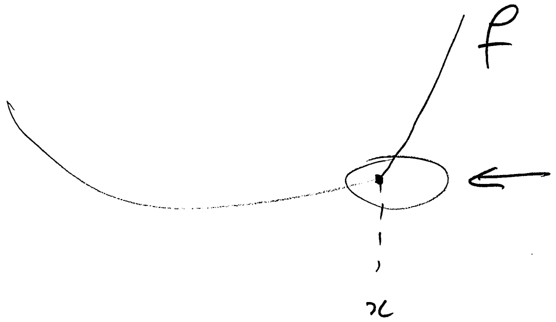
Conjugaison

1) Motivation

- On rencontre beaucoup de fonctions non différentiables (au sens de Fréchet) en optimisation et on aimerait introduire une notion de différentiabilité plus faible pour pouvoir décrire le comportement de la fonction autour du point considéré.

- Ce n'est pas trop difficile si la fonction $f : \mathbb{E} \to \overline{\mathbb{R}}$ est <u>convexe</u> (extensions possibles et très utilisée lorsque la fonction est localement lipschitzienne)



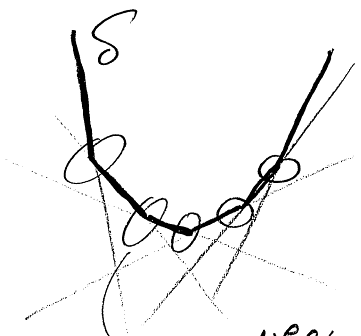
<u>but</u> : décrire comment varie f autour de x

- <u>Exemples de fonctions convexes non différentiables</u>

(1) <u>En dualité lagrangienne</u>

$$(P) \quad \inf_{\substack{x\in X \\ x \in \mathbb{R}^n}} \ \sup_{\substack{\lambda\in\Lambda \\ \lambda \geq 0}} \ \underset{f(x) + \lambda^T c(x)}{\ell(x,\lambda)} \quad \equiv (P) \begin{cases} \min f(x) \\ c(x) \leq 0 \end{cases}$$

$$(D) \quad \sup_{\lambda\in\Lambda} \ \inf_{x\in X} \ \ell(x,\lambda) = - \inf_{\lambda\in\Lambda} \underbrace{\sup_{x\in X} -\ell(x,\lambda)}_{\delta(\lambda) \text{ fonction duale}}$$



$$\delta(\lambda) = \underbrace{\sup_{x\in X} -\ell(x,\lambda)}_{\text{enveloppe supérieure de fonctions affines}} = -f(x) - \lambda^T c(x)$$

$\Rightarrow$ c'est une fonction <u>convexe</u> <u>fermée</u> souvent <u>non</u> différentiable

$\Rightarrow$ il faut minimiser δ qui est convexe fermée <u>non</u> différentiable : comment faire cela ?

② Etude de la fonction valeur

$$(P)\ \begin{cases} \inf f(x) \\ c(x) \leq 0 \end{cases} \qquad f : E \to \mathbb{R}, \quad c : E \to \mathbb{R}^m$$

La fonction valeur $v : \mathbb{R}^m \to \overline{\mathbb{R}}$ est définie en $p \in \mathbb{R}^m$ par

$$\boxed{v(p) = \inf_{\substack{x \in E \\ c(x) + p \leq 0}} f(x)}$$

Si f est convexe et les c_i sont convexes, alors v est convexe, mais en général non différentiable

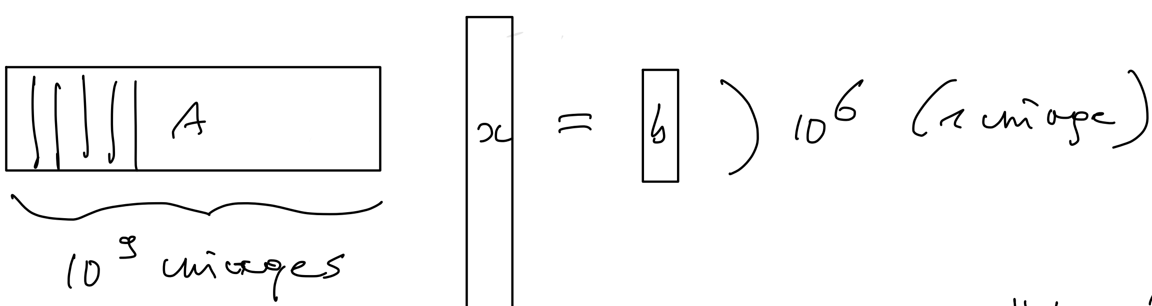
On peut montrer que

$$\boxed{\partial v(0) = \Lambda_*}$$

→ ensemble des multiplicateurs optimaux de (P)

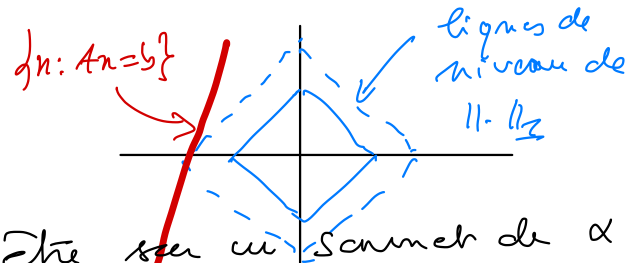
→ le sous-différentiel de v en 0

③ Données massives (recouvrement ℓ_1, poursuite de base)



On cherche à représenter l'image b dans la "base" des colonnes de A (des images) avec un x ayant le plus de 0 possible (c'est lui l'image "compressée")

$$\longrightarrow \begin{cases} \min \|x\|_1 \\ Ax = b \end{cases}$$



les solutions a tendance à être sur un sommet de $\alpha \overline{B}_1$ donc avec beaucoup de zéros.

Conclusion

- En optimisation, on rencontre beaucoup de fonctions convexes non différentiables.
- But des 2 séances suivantes : définir une notion de différentiabilité pour ces fonctions (utile pour les CO, algos, ...)
- séance 8 : conjugaison (1 outil très utile)

  séance 9 : sous-différentiabilité

2) Enveloppe convexe fermée

A Définitions

$\mathbb{E}$: espace vectoriel (dim finie), $P \subset \mathbb{E}$

Enveloppe convexe de P

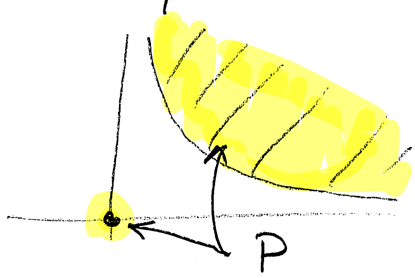
= co P (notation)

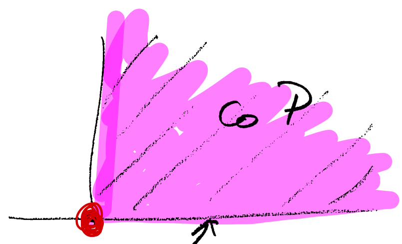
= le plus petit convexe contenant P (sens ?)

= $\bigcap$ de tous les convexes contenant P
(donc c'est un convexe)

- P fermé $\not\Rightarrow$ co P fermé

ex :





Enveloppe convexe fermée de P

= $\overline{\text{co}}\, P$ (notation)

= le plus petit convexe fermé contenant P

= $\bigcap$ de tous les convexes fermés contenant P
(donc c'est un convexe fermé)

- on peut montrer que

$$\boxed{\overline{\text{co}}\, P = \overline{\text{co}\, P}}$$

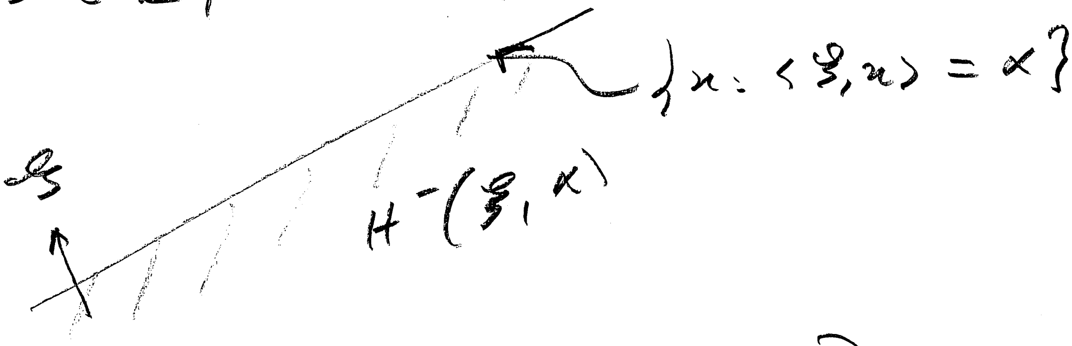
mais ce n'est pas ce qui nous intéresse ici et il est préférable de définir $\overline{\text{co}}\, P$ par une propriété de minimalité.

$\boxed{B}$ Description extérieure d'un convexe

- Soit $\langle \cdot, \cdot \rangle$ un produit scalaire sur $\mathbb{E}$
- Un <u>demi-espace fermé</u> est une partie de $\mathbb{E}$ de la forme

$$H^-(\xi, \alpha) := \{ x \in \mathbb{E} : \langle \xi, x \rangle \leq \alpha \}$$

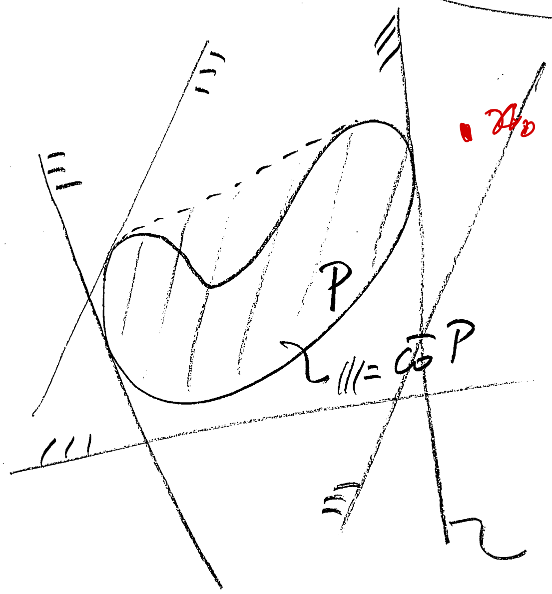
où $\xi \in \mathbb{E}$, $\xi \neq 0$, $\alpha \in \mathbb{R}$



- <u>prop</u> (fondamentale)

$$\boxed{\overline{co}\, P = \underbrace{\bigcap \text{ de tous les demi-espaces fermés contenant } P}_{D}}$$

donc il n'est pas néc. de prendre tous les convexes fermés pour obtenir $\overline{co}\, P$, il suffit de prendre les demi-espaces fermés



on $\bigcap$ tous les $H^-(\xi, \alpha)$ qui contiennent P.

Dém ⊂ clair

⊃ (par l'absurde)

Si ce n'est pas vrai : $D \not\subset \overline{co}\, P$

$\Rightarrow \exists\, x_0 \in D \setminus \overline{co}\, P$

$\underbrace{\{x_0\}}_{\text{convexe compact}} \cap \underbrace{\overline{co}\, P}_{\text{convexe fermé}} = \emptyset$

HB $\Rightarrow \exists\, \xi \in E$:

$\underbrace{\sup_{x \in \overline{co}\, P} \langle \xi, x \rangle =: \alpha}$ $<$ $\underbrace{\langle \xi, x_0 \rangle}$

$x_0 \notin H^-(\xi, \alpha)$

$\Downarrow$

$P \subset H^-(\xi, \alpha)$

$\hookrightarrow$ $H^-(\xi, \alpha)$ est un $\frac{1}{2}$-espace fermé qui contient P

$D \subset H^-(\xi, \alpha)$

$x_0 \in D$ $\qquad$ $x_0 \notin$

Contradiction

$\square$

Corol

P convexe fermé (i)

$\Longleftrightarrow \quad P = \overline{co}\ P$ (ii)

$\Longleftrightarrow \quad P = \bigcap$ des demi-espaces fermés qui contiennent P (iii)

Dém

(i) $\Rightarrow$ (ii)

$\underset{\text{toujours}}{P \subset} \overline{co}\ P \underset{\text{car } P \text{ est un convexe fermé qui contient } P}{\subset} P$

$\Longrightarrow$ = partout

(i) $\Leftarrow$ (ii) car $\overline{co}\ P$ est un convexe fermé

(ii) $\Leftrightarrow$ (iii) par la proposition

$\square$

3) Conjuguée de Fenchel

$\boxed{A}$ Motivation

- $\mathbb{E}$ e.v. euclidien (p.s. $\langle \cdot , \cdot \rangle$)

  $f : \mathbb{E} \to \overline{\mathbb{R}}$ non néc. convexe

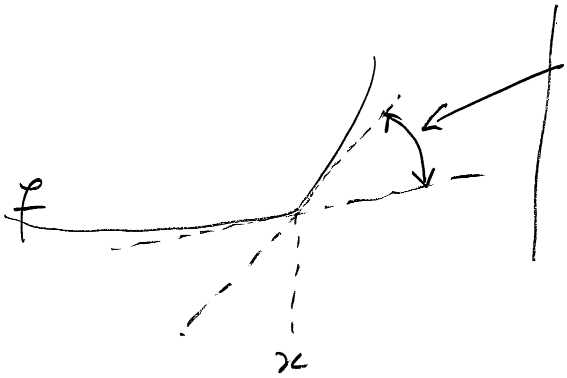
- Une minorante affine de f est une fonction affine $a : \mathbb{E} \to \mathbb{R}$ qui minore f :

$$\forall x \in \mathbb{E} : \quad a(x) \leqslant f(x) .$$

  Elle est dite exacte en x_0 si, de plus,

$$a(x_0) = f(x_0)$$

- Notre motivation principale pour introduire la conjugaison vient de ce que c'est un concept utile pour calculer le sous-différentiel $\partial f(x)$ de f en x :



$\partial f(x) :=$ ensemble des pentes des minorantes affines de f exactes en x

Donc

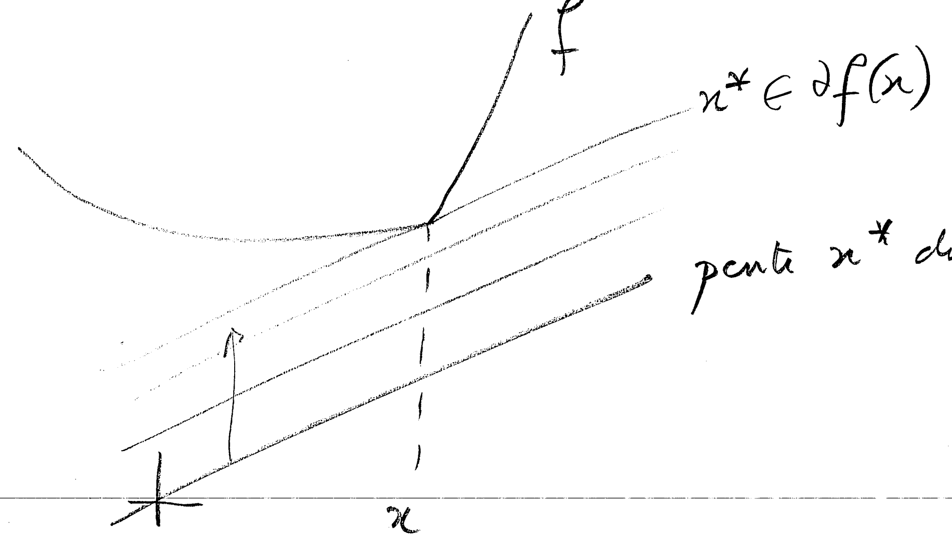
$$\partial f : \mathbb{E} \multimap \mathbb{E}$$

est une multifonction : à $x \in \mathbb{E}$, elle associe une partie $\partial f(x)$ de $\mathbb{E}$

— C'est difficile de calculer $\partial f(x)$; on préfère parfois calculer sa multifonction réciproque :

$$\partial f^{-1} : \mathbb{E} \multimap \mathbb{E}$$
$$x^* \to \{x \in \mathbb{E} : x^* \in \partial f(x)\}$$

Autrement dit on se donne une <u>pente</u> $x^* \in \mathbb{E}$ et on cherche tous les x tels que $x^* \in \partial f(x)$.
Comment faire ?



Pour x^* donné, on cherche le plus petit $\alpha \in \mathbb{R}$ (ou le plus grand $-\alpha \in \mathbb{R}$) tel que $x \mapsto \langle x^*, x\rangle - \alpha$ est une minorante affine de f

$$\langle x^*, x\rangle - \alpha \leq f(x), \quad \forall x \in \mathbb{E}$$
$$\alpha \geq \langle x^*, x\rangle - f(x), \quad \forall x \in \mathbb{E}$$

$\Rightarrow$ le plus petit α est

$$\boxed{f^*(x^*) := \sup_{x \in \mathbb{E}} \langle x^*, x\rangle - f(x)}$$

c'est la valeur en x^* de la <u>conjuguée</u> de f

[B] Conjugué

– Soit $f : \mathbb{E} \to \overline{\mathbb{R}}$ (non nécessairement convexe)

– La conjugué de f est l'application

$$f^* : \mathbb{E} \to \overline{\mathbb{R}}$$

définie en $x^* \in \mathbb{E}$ (une "pente") par

$$\boxed{f^*(x^*) = \sup_{x \in \mathbb{E}} \left(\langle x^*, x \rangle - f(x) \right)} \qquad (1)$$

– Retour au calcul de ∂f^{-1}.

- soit $x^* \in \mathbb{E}$ donné
- si x est 1 solution du problème en (1) et si $f(x) \in \mathbb{R}$, alors $x^* \in \partial f(x)$; en effet

$$\langle x^*, x \rangle - \underbrace{f(x)}_{\in \mathbb{R}} \geqslant \langle x^*, y \rangle - f(y), \quad \forall y \in \mathbb{E}$$

$$f(y) \geqslant \underbrace{f(x) + \langle x^*, y - x \rangle}_{\substack{\text{une minorante affine de } f \\ \text{qui est exacte en } x \text{, donc} \\ x^* \in \partial f(x)}}, \quad \forall y \in \mathbb{E}$$

– On rappelle que $f^* : \mathbb{E} \to \overline{\mathbb{R}}$ est propre si f^* ne prend pas la valeur $-\infty$ (noté $f^* > -\infty$) et n'est pas identiquement égale à $+\infty$ (noté $f^* \not\equiv +\infty$)

$$\boxed{f^* > -\infty \iff f \not\equiv +\infty}$$

$$\iff \forall x^* \in \mathbb{E}, \quad f^*(x^*) > -\infty$$

$$\iff \forall x^* \in \mathbb{E}, \ \exists x_0 : \langle x^*, x_0 \rangle - f(x_0) > -\infty$$

$$\iff \exists x_0 \in \mathbb{E} : \ f(x_0) < +\infty$$

$$\iff f \not\equiv +\infty \qquad \square$$

$$\boxed{f^* \not\equiv +\infty \iff f \text{ a une minorante affine}}$$

$$\iff \exists x_0^* \in \mathbb{E} : \underbrace{f^*(x_0^*)}_{\sup_x \langle x_0^*, x \rangle - f(x)} < +\infty$$

$$\iff \exists x_0^* \in \mathbb{E}, \ \exists \alpha_0 \in \mathbb{R} : \ \langle x_0^*, x \rangle - f(x) \leq \alpha_0, \ \forall x \in \mathbb{E}$$

$$f(x) \geq \langle x_0^*, x \rangle - \alpha_0$$

$$\iff f \text{ a une minorante affine} \qquad \square$$

$$\boxed{f^* \text{ propre} \iff \begin{cases} f \not\equiv \infty \\ f \text{ a une minorante affine} \end{cases}}$$

On admet que c'est vrai si $f \in \mathrm{Conv}(\mathbb{E})$, i.e., f est <u>convexe</u> et <u>propre</u>

Propriété

- Soit $f : \mathbb{E} \to \overline{\mathbb{R}}$

  $\operatorname{dom} f = \{x \in \mathbb{E} : f(x) < +\infty\}$ domaine

  $\operatorname{epi} f = \{(x, \alpha) \in \mathbb{E} \times \mathbb{R} : f(x) \leq \alpha\}$ épigraphe

- $\boxed{\operatorname{Conv}(\mathbb{E})} \supset \boxed{\overline{\operatorname{Conv}}(\mathbb{E})}$

  $\operatorname{Conv}(\mathbb{E})$: ens. des fonctions convexes, propres

  $\overline{\operatorname{Conv}}(\mathbb{E})$: ens. des fonctions convexes, propres, fermées

- (rappel séance 6 sur la dualité)

  L'enveloppe supérieure d'une famille de $f_i : \mathbb{E} \to \overline{\mathbb{R}}$ pour $i \in I$ (ens. d'indices quelconques) est la fonction notée $\sup_{i \in I} f_i$, prenant en $x \in \mathbb{E}$ la valeur

$$\boxed{\left(\sup_{i \in I} f_i\right)(x) := \sup_{i \in I} \left(f_i(x)\right)}$$

On a

- $\operatorname{epi}\left(\sup_{i \in I} f_i\right) = \bigcap_{i \in I} (\operatorname{epi} f_i)$
- f_i convexe, $\forall i \in I \Rightarrow \left(\sup_{i \in I} f_i\right)$ convexe
- f_i fermée, $\forall i \in I \Rightarrow \left(\sup_{i \in I} f_i\right)$ fermée

(P_1) $\boxed{f \text{ propre avec minorante affine} \implies f^* \in \text{Conv}(\mathbb{E})}$

Etonnant ! car f n'est pas néc. convexe

Dem

- On a vu que f^* est propre
- $$f^*(x^*) = \sup_{x \in \mathbb{E}} \underbrace{\langle x^*, x \rangle - f(x)}_{\substack{\text{affine} \\ \text{donc convexe fermée}}}$$

$\implies$ convexe fermée

$\square$

C Biconjuguée

C'est la conjuguée de la conjuguée

$$f^{**} : E \longrightarrow \overline{\mathbb{R}}$$

qui prend en $x \in E$ la valeur

$f^* : E \to \overline{\mathbb{R}}$
$x^* \mapsto \sup_{x \in E} \langle x^*, x\rangle - f(x)$

$$\boxed{f^{**}(x) = \sup_{x^* \in E} \langle x^*, x\rangle - f^*(x^*)}$$

P₂ f propre avec minorante affine $\Rightarrow f^{**} \in \mathrm{Conv}(E)$

Dém

P₁

- $f^* \in \mathrm{Conv}(E) \Rightarrow f^*$ est propre
- f^* a une minorante affine car

f propre $\Rightarrow \exists x_0 : f(x_0) \in \mathbb{R}$

$$\Rightarrow f^*(x^*) \geq \underbrace{\langle x^*, x_0\rangle - f(x_0)}_{\text{c'est une minorante affine (fct de } x^*\text{) de } f^*}$$

$f^{**} \in \mathrm{Conv}(E)$ par P₁

$\square$

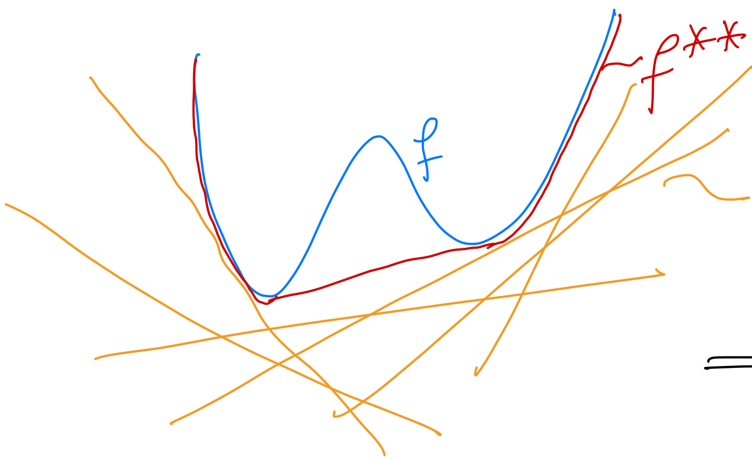
f est définie sur $\mathbb{E}$

f^* est définie sur $\mathbb{E}'$ (espace des pentes, c-à-d le dual $\mathbb{E}'$ de $\mathbb{E}$, identifié à $\mathbb{E}$ en dimension finie)

f^{**} est définie sur $\mathbb{E}$

<u>question</u> : $f^{**} \overset{?}{=} f$

P_3 $\boxed{f^{**} \text{ est l'enveloppe supérieure des minorantes affines de } f}$



toutes les minorantes affines de f

en particulier $\boxed{f^{**} \le f}$

$\Longrightarrow$ c'est un moyen de <u>convexifier f</u>

<u>Dém</u> (facile !!)

Soit $x \in \mathbb{E}$ fixé. La valeur en x de l'enveloppe supérieure des minorantes affines de f est

$$\sup_{\substack{(x^*, \alpha) \in \mathbb{E} \times \mathbb{E} \\ \langle x^*, y\rangle + \alpha \le f(y),\ \forall y \in \mathbb{E}}} \langle x^*, x\rangle + \alpha$$

$$= \sup_{x^* \in \mathbb{E}} \ \sup_{\substack{\alpha \in \mathbb{E} \\ \langle x^*, y\rangle + \alpha \le f(y),\ \forall y \in \mathbb{E}}} \langle x^*, x\rangle + \alpha$$

$$\alpha \le f(y) - \langle x^*, y\rangle, \ \forall y \in \mathbb{E}$$

le meilleur $\alpha = \inf_{y \in \mathbb{E}} f(y) - \langle x^*, y\rangle = -f^*(x^*)$

$$= \sup_{x^* \in \mathbb{E}} \langle x^*, x\rangle - f^*(x^*)$$

$$= f^{**}(x)$$

$\square$

(P_4)

Si f est propre et a une minorante affine **alors** $f^{**} = f \iff f \in \mathrm{Conv}(E)$

$\in \mathrm{Conv}(E)$

Dém

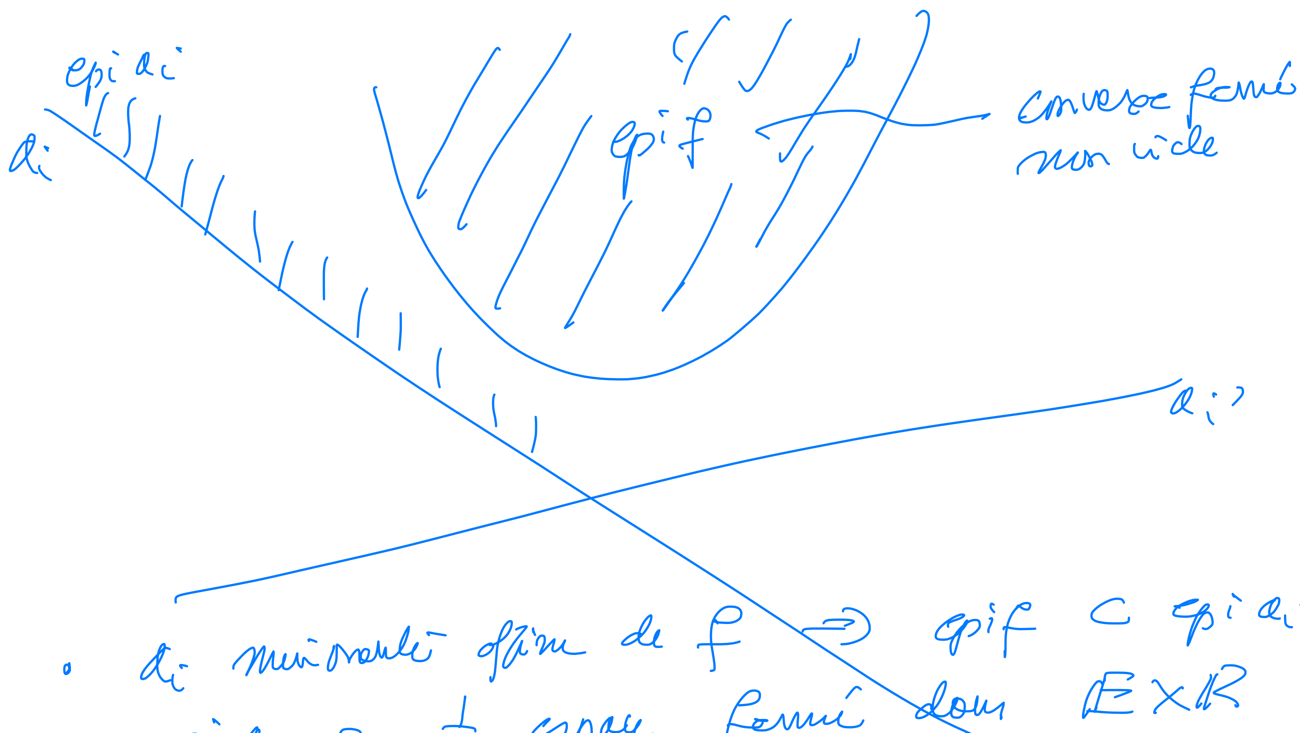
$(\Rightarrow)$ Clair

$(\Leftarrow)$ Il suffit de montrer que f = l'enveloppe supérieure de ses minorantes affines

$\{a_i\}_{i\in I}$ = ens des minorantes affines de f

$$f = \sup_{i\in I} a_i \;?$$

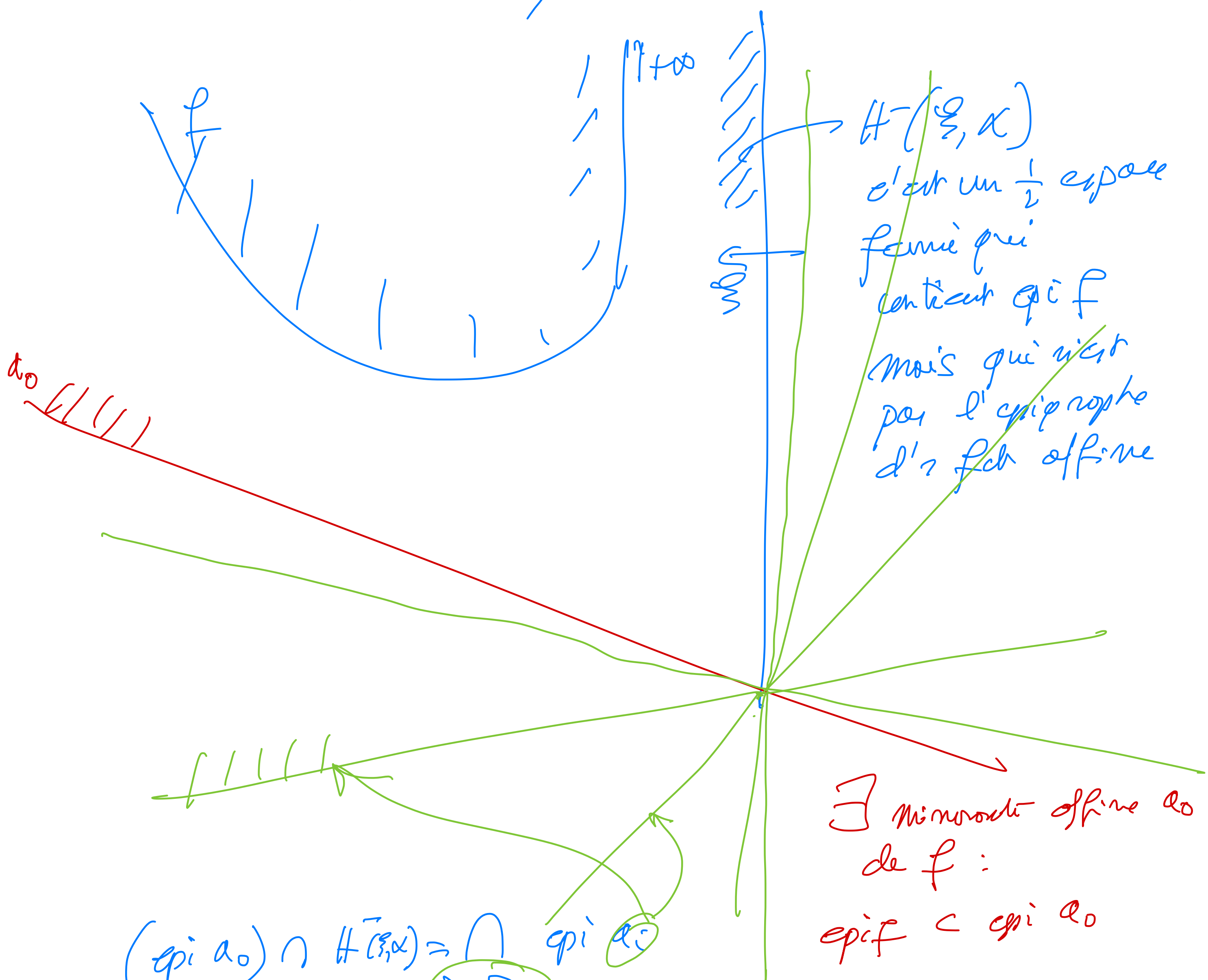
$$\iff \operatorname{epi} f = \bigcap_{i\in I} (\operatorname{epi} a_i) \;?$$



- a_i minorante affine de $f \Rightarrow \operatorname{epi} f \subset \operatorname{epi} a_i$
- $\operatorname{epi} a_i = \frac{1}{2}$ espace fermé dans $E\times\mathbb{R}$

$\Rightarrow$ c'est fini si on montre que tout $\frac{1}{2}$ espace fermé qui contient $\operatorname{epi} f$ = epi d'1 minorante affine de f

Ce n'est pas vrai ; par exemple



$H^-(\xi,\alpha)$ c'est un $\frac{1}{2}$ espace fermé qui contient epi f mais qui n'est pas l'épigraphe d'1 fct affine

$\exists$ minorante affine a_0 de f :

$\text{epi} f \subset \text{epi}\, a_0$

$(\text{epi}\, a_0) \cap H^-(\xi,\alpha) = \bigcap_{i \in I_0} \text{epi}\, a_i$

fin de la démonstration)