$\partial_i \partial_j f = \partial_j \partial_i f$

3 Optimisation : étude au second ordre ordre

3.1 Matrice hessienne

Définition. Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction de classe $\mathcal{C}^2$ sur U ouvert. Pour $a \in U$, on définit la **matrice hessienne de f en a** :

$$H_f(a) = \left(\frac{\partial^2 f}{\partial x_j \partial x_i}(a) \right)_{1 \leqslant i,j \leqslant n}$$

Remarque. *Ainsi, pour $n = 2$:*

$$H_f(a) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

Proposition. Pour tout $a \in U$, $H_f(a) \in \mathcal{S}_n(\mathbb{R})$.

Preuve : Th de Schwarz.

$f : I \to \mathbb{R}$

$f(a+h) = f(a) + f'(a) h + \frac{1}{2} f''(a) h^2 + o(h^2)$

3.2 Formule de Taylor-Young à l'ordre 2

Théorème.

Soit $f : \mathbb{R}^n \to \mathbb{R}$ de classe $\mathcal{C}^2$ sur U ouvert, et $a \in U$. Alors :

$$
\begin{aligned}
f(a+h) &= f(a) + \langle \nabla f(a), h\rangle + \frac{1}{2}\langle H_f(a)h, h\rangle + \underset{h\to 0}{o}(\|h\|^2) \\
&= f(a) + \nabla f(a)^\top h + \frac{1}{2} h^\top H_f(a)\, h + \underset{h\to 0}{o}(\|h\|^2)
\end{aligned}
$$

$df(a).h$

Remarque. *Sous forme développée avec les dérivées partielles, l'expression s'écrit :*

$$f(a+h) = f(a) + \sum_{1\leqslant i\leqslant n} \frac{\partial f}{\partial x_i}(a)\, h_i + \frac{1}{2} \sum_{1\leqslant i,j\leqslant n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\, h_i h_j + \underset{h\to 0}{o}(\|h\|^2)$$

On reconnaît la différentielle de f en a dans le terme linéaire. Le terme suivant s'appelle « quadratique ».

$$\nabla f(a) = \begin{pmatrix} \partial_1 f(a) \\ \vdots \\ \partial_n f(a) \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{R}) \qquad h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}$$

$$\langle \nabla f(a), h \rangle = \nabla f(a)^T h$$

$$
\begin{aligned}
\langle H_f(a)h, h\rangle &= (H_f(a)h)^T h \\
&= h^T H_f(a)\, h
\end{aligned}
$$

${}^T X S X$ où $S \in S_n(\mathbb{R})$

Étude : a extrem ? $\longrightarrow \nabla f(a) = 0$

$$f(a+h) = f(a) + \cancel{\langle \nabla f(a), h\rangle} + \frac{1}{2}\langle H_f(a)h, h\rangle + o(h^2)$$

$\forall h$ $f(a+h) \geqslant f(a)$?

$f(a+h) \leqslant f(a)$?

$$f(a+h) - f(a) \text{ “}\sim\text{” } \frac{1}{2}\underbrace{\langle H_f(a)h, h\rangle}_{\text{signe ?}}$$

3.3 Conditions du second ordre

Théorème.

> Soit $f : \mathbb{R}^n \to \mathbb{R}$ de classe $\mathcal{C}^2$ sur U ouvert, et $a \in U$.
> Si :
>
> - f admet un minimum local en a
>
> alors :
>
> - $\nabla f(a) = 0$
> - $H_f(a) \in \mathcal{S}_n^+(\mathbb{R})$

Remarque. *De façon équivalente, on peut conclure que les valeurs propres de $H_f(a)$ sont $\geqslant 0$.*

Remarque. *On peut adapter le résultat dans le cas d'un maximum : les valeurs propres de $H_f(a)$ sont $\leqslant 0$.*

Théorème.

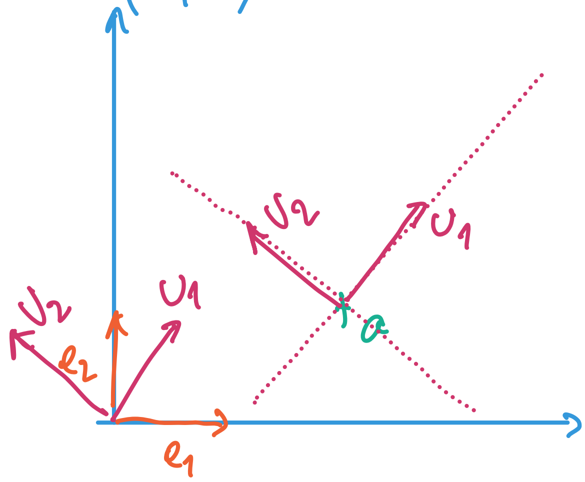
> Soit $f : \mathbb{R}^n \to \mathbb{R}$ de classe $\mathcal{C}^2$ sur U ouvert, et $a \in U$.
> Si :
>
> - a est un point critique de f
> - $H_f(a) \in \mathcal{S}_n^{++}(\mathbb{R})$ ie $\forall X \neq 0 \quad X^T H_f(a) X > 0$ ou $Sp(H_f(a)) \subset \mathbb{R}_+^*$
>
> alors :
>
> - f admet un minimum local strict en a

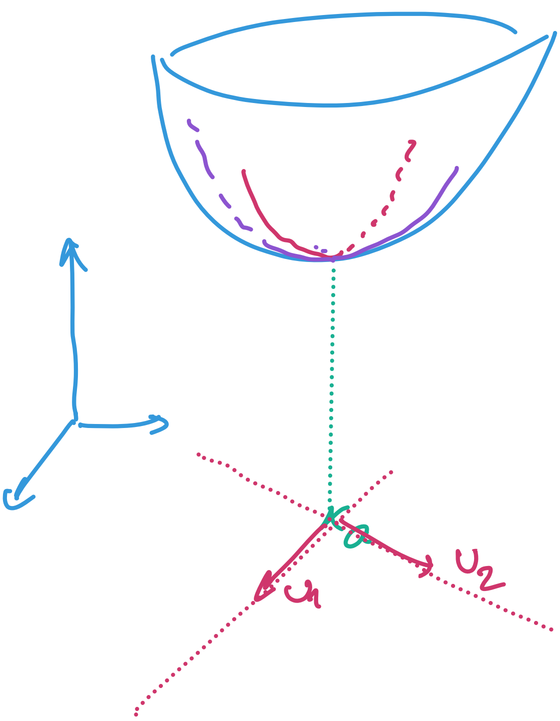
Remarque. *De façon équivalente, on peut supposer que les valeurs propres de $H_f(a)$ sont > 0.*

Remarque. *On peut adapter le résultat dans le cas d'un maximum : les valeurs propres de $H_f(a)$ sont < 0.*

$$H_f(a) = P\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} P^T$$

$(U_1 | U_2)$





$t \to 0$

fct d'1 variable

$$f(a + tU_1) = f(a) + \underbrace{\cancel{\langle \nabla f(a), tU_1 \rangle}}_{=0} + \frac{1}{2}\langle H_f(a)(tU_1), tU_1 \rangle + o(\|tU_1\|^2)$$

$$= f(a) + \frac{1}{2} t^2 \lambda_1 \|U_1\|^2 + o(t^2 \|U_1\|^2)$$

$$= f(a) + \frac{1}{2} t^2 \underbrace{\lambda_1 \|U_1\|^2}_{>0} + o(t^2)$$

au voisinage de $t \to 0$, $f(a+tU_1) > f(a)$

Dans la direction de U_2, idem.

Preuve: On suppose $H_f(a) \in S_n^{++}(\mathbb{R})$

f admet un point critique en a.

Mq f admet un minimum en a.

$$f(a+h) - f(a) = \frac{1}{2} h^T H_f(a) h + o(\|h\|^2)$$

$H_f(a) \in S_n^{++}$ donc $\langle x, y \rangle = x^T H_f(a) y$

est un produit scalaire.

donc $h^T H_f(a) h$ est la norme (au carré)

euclidienne associée.

on peut noter $h^T H_f(a) h = \|h\|_2^2$

$$f(a+h) - f(a) = \frac{1}{2} \|h\|_2^2 + o(\underbrace{\|h\|^2}_{\text{par équivalence}})$$

$$+ o(\|h\|_2^2)$$

$$\underset{h\to 0}{\sim} \frac{1}{2}\|h\|_2^2$$

$$\underset{h\to 0}{\sim} \frac{1}{2} h^T H_f(a)\, h$$

$$> 0$$

donc f admet un minimum en a.

<u>Remarque:</u>

notation qui n'existe pas

et si $H_f(a) \in S_n^{--}(\mathbb{R})$

ie $-H_f(a) \in S_n^{++}(\mathbb{R})$

$\Big($ ie $\forall X \neq 0 \quad X^T H_f(a) X < 0$

<u>ou</u> $\quad Sp(H_f(a)) \subset \mathbb{R}_-^* \Big)$

$\longrightarrow$ maximum local.

<u>On continue</u>

On suppose a point critique de f

$$H_f(a) \notin S_n^+(\mathbb{R})$$

donc $\exists \lambda < 0$ tq λ vp de $H_f(a)$

donc $\exists v$ vecteur propre associé.

On a alors:

$$f(a+tv) - f(a) = 0 + \frac{1}{2}\langle H_f(a)(tv), tv\rangle + o(t^2)$$

$$= \frac{1}{2} t^2 \langle \lambda v, v \rangle + o(t^2)$$

$$= \frac{1}{2} \lambda t^2 \|v\|^2 + o(t^2)$$

$$\sim \frac{1}{2} \lambda t^2 \|v\|^2$$

$$< 0 \quad \text{au voisinage de } t \to 0$$

donc $f(a)$ n'est pas un minimum.

<u>Bref</u> : Si que des vp > 0 $\longrightarrow$ minimum

Si que des vp < 0 $\longrightarrow$ maximum

Si vp < 0 et vp > 0 $\longrightarrow$ point col.

Si vp > 0 et vp $= 0$ $\longrightarrow$ on ne peut pas conclure

Si vp < 0 et vp $= 0$ $\longrightarrow$ "

Corollaire. Dans le cas où $n = 2$, on note :

$$H_f(a) = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$$

où $r = \dfrac{\partial^2 f}{\partial x^2}(a)$, $s = \dfrac{\partial^2 f}{\partial x \partial y}(a) = \dfrac{\partial^2 f}{\partial y \partial x}(a)$ et $t = \dfrac{\partial^2 f}{\partial y^2}(a)$.

On suppose que a est un point critique : $\underset{p=}{\dfrac{\partial f}{\partial x}}(a) = \underset{q=}{\dfrac{\partial f}{\partial y}}(a) = 0$.

- Si $\det H_f(a) = rt - s^2 > 0$ et $\operatorname{tr} H_f(a) = r + t > 0$, alors $H_f(a) \in \mathbb{S}_n^{++}(\mathbb{R})$
  donc f présente un minimum local strict en a.
- Si $\det H_f(a) = rt - s^2 > 0$ et $\operatorname{tr} H_f(a) = r + t < 0$, alors $-H_f(a) \in \mathbb{S}_n^{++}(\mathbb{R})$
  donc f présente un maximum local strict en a.
- Si $\det H_f(a) = rt - s^2 < 0$, f présente un point col en a.
- Si $\det H_f(a) = rt - s^2 = 0$, on ne peut rien dire.

uniquement si $n=2$

$$H_f(a) = P \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} P^T$$

$$\det H_f(a) = \lambda\mu$$

$$\operatorname{tr}(H_f(a)) = \lambda + \mu$$

Calcul différentiel

$$f(x,y) = g(r,\theta)$$

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$$ à exprimer à l'aide de g

Exemple. Dans le plan euclidien usuel, exprimer le gradient en coordonnées polaires.

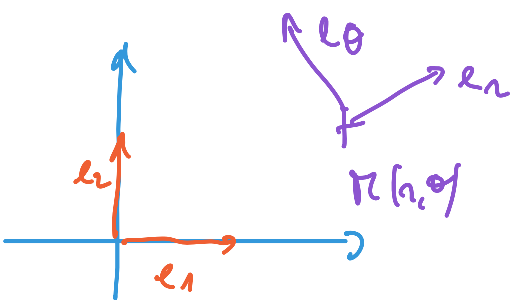
et le laplacien $\Delta f = \dfrac{\partial^2 f}{\partial x^2} + \dfrac{\partial^2 f}{\partial y^2}$

- $g(r,\theta) = f(r\cos\theta, r\sin\theta)$

donc
$$\begin{cases} \dfrac{\partial g}{\partial r} = \cos\theta \dfrac{\partial f}{\partial x} + \sin\theta \dfrac{\partial f}{\partial y} \\[2ex] \dfrac{\partial g}{\partial \theta} = -r\sin\theta \dfrac{\partial f}{\partial x} + r\cos\theta \dfrac{\partial f}{\partial y} \quad \times \dfrac{1}{r} \end{cases}$$

$\sin\theta\, L_1 + \frac{1}{r}\cos\theta\, L_2$ $$\frac{\partial f}{\partial y} = \sin\theta \frac{\partial g}{\partial r} + \frac{1}{r}\cos\theta \frac{\partial g}{\partial \theta}$$

$\cos\theta\, L_1 - \frac{1}{r}\sin\theta\, L_2$ $$\frac{\partial f}{\partial x} = \cos\theta \frac{\partial g}{\partial r} - \frac{1}{r}\sin\theta \frac{\partial g}{\partial \theta}$$



$$e_r = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$$

$$e_\theta = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$$

$$\overrightarrow{\nabla f} = \frac{\partial f}{\partial x} e_1 + \frac{\partial f}{\partial y} e_2$$

$$= \left(\cos\theta \frac{\partial g}{\partial r} - \frac{1}{r}\sin\theta \frac{\partial g}{\partial \theta}\right) e_1$$

$$+ \left(\sin \frac{\partial g}{\partial r} + \frac{1}{r}\cos\theta \frac{\partial g}{\partial \theta}\right) e_2$$

$$= (\underbrace{\cos\theta e_1 + \sin\theta e_2}_{e_r}) \frac{\partial g}{\partial r}$$

$$+ \frac{1}{r}(\underbrace{-\sin\theta e_1 + \cos\theta e_2}_{e_\theta}) \frac{\partial g}{\partial \theta}$$

$$= \frac{\partial g}{\partial r}\vec{e_r} + \frac{1}{r}\frac{\partial g}{\partial \theta}\vec{e_\theta}$$

$$\boxed{(r\cos\theta, r\sin\theta)}$$

<u>Laplacien:</u>

$$\begin{cases} \dfrac{\partial g}{\partial r} = \cos\theta \dfrac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) + \sin\theta \dfrac{\partial f}{\partial y}(r\cos\theta, r\sin\theta) \\ \dfrac{\partial g}{\partial \theta} = -r\sin\theta \dfrac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) + r\cos\theta \dfrac{\partial f}{\partial y}(r\cos\theta, r\sin\theta) \end{cases}$$

$$\frac{\partial^2 g}{\partial r^2} = 4 \text{ termes}$$

$$= \cos^2\theta \frac{\partial^2 f}{\partial x^2} + 2\cos\theta\sin\theta \frac{\partial^2 f}{\partial x\partial y} + \sin^2\theta \frac{\partial^2 f}{\partial y^2}$$

$$\frac{1}{r^2} \times \frac{\partial^2 g}{\partial \theta^2} = -r\cos\theta \frac{\partial f}{\partial x} - r\sin\theta \frac{\partial f}{\partial y}$$

$$+ r^2\sin^2\theta \frac{\partial^2 f}{\partial x^2} + r^2\cos^2\theta \frac{\partial^2 f}{\partial y^2}$$

$$+ \left(-r^2\sin\theta\cos\theta - r^2\cos\theta\sin\theta\right)\frac{\partial^2 f}{\partial x\partial y}$$

$$\frac{\partial^2 g}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 g}{\partial \theta^2} = \Delta f - \frac{1}{r}\cos\theta\,\frac{\partial h}{\partial x} - \frac{1}{r}\sin\theta\,\frac{\partial h}{\partial y}$$

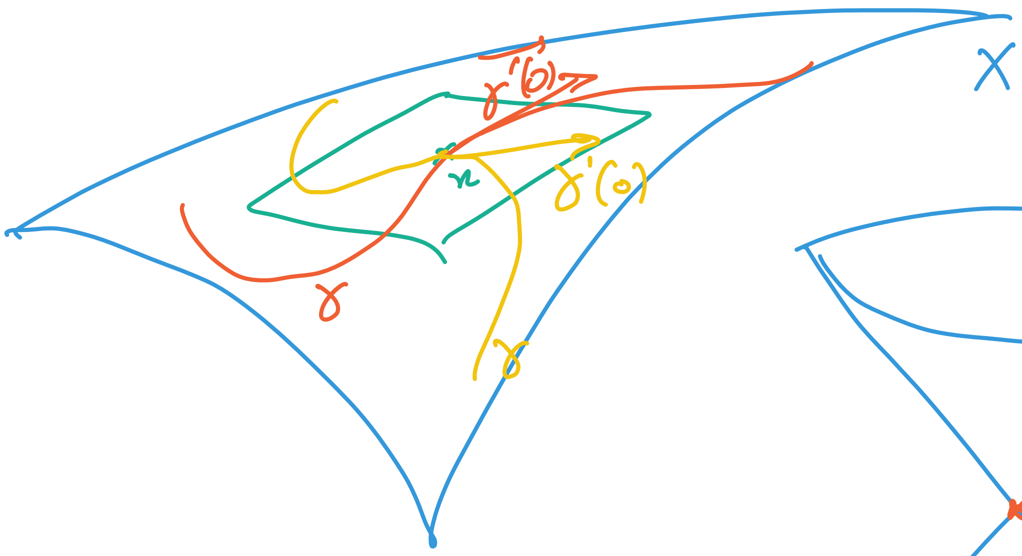
$$= \Delta f - \frac{1}{r}\frac{\partial g}{\partial r}$$

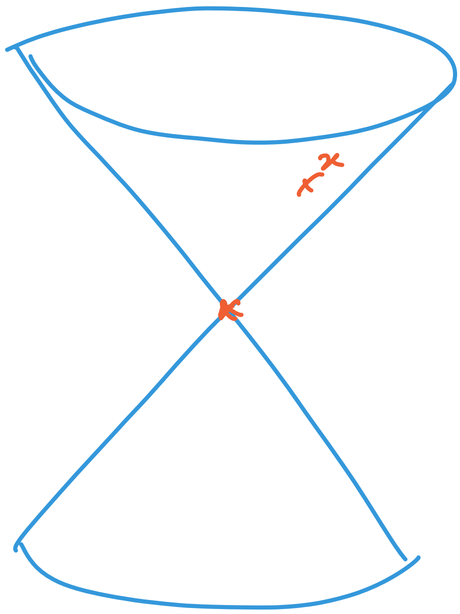
5 Vecteurs tangents à une partie d'un espace normé de dimension finie

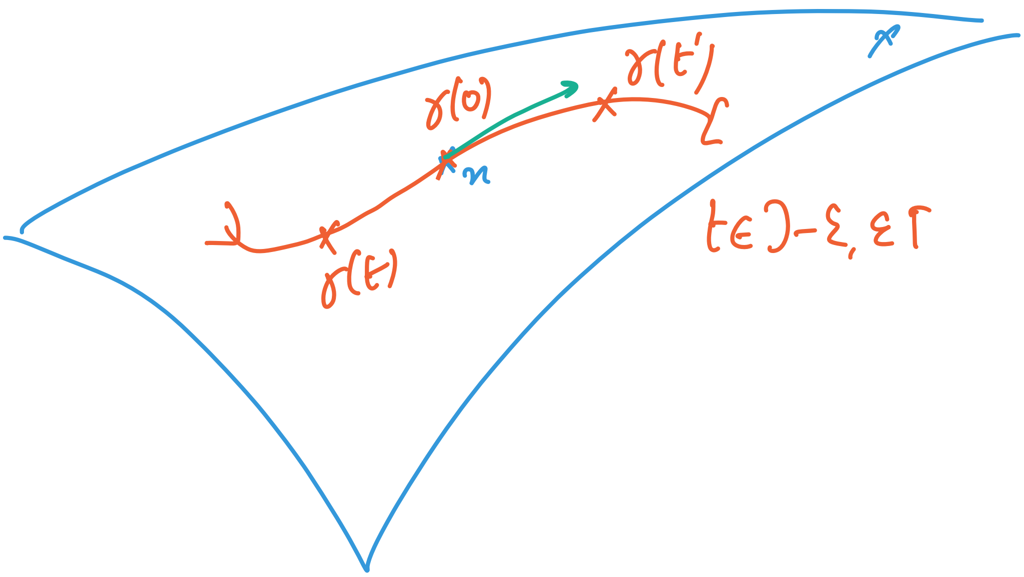
Définition. Soit X une partie non vide de E et $x \in X$.

On dit qu'un vecteur $v \in E$ est **tangent à** X **en** x lorsqu'il existe $\varepsilon > 0$ et un arc $\gamma :]-\varepsilon, \varepsilon[\to E$, à valeurs dans X, dérivable en 0, et tel que $\gamma(0) = x$ et $\gamma'(0) = v$.

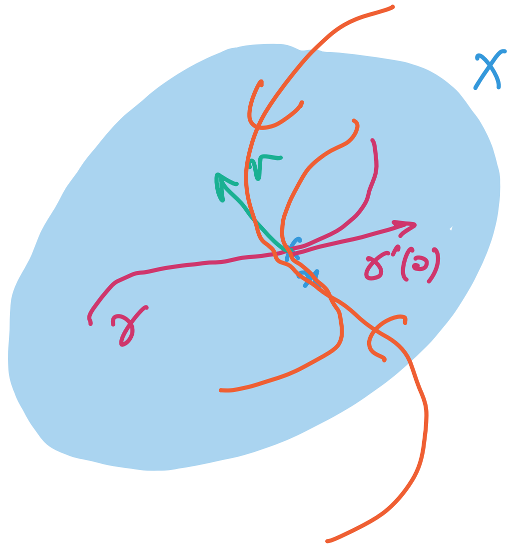
On note $T_x X$ l'ensemble des vecteurs tangents à X en x.







Exemple. Pour X est un ouvert de E et $x \in X$, déterminer $T_x X$.



Mq que $T_x X = E$

$\boxed{\subset}$

$\boxed{\supset}$ Soit $v \in E$

X ouvert donc

$\exists \eta > 0$ tq $B(x, \eta) \subset X$

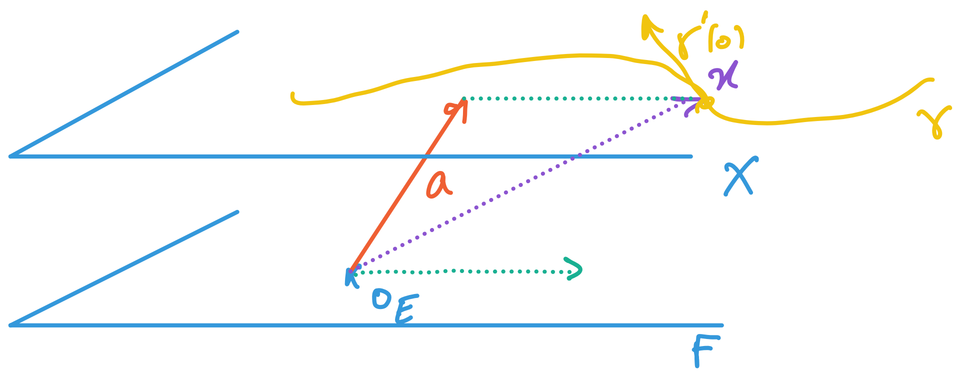
On note $\gamma : \left]-\frac{\eta}{\|v\|}, \frac{\eta}{\|v\|}\right[\to E$

$t \longmapsto x + tv$

- $\forall t$, $\gamma(t) \in B(x, \eta) \subset X$
- $\gamma(0) = x$
- $\gamma'(t) = v$ donc $\gamma'(0) = v$

donc $v \in T_x X$

Exemple. Pour $X = a + F$ sous-espace affine de E et $x \in X$, déterminer $T_x X$.



Mq $T_x X = F$

$\boxed{\supset}$ Soit $v \in F$

On note $\gamma(t) = x + tv \quad \forall t \in \mathbb{R}$

$\gamma(0) = x$

$\gamma'(t) = v$ donc $\gamma'(0) = v$

donc $v \in T_x X$

$\boxed{\subset}$ Soit $v \in T_x X$

donc $\exists \gamma :]-\varepsilon, \varepsilon[\longrightarrow E$

$t \longmapsto \gamma(t)$

tq $\gamma(0) = x$ et $\gamma'(0) = v$

$\frac{1}{t}\left(\gamma(t) - \gamma(0)\right) \in F$ (car $\gamma(t), \gamma(0) \in X$)

$\downarrow t \to 0$

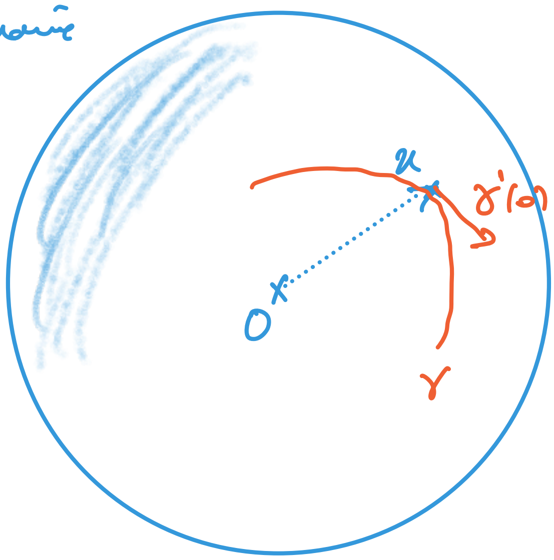
$\gamma'(0) \in F$ car F fermé, car ev de dim finie

donc $v \in F$.

Exemple. Lorsque E est euclidien, $X = S(0,1)$ la sphère unité et $x \in X$, déterminer $T_x X$.

$\mathcal{B} = (e_1 \dots e_n)$ base orthonormée

$S(0,1) : \sum_{i=1}^{n} x_i^2 = 1$



Mq $\underline{T_x X = \mathrm{Vect}(x)^\perp}$

$\boxed{\subset}$ Soit $v \in T_x X$

ie $\exists \gamma :]-\varepsilon, \varepsilon[\longrightarrow E$ tq

$\gamma(0) = x$

$\gamma'(0) = v$

$\forall t \ \gamma(t) \in S(0,1)$

donc : $\forall t \quad \|\gamma(t)\|^2 = 1$

ie $\quad \langle \gamma(t), \gamma(t) \rangle = 1$

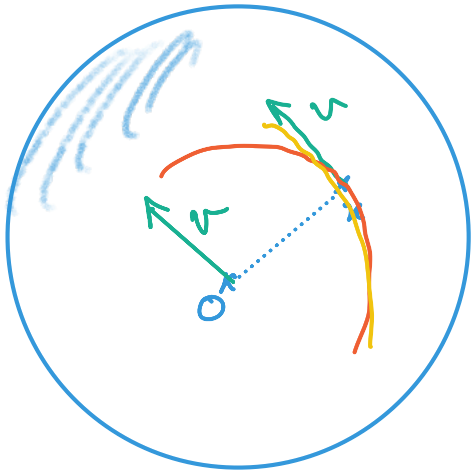
On dérive : $\quad 2 \langle \gamma'(t), \gamma(t) \rangle = 0$

en part $\langle \gamma'(0), \gamma(0) \rangle = 0$

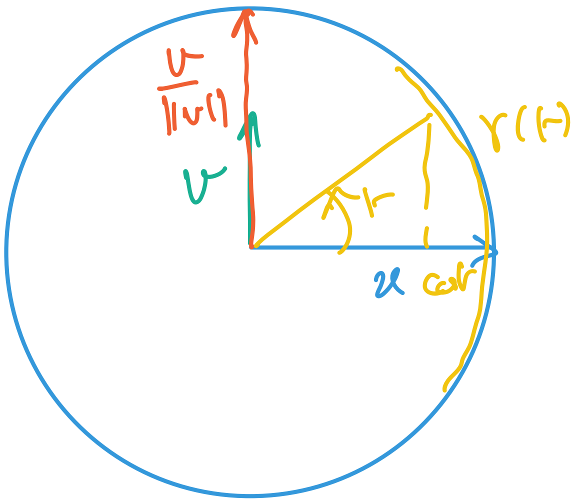
donc $\quad \langle v, x \rangle = 0$

donc $\quad \underline{v \in \mathrm{Vect}(x)^\perp}$

$\boxed{2}$ Soit $v \in \mathrm{Vect}(x)^{\perp}$



On refait la figure dans le plan v, x.



On pose : $\gamma(t) = \cos t \cdot \vec{x} + \sin t \dfrac{\vec{v}}{\|v\|}$

$$\|\gamma(t)\|^2 = \|\cos t\, \vec{x}\|^2 + \left\|\sin t \frac{v}{\|v\|}\right\|^2 \quad \text{car } x \perp v$$

$$= \cos^2 t + \sin^2 t \quad \text{par le th de Pythagore}$$

$$= 1$$

donc γ est tracée sur $S(0,1)$

$\gamma(0) = x$

$\gamma'(t) = -\sin t \, x + \cos t \, \dfrac{v}{\|v\|}$

donc $\gamma'(0) = \dfrac{v}{\|v\|}$

c'est bon.

Exemple. Pour $E = \mathbb{R}^3$, X graphe d'une fonction numérique $f : \mathbb{R}^2 \to \mathbb{R}$ définie sur un ouvert U et $x \in X$, montrer que $T_x X$ est un plan vectoriel.

Exemple. Déterminer l'ensemble des vecteurs tangents à $\mathrm{SO}_n(\mathbb{R})$ en I_n.

Théorème.

Soit $g:\ E \to \mathbb{R}$ une fonction numérique de classe $\mathcal{C}^1$ sur un ouvert U. On considère l'ensemble :

$$X = \{x \in U,\ g(x) = 0\}$$

Si $x \in X$ et $\mathrm{d}g(x) \neq 0_{\mathcal{L}(E,\mathbb{R})}$, alors :

$$T_x X = \mathrm{Ker}\big(\mathrm{d}g(x)\big)$$

C'est un hyperplan, comme noyau d'une forme linéaire non nulle.

Preuve. Démonstration hors programme. □

Corollaire. Lorsque E est euclidien, la différentielle peut être représentée par le gradient :
Si $x \in X$ et $\nabla g(x) \neq 0_E$, alors :

$$T_x X = \big(\nabla g(x)\big)^{\perp}$$

C'est un hyperplan, et $\nabla g(x)$ en est un vecteur orthogonal.

Corollaire. Lorsque $E = \mathbb{R}^3$ et :

$$X = \{(x, y, z),\ g(x, y, z) = 0\}$$

alors, pour $m \in X$, si $\nabla(g)(m) \neq (0, 0, 0)$, l'ensemble $T_m X$ des vecteurs tangents à X en m est l'hyperplan d'équation :

$$\frac{\partial g}{\partial x}(m)x + \frac{\partial g}{\partial y}(m)y + \frac{\partial g}{\partial z}(m)z = 0$$

Exemple. Dans $\mathbb{R}^2$, déterminer l'espace tangent à un point l'ensemble :

$$X = \{(x, y),\ x^2 + y^2 - 1 = 0\}$$

Exemple. Dans $\mathbb{R}^2$, déterminer l'espace tangent à un point l'ensemble :

$$X = \{(x, y),\ y^2 - 4(1 - x^2)x^2 = 0\}$$

Exemple. Dans $\mathbb{R}^3$, déterminer l'espace tangent à un point l'ensemble :

$$X = \{(x, y, z),\ x^2 + y^2 - z^2 + 1 = 0\}$$