$v \in \mathbb{R}^n$, $\|v\| = 1$,

then
$$\underbrace{Df(p)}_{\Lambda}[v] = \frac{\partial f}{\partial v}(p) = \lim_{t \to 0} \underbrace{\frac{f(p+tv) - f(p)}{t}}_{\in \mathbb{R}^m}$$

$$= \left(\lim_{t \to 0} \frac{f^1(p+tv) - f^1(p)}{t}, \ldots \right)$$

1-D analysis.

Any linear map $\Lambda \in L(\mathbb{R}^n ; \mathbb{R}^m)$ is determined by its values on the set $\{e_i\}_{i=1}^n$, which is the standard bases for $\mathbb{R}^n$.

In particular

$$D_i f = Df(p)[e_i] = \frac{\partial f}{\partial e_i}(p) = \lim_{t \to 0} \frac{f(p+te_i) - f(p)}{t}$$

W3, L2

If $f: \mathbb{R}^3 \to \mathbb{R}$. differentiable at (x,y,z)

then

$$D_1 f(x,y,z) = \lim_{t \to 0} \frac{f(x+t, y, z) - f(x,y,z)}{t}$$

" fix y, z, differentiate w.r.t. to x."

Thm 1.9. Suppose $\Omega \subseteq \mathbb{R}^n$ is open, and $f: \Omega \to \mathbb{R}^m$ is of the form

$$f(x) = \begin{pmatrix} f^1(x) \\ f^2(x) \\ f^m(x) \end{pmatrix}.$$

If f is differentiable at some $p \in \Omega$, then the Jacobian of f at p is

$$Df(p) = \begin{pmatrix} D_1 f^1(p) & D_2 f^1(p) & \cdots & D_n f^1(p) \\ D_1 f^2(p) & D_2 f^2(p) & \cdots & D_n f^2(p) \\ \vdots & \vdots & & \vdots \\ D_1 f^m(p) & D_2 f^m(p) & \cdots & D_n f^m(p) \end{pmatrix}$$

1.4.2 Relation between partial derivatives & differentiability.

Example 1.11. Consider the map $f: \mathbb{R}^2 \to \mathbb{R}$, defined as

$$f(x,y) = \begin{cases} 0 & x = y = 0 \\ \dfrac{xy}{\sqrt{x^2+y^2}} & \text{otherwise} \end{cases}$$

f is continuous at $(0,0)$, because

$$|xy| \leq |x|\,|y| \leq \|(x,y)\|\,\|(x,y)\| = \|(x,y)\|^2.$$

Then

$$|f(x,y)| = \frac{|xy|}{\|(x,y)\|} \leq \|(x,y)\|.$$

$$\Longrightarrow \lim_{(x,y)\to 0} f(x,y) = 0.$$

W's, L2

4

$$D_1 f(0,0) = \lim_{t\to 0} \frac{f((0,0)+t(1,0)) - f(0,0)}{t}$$

$$= \lim_{t\to 0} \frac{0-0}{t} = 0$$

$$D_2 f(0,0) = \lim_{t\to 0} \frac{f((0,0)+t(0,1)) - f(0,0)}{t} = 0.$$

f has partial derivatives at $(0,0)$ in directions e_1 & e_2.

$\Longrightarrow$ If f is differentiable, $\forall v \in \mathbb{R}^2$, we have $Df(0,0)[v] = 0$.

$$V = \frac{1}{\sqrt{2}}(1,1)$$

$$Df(0,0)[v] = \frac{\partial f}{\partial v}(0,0) = \lim_{t\to 0} \frac{f((0,0)+(\frac{t}{\sqrt{2}}, \frac{t}{\sqrt{2}}))}{t} = \lim_{t\to 0} \frac{\frac{1}{2}t^2}{t^2} = \frac{1}{2}$$

Thus, f is not differentiable at $(0,0)$.

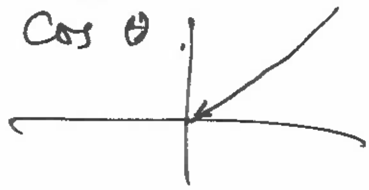
Let $(x,y) \neq (0,0)$

$$D_1 f(x,y) = \text{---} \text{--} = \frac{y^3}{(x^2+y^2)^{3/2}}.$$

$$D_2 f(x,y) = \text{---} \text{---} = \frac{x^3}{(x^2+y^2)^{3/2}}.$$

Let $\theta \in [0, 2\pi]$, let us look at

$$\underset{t \to 0}{D_2 f(t\cos\theta, t\sin\theta)} = \text{--} = \cos^3\theta.$$



$\Rightarrow D_2 f$ is not continuous at $(0,0)$.

Thm 1.12 Let $\Omega \subseteq \mathbb{R}^n$ be an open set, and $f : \Omega \to \mathbb{R}^m$. Assume that the partial derivatives

$$D_i f(x), \quad \text{for } i = 1, 2, \dots, n$$

exist at every $x \in \Omega$, and the maps

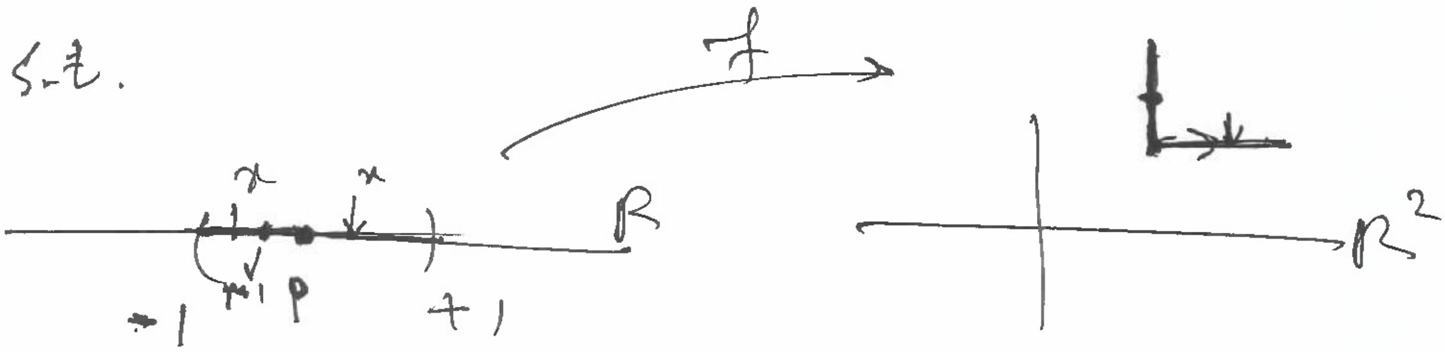
$$\underset{\in \mathbb{R}}{x} \mapsto \underset{\in \mathbb{R}^m}{D_i f(x)} \quad \text{for } i = 1, \dots, n$$

are continuous at some $p \in \Omega$. Then, f is differentiable at p

proof*: see typed notes for the proof.

Let $f : (-1, +1) \longrightarrow \mathbb{R}^2$ be a map

s.t.



Is f differentiable at 0?

See example 1.12 in typed notes.

$$\text{Lt} \frac{\left\| \left(f(p+h) - f(p) \right) - A(h) \right\|}{\| h \|} = 0.$$

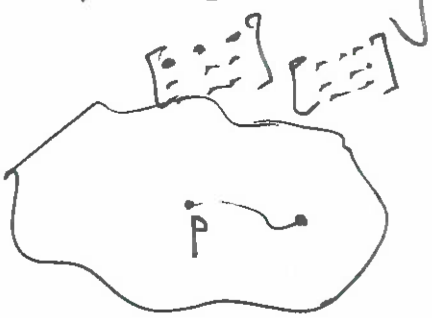
1.5 Higher derivatives

1.5.1 higher derivatives as linear maps

$\Omega \subseteq \mathbb{R}^n$ is open, $f: \Omega \to \mathbb{R}^m$ is differentiable,

$$\begin{array}{ccc} Df: \Omega & \longrightarrow & L(\mathbb{R}^n; \mathbb{R}^m) \\ \Downarrow & & \Downarrow \\ p & \longmapsto & Df(p)\,(: \mathbb{R}^n \to \mathbb{R}^m) \end{array}$$

We would like to study the dependence of Df to p.



every element of $L(\mathbb{R}^n; \mathbb{R}^m)$ ~~is an~~ may be expressed as an $m \times n$ matrix.

every $m \times n$ matrix may be considered as an element of $\mathbb{R}^{mn}$

$$(a_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le m}} \longmapsto (a_{1,1}, a_{1,2}, a_{1,3}, \dots, a_{1,m}, a_{2,1}, a_{2,2}, \dots, a_{3m}, \dots, a_{n,1}, a_{n,2}, \dots, a_{n,m})$$

Then $L(\mathbb{R}^n; \mathbb{R}^m) \simeq \mathbb{R}^{mn}$ W4, L1

Thus $Df : \Omega \to \mathbb{R}^{mn}$

If $Df : \Omega \to \mathbb{R}^{mn}$ is continuous, we say $f : \Omega \to \mathbb{R}^m$ is continuously differentiable.

If $Df : \Omega \to \mathbb{R}^{mn}$ is differentiable, we say f is two times differentiable.

$$DDf : \mathbb{R}^n \longrightarrow L(\mathbb{R}^n; \mathbb{R}^{mn})$$
$$\| $$
$$L(\mathbb{R}^n; L(\mathbb{R}^n; \mathbb{R}^m))$$
$$(\, v \longmapsto \langle v, \cdot \rangle) \text{ if } m=1$$

for $p \in \Omega$, $DDf(p)$ is an element $\mathcal{L} \in L(\mathbb{R}^n; L(\mathbb{R}^n; \mathbb{R}^m))$ s.t.

$$\lim_{x \to p} \frac{\| Df(x) - Df(p) - \mathcal{L}(x-p) \|}{\| x-p \|} = 0.$$

We may continue to look at higher derivatives and say f is k-times differentiable, if

$$\underbrace{DD\cdots D}_{k\text{ times}}f \quad \text{is defined.}$$

This is formally difficult, it requires multi-linear maps.

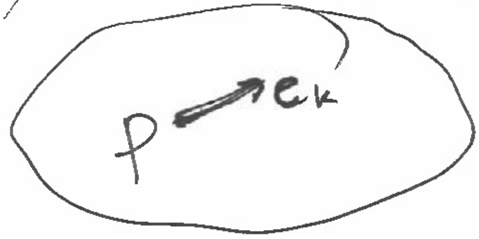
It is easier to look at higher partial derivatives, and ask if they are continuous on ~~any~~ an open set. If $f=(f^1,f^2,\ldots,f^m):\Omega\to\mathbb{R}^m$, then

$$D_i f^j : \Omega \longrightarrow \mathbb{R}^1$$
$$x \longmapsto D_i f^j(x)$$

Then

$$D_k \underbrace{D_i f^j(p)} = \lim_{t\to 0} \frac{D_i f^j(p+te_k)-D_i f^j(p)}{t} \in \mathbb{R}^1$$

If the k-th partial derivative of f exist and are continuous, then f is k-times differentiable.



Example 1.14. Let $f: \mathbb{R}^2 \to \mathbb{R}$,

$$f(x,y) = x^3 + y^3 + 5x^2y$$

$$D_1 f(x,y) = 3x^2 + 10x$$

$$D_2 f(x,y) = 3y^2 + 5x^2$$

These are continuous, then f is 1 times differentiable.

$$D_1 D_1 f(x,y) = 6x + 10 \qquad D_2 D_1 f = 10x.$$

$$D_1 D f(x,y) = 10x \qquad D_2 D_2 f = 6y$$

$\Rightarrow$ f is 2-times differentiable.

1.5.2 symmetry of partial derivatives.

Thm 1.13. Let $\Omega \subseteq \mathbb{R}^n$ be open, $f: \Omega \to \mathbb{R}^m$ is differentiable at every $p \in \Omega$. Assume that, for some i,j in $\{1,2,\dots,n\}$ the second partial derivatives

$$D_i D_j f \quad \& \quad D_j D_i f$$

exist and are continuous at all points in Ω. Then

$$D_i D_j f(p) = D_j D_i f(p), \ \forall p \in \Omega.$$

1.5.3 Taylor's thm.

A multi index α is an element of the form $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ for some $n \geq 1$,

$$\alpha_i \in \{0, 1, 2, \ldots\}$$

$$|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n$$

given $f : \Omega \to \mathbb{R}$ which is k times differentiable

for any α with $|\alpha| \leq k$, we define

$$D^\alpha f(p) = D_1^{\alpha_1} D_2^{\alpha_2} \cdots D_n^{\alpha_n} f(p).$$

For $h \in \mathbb{R}^n$, $h^\alpha = h_1^{\alpha_1} h_2^{\alpha_2} \cdots h_n^{\alpha_n}$

$$\alpha! = \alpha_1! \, \alpha_2! \cdots \alpha_n!$$

Example $D^{(0,3,0)} f(p) = D_2^3 f(p) = D_2 D_2 D_2 f(p)$

$$D^{(1,0,1)} f(p) = D_1 D_3 f(p)$$

$$(x, y, z)^{(2,1,5)} = x^2 y \cdot z^5$$

W4, L1

Thm 1.14. suppose $p \in \mathbb{R}^n$, and $f: B_r(p) \to \mathbb{R}$ $\in$

is k-times differentiable at all points in $B_r(p)$, for some $r > 0$, and $k \geq 1$.

Then, for any $h \in \mathbb{R}^n$ with $\|h\| < r$, we have

$$f(p+h) = \sum_{\alpha;\, |\alpha| \leq k-1} D^\alpha f(p) \cdot \frac{h^\alpha}{\alpha!} + R_k(p,h)$$

where there is $x_h \in \mathbb{R}^n$ with $0 < \|x_h - p\| \leq \|h\| < r$.

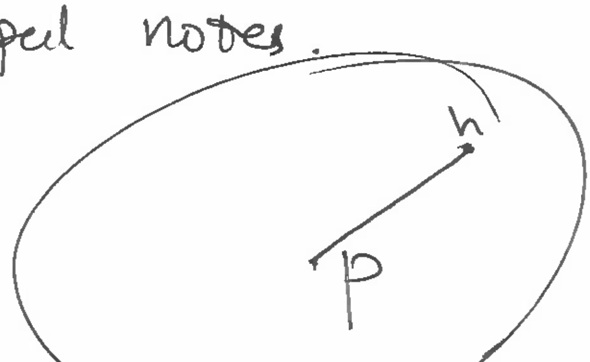
$$R_k(p,h) = \sum_{\alpha;\, |\alpha| = k} D^\alpha f(x_h) \cdot \frac{h^\alpha}{\alpha!}.$$

Evidently,
$$\lim_{h \to 0} \frac{|R_k(p,h)|}{\|h\|^{k-1}} = 0$$

$h = 1/10$

$1/10^5$

proof*; see typed notes.



$t \mapsto f(p+th)$

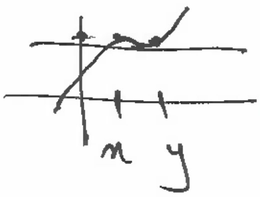
1.6.1 Inverse function theorem

Let $f : (a, b) \longrightarrow \mathbb{R}$ be a continuously differentiable function, and assume that there is $c \in (a, b)$ s.t. $f'(c) \neq 0$.

WLG, $f'(c) > 0$

Because f is continuously differentiable, there is an interval $I \subseteq (a, b)$ s.t. $\forall x \in I$, $f'(x) > 0$.

By the mean value theorem, this implies that f is strictly increasing on I.



In particular, $f : I \to f(I)$ is a bijection, and

(i) $f^{-1} : f(I) \longrightarrow I$ is continuously differentiable,

(ii) $$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}.$$

Thm 1.15. (Inverse function theorem)

Let $\Omega \subseteq \mathbb{R}^n$ be an open set, $q \in \Omega$, and $f : \Omega \to \mathbb{R}^n$ is a map. If

(i) f is continuously differentiable on Ω, and,

(ii) $Df(q)$ is invertible,

then, there are open sets $U \subseteq \Omega$, and $V \subseteq \mathbb{R}^n$ with $q \in U$, $f(q) \in V$, s.t.

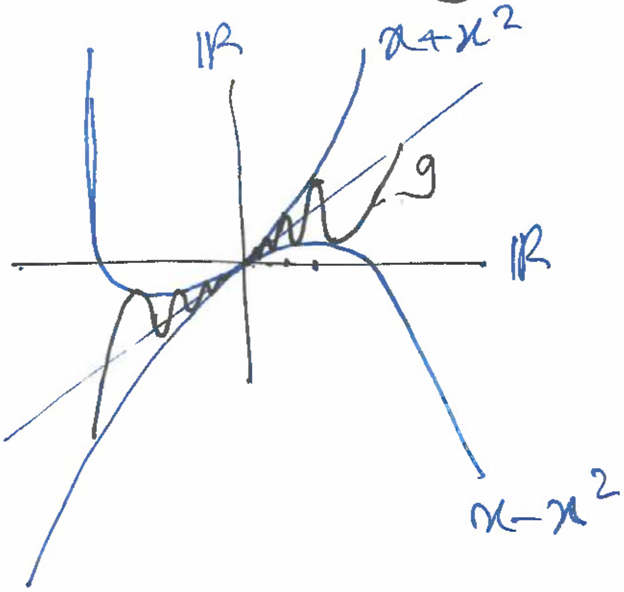
(i) $f : U \to V$ is a bijection

(ii) $f^{-1} : V \to U$ is continuously differentiable.

(iii) for any $y \in V$,

$$Df^{-1}(y) = [Df(f^{-1}(y))]^{-1}.$$

Remark 1.



g is differentiable at 0, but it is not cont. diff. on $\mathbb{R}$. So the theorem may not be used.

Remark 2. The function $f(x) = x^3$.

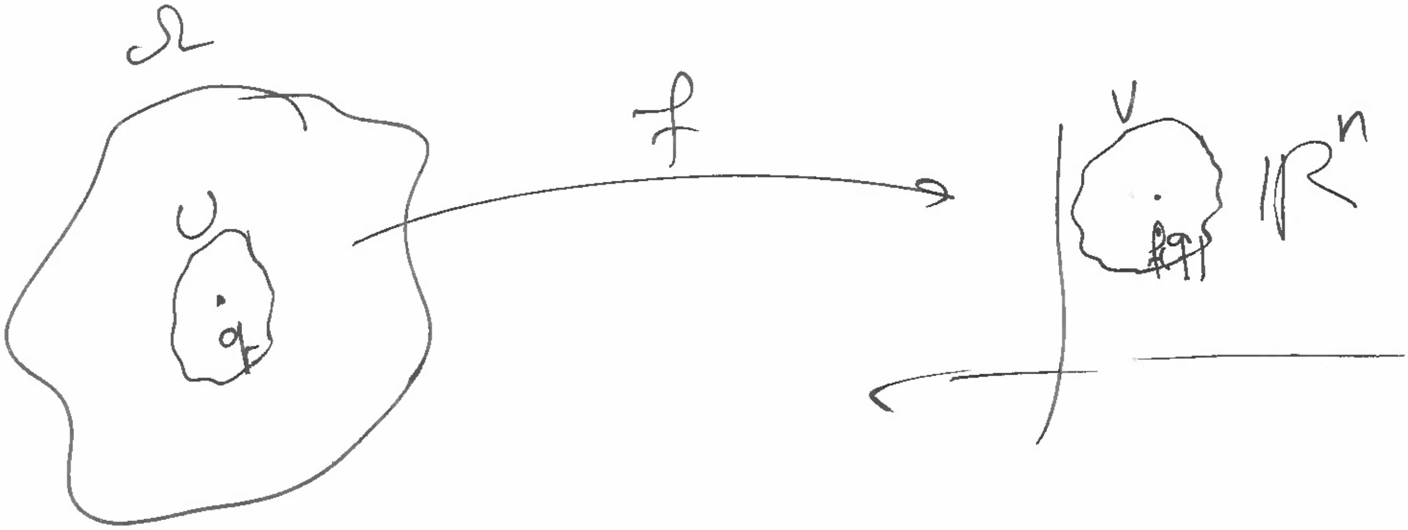
$f'(x) = 3x^2$, f is ~~continuously~~ continuously differentiable,

$f^{-1}(x) = x^{1/3}$

but f^{-1} is not even differentiable at 0.

The problem here is that $f'(0)$ ~~≠ 0~~

= 0 it's not invertible.



Example: 1.15 Consider the map $f : \mathbb{R}^2 \to \mathbb{R}^2$

defined as $f(x,y) = (x+y+5xy,\ y-x^2)$

$f(0,0) = (0,0)$.

The first partial derivatives

$D_1 f(x,y) = (1+5y, -2x)$, $D_2 f(x,y) = (1+5x, 1)$

are continuous, then by Thm 1.12, f is differentiable.

W4, L2

4

$$Df(x,y) = \begin{pmatrix} 1+5y & 1+5x \\ -2x & 1 \end{pmatrix}$$

all entries are continuous, then Df is continuous. Then f is continuously differentiable

$$Df(0,0) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$\det Df(0,0) = 1 \neq 0$, hence Df is invertible at $(0,0)$.

By the IFT, there is an open set $U \subseteq \mathbb{R}^2$, an open set $V \subseteq \mathbb{R}^2$, with $(0,0) \in U$, $f(0,0) \in V$,

$f: U \to V$ is a bijection.

$$(Df^{-1})(0,0) = \begin{pmatrix} 1 & -1 \\ 0 & +1 \end{pmatrix}$$

An important application of IFT is to the system of n non-linear equations in n-unknowns.

$$f^1(x^1, x^2, \ldots, x^n) = y^1,$$
$$f^2(x^1, x^2, \ldots, x^n) = y^2,$$
$$\vdots$$
$$f^n(x^1, x^2, \ldots, x^n) = y^n.$$

let $f = (f^1, f^2, \ldots, f^n) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$,

for $(x_0^1, x_0^2, \ldots, x_0^n)$ we get $(y_0^1, y_0^2, \ldots, y_0^n) = f(x_0^1, x_0^2, \ldots, x_0^n)$

If f is cont diff. & $Df(x_0^1, x_0^2, \ldots, x_0^n)$ is invertible, then, for any $(y^1, y^2, \ldots, y^n)$ close to $(y_0^1, y_0^2, \ldots, y_0^n)$, the system has a unique solution $(x^1, x^2, \ldots, x^n)$ close to $(x_0^1, \ldots, x_0^n)$.

W4, L2

$\underline{6}$

$$\begin{cases} x^2 y^7 + \sin x^5 \cdot \cos xy = 5 \\ x y^2 + \tan(xy) = 3 \end{cases}$$

1.6.2 Implicit function theorem

Can we solve a system of non-linear equations with more unknowns than equations.

$$f^1(x^1, x^2, \ldots, x^n) = y^1$$
$$f^2(x^1, x^2, \ldots, x^n) = y^2$$
$$\vdots$$
$$f^m(x^1, x^2, \ldots, x^n) = y^m$$

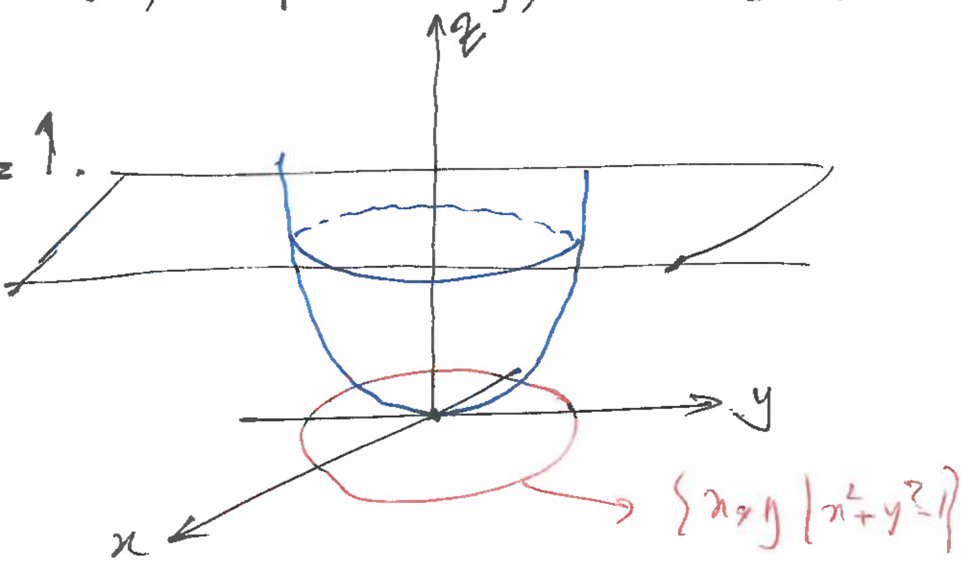
and $m < n$.

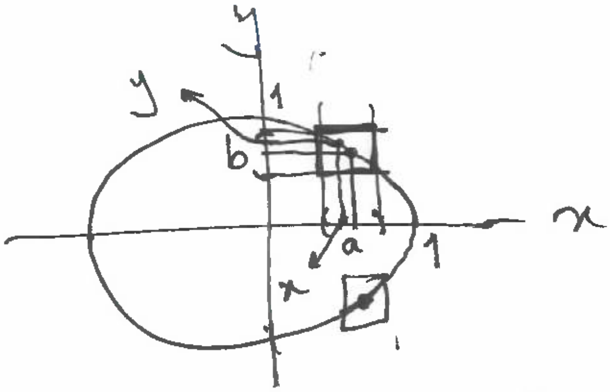
Consider the example

$$x^2 + y^2 - 1 = 0$$

let $f(x,y) = x^2 + y^2$; equivalently, we consider

$$f(x,y) = 1.$$





Let $(a,b) \in \mathbb{R}^2$ be s.t. $f(a,b)=1$, and $(a,b) \neq (\pm 1, 0)$

There are open sets $A \subseteq \mathbb{R}$ and $B \subseteq \mathbb{R}$ with $a \in A$, $b \in B$ which satisfies the following

for any $x \in A$, there is a unique $y \in B$ s.t.

$$f(x,y)=1$$

equivalently, there is a function $g: A \to B$ s.t.

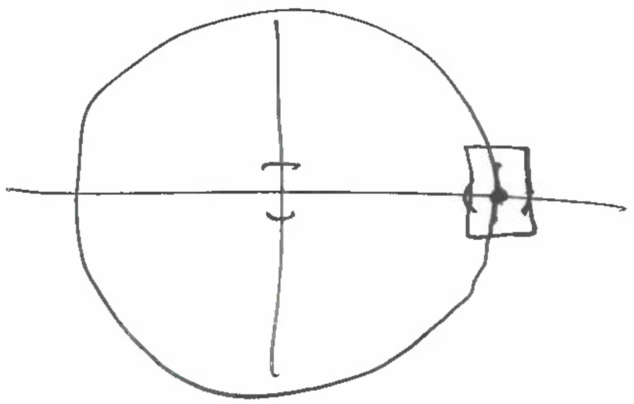
$$\{(x,y) \in A \times B \mid f(x,y)=1\} = \{(x, g(x)) \mid x \in A\}$$

Similarly, if $a \neq \pm 1$, $b<0$, there are open sets $A' \subseteq \mathbb{R}$, $B' \subseteq \mathbb{R}$ with $a \in A'$, $b \in B'$, and a function $h: A' \to B'$ s.t.

$$\{(x,y) \in A' \times B' \mid f(x,y)=1\} = \{(x, h(x)) \mid x \in A'\}.$$

$g(x) = \sqrt{1-x^2}$, $\quad h(x) = -\sqrt{1-x^2}$.

note that the uniqueness is only true after restricting to the open set $A \times B$.

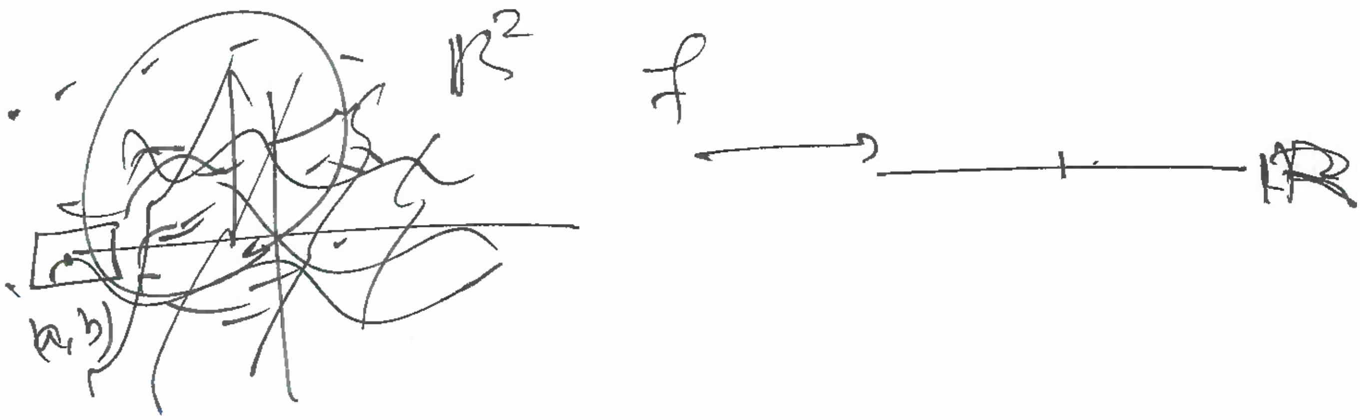


if $(a,b) = (+1, 0)$ there are no open sets $A \times B$ with the above property.

if $1 \in A$, $\exists \delta > 0$ s.t. $(1-\delta, 1+\delta) \subseteq A$.

If $0 \in B$, $\exists \varepsilon > 0$ s.t. $(-\varepsilon, \varepsilon) \subseteq B$.

for any $x \in (1-\delta, 0)$, there are two y satisfying $f(x,y) = 1$.



Thm 1.16 (Implicit function theorem — low dimensional version)

Assume that $\Omega \subseteq \mathbb{R}^2$ is open, $F: \Omega \to \mathbb{R}^1$ is continuously differentiable, and there a point $(x^1, x^2) \in \Omega$ s.t.

(i) $F(x^1, x^2) = 0$

(ii) $D_2 F(x^1, x^2) \neq 0$.

Then, there are open sets $A \subseteq \mathbb{R}^1$, $B \subseteq \mathbb{R}^1$ with $x^1 \in A$, $x^2 \in B$, and a function $f: A \to B$, s.t.

$$\{(y^1, y^2) \in A \times B \mid F(y^1, y^2) = 0\} = \{(y, f(y)) \mid y \in A\}.$$

More over, $f: A \to B$ is continuously differentiable.

~~Lm~~ Thm 1.17 (Implicit function theorem)

Let $\Omega \subseteq \mathbb{R}^n$, $\Omega' \subseteq \mathbb{R}^m$ be open sets, and

$F : \Omega \times \Omega' \longrightarrow \mathbb{R}^m$ be continuously

differentiable. Suppose that there is

$P = (a,b) \in \Omega \times \Omega'$ satisfies.

(i) $F(a,b) = 0$

(ii) the matrix $\left(D_{n+j} F^i (a,b) \right)_{\substack{1 \le j \le m \\ 1 \le i \le m}}$

is invertible.

Then, there are open sets $A \subseteq \Omega$ & $B \subseteq \Omega'$

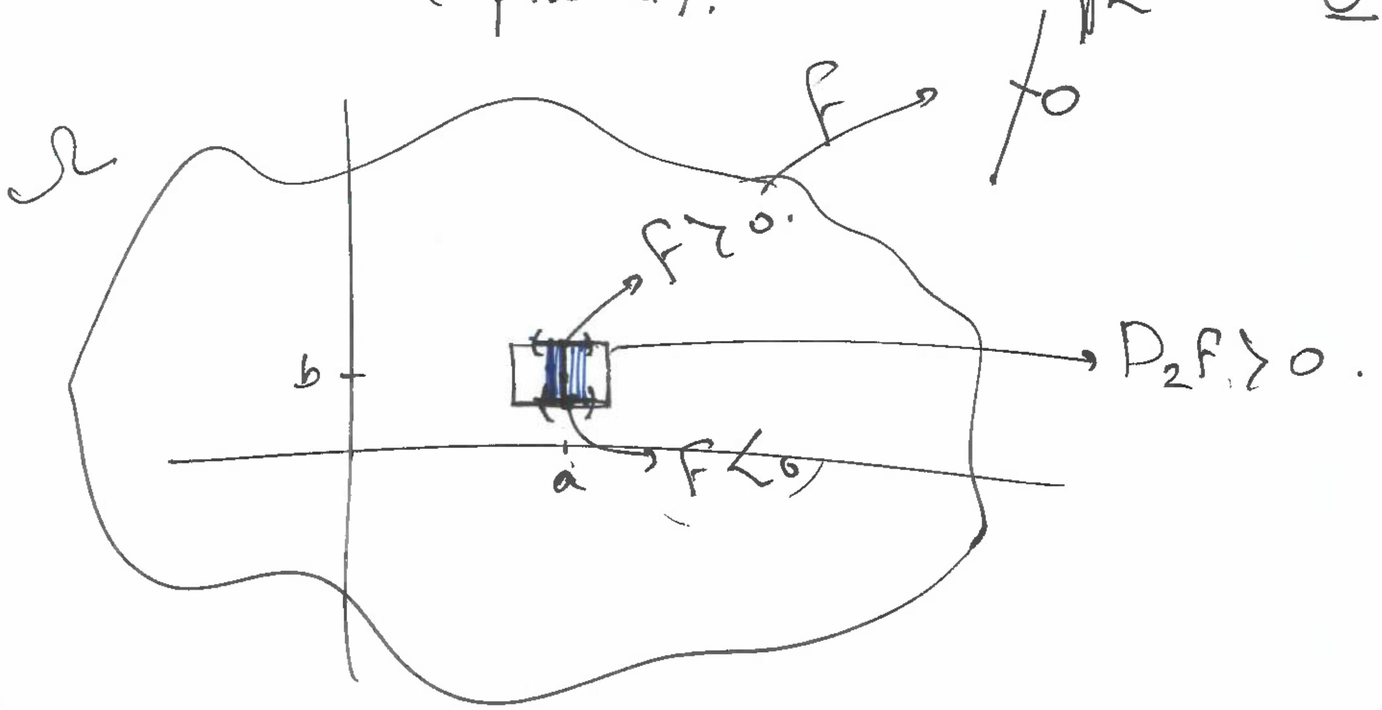
with $a \in A$, $b \in B$, ~~s.t.~~ and a function $g : A \to B$, s.t.

$$\{ (x,y) \in \Omega \times \Omega' \mid F(x,y) = 0 \} = \{ (x, g(x)) \mid x \in A \}.$$

Moreover, g is continuously differentiable on A.

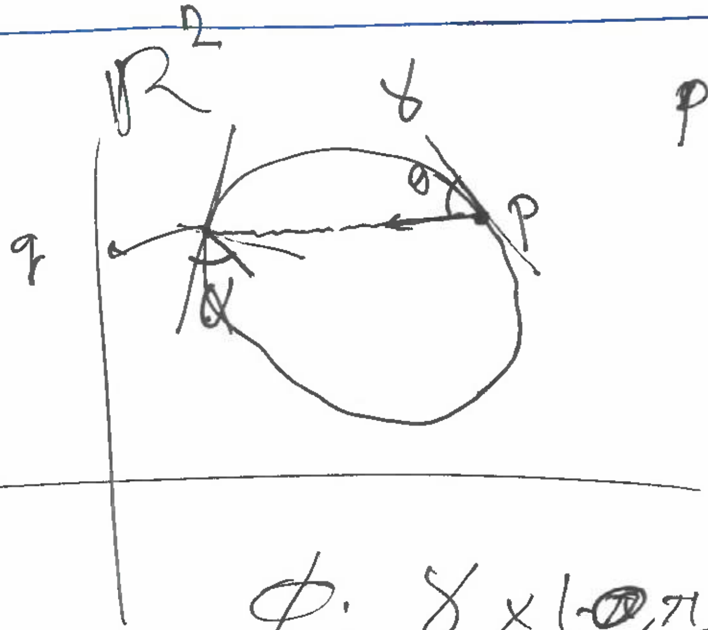
$$F(x, g(x)) = 0$$

proof of thm 1.16 (optional).



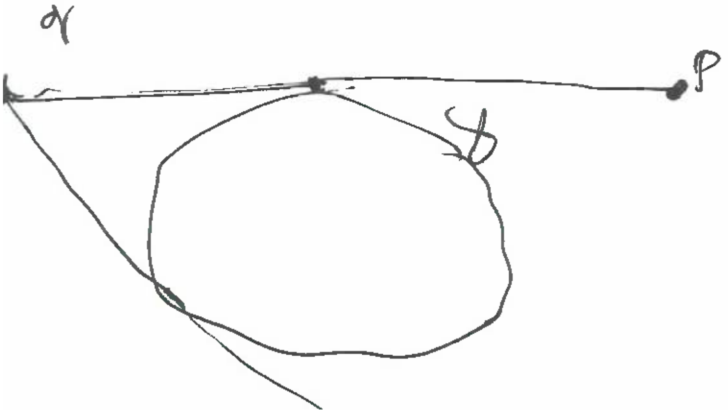
WLG

$$D_2 F(a,b) > 0.$$



$p \in \gamma,\ \theta \in (0, \pi)$

$\phi : \gamma \times (0,\pi) \longrightarrow \gamma \times (0,\pi)$

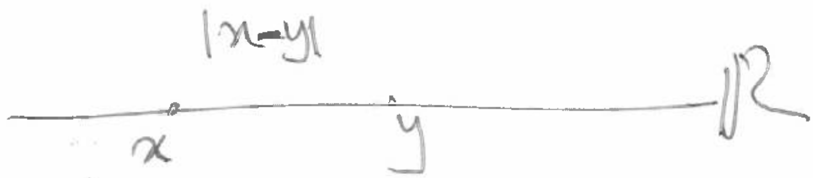


$(p, \theta) \longmapsto (q, \alpha)$

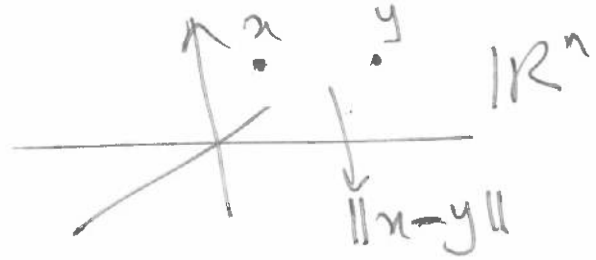
Chapter 2.

metric & topological spaces.

- $\mathbb{R}$, $|\cdot|$



- $\mathbb{R}^n$, $\|\cdot\|$



Def 2.1 Let X be an arbitrary set. A metric on X is a function

$$d: X \times X \to \mathbb{R}$$

satisfying the following 3 properties:

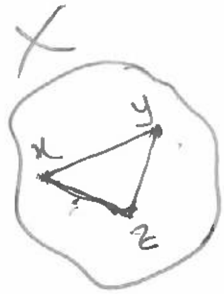
M1) $\forall x, y \in X$, $d(x,y) \geq 0$, and $d(x,y) = 0 \Leftrightarrow x = y$.

M2) $\forall x, y \in X$, $d(x,y) = d(y,x)$.

M3) $\forall x, y, z \in X$,

$$d(x,y) \leq d(x,z) + d(z,y)$$

triangle inequality.



Def 2.2 By a metric space, we mean a set X and a metric $d: X \times X \to \mathbb{R}$.

let M be a metric space,

$$M = (X, d)$$

In a metric space $M = (X, d)$, any element of X is called a point.

For points $x, y \in X$, $d(x, y)$ is called the "distance" between x & y.

Example 2.1: if $X = \mathbb{R}$, ad $d_1 : \mathbb{R}^1 \times \mathbb{R}^1 \to \mathbb{R}$

$$d_1(x, y) = |x - y|.$$

(M2) $\quad d(x, y) = |x - y| = |y - x| = d(y, x)$.

Example 2.2. let $X = \mathbb{R}^n$, and $d_2 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$,

$$d_2(x, y) = \|x - y\|.$$

Euclidean metric on $\mathbb{R}^n$.

Example 2.6. Let X be an arbitrary set

define $d_{disc} : X \times X \longrightarrow \mathbb{R}$, as

$$d_{disc}(x,y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{if } x \neq y \end{cases}$$

This is called the discrete metric on X.

Let $a < b$ be real numbers.

define $C[a,b] = \{ f : [a,b] \longrightarrow \mathbb{R} \mid f \text{ is continuous on } [a,b]. \}$

Rem. $(\mathbb{R} \times \mathbb{R} \times \mathbb{R} \dots \times \mathbb{R} \times \dots)$

Example 2.9 For f and g in $C[a,b]$ define

$$d_1(f,g) = \int_a^b |f(t) - g(t)| \, dt \in \mathbb{R}.$$

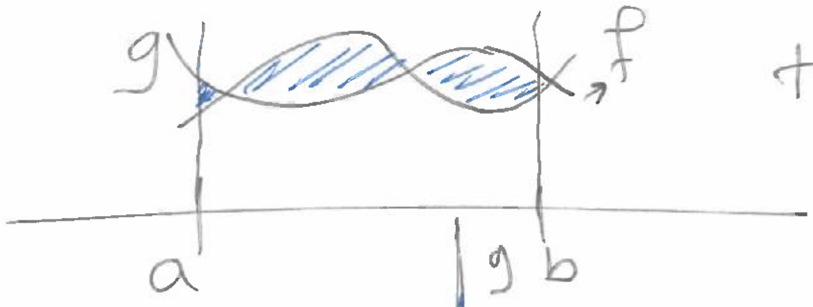
For example for (M3) let $f, g, h \in C[a,b]$ be arbitrary,

For any $t \in [a,b]$,

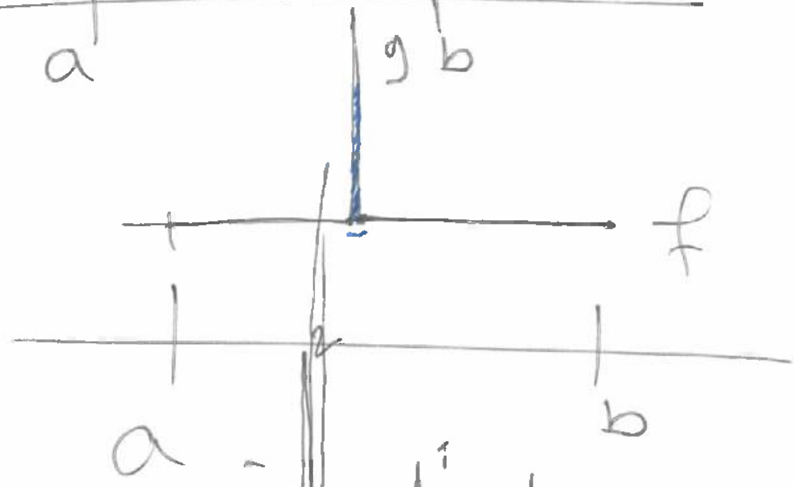
$$|f(t)-g(t)| \leq |f(t)-h(t)| + |h(t)-g(t)|$$

$$\int_a^b \downarrow$$

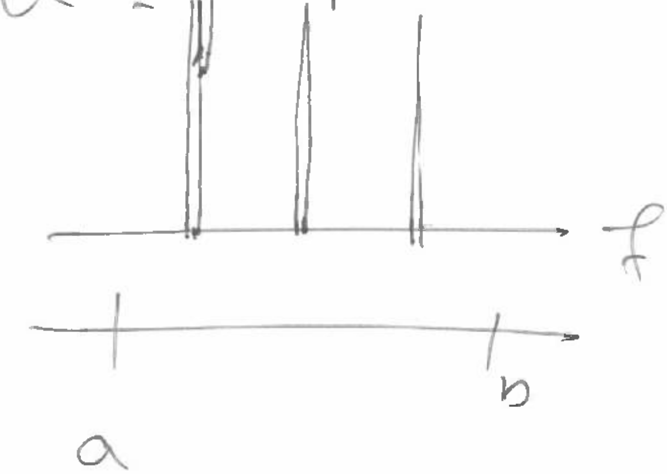
$$d(f,g) \leq d(f,h) + d(h,g).$$



total area $= d(f,g)$.



$d(f,g)$ is small.



Remark.

$$\varepsilon > \varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \cdots \qquad \varepsilon_i > 0$$

Def 2.3. Let (X,d) be a metric space and $Y \subseteq X$. The function $d: X \times X \to \mathbb{R}$,

$$d|_Y : Y \times Y \to \mathbb{R}^1,$$

$$d|_Y (y_1, y_2) = d(y_1, y_2).$$

The pair $(Y, d|_Y)$ is called a metric subspace of (X,d). The metric $d|_Y$ is called the induced metric on Y from X.

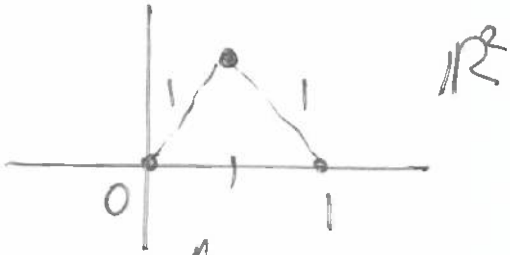
Example $\mathbb{Q} \subseteq \mathbb{R}$, d_1 on $\mathbb{R}^1$ induces

$$d|_{\mathbb{Q}} \text{ on } \mathbb{Q}, \quad d|_{\mathbb{Q}}(p/q, m/n) = |p/q - m/n|.$$

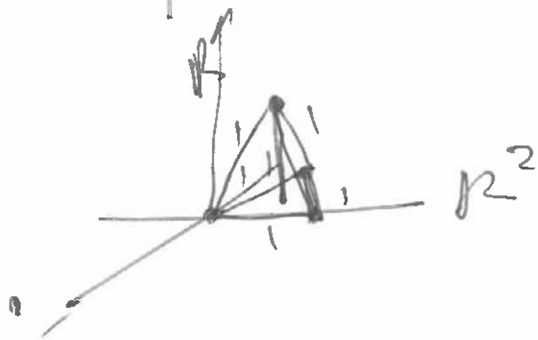
If $X = \{a_1\}$,

if $(X = \{a_1, a_2\}, d_{disc}) \cong$

If $X = \{a_1, a_2, a_3\}, d_{disc} \simeq$



If $X = \{a_1, a_2, a_3, a_4\}\, d_{disc} \simeq$



$\vdots$

$X = \{1, 2, 3, \ldots\} = \mathbb{N}$, d_{disc}.