CONCEPTION OPTIMALE DE STRUCTURES

G. ALLAIRE

13 Janvier 2010

CHAPITRE III

RAPPELS D'OPTIMISATION

DEFINITIONS

Soit V un espace de Banach, i.e., un espace vectoriel muni d'une norme qui est complet (toute suite de Cauchy est convergente dans V).

Soit $K \subset V$ un sous-ensemble non-vide. Soit $J : V \to \mathbb{R}$. On considère

$$\inf_{v \in K \subset V} J(v).$$

Définition. On dit que u est un minimum local de J sur K si

$$u \in K \quad \text{et} \quad \exists \delta > 0 \; , \; \forall v \in K \; , \; \|v - u\| < \delta \Longrightarrow J(v) \geq J(u) \; .$$

On dit que u est un minimum global de J sur K si

$$u \in K \quad \text{et} \quad J(v) \geq J(u) \quad \forall v \in K \; .$$

(différence: théorie $\leftrightarrow$ global / numérique $\leftrightarrow$ local)

Définition. Une suite minimisante du critère J sur l'ensemble K est une suite $(u^n)_{n\in\mathbb{N}}$ telle que

$$u^n \in K \;\forall\, n \quad \text{et} \quad \lim_{n\to+\infty} J(u^n) = \inf_{v\in K} J(v).$$

Par définition de l'infimum de J sur K **il en existe toujours !**

Optimisation en dimension finie $V = \mathbb{R}^N$

Théorème. Soit K un ensemble fermé non vide de $\mathbb{R}^N$, et J une fonction continue sur K à valeurs dans $\mathbb{R}$ vérifiant la propriété, dite "infinie à l'infini",

$$\forall (u^n)_{n\geq 0} \text{ suite dans } K \; , \; \lim_{n\to+\infty} \|u^n\| = +\infty \Longrightarrow \lim_{n\to+\infty} J(u^n) = +\infty \; .$$

Alors il existe au moins un point de minimum de J sur K. De plus, on peut extraire de toute suite minimisante de J sur K une sous-suite convergeant vers un point de minimum sur K.

(Idée: les fermés bornés sont compacts en dimension finie.)

Optimisation en dimension infinie

Problème: une fonction continue sur un fermé borné n'atteint pas toujours son minimum !

Exemple de non-existence: soit l'espace de Sobolev $H^1(0,1)$ muni de la norme $\|v\| = \left(\int_0^1 \left(v'(x)^2 + v(x)^2 \right) dx \right)^{1/2}$. Soit

$$J(v) = \int_0^1 \Big((|v'(x)| - 1)^2 + v(x)^2 \Big) dx \; .$$

On vérifie que l'application J est continue et "infinie à l'infini", et pourtant le problème de minimisation

$$\inf_{v \in H^1(0,1)} J(v)$$

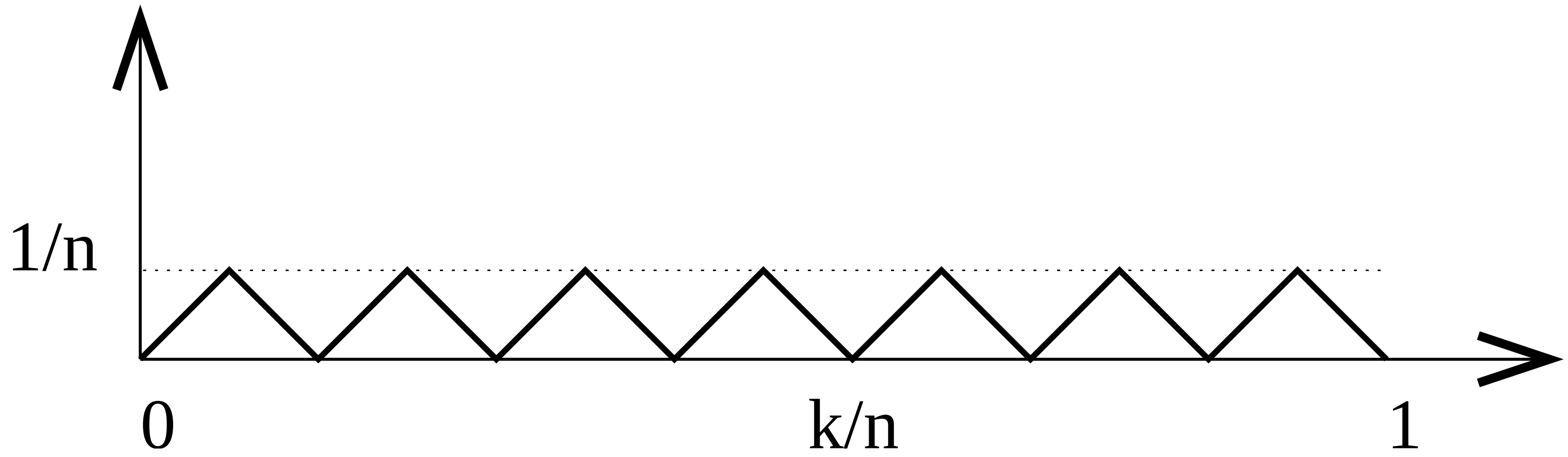
n'admet pas de solution. (Difficulté indépendante du choix de l'espace.)

Démonstration

Il n'existe aucun $v \in H^1(0,1)$ tel que $J(v) = 0$, et pourtant

$$\left(\inf_{v \in H^1(0,1)} J(v) \right) = 0,$$

car si on définit la suite u^n telle que $(u^n)' = \pm 1$



on a $J(u^n) = \int_0^1 u^n(x)^2 dx = \frac{1}{4n} \to 0$.

On voit dans cet exemple que la suite minimisante u^n "oscille" de plus en plus et n'est pas compacte dans $H^1(0,1)$ (bien qu'elle soit bornée).

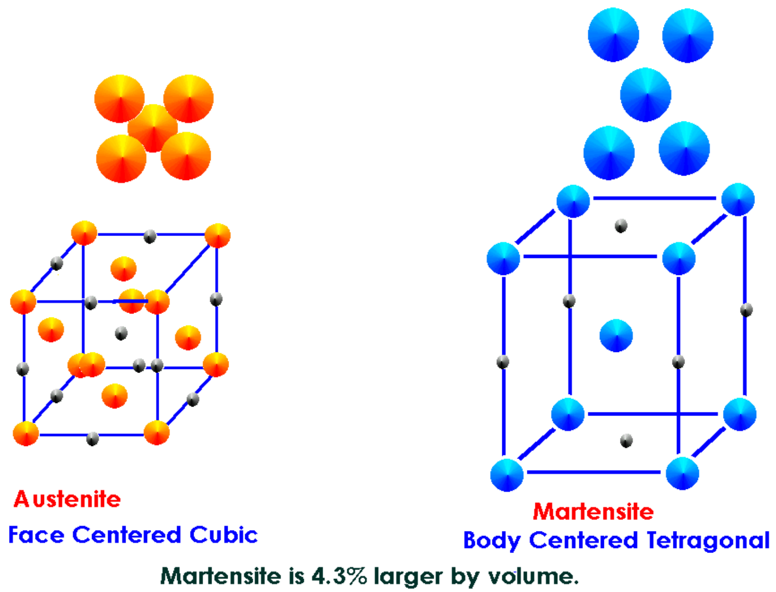
Une parenthèse en sciences des matériaux

La non-existence de solution à des problèmes de minimisation est utile en sciences des matériaux !

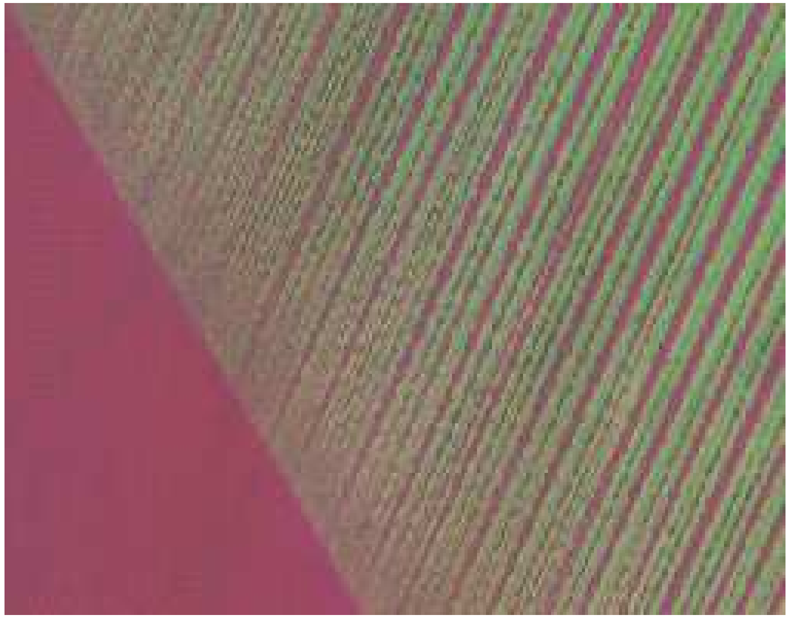
Théorie de Ball-James (1987).

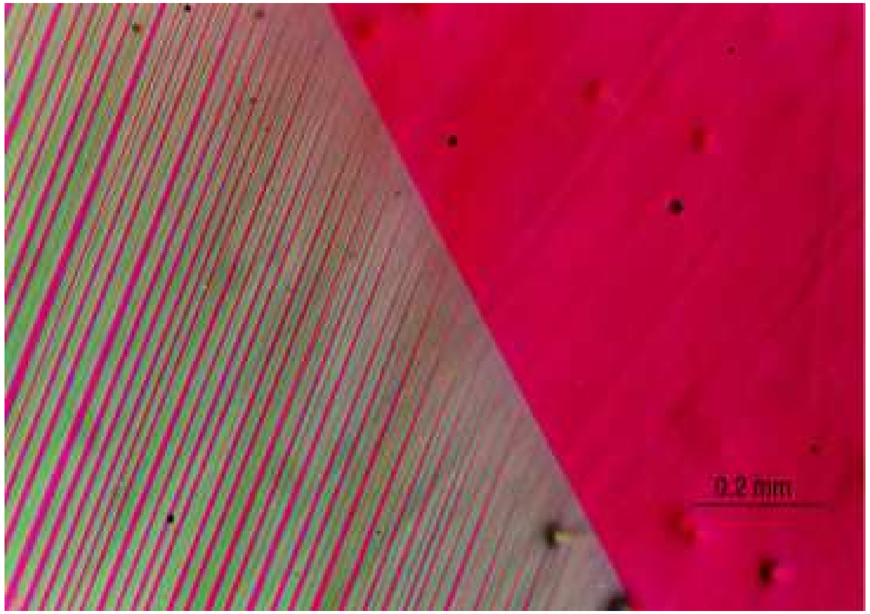
Matériaux à mémoire de forme = alliages avec des changements de phase.

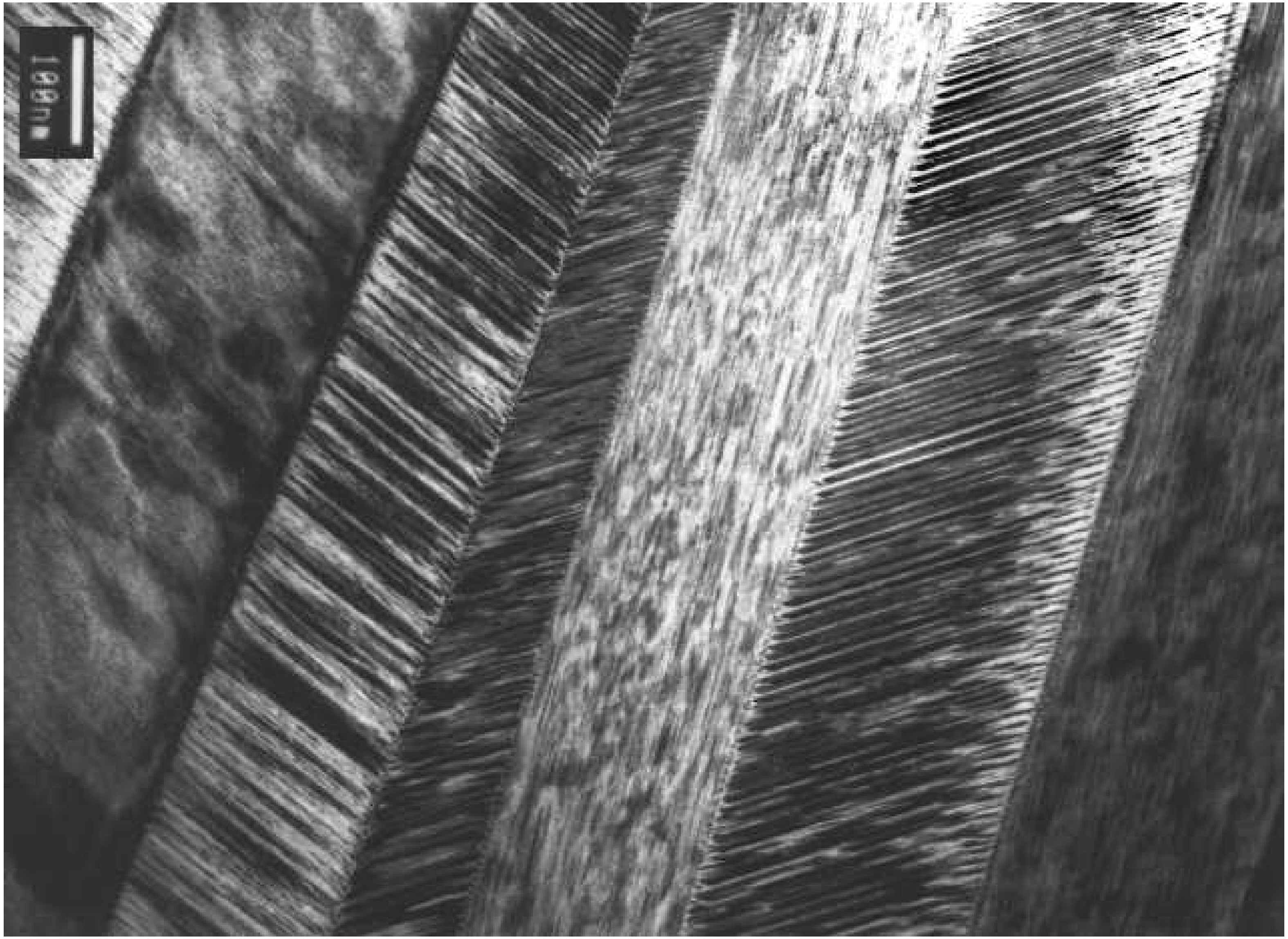
Co-existence de plusieurs phases cristallines: austenite et martensite.



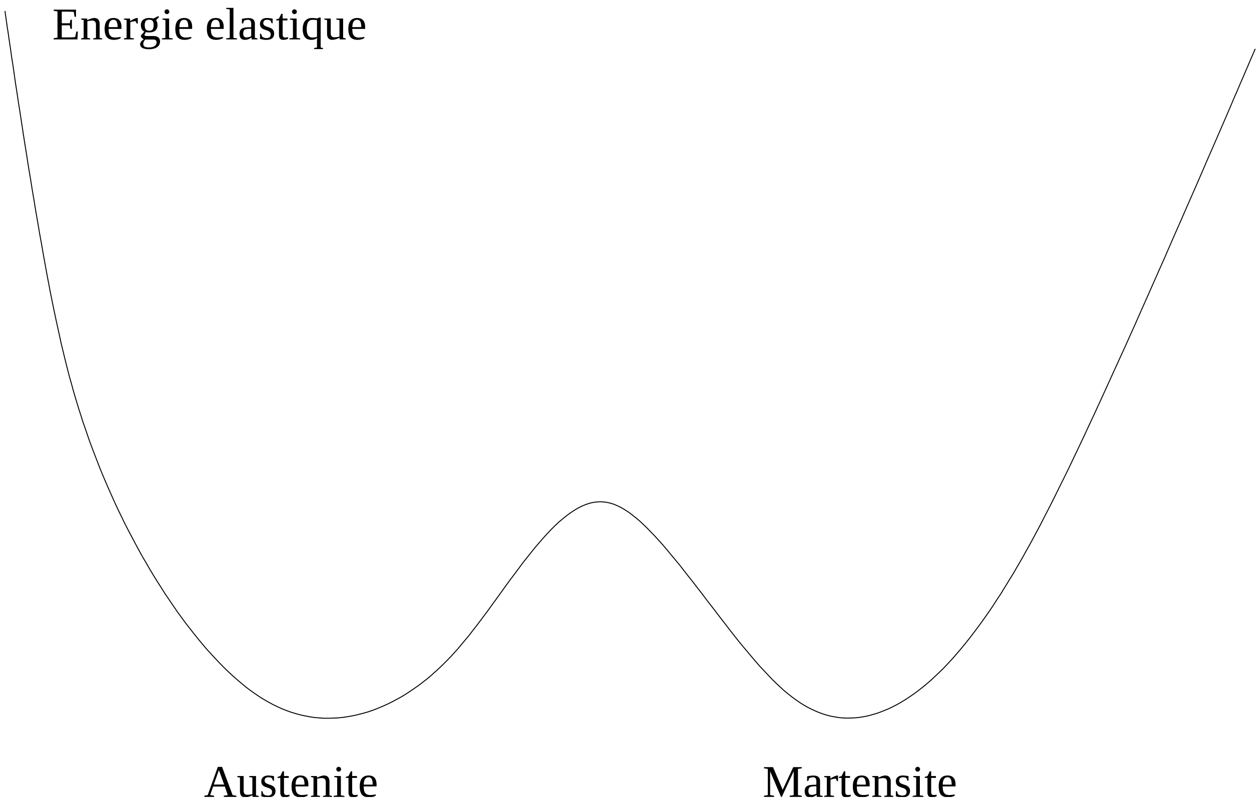
Alliage Cu-Al-Ni (clichés YONG S. CHU)






100nm

Mécanisme proposé par J. Ball et R. James: pour s'adapter aux efforts appliqués, l'alliage a intérêt à coexister sous plusieurs phases convenablement alignées et qui minimisent l'énergie $\Rightarrow$ **suite minimisante !**



Analyse convexe

Pour obtenir des résultats d'existence on rajoute une hypothèse de convexité.

Définition. Un ensemble $K \subset V$ est dit **convexe** si, pour tout $x, y \in K$ et tout réel $\theta \in [0,1]$, l'élément $(\theta x + (1-\theta)y)$ appartient à K.

Définition. On dit qu'une fonction J définie sur un ensemble convexe non vide $K \in V$ et à valeurs dans $\mathbb{R}$ est **convexe** sur K si et seulement si

$$J(\theta u + (1-\theta)v) \leq \theta J(u) + (1-\theta)J(v) \quad \forall\, u, v \in K \ , \ \forall\, \theta \in [0,1] \ .$$

De plus, J est dite **strictement convexe** si l'inégalité est stricte lorsque $u \neq v$ et $\theta \in]0,1[$.

Résultats d'existence

Théorème. Soit K un convexe fermé non vide d'un espace de Banach réflexif V, et J une fonction convexe continue sur K, qui est "infinie à l'infini" dans K, à savoir,

$$\forall (u^n)_{n\geq 0} \text{ suite dans } K \ , \ \lim_{n\to+\infty} \|u^n\| = +\infty \Longrightarrow \lim_{n\to+\infty} J(u^n) = +\infty \ .$$

Alors il existe un minimum de J sur K.

Remarques:

1. V Banach réflexif $\Leftrightarrow (V')' = V$ (V' est le dual de V)

2. Ce théorème reste vrai si l'on suppose simplement que V est le dual d'un espace de Banach séparable.

3. En pratique, **c'est vrai pour tous les espaces que l'on va rencontrer**: par exemple, $L^p(\Omega)$ avec $1 < p \leq +\infty$.

Unicité

Proposition. Si de plus J est strictement convexe, alors il existe au plus un point de minimum.

Proposition. Si J est une fonction convexe sur un ensemble convexe K, tout point de minimum local de J sur K est un minimum global.

Remarque. Pour les fonctions convexes il n'y a donc pas de différence entre minima locaux et globaux .

Exemple. Minimisation de l'énergie

$$\min_{u \in H_0^1(\Omega)} \frac{1}{2} \int_\Omega |\nabla u|^2 dx - \int_\Omega f\, u\, dx$$

On retrouve

$$-\Delta u = f \text{ dans } \Omega, \quad u = 0 \text{ sur } \partial\Omega.$$

Remarque. Autre méthode d'existence: compacité.

Différentiabilité

Définition. Soit V un espace de Banach. On dit que la fonction J, définie sur un voisinage de $u \in V$ à valeurs dans $\mathbb{R}$, est différentiable au sens de Fréchet en u s'il existe une forme linéaire continue sur V, $L \in V'$, telle que

$$J(u+w) = J(u) + L(w) + o(w) \quad , \quad \text{avec} \quad \lim_{w \to 0} \frac{|o(w)|}{\|w\|} = 0 \, .$$

On appelle L la différentielle (ou la dérivée, ou le gradient) de J en u et on note $L = J'(u)$, ou $L(w) = \langle J'(u), w \rangle_{V',V}$.

- ☞ Si V est un espace de Hilbert, on peut identifier V et son dual V' grâce au Théorème de représentation de Riesz. Il existe donc un unique $p \in V$ tel que $\langle p, w \rangle = L(w)$. On note aussi $p = J'(u)$.
- ☞ On utilise l'identification $V = V'$ si $V = \mathbb{R}^n$ ou $V = L^2(\Omega)$.
- ☞ En pratique, il est plus facile de calculer la dérivée directionnelle $j'(0) = \langle J'(u), w \rangle_{V',V}$ avec $j(t) = J(u + tw)$.

Exemple classique à connaitre absolument

Soit $J(v) = \frac{1}{2}a(v,v) - L(v)$ avec a symétrique. On pose $j(t) = J(u+tw)$

$$j(t) = \frac{t^2}{2}a(w,w) + t\Big(a(u,w) - L(w)\Big) + J(u)$$

et on dérive $t \to j(t)$

$$j'(t) = ta(w,w) + \Big(a(u,w) - L(w)\Big).$$

Par définition, $j'(0) = \langle J'(u), w\rangle_{V',V}$, donc

$$\langle J'(u), w\rangle_{V',V} = a(u,w) - L(w).$$

La condition $J'(u) = 0$ est une formulation variationnelle. On peut démontrer ainsi l'équivalence entre la minimisation de l'énergie $J(v)$ et la résolution de la F.V.

Exemples: (on utilise le produit scalaire "usuel" de L^2)

1. $J(v) = \int_\Omega \left(\frac{1}{2}v^2 - fv\right) dx$ avec $v \in L^2(\Omega)$

$$\langle J'(u), w\rangle = \int_\Omega (uw - fw)\, dx.$$

   Donc

$$J'(u) = u - f \ \in L^2(\Omega) \text{ (identifié à son dual)}$$

2. $J(v) = \int_\Omega \left(\frac{1}{2}|\nabla v|^2 - fv\right) dx$ avec $v \in H_0^1(\Omega)$

$$\langle J'(u), w\rangle = \int_\Omega (\nabla u \cdot \nabla w - fw)\, dx.$$

   Donc, après intégration par parties,

$$J'(u) = -\Delta u - f \ \in H^{-1}(\Omega) = \left(H_0^1(\Omega)\right)' \text{ (non identifié à son dual)}$$

Conditions d'optimalité

Théorème (Inéquation d'Euler). Soit $u \in K$ convexe. On suppose que J est différentiable en u. Si u est un point de minimum local de J sur K, alors

$$\langle J'(u), v - u \rangle \geq 0 \quad \forall\, v \in K \ .$$

Si $u \in K$ vérifie cette inéquation et si J est convexe, alors u est un minimum global de J sur K.

Remarques.

- ☞ Si u est intérieur à K, on obtient l'équation d'Euler $J'(u) = 0$.
- ☞ L'inéquation d'Euler est seulement nécessaire en général. Pour les fonctions convexes, elle est nécessaire et suffisante.

Minimisation avec contraintes d'égalité

$$\inf_{v \in V,\ F(v)=0} J(v)$$

avec $F(v) = (F_1(v), ..., F_M(v))$ une application dérivable de V dans $\mathbb{R}^M$.

Définition. On appelle **Lagrangien** du problème la fonction

$$\mathcal{L}(v, \mu) = J(v) + \sum_{i=1}^{M} \mu_i F_i(v) = J(v) + \mu \cdot F(v) \qquad \forall (v, \mu) \in V \times \mathbb{R}^M.$$

La nouvelle variable $\mu \in \mathbb{R}^M$ est appelée **multiplicateur de Lagrange** pour la contrainte $F(v) = 0$.

Lemme. Le problème de minimisation est équivalent à

$$\inf_{v \in V,\ F(v)=0} J(v) = \inf_{v \in V} \sup_{\mu \in \mathbb{R}^M} \mathcal{L}(v, \mu).$$

Stationnarité du Lagrangien

Théorème. On suppose que J et F sont continûment dérivables au voisinage de $u \in V$ tel que $F(u) = 0$. Si u est un minimum local, et si les vecteurs $\big(F_i'(u)\big)_{1 \le i \le M}$ sont linéairement indépendants, alors il existe des multiplicateurs de Lagrange $\lambda_1, \ldots, \lambda_M \in \mathbb{R}$, tels que

$$\frac{\partial \mathcal{L}}{\partial v}(u, \lambda) = J'(u) + \lambda \cdot F'(u) = 0 \quad \text{et} \quad \frac{\partial \mathcal{L}}{\partial \mu}(u, \lambda) = F(u) = 0 \; .$$

Minimisation avec contraintes d'inégalité

$$\inf_{v \in V,\ F(v) \leq 0} J(v)$$

où $F(v) \leq 0$ signifie que $F_i(v) \leq 0$ pour $1 \leq i \leq M$, avec $F_1, \ldots, F_M$ fonctions dérivables de V dans $\mathbb{R}$.

Définition. Soit u tel que $F(u) \leq 0$. L'ensemble $I(u) = \{i \in \{1, \ldots, M\}\ ,\ F_i(u) = 0\}$ est appelé l'ensemble des contraintes actives en u. On dit que les contraintes d'inégalité sont qualifiées en $u \in K$ si la famille $(F_i'(u))_{i \in I(u)}$ est libre.

Définition. On appelle Lagrangien du problème la fonction

$$\mathcal{L}(v,\mu) = J(v) + \sum_{i=1}^{M} \mu_i F_i(v) = J(v) + \mu \cdot F(v) \qquad \forall (v,\mu) \in V \times (\mathbb{R}^+)^M.$$

La nouvelle variable **positive** $\mu \in (\mathbb{R}^+)^M$ est appelée multiplicateur de Lagrange pour la contrainte $F(v) \leq 0$.

Lemme. Le problème de minimisation est équivalent à

$$\inf_{v \in V,\ F(v) \leq 0} J(v) = \inf_{v \in V} \sup_{\mu \in (\mathbb{R}^+)^M} \mathcal{L}(v,\mu).$$

Stationnarité du Lagrangien

Théorème. On suppose que les contraintes sont qualifiées en u tel que $F(u) \leq 0$. Alors, si u est un minimum local, il existe $\lambda_1, \ldots, \lambda_M \geq 0$, appelés multiplicateurs de Lagrange, tels que

$$J'(u) + \sum_{i=1}^{M} \lambda_i F_i'(u) = 0 \,, \quad \lambda_i \geq 0 \,, \quad \lambda_i = 0 \text{ si } F_i(u) < 0 \quad \forall\, i \in \{1, \ldots, M\} \,.$$

Cette condition est bien la stationnarité du Lagrangien puisque

$$\frac{\partial \mathcal{L}}{\partial v}(u, \lambda) = J'(u) + \lambda \cdot F'(u) = 0,$$

et que la condition $\lambda \geq 0$, $F(u) \leq 0$, $\lambda \cdot F(u) = 0$ est équivalente à l'inéquation d'Euler pour la **maximisation** par rapport à μ dans le convexe fermé $(\mathbb{R}^+)^M$

$$\frac{\partial \mathcal{L}}{\partial \mu}(u, \lambda) \cdot (\mu - \lambda) = F(u) \cdot (\mu - \lambda) \leq 0 \quad \forall \mu \in (\mathbb{R}^+)^M.$$

Une interprétation des multiplicateurs de Lagrange

Soit le Lagrangien pour la minimisation de $J(v)$ sous la contrainte $F(v) = c$

$$\mathcal{L}(v, \mu, c) = J(v) + \mu \cdot (F(v) - c)$$

On étudie la sensibilité du minimum à la variation de c.

On note $u(c)$ et $\lambda(c)$ le point de minimum et le multiplicateur de Lagrange correspondant. On suppose qu'ils sont dérivables par rapport à c. Alors

$$\nabla_c \Big(J(u(c)) \Big) = -\lambda(c).$$

λ donne la dérivée (sans la calculer) du minimum par rapport à c !

En effet

$$\nabla_c \Big(J(u(c)) \Big) = \nabla_c \Big(\mathcal{L}(u(c), \lambda(c), c) \Big) = \frac{\partial \mathcal{L}}{\partial c}(u(c), \lambda(c), c) = -\lambda(c)$$

car

$$\frac{\partial \mathcal{L}}{\partial v}(u(c), \lambda(c), c) = 0 \quad , \quad \frac{\partial \mathcal{L}}{\partial \mu}(u(c), \lambda(c), c) = 0 \, .$$

Dualité et point selle

Définition. Soit un Lagrangien $\mathcal{L}(v,q)$. On dit que $(u,p) \in U \times P$ est un point-selle (ou col, ou min-max) de $\mathcal{L}$ sur $U \times P$ si

$$\forall\, q \in P \qquad \mathcal{L}(u,q) \leq \mathcal{L}(u,p) \leq \mathcal{L}(v,p) \quad \forall\, v \in U\ .$$

Pour $v \in U$ et $q \in P$, posons $\mathcal{J}(v) = \sup_{q \in P} \mathcal{L}(v,q)$ et $\mathcal{G}(q) = \inf_{v \in U} \mathcal{L}(v,q)$. On appelle problème primal

$$\inf_{v \in U} \mathcal{J}(v)\ ,$$

et problème dual

$$\sup_{q \in P} \mathcal{G}(q)\ .$$

Exemple. $U = V$, $P = \mathbb{R}^M$ ou $\mathbb{R}^M_+$, et $\mathcal{L}(v,q) = J(v) + q \cdot F(v)$. Dans ce cas $\mathcal{J}(v) = J(v)$ si $F(v) = 0$ et $\mathcal{J}(v) = +\infty$ sinon, tandis qu'il n'y a pas de contraintes pour le problème dual (autres que $q \in P$).

Lemme (dualité faible). On a toujours

$$\inf_{v \in U} \mathcal{J}(v) \geq \sup_{q \in P} \mathcal{G}(q).$$

Preuve: $\inf \sup \mathcal{L} \geq \sup \inf \mathcal{L}$.

Théorème (dualité forte). Le couple (u, p) est un point-selle de $\mathcal{L}$ sur $U \times P$ si et seulement si

$$\mathcal{J}(u) = \min_{v \in U} \mathcal{J}(v) = \max_{q \in P} \mathcal{G}(q) = \mathcal{G}(p) \; .$$

Remarque. Le problème dual est souvent plus simple que le primal (pas de contraintes). Si on peut résoudre le dual, on obtient la solution du primal grâce à une minimisation sans contrainte.

Application: énergie complémentaire ou duale

Très important pour la suite !

Soit $f \in L^2(\Omega)$. On sait que pour résoudre

$$\begin{cases} -\Delta u = f & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\Omega \end{cases}$$

il suffit de minimiser l'énergie dite primale

$$\min_{v \in H^1_0(\Omega)} \left\{ J(v) = \frac{1}{2} \int_\Omega |\nabla v|^2 dx - \int_\Omega f v \, dx \right\}$$

On introduit une énergie duale (ou complémentaire)

$$\max_{\substack{\tau \in L^2(\Omega)^N \\ -\operatorname{div}\tau = f \text{ dans } \Omega}} \left\{ G(\tau) = -\frac{1}{2} \int_\Omega |\tau|^2 dx \right\}.$$

J est convexe et G est concave.

Proposition. Soit $u \in H_0^1(\Omega)$ la solution unique de l'équation. On pose $\sigma = \nabla u$. Alors,

$$J(u) = \min_{v \in H_0^1(\Omega)} J(v) = \max_{\substack{\tau \in L^2(\Omega)^N \\ -\operatorname{div}\tau = f \text{ dans } \Omega}} G(\tau) = G(\sigma),$$

et σ est l'unique point de maximum de G.

Preuve. On introduit le Lagrangien dans $H_0^1(\Omega) \times L^2(\Omega)^N$

$$\mathcal{L}(v, \tau) = -\frac{1}{2} \int_\Omega |\tau|^2 dx - \int_\Omega (f + \operatorname{div}\tau) v \, dx.$$

Par intégration par parties

$$\mathcal{L}(v, \tau) = -\frac{1}{2} \int_\Omega |\tau|^2 dx - \int_\Omega f v \, dx + \int_\Omega \tau \cdot \nabla v \, dx.$$

v **est le multiplicateur de Lagrange pour la contrainte** $-\operatorname{div}\tau = f$**.**

On vérifie que le dual du dual est le primal !

$$\max_\tau \mathcal{L}(v, \tau) = J(v).$$

Fin de la démonstration

Par définition, si τ vérifie la contrainte $-\operatorname{div}\tau = f$, on a

$$G(\tau) = \mathcal{L}(v,\tau) \quad \forall v$$

D'autre part,

$$\mathcal{L}(v,\tau) \leq \max_{\tau} \mathcal{L}(v,\tau) = J(v).$$

Or, par intégration par parties, on a $\int_\Omega |\nabla u|^2 dx = \int_\Omega f u \, dx$, donc

$$J(u) = \frac{1}{2}\int_\Omega |\nabla u|^2 dx - \int_\Omega f u \, dx = -\frac{1}{2}\int_\Omega |\nabla u|^2 dx = G(\nabla u).$$

Autrement dit, pour tout τ vérifiant $-\operatorname{div}\tau = f$,

$$G(\tau) = \mathcal{L}(u,\tau) \leq J(u) = G(\sigma)$$

c'est-à-dire que $\sigma = \nabla u$ maximise G parmi tous les τ tels que $-\operatorname{div}\tau = f$.

Algorithmes numériques de minimisation

Panorama simpliste:

☞ Algorithmes stochastiques: minimum global. Exemples: Monte-Carlo, recuit simulé, génétique. Voir le dernier chapitre et le dernier amphi.

Inconvénient: très couteux en CPU.

☞ Algorithmes déterministes: minimum local. Exemples: méthodes de gradient, Newton.

Inconvénient: il faut pouvoir disposer du gradient.

Gradient à pas optimal

On veut résoudre

$$\inf_{v \in V} J(v) \ .$$

Initialisation: choisir u^0 dans V. Itérations: pour $n \geq 0$

$$u^{n+1} = u^n - \mu^n \, J'(u^n) \ ,$$

où $\mu^n \in \mathbb{R}$ est choisi à chaque étape tel que

$$J(u^{n+1}) = \inf_{\mu \in \mathbb{R}^+} J\Big(u^n - \mu \, J'(u^n)\Big) \ .$$

Idée principale: si $u^{n+1} = u^n - \mu w^n$ avec $\mu > 0$ petit, alors

$$J(u^{n+1}) = J(u^n) - \mu \langle J'(u^n), w^n \rangle + \mathcal{O}(\mu^2),$$

donc pour descendre il vaut mieux prendre w^n colinéaire à $J'(u^n)$.

Convergence

Théorème On suppose que J est fortement convexe différentiable et que J' est Lipschitzien sur tout borné de V, c'est-à-dire que

$$\forall\, M > 0\,, \quad \exists\, C_M > 0\,, \quad \|v\| + \|w\| \leq M \;\Rightarrow\; \|J'(v) - J'(w)\| \leq C_M \|v - w\|\,.$$

Alors l'algorithme de gradient à pas optimal converge: quel que soit u^0, la suite (u^n) converge vers la solution u.

Remarque. Si J n'est pas fortement convexe:

- ☞ l'algorithme peut ne pas converger (il oscille entre plusieurs solutions),
- ☞ l'algorithme peut converger vers un minimum local,
- ☞ le résultat de l'algorithme peut varier selon l'initialisation.

Variante: gradient à pas fixe

On veut résoudre

$$\inf_{v \in V} J(v) .$$

Initialisation: choisir u^0 dans V. Itérations: pour $n \geq 0$

$$u^{n+1} = u^n - \mu \, J'(u^n) ,$$

Théorème. On suppose que J est fortement convexe, différentiable et que J' est Lipschitzien sur V. Alors, si $\mu > 0$ est suffisamment petit, l'algorithme de gradient à pas fixe converge : quel que soit u^0, la suite (u^n) converge vers la solution u.

Remarque. Variante intermédiaire possible: on ne garde pas le pas fixe mais on ne le prend pas optimal non plus ; on se contente d'augmenter le pas $\mu_{n+1} = 1.1\mu_n$ si J décroit, et de le réduire $\mu_{n+1} = 0.5\mu_n$ si J croit.

Gradient projeté

Soit K convexe fermé non vide de V. On veut résoudre

$$\inf_{v \in K} J(v) .$$

Initialisation: choisir u^0 dans K. Itérations: pour $n \geq 0$

$$u^{n+1} = P_K\big(u^n - \mu\, J'(u^n)\big) ,$$

avec P_K projection sur K.

Théorème. On suppose que J est fortement convexe différentiable et que J' est Lipschitzien sur V. Alors, si $\mu > 0$ est suffisamment petit, l'algorithme de gradient à pas fixe avec projection converge.

Remarque. On peut aussi **pénaliser** les contraintes, i.e. on remplace

$$\inf_{v \in V,\ F(v) \leq 0} J(v) \quad \text{par} \quad \inf_{v \in V} \left(J(v) + \frac{1}{\epsilon} \sum_{i=1}^{M} \left[\max\left(F_i(v), 0\right)\right]^2 \right) .$$

Exemples d'opérateurs de projection P_K

☞ Si $V = \mathbb{R}^M$ et $K = \prod_{i=1}^{M} [a_i, b_i]$, alors pour $x = (x_1, x_2, \ldots, x_M) \in \mathbb{R}^M$

$$P_K(x) = y \quad \text{avec} \quad y_i = \min\left(\max\left(a_i, x_i\right), b_i\right) \quad \text{pour} \quad 1 \leq i \leq M \; .$$

☞ Si $V = \mathbb{R}^M$ et $K = \{x \in \mathbb{R}^M \;\; \sum_{i=1}^{M} x_i = c_0\}$, alors

$$P_K(x) = y \quad \text{avec} \quad y_i = x_i - \lambda \quad \text{et} \quad \lambda = \frac{1}{M}\left(-c_0 + \sum_{i=1}^{M} x_i\right).$$

☞ Même chose si $V = L^2(\Omega)$ et $K = \{\phi \in V \mid a(x) \leq \phi(x) \leq b(x)\}$ ou $K = \{\phi \in V \mid \int_\Omega \phi \, dx = c_0\}$.

Pour des convexes fermés K plus généraux, P_K peut être très difficile à déterminer. Dans ce cas on utilise l'algorithme d'Uzawa qui recherche un point selle du Lagrangien.

Algorithmes de Newton (ordre 2)

Idée principale: si $V = \mathbb{R}^N$ et si $J'' \geq 0$

$$J(w) \approx J(v) + J'(v) \cdot (w - v) + \frac{1}{2} J''(v)(w - v) \cdot (w - v),$$

dont le minimum est $w = v - \left(J''(v)\right)^{-1} J'(v)$.

Algorithme: $u^{n+1} = u^n - \left(J''(u^n)\right)^{-1} J'(u^n)$.

☞ Converge plus vite si u^0 est proche du minimum u

$$\|u^{n+1} - u\| \leq C \|u^n - u\|^2 \ .$$

☞ Nécessite d'inverser un système linéaire de matrice $J''(u^n)$.

☞ Peut se généraliser en méthode de quasi-Newton (sans calculer J'') ou au cas avec contraintes.

CHAPITRE IV

CONTROLE OPTIMAL

Optimisation des systèmes distribués:

Calcul du gradient par état adjoint

Contrôle d'une membrane élastique

Pour $f \in L^2(\Omega)$, le déplacement vertical de la membrane u est solution de

$$\begin{cases} -\Delta u = f + v & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\Omega, \end{cases}$$

où v est une **force de contrôle** qui sera la variable d'optimisation (par exemple, un actionneur piézzo-électrique). On définit alors l'ensemble des contrôles admissibles

$$K = \left\{ v \in L^2(\omega) \mid v_{min}(x) \le v(x) \le v_{max}(x) \text{ dans } \omega \text{ et } v = 0 \text{ dans } \Omega \setminus \omega \right\}.$$

On veut contrôler la membrane pour qu'elle adopte un déplacement $u_0 \in L^2(\Omega)$ en minimisant ($c > 0$)

$$\inf_{v \in K} \left\{ J(v) = \frac{1}{2} \int_\Omega \left(|u - u_0|^2 + c|v|^2 \right) dx \right\}.$$

Existence d'un contrôle optimal

Proposition.

Il existe un unique contrôle optimal $\overline{v} \in K$.

Preuve. $v \rightarrow J(v)$ est fortement convexe et K est convexe (voir le livre de cours).

Remarque. L'existence est souvent beaucoup plus dure, mais ce qui importe surtout c'est le calcul du gradient $J'(v)$ pour mettre en oeuvre une méthode numérique.

Attention: la solution u de l'e.d.p. dépend du contrôle v.

Gradient et condition d'optimalité

La méthode la plus sûre et la plus simple pour calculer le gradient est

$$\lim_{\epsilon \to 0} \frac{J(v+\epsilon w) - J(v)}{\epsilon} = \langle J'(v), w \rangle = \int_\Omega J'(v) w \, dx \, .$$

Par linéarité, on a $u(v+\epsilon w) = u(v) + \epsilon \tilde{u}(w)$ avec

$$\begin{cases} -\Delta \tilde{u}(w) = w & \text{dans } \Omega \\ \tilde{u}(w) = 0 & \text{sur } \partial\Omega. \end{cases}$$

Autrement dit, $\tilde{u}(w) = \langle u'(v), w \rangle$.

Comme $J(v)$ est quadratique le calcul est très simple et on obtient

$$\int_\Omega J'(v) w \, dx = \int_\Omega \Big((u(v) - u_0) \tilde{u}(w) + c v w \Big) dx,$$

Malheureusement $J'(v)$ n'est pas explicite !

Etat adjoint

Pour simplifier l'expression du gradient on utilise l'état adjoint p défini comme l'unique solution dans $H_0^1(\Omega)$ de

$$\begin{cases} -\Delta p = u - u_0 & \text{dans } \Omega \\ p = 0 & \text{sur } \partial\Omega. \end{cases}$$

On multiplie l'équation de $\tilde{u}(w)$ par p et réciproquement, puis on intègre par parties

$$\text{équation pour } p \times \tilde{u}(w) \quad \Rightarrow \quad \int_\Omega \nabla p \cdot \nabla \tilde{u}(w)\, dx = \int_\Omega (u - u_0)\tilde{u}(w)\, dx$$

$$\text{équation pour } \tilde{u}(w) \times p \quad \Rightarrow \quad \int_\Omega \nabla \tilde{u}(w) \cdot \nabla p\, dx = \int_\Omega wp\, dx$$

Par comparaison de ces deux égalités on en déduit que

$$\int_\Omega J'(v)w\, dx = \int_\Omega (p + cv)w\, dx.$$

Conclusion l'état adjoint

On a donc trouvé une formule explicite du gradient

$$J'(v) = p + cv.$$

- Méthode adjointe: calcul du gradient grâce à la résolution de 2 problèmes aux limites (u et p).
- Si on ne calcule pas l'adjoint: pour **chaque** direction w il faut calculer 2 problèmes aux limites (u et $\tilde{u}(w)$) pour connaitre $\langle J'(v), w \rangle$.

  Par exemple, si $J'(v)$ est un vecteur en dimension n, pour obtenir ses n composantes il faut résoudre $(n+1)$ problèmes !
- Très efficace en pratique: c'est la meilleure méthode possible.
- Inconvénient: si on utilise un solveur numérique pour u en boite noire, il peut être très compliqué de le modifier pour obtenir la solution p de l'état adjoint.

- ☞ Si l'équation d'état est non-autoadjointe (forme bilinéaire non symétrique), l'opérateur principal de l'équation adjointe est le transposé ou adjoint de celui de l'équation d'état.
- ☞ Si l'équation d'état est non-linéaire, l'équation adjointe est linéaire.

Méthode générale pour trouver l'état adjoint

On considère l'équation d'état comme une contrainte et, pour tout $(\hat{v}, \hat{u}, \hat{p}) \in L^2(\Omega) \times H_0^1(\Omega) \times H_0^1(\Omega)$, on introduit le Lagrangien du problème de minimisation

$$\mathcal{L}(\hat{v}, \hat{u}, \hat{p}) = \frac{1}{2} \int_\Omega \left(|\hat{u} - u_0|^2 + c|\hat{v}|^2 \right) dx + \int_\Omega \hat{p}(\Delta \hat{u} + f + \hat{v}) \, dx,$$

où $\hat{p}$ est le multiplicateur de Lagrange pour la contrainte qui relie les deux variables indépendantes $\hat{v}$ et $\hat{u}$.

Par intégration par parties

$$\mathcal{L}(\hat{v}, \hat{u}, \hat{p}) = \frac{1}{2} \int_\Omega \left(|\hat{u} - u_0|^2 + c|\hat{v}|^2 \right) dx + \int_\Omega (-\nabla \hat{p} \cdot \nabla \hat{u} + f\hat{p} + \hat{v}\hat{p}) \, dx.$$

Proposition. Les conditions d'optimalité sont équivalentes à la stationnarité du Lagrangien, c'est-à-dire que

$$\frac{\partial \mathcal{L}}{\partial v} = \frac{\partial \mathcal{L}}{\partial u} = \frac{\partial \mathcal{L}}{\partial p} = 0.$$

Démonstration

- $\frac{\partial \mathcal{L}}{\partial p} = 0 \Rightarrow$ par construction, on trouve l'équation vérifiée par l'état u.
- $\frac{\partial \mathcal{L}}{\partial u} = 0 \Rightarrow$ équation vérifiée par l'état adjoint p. En effet,

$$\lim_{\epsilon \to 0} \frac{\mathcal{L}(u + \epsilon w) - \mathcal{L}(u)}{\epsilon} = \langle \frac{\partial \mathcal{L}}{\partial u}, w \rangle = \int_\Omega \left((u - u_0) w - \nabla p \cdot \nabla w \right) dx$$

ce qui est la formulation variationnelle de l'équation adjointe.

- $\frac{\partial \mathcal{L}}{\partial v} = 0 \Rightarrow$ formule pour $J'(v)$. En effet,

$$\lim_{\epsilon \to 0} \frac{\mathcal{L}(v + \epsilon w) - \mathcal{L}(v)}{\epsilon} = \langle \frac{\partial \mathcal{L}}{\partial v}, w \rangle = \int_\Omega (cv + p) w \, dx$$

Formule simple de la dérivée

On a obtenu dans la démonstration précédente

$$\boxed{J'(v) = \frac{\partial \mathcal{L}}{\partial v}(v, u, p)}$$

avec u l'état, et p l'adjoint.

Cela n'est pas une surprise ! En effet,

$$J(v) = \mathcal{L}(v, u, \hat{p}) \quad \forall \hat{p}$$

car u est l'état. Donc, si $u(v)$ est dérivable, on a

$$\langle J'(v), w \rangle = \langle \frac{\partial \mathcal{L}}{\partial v}(v, u, \hat{p}), w \rangle + \langle \frac{\partial \mathcal{L}}{\partial u}(v, u, \hat{p}), \frac{\partial u}{\partial v}(w) \rangle$$

On prend alors $\hat{p} = p$, l'adjoint, pour obtenir

$$\langle J'(v), w \rangle = \langle \frac{\partial \mathcal{L}}{\partial v}(v, u, p), w \rangle$$

Une (autre) interprétation de l'état adjoint

L'état adjoint p est le multiplicateur de Lagrange pour la contrainte de l'équation d'état. Mais c'est aussi une fonction de sensibilité.

Soit le Lagrangien

$$\mathcal{L}(\hat{v}, \hat{u}, \hat{p}, f) = \frac{1}{2} \int_\Omega \left(|\hat{u} - u_0|^2 + c|\hat{v}|^2 \right) dx + \int_\Omega (-\nabla \hat{p} \cdot \nabla \hat{u} + f\hat{p} + \hat{v}\hat{p}) \, dx.$$

On étudie la sensibilité du minimum à la variation de f.

On note $v(f)$, $u(f)$ et $p(f)$ les valeurs optimales, dépendant de f. On suppose qu'ils sont dérivables par rapport à f. Alors

$$\nabla_f \Big(J(v(f)) \Big) = p(f).$$

p donne la dérivée (sans la calculer) du minimum par rapport à f !

En effet $J(v(f)) = \mathcal{L}(v(f), u(f), p(f), f)$ et $\frac{\partial \mathcal{L}}{\partial v} = \frac{\partial \mathcal{L}}{\partial u} = \frac{\partial \mathcal{L}}{\partial p} = 0$.