1

# 3. Differential Calculus for functions of several variables II: Differentiability & Properties

$f : \mathbb{R}^d \to \mathbb{R}, \quad f = f(x_1, \dots, x_d)$

$x = (x_1, \dots, x_d), \quad y = (y_1, \dots, y_d) \in \mathbb{R}^d$

$x \cdot y = x_1 y_1 + x_2 y_2 + \dots + x_d y_d$

$\|x\| = \sqrt{x \cdot x}$

Def (continuity) $f : \mathbb{R}^d \to \mathbb{R}$ is cont. at $x \in \mathbb{R}^d$ if $\forall \varepsilon > 0 \ \exists \delta > 0$ s.t. $|f(x) - f(y)| < \varepsilon$ $\forall y \in \mathbb{R}^d$ with $\|x - y\| < \delta$.

Partial derivatives = normal/standard derivatives but w.r.t. one of the $x_1, \dots, x_d$ (the rest being fixed)

Ex. $f(x_1, x_2) = x_1 x_2^3$ $\qquad g(x_1, x_2) = x_1 x_2 + x_2$

$\frac{\partial f}{\partial x_1}(x_1, x_2) = 1 x_2^3 \qquad \frac{\partial f}{\partial x_2}(x_1, x_2) = x_1 3 x_2^2$

$\frac{\partial g}{\partial x_1}(x_1, x_2) = 1 x_2 + 0 \qquad \frac{\partial g}{\partial x_2}(x_1, x_2) = x_1 \cdot 1 + 1$

Partial derivative = partial info // Full info : The Gradient

$\nabla f(x_1, \dots, x_d) = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_d}\right)$

Th 1. (CHAIN rule)

$f : \mathbb{R}^d \to \mathbb{R}$, $\quad x_1, \dots, x_d : [a,b] \subset \mathbb{R} \to \mathbb{R}$

has cont $\frac{\partial}{\partial x_i}$ $\quad$ all differentiable

Then $F : [a,b] \to \mathbb{R}$, $\quad F(t) = f(x_1(t), \dots, x_d(t))$

"$F = f \circ (x_1, \dots, x_d)$" is diffable

and $$\frac{d}{dt} F(t) = \nabla f \underbrace{(x_1(t), \dots, x_d(x))}_{x(t)} \cdot \frac{d}{dt} x(t)$$

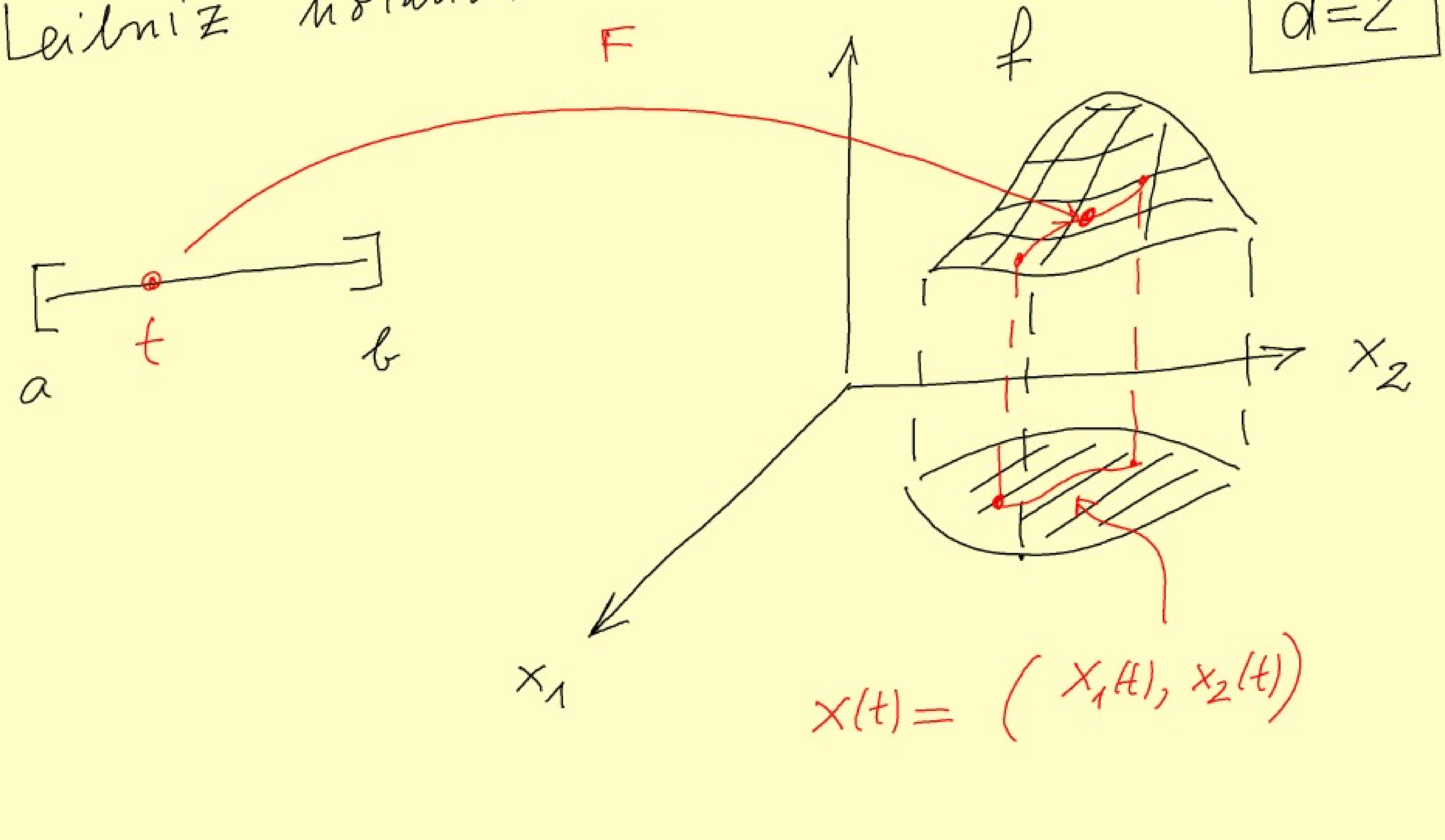
$\frac{d}{dt}$ = standard derivative w.r.t. one variable $(t)$

$$\frac{d}{dt} x(t) = \left(\frac{d}{dt} x_1(t), \dots, \frac{d}{dt} x_d(t)\right) = (x_1'(t), \dots, x_d'(t))$$

$$x(t) = (x_1(t), \dots, x_d(t))$$

Leibniz notation for a derivative

$\boxed{d=2}$



Th $(d=1)$ $\quad f:[a,b]\to\mathbb{R}$ **LAGRANGE**

If
- $f$ cont. on $[a,b]$
- $f$ diffable $(a,b)$

Then $\exists\, \xi\in(a,b)$ s.t.

$$f'(\xi)=\frac{f(b)-f(a)}{b-a}$$

(see Lecture 1.) $\quad \Leftrightarrow\ f(b)-f(a)=f'(\xi)(b-a)$

Th 2. (Lagrange for $d>1$)

If $D\subseteq\mathbb{R}^d$ convex, $\quad a,b\in D \quad (a\neq b)$

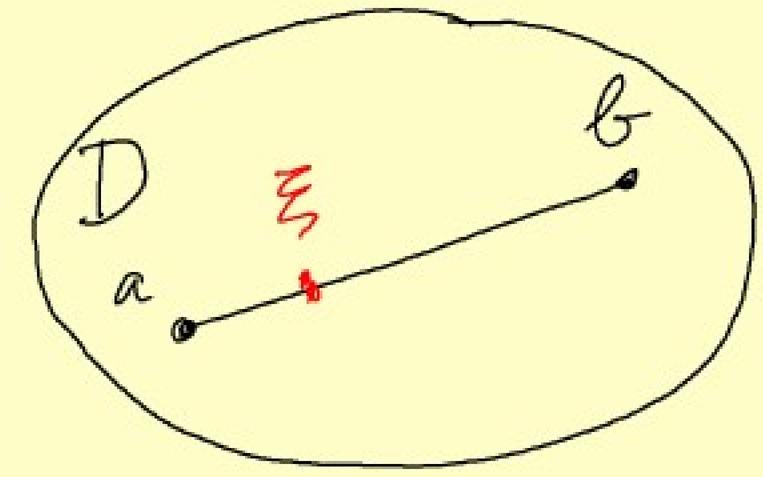
$f:D\to\mathbb{R}$ (or $\mathbb{R}^d$) has cont. $\frac{\partial}{\partial x_i}$

Then $\exists\,\xi\in(a,b)$ s.t.
open segment

$$f(b)-f(a)=\nabla f(\xi)\cdot(b-a)$$

IDEA of the proof: apply the $d=1$ Lagrange Th to $F:[0,1]\to\mathbb{R}$, $\quad F(t)=f(a+t(b-a))$

(use Th CHAIN) $\quad \boxed{\text{HW}}$




Key point: segment = one-dimens.

you can use $d=1$ Lagrange on the segment

Def (directional derivati

$v = (v_1, \dots, v_d) \in \mathbb{R}^d$ "the direction"

$f : \mathbb{R}^d \to \mathbb{R}$

$$\frac{\partial f}{\partial v}(x_1, \dots, x_d) = \nabla f(x_1, \dots, x_d) \cdot v$$

(geometrically $\to$ = the projection of $\nabla f$ onto $v$)

Rk: Partial derivatives are directional derivatives w.r.t. (special directions)

$$\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial v} \quad \text{for } v = (1, 0, \dots, 0)$$

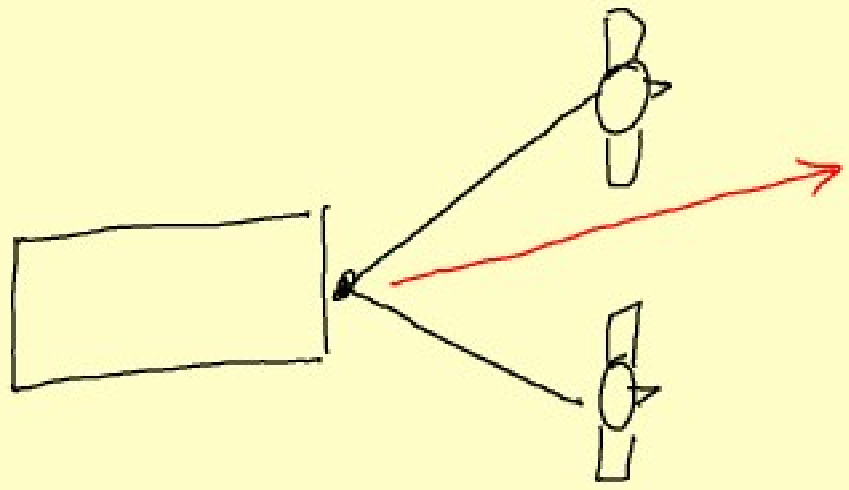
$$\frac{\partial f}{\partial x_i} = \frac{\partial f}{\partial v} \quad \text{for } v = (0, \dots, 0, \underset{i\text{-th pos}}{1}, 0, \dots, 0)$$

Rk: any $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ can be decomposed w.r.t. these special (CANONICAL) directions

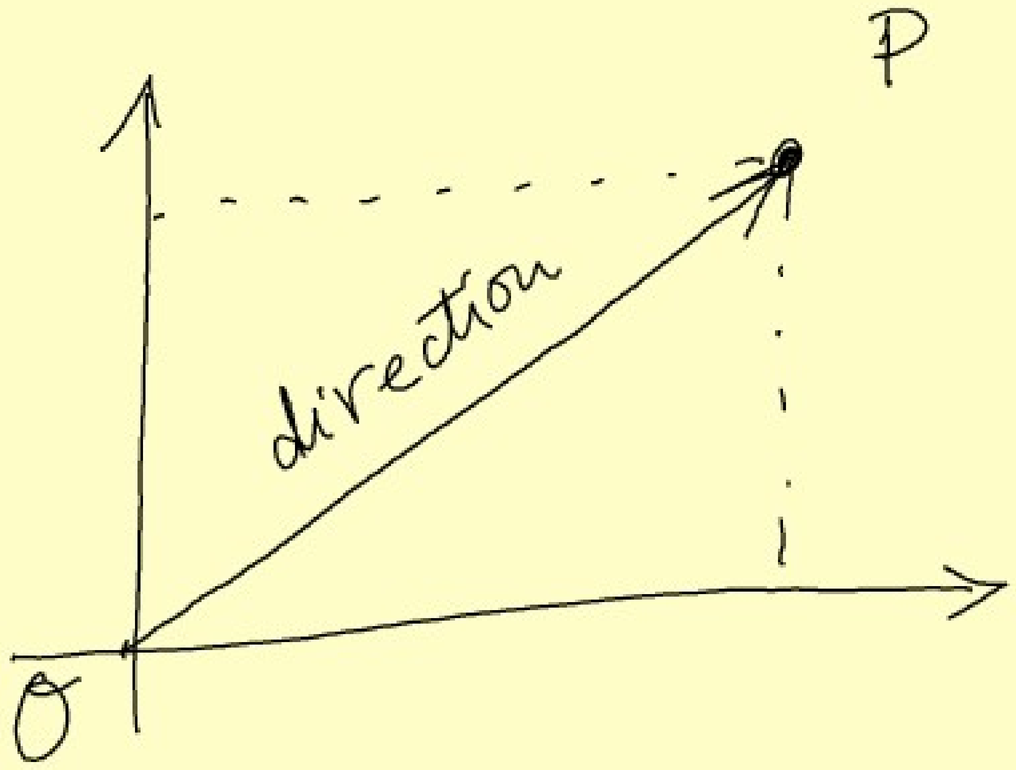
$$\begin{array}{l} x_1 (1, 0, 0, \dots, 0, 0) \\ x_2 (0, 1, 0, \dots, 0, 0) \\ \dots \\ x_{d-1} (0, \dots, 1, 0) \\ x_d (0, \dots, 0, 1) \\ \hline (x_1, x_2, \dots, x_{d-1}, x_d) \end{array}$$

$$x_i = x \cdot (0, \dots, \underset{i\text{-th}}{1}, 0, \dots 0)$$

$\uparrow$ HW



"direction" = "vector/arrow"




Any Point (away of O) defines a "direction" = "position vector"

---

## § 3.2. Higher order partial derivatives

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)$$

**Th 3.** (H. A. Schwarz) $f : \mathbb{R}^d \to \mathbb{R}$ admits continuous mixed second order partial derivatives (on a small ball around $x$) then

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x)$$

"order doesn't matter if cont."

$\frac{\partial^2}{\partial x_i \partial x_j}$ only partial info  Full info : Hesse matrix

$$H_f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right)_{i,j=\overline{1,d}}$$

## § 3.3. The Fréchet differential

- Linear functions of several variables

$T: \mathbb{R}^d \to \mathbb{R}$ is called linear if

(i) $T(x+y) = T(x) + T(y) \quad \forall x, y \in \mathbb{R}^d$

(ii) $T(\alpha x) = \alpha T(x) \quad \forall x \in \mathbb{R}^d, \forall \alpha \in \mathbb{R}$

((i) & (ii) $\Longleftrightarrow$ (iii) $T(\underbrace{\alpha x + \beta y}_{\text{lin. comb}}) = \underbrace{\alpha T(x) + \beta T(y)}_{\text{lin. comb of } T\text{s}}$ )

Example: $a = (a_1, \dots, a_d) \in \mathbb{R}^d$

$T_a(x) := a \cdot x$ is linear (i) & (ii) ✓ check this HW

Th 4. $\forall T: \mathbb{R}^d \to \mathbb{R}$ linear

there exists a unique $a_T \in \mathbb{R}^d$ such that

$$T(x) = a_T \cdot x.$$

"All linear functions are of the form $a \cdot x$"

(unsurprising if you think about $f(x) = ax$ in $\mathbb{R}$)

Prove this for 2+ bonus points (first correct proof)

~

- Quadratic functions

$Q : \mathbb{R}^d \to \mathbb{R}$ quadratic if

$$Q(x) = \sum_{i,j=1}^{d} a_{ij} x_i x_j , \quad a_{ij} = a_{ji} \quad i,j=\overline{1,d}$$
$$a_{ij} \in \mathbb{R}$$

$x = (x_1, \dots, x_d)$

$A = (a_{ij})_{i,j=\overline{1,d}}$ the matrix of the quadratic function

Examples:

$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\quad Q_1(x_1, x_2) = x_1^2 + x_2^2$

$A_2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ $\quad Q_2(x_1, x_2) = x_1^2 + x_1 x_2 + x_2 x_1 + x_2^2$
$= x_1^2 + 2x_1 x_2 + x_2^2$

If the $\nabla f$ is the $\mathbb{R}^d$ version of $f'$, how should we define higher order derivative equivalents in $\mathbb{R}^d$?

Def. $f : \mathbb{R}^d \to \mathbb{R}$ Fréchet diffable at $x$ if there exist a linear function $T : \mathbb{R}^n \to \mathbb{R}$ s.t.

$$\lim_{y \to x} \frac{|f(y) - f(x) - T(y-x)|}{\|y - x\|} = 0 .$$

Not $\underbrace{df(x)}_{T \text{ in def.}} (z)$

— Def $d^{k+1} f(x)(r) = \frac{d}{dt} \left(d^k f(x + tr)(r)\right)\Big|_{t=0}$