Supplement to the derivation of Lagranges method:

For a C^1-function $G:\mathbb{R}^n \to \mathbb{R}$, consider the $\boxed{\text{level sets } S_c = \{\vec{x} \in \mathbb{R}^n : G(\vec{x}) = c\}}$, for any $c \in \mathbb{R}$ (all $\vec{x}$'s where G has the constant value c; geometrically for $G:\mathbb{R}^2 \to \mathbb{R}$ this means all $\vec{x}$ where $G(\vec{x})$ has the same height w.r.t. the x-y-plane).

Now let $\vec{h}(t)$ be any curve in S_c ($\vec{h}(t) \in S_c \ \forall t \in \mathbb{R}$).

Claim: $\vec{\nabla} G(\vec{h}(t)) \cdot \frac{d\vec{h}}{dt} = 0$, i.e., $\boxed{\vec{\nabla} G \text{ is orthogonal to } S_c \text{ (at any point)}}$

Why? $G(\vec{h}(t)) = c \ \forall t \in \mathbb{R}$ by definition ($\vec{h}$ is a curve in S_c).

$$\Rightarrow \ 0 = \frac{d}{dt} G(\vec{h}(t)) = \vec{\nabla} G(\vec{h}(t)) \cdot \frac{d\vec{h}}{dt}$$

Example: see geogebra picture on the last page

$\hookrightarrow$ $G(x,y) = x^2 + y^3 + 2$, $c = 3$, i.e., the level set is the intersection of G with the $z=3$-plane

generally: $\vec{\nabla} G = \begin{pmatrix} 2x \\ 3y^2 \end{pmatrix}$

$\hookrightarrow$ e.g., the points $(1,0)$, $(0,1)$, and $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt[3]{2}}\right)$ $\overbrace{\text{are in the level set}}^{G=3 \text{ at all three points}}$,

there, we have $\vec{\nabla} G(1,0) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$, $\vec{\nabla} G(0,1) = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$, $\vec{\nabla} G\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt[3]{2}}\right) = \begin{pmatrix} 2/\sqrt{2} \\ 3/\sqrt[3]{4} \end{pmatrix}$

In the picture one can clearly see how $\vec{\nabla} G$ is orthogonal to the level set $G=3$ at these points.

Back to vector-valued functions...

Recall that for $\vec{f}: \mathbb{R} \to \mathbb{R}^n$ we define the derivative component-wise: $\frac{d\vec{f}(t)}{dt} = \begin{pmatrix} \frac{df_1(t)}{dt} \\ \vdots \\ \frac{df_n(t)}{dt} \end{pmatrix}$

Note: integration is also defined component-wise: $\int \vec{f}(t)\,dt := \begin{pmatrix} \int f_1(t)\,dt \\ \vdots \\ \int f_n(t)\,dt \end{pmatrix}$

Newton's law (mass $m=0$) force

Ex.: Conservation of angular momentum: let $\frac{d^2\vec{x}(t)}{dt^2} = \vec{F} = \underbrace{f(|\vec{x}|)}_{\text{e.g., Coulomb force } \vec{F}(\vec{x}) = \text{const.}\cdot\frac{\vec{x}}{|\vec{x}|^3}}\vec{x}$,

$\vec{x}(t)$ = particle position, $\vec{p}(t) = \frac{d\vec{x}(t)}{dt}$ momentum (mass = 1)

$$\Rightarrow \frac{d}{dt}\underbrace{\left(\vec{x}(t) \times \vec{p}(t)\right)}_{\text{cross product}} = \underbrace{\left(\underbrace{\frac{d\vec{x}}{dt}}_{=\vec{p}} \times \vec{p}\right)}_{\vec{p}\times\vec{p}=0 \text{ (cross product of parallel vectors)}} + \underbrace{\left(\vec{x} \times \overbrace{f(|\vec{x}|)\,\vec{x}}^{\frac{d\vec{p}}{dt}=\vec{F}}\right)}_{=0} = \vec{0}$$

i.e., $\vec{x}(t) \times \vec{p}(t)$ does not change in time

$\Rightarrow \vec{x}(t) \times \vec{p}(t) = \vec{x}(0) \times \vec{p}(0)$ is a constant of motion, it is called angular momentum

But in the most general case, we need to consider

$$\vec{f}: \mathbb{R}^n \to \mathbb{R}^m, \quad (x_1, \dots, x_n) \longmapsto \vec{f}(\vec{x}) = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix}$$

Ex.: $\vec{f}(x,y) = \begin{pmatrix} x^2 - y^4 - 4 \\ 2xy \end{pmatrix}$, see geogebra picture at the end of notes

First, some terminology:

- a fct. $f: \mathbb{R} \to \mathbb{R}^m$ $(m \geq 2)$ is called a curve in $\mathbb{R}^m$ (e.g., particle trajectory)
- a fct. $f: \mathbb{R}^n \to \mathbb{R}$ (usually $n=3$) is called a scalar field (e.g., temperature)
- a fct. $f: \mathbb{R}^n \to \mathbb{R}^m$ (usually $n=3$; $m \geq 2$) is called a vector field (e.g., electric field)

Next: What about derivatives of $\vec f : \mathbb{R}^n \to \mathbb{R}^m$?

Slightly abstract point of view:

- as usual, we need a linear approximation of $\vec f$ near $\vec a$

- $\boxed{\text{\underline{Def.}: A map } \vec L : \mathbb{R}^n \to \mathbb{R}^m \text{ is called linear if } \vec L(a\vec x + b\vec y) = a\vec L(\vec x) + b\,\vec L(\vec y) \text{ for all } \vec x, \vec y \in \mathbb{R}^n,\ a, b \in \mathbb{R}.}$

Writing $\vec x = \sum_{j=1}^{n} \overbrace{x_j}^{\text{j-th component of the vector } \vec x \text{ (in standard basis)}} \underbrace{\vec e_j}_{\text{unit vector in j-direction}}$, we find: $\underbrace{\big(\vec L(\vec x)\big)_i}_{\text{i-th component of the vector } \vec L(\vec x)} = \Big(\vec L\big(\sum_{j=1}^{n} x_j \vec e_j\big)\Big)_i \overset{\text{linearity}}{=} \sum_{j=1}^{n} x_j \underbrace{\big(\vec L(\vec e_j)\big)_i}_{=:\, \underbrace{L_i(e_j)}_{=\text{ i-th component of the vector } \vec L(e_j)}}$

$$= \sum_{j=1}^{n} \underbrace{L_i(\vec e_j)}_{\text{matrix}} x_j$$

$\Rightarrow$ $\boxed{\text{a linear map is represented by a matrix}}$

Thus, we know what differentiability means:

$\boxed{\begin{array}{l}\underline{\text{Definition:}} \text{ A function } \vec f : \mathbb{R}^n \to \mathbb{R}^m \text{ is called } \boxed{\text{(totally) differentiable}} \\ \text{at } \vec a \in \mathbb{R}^n \text{ if there is an } m \times n \text{ matrix } L_{\vec a} \text{ such that} \\ \qquad \lim_{\vec h \to 0} \dfrac{|\vec f(\vec a + \vec h) - \vec f(\vec a) - L_{\vec a}\vec h|}{|\vec h|} = 0. \\ L_{\vec a} = D\vec f(\vec a) = \vec f\,'(\vec a) \text{ is called } \boxed{\text{(total) derivative of } \vec f \text{ at } \vec a.}\end{array}}$

As before, we can figure out what $L_{\vec a}$ is:

e.g., choose $\vec h = \begin{pmatrix} h_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = h_1 \vec e_1$ with, say, $h_1 > 0$; then $\lim\limits_{h_1 \to 0} \dfrac{|\vec f(\vec a + \vec h) - \vec f(\vec a) - L_{\vec a}\vec h|}{h_1} = 0$

$$L_{\vec a}\vec h = \begin{pmatrix} L_{11} & \cdots & L_{1n} \\ \vdots & & \vdots \\ L_{m1} & \cdots & L_{mn} \end{pmatrix} \begin{pmatrix} h_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = h_1 \begin{pmatrix} L_{11} \\ \vdots \\ L_{m1} \end{pmatrix} \implies \lim_{h_1 \to 0} \frac{\vec f(\vec a + \vec h) - \vec f(\vec a) - h_1 \begin{pmatrix} L_{11} \\ \vdots \\ L_{m1} \end{pmatrix}}{h_1} = \vec 0,$$

recall def. of partial derivatives

i.e., $$\begin{pmatrix} L_{11} \\ \vdots \\ L_{m1} \end{pmatrix} = \lim_{h_1 \to 0} \frac{1}{h_1} \begin{pmatrix} f_1(\vec a + h_1 \vec e_1) - f_1(\vec a) \\ \vdots \\ f_m(\vec a + h_1 \vec e_1) - f_m(\vec a) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\vec a) \\ \vdots \\ \frac{\partial f_m}{\partial x_1}(\vec a) \end{pmatrix}$$

$\implies$ generally $L_{ij} = \dfrac{\partial f_i}{\partial x_j}(\vec a)$

<u>Theorem:</u> If $\vec f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $\vec a$, the derivative at $\vec a$ is the $m \times n$ matrix

$$L_{\vec a} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\vec a) & \cdots & \frac{\partial f_1}{\partial x_n}(\vec a) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\vec a) & \cdots & \frac{\partial f_m}{\partial x_n}(\vec a) \end{pmatrix}, \text{ i.e., } (L_{\vec a})_{ij} = \frac{\partial f_i}{\partial x_j}(\vec a).$$

$L_{\vec a} = J_{\vec f}(\vec a)$ is called $\boxed{\text{Jacobian matrix}}$ (of $\vec f$ at $\vec a$)

Ex.: $\vec f(x,y) = \begin{pmatrix} x^2 - y^4 - 4 \\ 2xy \end{pmatrix}$ from before

$$\implies J_{\vec f}(x,y) = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix} = \begin{pmatrix} 2x & -4y^3 \\ 2y & 2x \end{pmatrix}$$

Note: • This is the most general notion of a derivative, covering any $\vec{f}: \mathbb{R}^n \to \mathbb{R}^m$, in particular the cases $n=1$ or $m=1$ we discussed before.

• There is also a chain rule: If $g: \mathbb{R}^k \to \mathbb{R}^n$ and $f: \mathbb{R}^n \to \mathbb{R}^m$ are differentiable, then

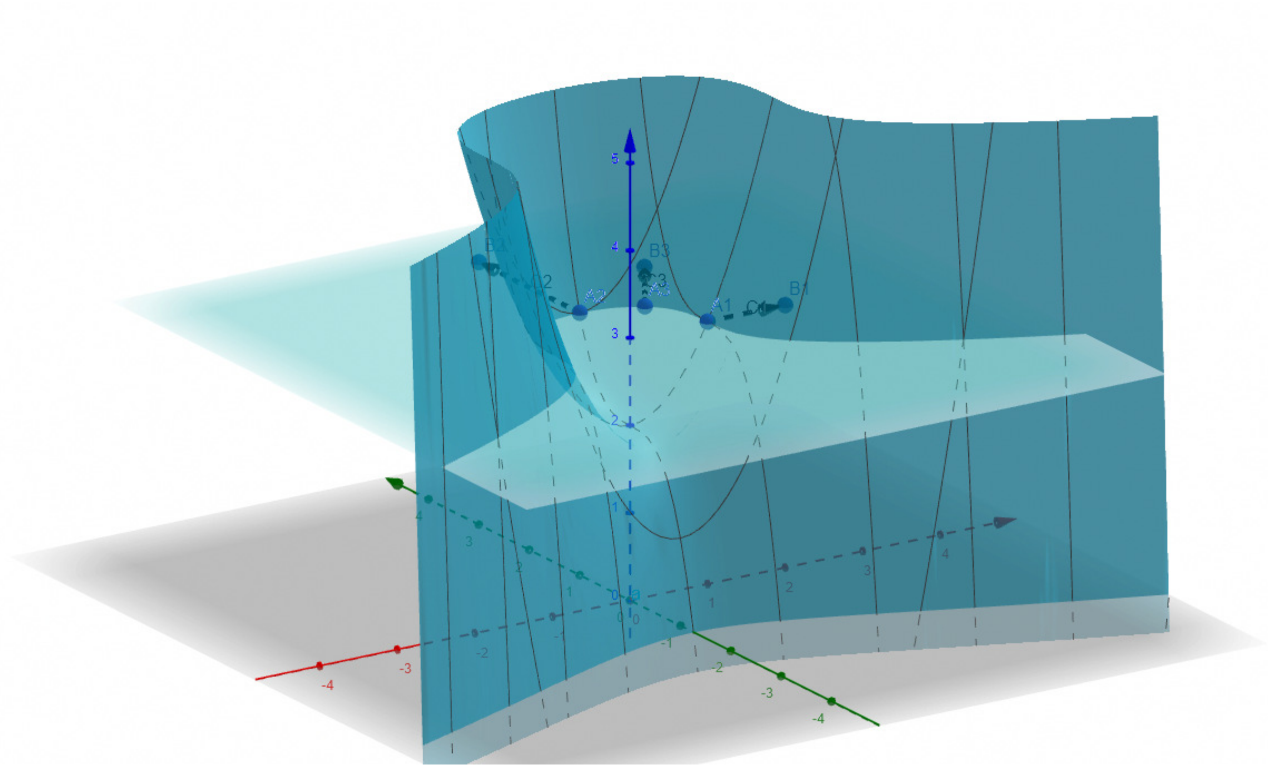
$$D(f(g(\vec{x}))) = \underbrace{\underbrace{Df(g(\vec{x}))}_{m\times n \text{ matrix}} \, \underbrace{Dg(\vec{x})}_{n\times k \text{ matrix}}}_{\text{matrix product} = m\times k \text{ matrix}} .$$

$\left(f\circ g: \mathbb{R}^k \to \mathbb{R}^m,\ \vec{x} \mapsto f(g(\vec{x})) \right)$

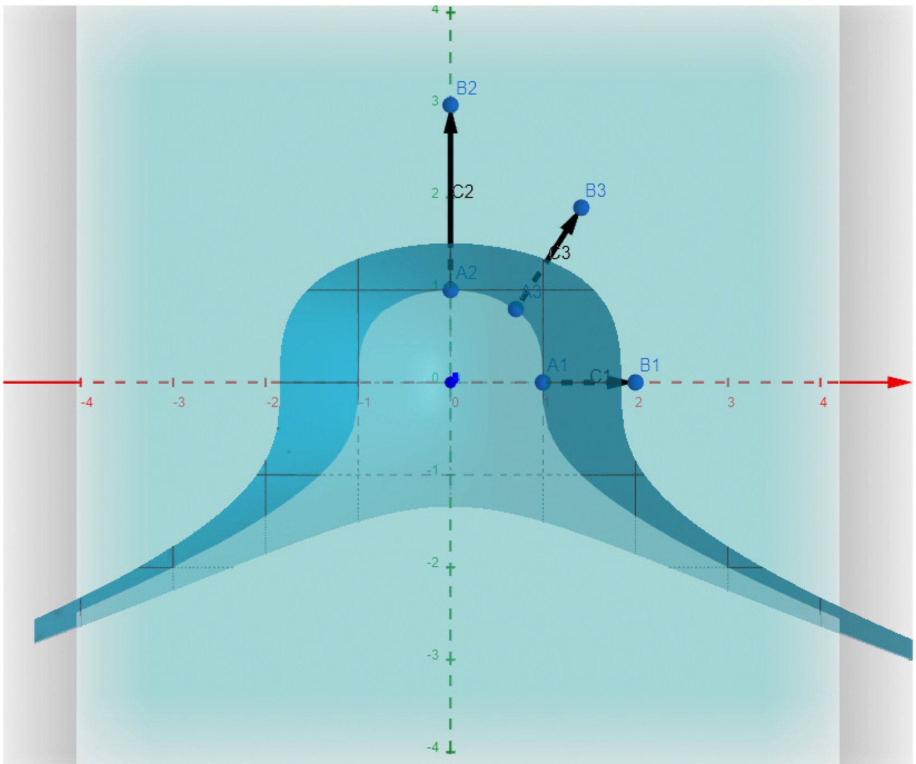
• As before, there is the theorem that says that if all $\frac{\partial f_i}{\partial x_j}$ exist and are continuous, then $\vec{f}$ is continuously differentiable.

Use https://www.geogebra.org/3d for generating the plots.

$G(x,y) = x^2 + y^3 + 2$, level set $G(x,y) = 3$, with three gradient vectors:



View from above to see orthogonality of nabla G and the level set:



Use https://www.geogebra.org/3d for generating the plots.

Vector field $f(x,y) = (x^2 - y^4 - 4, 2xy)$