$\boxed{\text{D.M. 16 : Optimisation, Centrale MP 2009, solution}}$

1 Produits d'autoadjoints positifs

Q1) Cette question 1 est très proche du cours, donc incontournable !

a) Question de cours. On sait comme $u \in S(E)$, que u est diagonalisable (dz) dans une base orthonormée (b.o.n).
Sens $\Rightarrow$: soit $\lambda \in \mathbb{R}$ une valeur propre de u et $x \neq 0$ un vecteur propre associé à λ.
Alors
$$(u(x)|x) = \lambda \|x\|^2 \quad (1)$$
et comme $u \in S^+(E)$, on sait que $(u(x)|x) \geq 0$ donc par (1) et $x \neq 0$, on conclut que $\lambda \geq 0$.
Sens $\Leftarrow$: soit $(e_1, \ldots, e_n)$ une b.o.n. de diagonalisation de u. Soit $x \in E$ quelconque qu'on décompose en $x = \sum\limits_{i=1}^{n} x_i e_i$.
Alors $(u(x)|x) = (\sum\limits_{i=1}^{n} \lambda_i x_i e_i | \sum\limits_{i=1}^{n} x_i e_i) = \sum\limits_{i=1}^{n} \lambda_i x_i^2$.
Comme tous les λ_i sont positifs, on conclut bien que $(u(x)|x) \geq 0$.
On a bien montré que $u \in S^+(E)$.

b) Soient $u \in S^{++}(E)$, $\mathcal{B} = (e_1, \ldots, e_n)$ une base orthonormale de vecteurs propres de u alors $\mathrm{Mat}_{\mathcal{B}}(u) = D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ et donc
$$\mathrm{Mat}_{\mathcal{B}}(u^{-1}) = D^{-1} = \mathrm{diag}\left(\lambda_1^{-1}, \ldots, \lambda_n^{-1}\right).$$
Donc u^{-1} a sa matrice dans une b.o.n. qui est symétrique (diagonale) donc $u^{-1} \in S(E)$.
De plus $\mathrm{Sp}\left(u^{-1}\right) = \left\{\lambda_i^{-1}, 1 \leq i \leq n\right\} \subset \mathbb{R}^{*+}$, donc $\boxed{u^{-1} \in S^{++}(E)}$.

c) (i) Encore un grand classique : existence d'une racine carrée (sym. pos.) pour un endomorphisme symétrique positif. Avec les mêmes notations qu'au b), soit $s \in \mathscr{L}(E)$ telle que $\mathrm{Mat}_{\mathcal{B}}(s) = \Delta := \mathrm{diag}\left(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}\right)$ dans cette base. Puisque Δ est symétrique et la base $\mathcal{B}$ est orthonormale, alors $s \in S(E)$. De plus $Sp(s) = \left\{\sqrt{\lambda_i}; 1 \leq i \leq n\right\} \subset \mathbb{R}^+$, alors $s \in S^+(E)$. De plus $\Delta^2 = D$ donc $s^2 = u$.
(ii) **(M1) (celle induite par l'énoncé)**
Soit $x \in E$ tel que : $(u(x) \mid x) = 0$, alors $\left(s^2(x) \mid x\right) = 0$.
Or comme s est symétrique $\left(s^2(x) \mid x\right) = (s(x) \mid s(x))$.
Donc ici $(s(x)|s(x)) = 0$ i.e. $\|s(x)\| = 0$ et $u(x) = s^2(x) = 0$. □

Remarque culturelle : le langage des formes quadratiques.
Pour tout $u \in S(E)$,
- on note $q : E \to \mathbb{R}$, $x \mapsto (u(x)|x)$ la *forme quadratique associée à u*
- on note $\varphi : E \times E \to \mathbb{R}$, $x \mapsto (u(x)|y)$ la *forme bilinéaire symétrique associée à u.*

On distingue alors, en général entre :
- les vecteurs $x \in E$ tel que $q(x) = 0$ qui s'appellent *vecteurs isotropes* pour q.
- les vecteurs $x \in E$ tels que $u(x) = 0$, qui sont également caractérisés par $\forall\, y \in E$, $\varphi(x, y) = 0$ appelés vecteurs du noyau de u, mais on dit aussi du *noyau* de φ.

Bien sûr tout vecteur du noyau est isotrope, mais en général si φ n'est pas une F.B.S. positive i.e. u n'est pas positif, la réciproque est fausse.
Par exemple si $E = \mathbb{R}^2$ et $u : e_1 \mapsto e_1, e_2 \mapsto -e_2$ alors $\ker(u) = \{0\}$ mais si $x = (x_1, x_2)$, alors $q(x) = x_1^2 - x_2^2$ donc l'ensemble des vecteurs isotropes (le *cône isotrope*) est la réunion des deux bissectrices d'équations $x_2 = \pm x_1$.
Ici on a montré que $\boxed{\text{pour } u \text{ F.B.S. } \textit{positive} \text{ le noyau est égal au cône isotrope.}}$ On peut en faire une autre démonstration avec l'inégalité de Cauchy-Schwarz pour φ

(M2) Si $u \in S^{+}(E)$, on sait que $\varphi : \quad (x,y) \mapsto (u(x)|y)$ est une F.B.S. positive, et donc on a l'inégalité de Cauchy-Schwarz :

$$\forall\,(x,y) \in E^2,\ |(u(x)|y)| \leq \sqrt{(u(x)|x)}.\sqrt{(u(y)|y)}$$

Donc si x est isotrope i.e. si $(u(x)|x) = 0$ alors on a pour tout $y \in E$, $(u(x)|y) = 0$ et donc $u(x) \in E^{\perp}$ i.e. $u(x) = 0$. □

Q2) (i) Montrons déjà que $u_1 \in S(\mathrm{Im}(u))$
Soit $(x,y) \in \mathrm{Im}(u)^2$: $(u_1(x)|y) = (u(x)|y) = (x|u(y)) = (x|u_1(y))$.
Ensuite, pour tout $x \in E \setminus \{0\}$, $(u(x)|x) > 0$ en particulier, pour tout $x \in \mathrm{Im}(u) \setminus \{0\}$; $(u_1(x))|x) = (u(x)|x) \geq 0$.
(ii) Attention :

Ici l'énoncé monte assez vite en difficulté, erreur fréquente : écrire u^{-1} au lieu de u_1^{-1}...
Notamment $u_1^{-1} \circ u$ n'est pas égal à l'identité, il est nul sur les vecteurs de $\ker(u)$.

Je suis ici la rédaction d'Adam :
Soit $(x,y) \in \mathrm{Im}(u)^2$, on veut montrer que $(u_1^{-1}(w(x))|y) = (u_1^{-1}(x)|w(y))$.
Or :

$$\begin{aligned}(u_1^{-1}(w(x))|y) &= (u_1^{-1}(u(v(x))|y)\\ &= (u((v(x))|u_1^{-1}(y)) \quad \text{car } u_1 \in S(\mathrm{Im}(u))\\ &= (v(x)|u(u_1^{-1}(y)) \quad \text{car } u \in S(E),\\ &= (v(x))|y) \quad (1)\end{aligned}$$

la dernière égalité étant vraie car pour tout $y \in \mathrm{Im}\,u$, $u_1^{-y} \in \mathrm{Im}\,u$ donc $u(u_1^{-1}(y)) = u_1(u_1^{-1}(y)) = y$.

Moralité, avec le p.s. on a pu transformer $u_1^{-1} \circ u$ qu'on ne connaît pas en $u \circ u_1^{-1} = \mathrm{id}_E$.

Mais alors par symétrie de v, on a :

$$\begin{aligned}(v(x)|y) &= (x|v(y))\\ &= (u(u_1^{-1}(x))|v(y))\\ &= (u_1^{-1}(x)|u(v(u)) \quad \text{car } u \in S(E) \qquad (2)\end{aligned}$$

Finalement avec (1) et (2), on a bien montré que $(u_1^{-1}(w(x))|y) = (u_1^{-1}(x)|w(y))$.
Reste à montrer que pour tout $x \in \mathrm{Im}(u)$, $(u_1^{-1}(w(x))|x) \geq 0$.
Or par la formule (1) précédente, avec $y = x$, on a : $(u_1^{-1}(w(x))|x) = (v(x)|x) \geq 0$ par hyp. sur v. □

Q3) Soit $y \in \mathrm{Im}(u \circ v) \cap \ker(u \circ v)$.
Alors $(u \circ v)(y) = 0$ et on a un $x \in E$ tel que $y = (u \circ v)(x)$.
Ainsi $(u \circ v)^2(x) = 0$.
Donc $(u_1^{-1}(x)|(u \circ v)^2(x)) = 0$ et comme $u \in S(E)$, on en déduit que :

$$(u(u_1^{-1}(x))|v \circ u \circ v(x)) = 0$$

autrement dit que :

$$(x|v \circ u \circ v(x)) = 0.$$

Comme $v \in S(E)$, on obtient :

$$(v(x)|u(v(x))=0$$

et par (1) de l'énoncé Q1) c) (ii), on en déduit que $v(x)=0$. Mais alors $y=u(v(x))=0$.
On a bien montré que $\mathrm{Im}(u\circ v)\cap\ker(u\circ v)=\{0\}$ ce qui, avec le théorème du rang donne bien :

$$E=\ker(u\circ v)\oplus\mathrm{Im}(u\circ v).$$

Q4) Comme on sait déjà que $u\circ v_{|\mathrm{Im}(u\circ v)}$ est dz on peut écrire $\mathrm{Im}(u\circ v)=\oplus_{i=1}^{r}E_{\lambda_i}$ où E_{λ_i} est formé de vecteurs propres de $u\circ v$ pour la v.p. λ_i.
Mais $\ker(u\circ v)=E_0(u\circ v)$ est le s.e.v. propre pour la v.p. 0.
Au total $E=E_0(u\circ v)\oplus\oplus_{i=1}^{r}E_{\lambda_i}$ ce qui donne bien la dz de $u\circ v$.
Remarque sur cette question 2 : L'idée derrière la démarche de l'énoncé est la suivante :

1ère idée : Dans le cas particulier où $u\in S^{++}(E)$, on montre en fait que $u\circ v$ est dz car elle est symétrique pour le p.s. $(x,y)\mapsto(x|u^{-1}(y))$.

En effet dans ce cas, $(u\circ v(x)|u^{-1}(y))=(u^{-1}\circ u\circ v(x)|y)=(v(x)|y)=(x|v(y))=(x|u^{-1}(u(v(x))))$
□

2ème idée : on se ramène à la première idée en restriction à $\mathrm{Im}\,u$, et là c'est plus technique.

Q5) **N.B.** La définition de l'adjoint de f demande une explication ici car

f n'est pas un endomorphisme, on sort du cadre normal du programme, l'énoncé est « gonglé »!

Normalement ils auraient dû donner la définition ici tâchons de la deviner.
Pour chaque $y\in F$, l'application $\varphi : x\in E\mapsto(f(x)|y)\ \in\mathbb{R}$ étant une forme linéaire de l'espace euclidien E, il existe un unique vecteur noté $f^*(y)\in E$ tel que $\forall\,x\in E$, $\varphi(x)=(f^*(y)|x)$.

Ainsi f^* est une application de F dans E, telle que $\forall\,(x,y)\in E\times F$, $$(f(x)|y)=(x|f^*(y)).$$

On vérifie comme dans le cours que f^* est bien linéaire.
a) On adopte la dém. du cours à ce cadre généralisé, elle reste la même. Vérifiez !
b) Question plus fine :
(M1) sans le « en déduire » Soit (z_k) comme dans l'énoncé. Soit $(e_1,\dots,e_d)$ une base de $\mathrm{Im}(f)$.
On note $e_i=f(\varepsilon_i)$ en fixant un antécédant $\varepsilon_i\in E$ de e_i.
Pour tout $i\in[\![1,d]\!]$, $(e_i|z_k)=(f(\varepsilon_i)|z_k)=(\varepsilon_i|f^*(z_k))$.
Par continuité du p.s. $(\varepsilon_i|f^*(z_k))\underset{k\to+\infty}{\longrightarrow}0$.
Donc pour tout $i\in[\![1,d]\!]$, $(e_i|z_k)\underset{k\to+\infty}{\longrightarrow}0$.
Or par équivalence des normes dans un e.v.n. de dim. finie, comme

$$\forall\,k\in\mathbb{N},\ z_k=\sum_{i=1}^{d}(z_k|e_i)e_i,$$

est l'écriture de z_k en coordonnées dans la base $(e_1,\dots,e_d)$), le fait que toutes les coordonnées tendent vers zéro dit que (z_k) tend vers zéro.
(M2) avec le « en déduire » Comme $\mathrm{Im}(f)$ est un supplémentaire de $\ker(f^*)$, par a), on sait que $f^*_{|\mathrm{Im}(f)} : \mathrm{Im}(f)\to\mathrm{Im}(f^*)$ est un isomorphisme d'espaces vectoriels.

Mais entre deux e.v.n. de dim. finie, un tel isomorphisme est aussi une application (bi)-continue. La continuité de $(f^*_{|\mathrm{Im}(f)})^{-1}$ donne immédiatement la propriété demandée pour $z_k = f^*_{|\mathrm{Im}(f)}(f^*(z_k))$.
c) Facile et standard (ici généralisé aux A.L. qui ne sont pas des endomorphismes)

Il *faut* savoir que $f^* \circ f^*$ (resp. $A^\top A$) est un endomorphisme (resp. une matrice) symétrique pos.

Soient $x, y \in E$: $(f^* \circ f(x) \mid y) = (f(x) \mid f(y)) = (x \mid f^* \circ f(y))$. Ainsi $f^* \circ f \in S(E)$.
De plus : $\forall x \in E; (f^* \circ f(x) \mid x) = (f(x) \mid f(x)) \geq 0$. D'où $f^* \circ f \in S^+(E)$.

Q6) On applique le résultat montré aux Q2 à 4, avec $u = a^{-1}$ et $v = f \circ f^*$.

Q7) Soit $x \in E$, $\|f(x)\|^2 = (f(x)|f(x)) = (f^* \circ f(x)|x)$.
Le résultat est donc une comparaison entre les deux formes quadratiques $x \mapsto (f^* \circ f(x)|x)$ et $x \mapsto (a(x)|x)$.
Pour les relier avec la Q6, on pose $y = a(x)$, et donc $x = a^{-1}(y)$, alors on peut écrire :

$$(f^* \circ f(x)|x) = (f^* \circ f(x)|a^{-1}(y)) = (a^{-1} \circ f^* \circ f(x)|y)$$

la dernière égalité venant du caractère symétrique de a^{-1}.
On veut donc montrer que :

$$(a^{-1} \circ f^* \circ f(x)|y) \leq \rho(x|y) \quad (\dagger)$$

pour $y = a(x)$.
Or, on sait que Q2 a) (ii) que $a^{-1} \circ f^* \circ f$ est un endomorphisme autoadjoint positif pour le p.s. Φ_a (voir aussi la remarque à la fin de la Q4 de ce corrigé, car on est ici dans le cas simplifié où a est inversible).
Or ($\dagger$) se réécrit :

$$\Phi_a((a^{-1} \ \circ f \circ f^*)(x), x) \leq \rho \Phi_a(x, x) \quad (\ddagger)$$

cette inégalité est maintenant immédiate en diagonalisant l'endomorphisme $g = a^{-1} \circ f \circ f^*$ dans une base Φ_a orthogonale, puisque cet endomorphisme est Φ_a-symétrique.

Minimisation d'une fonctionnelle quadratique

Q8) **Argument standard de minoration :** Soient $0 < \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ les valeurs propres de a, en considérant une base orthonormale de vecteurs propres de a, et en décomposant chaque $x \in E$ dans cette base, on obtient : $(a(x) \mid x) \geq \lambda_1 \|x\|^2$.
Puisque $(b \mid x) \leq \|b\| \cdot \|x\|$ alors :

$$J(x) = \frac{1}{2}(a(x) \mid x) - (b \mid x) \geq \frac{1}{2}\lambda_1 \|x\|^2 - \|b\| \cdot \|x\|.$$

Par minoration, on en déduit bien que : $\boxed{\lim_{\|x\| \to +\infty} J(x) = +\infty}$.

Q9) **Argument standard de coercivité :** Le cas $V = 0$ est évident, supposons que $V \neq \{0\}$.
Soit $x_0 \in V$
Par déf. de la limite quand $\|x\| \to \infty$, on sait que :

$$\exists r > 0; \forall x \in V; [\|x\| > r \Rightarrow J(x) > J(x_0)]$$

Bien sûr, $x_0 \in \bar{B}_V(0, r)$, où $\bar{B}_V(0, r)$ est la boule fermée de V de centre O et de rayon r.
La fonction J est continue sur cette boule qui est compacte, puisque V est de dimension finie.
Alors J est minorée et atteint sa borne inf sur $\bar{B}_V(0, r)$.
$\exists y_0 \in \bar{B}_V(0, r)$ tel que $J(y_0) = \min_{x \in \bar{B}_V(0,r)} J(x) \leq J(x_0)$.
Et pour tout $x \in V \setminus \bar{B}_V(0, r)$, $J(x) > J(x_0) \geq J(y_0)$.
Ainsi $\boxed{J \text{ atteint en } y_0 \text{ son minimum sur } V.}$

Q10) Stricte convexité et conséquence : Soit $(x,y)\in V^2$ tel que : $x\neq y$.

a) **(M1) Par calcul brut :** Par linéarité de $x\mapsto (b|x)$, il suffit de montrer que si on pose $q(x)=(a(x)|x)$, alors :

$$q(\frac{x+y}{2}) < \frac{1}{2}(q(x)+q(y))$$

N.B. Cette propriété est un cas particulier de la *stricte convexité* de q, mais la même preuve donnerait que pour tout $t\in]0,1[$,

$$q((1-t)x+ty)\leq (1-t)q(x)+tq(y).$$

(M1) Par calcul à partir de la déf. de q et polarisation :

$$\begin{aligned}
q\left(\frac{x+y}{2}\right) &= \left(a\left(\frac{x+y}{2}\right) \mid \frac{x+y}{2}\right) \quad \text{en dév. par bilinéarité :}\\
&= \frac{1}{4}\left((a(x)\mid x)+(a(y)\mid y)+(a(x)\mid y)+(a(y)\mid x)\right).\\
&= \frac{1}{4}\left(q(x)+q(y)+2(a(x)\mid y)\right) \quad \text{par symétrie de } a\\
&= \frac{1}{4}\left(q(x)+q(y)+q(x)+q(y)-q(x-y)\right) \quad \text{par polarisation}\\
&= \frac{1}{2}(q(x)+q(y))-\frac{1}{4}q(x-y) < \frac{1}{2}(q(x)+q(y))
\end{aligned}$$

la dernière inégalité étant stricte pour $x-y\neq 0$ par caractère défini positif de a

(M2) Par l'écriture de q dans une b.o.n. de dz : dans une b.o.n. $(\varepsilon_1,\dots,\varepsilon_n)$ de dz de a : si $x=\sum_{i=1}^n x_i\varepsilon_i$ alors $q(x)=\sum_{i=1}^n \lambda_i x_i^2$. Alors

$$q(\frac{x+y}{2})=\sum_{i=1}^n \lambda_i(\frac{a_i+b_i}{2})^2$$

et on applique la stricte convexité de $x\mapsto x^2$ de $\mathbb{R}$ dans $\mathbb{R}$ à chaque terme de la somme sachant qu'il y a au moins un indice i tel que $a_i\neq b_i$ donc pour lequel l'inégalité est stricte. □

b) *Par l'absurde*, supposons que J atteint son minimum sur V en deux points distincts x,y. D'après la question précédente : $J\left(\frac{x+y}{2}\right) < \frac{J(x)+J(y)}{2} = J(x)=J(y)$. *contradiction.*
Ainsi le point où J atteint son minimum sur V est unique.

Q11) a) Question un peu vague, on ne sait pas exactement où s'arrêter, mais si on regarde la question suivante, *ce qu'il faut toujours faire dans ce cas là*, on a intérêt à faire apparaitre les choses dans l'ordre d'un D.L.

Toujours en notant $q(x)=(a(x)|x)$, on sait, par bilinéarité et symétrie que :

$$\begin{aligned} q(x+th) = (a(x)|x)+(a(x)|th)+(a(th)|x)+(a(th)|th),\\ q(x)+2t(a(x)|h)+t^2q(h)\end{aligned}$$

Grâce à la linéarité de $x\mapsto (b|x)$, on en déduit immédiatement que :

$$\begin{aligned}
J(x+th) &= \frac{1}{2}q(x)+t(a(x)|h)+\frac{1}{2}t^2q(h)-(b|x)-t(b|h),\\
&= J(x)+t(a(x)-b|h)+\frac{1}{2}t^2q(h) \quad (\dagger)
\end{aligned}$$

Cette écriture $(\dagger)$ sous la forme d'un D.L.$_2$ sans reste semble pertinente pour le b). Pour répondre à la question, on écrira :

$$\boxed{J(x+th)-J(x)=t(a(x)-b|h)+\frac{1}{2}t^2(a(h)|h)}$$

b) On considère la fonction $J_{|V}$ qui est donc définie sur un e.v. V

Sens $\Rightarrow$: comme $J_{|V}$ est une fonction définie sur un e.v. entier, si elle admet un min. global sur V en x, en particulier c'est un min. local de $J_{|V}$ donc $d(J_{|V})=0$.

Or par le résultat du a), comme $|(a(h)|h)|\leq\|a(h)\|.\|h\|\leq|||a|||.\ \|h\|^2=O(\|h\|^2)$, on sait que :

$$d(J_{|V})(x).h=((a(x)-b)|h)$$

Donc comme $d(J_{|V})(x).h=0$ pour tout $h\in V$, on obtient donc :

$$\forall\, h\in V,\ ((a(x)-b)|h)=0$$

autrement dit : $\boxed{a(x)-b\in V^{\perp}.}$

Sens $\Leftarrow$: Cette fois un fait un raisonnement global, avec l'égalité du a). Si $(a(x)-b)\in V^{\perp}$ alors avec le a), pour tout $h\in V$, on a :

$$J(x+th)-J(x)=t(a(x)-b|h)+\frac{1}{2}t^2(a(h)|h)=\frac{1}{2}t^2(a(h)|h)\ \geq 0$$

Donc pour tout $y\in V$, en écrivant $y=x+h$ avec $h\in V$

$$J(y)-J(x)\geq 0$$

Q12)

Q13) Soit $x\in E$. Notons $L_{r,x}=L_r(x,\ \cdot)$ l'application $p\mapsto L(x,p)$.

Pour tout $p,h\in F$ alors par linéarité du p.s.,

$$L_{r,x}(p+h)=L_{r,x}(p)+(h|f(x)).$$

Ceci montre que $L_{r,x}$ est différentiable en tout point $p\in F$ et que

$$\boxed{\nabla L_{r,x}(p)=f(x).}$$

Avec cette remarque tout devient évident !

(i) $\Rightarrow$ (ii) si $L_{r,x}$ admet un maximum en un point p, on a, avec l'encadré, $f(x)=0$ i.e. $x\in\ker(f)$.

(ii) $\Rightarrow$ (iii) Si $x\in\ker f$, par définition de L_r, on a pour tout $p\in F$,

$$L_{r,x}(p)=J(x)+\frac{r}{2}\|x\|^2+0$$

le second membre ne dépend pas de h, donc $L_{r,x}$ est constante.

(iii) $\Rightarrow$ (i) une fonction constante atteint son maximum en tout point.

Q14) En transformant l'écriture grâce à l'adjoint de f (défini Q5) dans ce cadre généralisé :

$$\begin{aligned}L_r(x,p) &= J(x)+\frac{r}{2}\|f(x)\|^2+(x|f^*(p)),\\ &= \frac{1}{2}\big(a(x)|x)+r(f^*\circ f(x)|x)\big)+(f^*(p)|x)-(b|x)\\ &= \frac{1}{2}(\alpha(x)|x)-(\beta|x)\end{aligned}$$

On a vu que $f^*\circ f\in S^+(E)$ et $a\in S^{++}(E)$ donc $\alpha:=a+rf^*\circ f\in S^{++}(E)$ et $\beta=f^*(p)-b\in E$

L'étude faite précédemment Q8 à 11) s'applique donc à $K\ :\ x\mapsto\dfrac{1}{2}(\alpha(x)|x)-(\beta|x)$.

Avec $V=E$ entier, on sait alors par Q11) que cette « fonctionnelle quadratique » K est minimale en x ssi $\alpha(x)-\beta=0$ ce qui donne exactement l'équation voulue :

$$\boxed{(a + rf^* \circ f)(x) + f^*(p) - b = 0.}$$

Q15) a) Par déf. (x,p) est un point-selle de L_r si et seulement si, $\begin{cases} L_r(x,\cdot) & \text{est maximale en } p \quad (1) \\ L_r(\cdot,p) & \text{est minimale en } x \quad (2). \end{cases}$

Par Q13) la condition (1) équivaut à $x \in \ker(f)$ et compte-tenu de cette condition $f(x) = 0$, la Q14) dit que (2) équivaut à $a(x) + 0 + f^*(p) = b$.

Ainsi on a bien $\begin{cases} (1) \\ (2) \end{cases} \Leftrightarrow (x \in \ker(f) \quad \text{et} \quad a(x) + f^*(p) = b).$

b) Par Q11 b) avec $V = \ker f$, on sait que $J_{|\ker f}$ est minimale en x, ssi, $(a(x) - b) \in \ker f^\perp$.

Or on a vu Q 5) a) que $\ker f^\perp = \operatorname{Im} f^*$.

Donc $J_{|\ker f}$ est minimale en x, ssi $\exists p \in F$, $(a(x) - b) = f^*(p)$ ce qui est la condition qu'on a introduite au a).

Compte-tenu du fait que $x \in \ker f$ par déf, la condition (1) du a) est automatique et donc par a), $J_{|\ker f}$ est minimale en x ssi il existe un $p \in F$ tel que (x,p) soit un point selle de L_r.

Q16) a) Si (x,p') et (x,p) sont deux points selle de L_r, alors par Q15) on doit avoir : $a(x)+f^*(p) = b$ et $a(x) + f^*(p') = b$, donc $f^*(p) = f^*(p')$ donc $p - p' \in \ker f^* = \operatorname{Im}(f)^\perp$ (par Q5).

Mais réciproquement, sachant que (x,p) est un point selle de L_r, on a $a(x) + f^*(p) = b$ et donc si $p - p' \in \operatorname{Im}(f)^\perp = \ker(f^*)$ on a $f^*(p') = f^*(p)$ donc aussi $a(x) + f^*(p') = b$. Cela en retour signifie que (x,p') est un point selle de L_r.

b) Par le a), si (x,p) est un point selle de L_r l'ensemble des $p' \in F$ tel que (x,p') soit un point selle est exactement le sous-espace affine passant par p de direction $\operatorname{Im}(f)^\perp$ autrement dit $\mathcal{F} := p + \operatorname{Im}(f)^\perp$.

On cherche l'élément $p' \in \mathcal{F}$ de plus petite norme.

(M1) pour les happy few connaissant les projections orthogonales sur un sous-espace affine.

Cela revient à déterminer le projeté orthogonal de 0 sur ce sous-espace affine.

Mais celui-ci s'obtient en prenant l'intersection de $\operatorname{Im} f$ avec $\mathcal{F}$.

(M2) le même résultat sans parler de projection sur un sous-espace affine.

En fait si on fait un dessin de la (M1) le point p' est aussi simplement le projeté orthogonal de p sur le s.e.v $\operatorname{Im}(f)$.

En effet, par Pythagore, on a $\|p\|^2 = \|p'\|^2 + \|p - p'\|^2$ donc $\|p\| > \|p'\|$ si $p \neq p'$.

Et comme p' est le projeté orthogonal sur $\operatorname{Im} f$ de *tous* les points de $\mathcal{F}$, on a bien $\|p'\| = \min(\|p\|)$ pour tout p point selle.

3 Algorithmes d'Uzawa et d'Arrow-Hurwicz

Remarque préliminaire sur l'existence des objets considérés : d'après la Q8) avec $V = \ker f$ il existe bien un élément $x \in \ker(f)$ en lequel $J_{|\ker f}$ est minimale.

Par Q 15) b) il existe alors aussi un point $p \in F$ tel que (x,p) est un point selle de L_r.

• Initialisation des suites : Pour $p_0 \in F$ fixé, il existe un élément $x_0 \in E$ tel que $L_r(\cdot,p_0)$ est minimale en x_0 car grâce à la Q14), comme $a + rf^* \circ f$ est symétrique définie positive, donc inversible, pour chaque $p_0 \in F$, il faut et il suffit de prendre l'unique $x_0 \in E$ tel que ;

$$x_0 = (a + rf^* \circ f)^{-1}(b - f^*(p_0)).$$

• Itération : si à l'étape k, on a défini, p_k, x_k, (et la suite (γ) étant donnée) on pose :

$$p_{k+1} = p_k + \gamma_k f(x_k),$$

et on définit alors

$$x_{k+1} = (a + rf^* \circ f)^{-1}(b - f^*(p_{k+1})).$$

Q17) a) Comme $x \in \ker f$, on peut écrire $p = p + \gamma_k f(x)$, et en retranchant cet égalité à la relation $p_{k+1} = p_k + \gamma_k f(x_k)$, on obtient, par linéarité de f :

$$p_{k+1} - p = p_k - p + \gamma_k f(x_k - x)$$

c'est-à-dire

$$\boxed{r_{k+1} = r_k + \gamma_k f(y_k).}$$

D'autre part, on a vu (cf. la remarque préliminaire), que pour tout $k \in \mathbb{N}$,

$$(a + rf^* \circ f)(x_k) + f^*(p_k) = b$$

En faisant la différence de cette égalité avec la même relation reliant x et p, et par linéarité des applications considérées, on obtient :

$$(a + rf^* \circ f)(x_k - x) + f^*(p_k - p) = b - b = 0$$

ce qui donne bien la relation demandée :

$$\boxed{(a + rf^* \circ f)(y_k) + f^*(r_k) = 0}$$

b) D'après le premier encadré du a) :

$$\begin{aligned} \|r_k\|^2 - \|r_{k+1}\|^2 &= (r_k - r_{k+1} \mid r_k + r_{k+1}) = (-\gamma_k f(y_k) \mid 2r_k + \gamma_k f(y_k)) \\ &= \gamma_k[-2(f(y_k)|r_k) + \gamma_k \|f(y_k)\|^2] \\ &= \gamma_k[-2(y_k|f^*(r_k)) + \gamma_k \|f(y_k)\|^2] \quad (1) \end{aligned}$$

D'autre part, par le second encadré du a) :

$$-f^*(r_k) = (a + rf^* \circ f)(y_k)$$

ce qui dans (1) donne :

$$\begin{aligned} \|r_k\|^2 - \|r_{k+1}\|^2 &= \gamma_k[2(y_k|a(y_k)) + 2r(y_k|f^* \circ f)(y_k) + \gamma_k \|f(y_k)\|^2] \\ &= \gamma_k\left[2(a(y_k) \mid y_k) + (2r - \gamma_k)\|f(y_k)\|^2\right] \quad (2) \end{aligned}$$

où (2) est exactement la relation demandée.
Pour l'inégalité demandée :

- par défintion : $\gamma_k \geq \alpha$ et $2r - \gamma_k \geq 2r - \beta$ car $\gamma_k \in [\alpha, \beta]$
- $(a(y_k) \mid y_k) \geq \frac{1}{\rho}\|f(y_k)\|$ (d'après la Q7 ce qui explique enfin le rôle de la partie I !)

En mettant ces deux inégalités dans (2), et comme $\left[2\left(r + \frac{1}{\rho}\right) - \beta\right] > 0$, on obtient bien :

$$\|r_k\|^2 - \|r_{k+1}\|^2 \geq \alpha\left[2\left(r + \frac{1}{\rho}\right) - \beta\right]\|f(y_k)\|^2.$$

c) L'inégalité démontrée au b), avec $\left[2\left(r + \frac{1}{\rho}\right) - \beta\right] > 0$ par déf. de β, donne que pour tout $k \in \mathbb{N}$, $\|r_k\|^2 - \|r_{k+1}\|^2 > 0$.
La suite $(\|r_k\|)$ est donc (strictement) décroissante et minorée par 0 donc <u>convergente</u>.
Toujours avec l'inégalité du b), on en déduit que la suite $(f(y_k))_k$ est convergente vers 0 .
Donc, puisque $\gamma_k \geq \alpha > 0$, l'inégalité de la question précèdente permet de déduire que : $\lim_{k \to \infty}(a(y_k) \mid y_k) = 0$.
Soit μ la plus petite valeur propre de a, on sait que $(a(y_k) \mid y_k) \geq \mu\|y_k\|^2 \geq 0$.
On déduit alors que la suite $(y_k)_k$ converge vers 0 , c'est à dire $\boxed{(x_k)_k \text{ converge vers } x}$.