Chapitre 4

Théorie variationnelle

4.1 Introduction

Comme nous l'avons vu au Chapitre précédent les équations que nous considérons ici peuvent se mettre sous la forme variationnelle

$$dF(u) = 0 \quad , \tag{4.1}$$

où F est une fonctionnelle définie sur un espace de fonctions adéquat. Autrement, dit les solutions sont les points critiques de F.

Comme il a été vu dans le cours de J.P. Bourguignon, l'étude des points critiques d'une fonctionnelle, peut fournir des informations intéressantes, même en dimension finie (de nature topologique par exemple, lorsqu'on regarde des fonctionnelles sur des variétés).

Le point de vue qui est privilégié ici est celui des *valeurs critiques.*

Définition 1. On appelle valeur critique, de la fonctionnelle F, de classe C^1 définie sur E, un nombre $\beta \in \mathbb{R}$, tel qu'il existe $u \in E$, tel que

$$F(u) = \beta \quad , dF(u) = 0 \quad .$$

Bien entendu, si on est capable de déterminer une valeur critique, on aura, ipso facto prouvé l'existence d'un point critique, et si on est capable de trouver k valeurs critiques distinctes, on aura k points critiques distincts (au moins).

Remarque: Il existe un point de vue sensiblement différent sur la théorie variationnelle, c'est celui du complexe de Thom-Smale (voir cours de F. Laudenbach). Ce point de vue a permis de résoudre récemment des problèmes difficiles en géométrie symplectique, grâce notamment aux travaux de A. Floers.

Afin de trouver des valeurs critiques, on considère les ensembles de niveau de F déterminés par

$$F^a = \{u \in E \quad , F(u) < a\} \quad .$$

L'idée de la théorie de Morse est la suivante :

Pour $a < b$, on compare les ensembles de niveau F^a et F^b ($F^a \subset F^b$). Si F^a et F^b n'ont pas la même "topologie", et si F satisfait certaines propriétés de compacité, on montre alors qu'il existe une valeur critique $c \in [a, b]$.

Bien entendu, l'énoncé précédent est très vague. Il faut en effet préciser ce que l'on entend par même "topologie" : nous donnerons des définitions plus précises lorsque nous parlerons du Lemme de Déformation qui est l'outil essentiel de cette démarche. Commençons par l'illustrer sur un dessin.

Exemple 1. $F : [0,1] \to \mathbb{R}$.

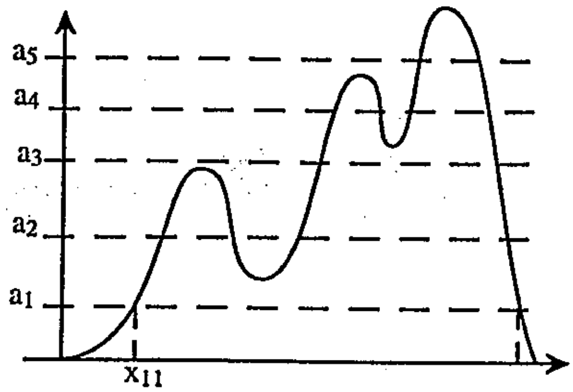


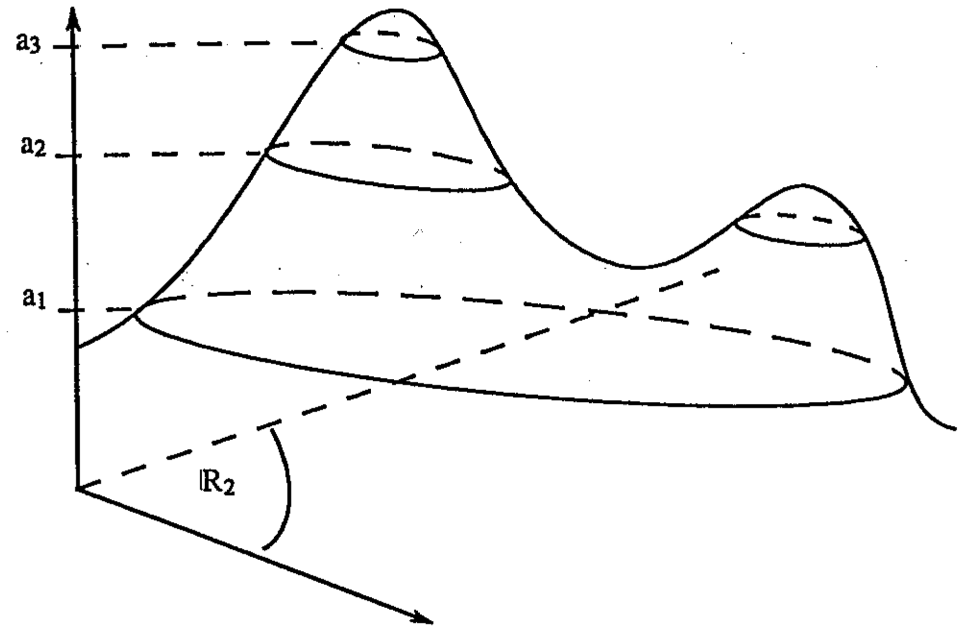
Figure 1

On vérifie sur ce dessin que

F^{a_1}	est composé de 2 intervalles connexes disjoints
F^{a_2}	est composé de 3 intervalles connexes disjoints
F^{a_3}	est composé de 2 intervalles connexes disjoints
F^{a_4}	est composé de 3 composantes connexes
F^{a_5}	est composé de 2 composantes connexes
F^b	est vide pour b grand.

On peut imaginer que pour des sous-ensembles "raisonnables" de [0,1] le nombre de composantes connexes est le même pour des ensembles qui ont la "même" topologie. On doit donc trouver 1 valeur critique dans $[a_1, a_2]$, une autre dans $[a_2, a_4]$, une autre dans $[a_3, a_4]$, une autre entre a_4 et a_5, et finalement une dans $[a_5, +\infty[$: en tout 5 valeurs critiques pour F. C'est bien ce que l'on vérifie sur le dessin.

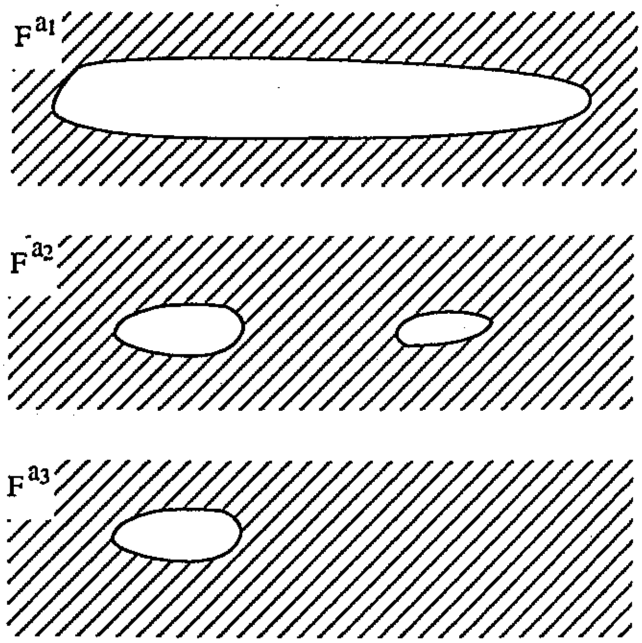
Exemple 2. $F : \mathbb{R}^2 \to \mathbb{R}$



Dessin du graphe de F

(Z! Le graphe de F vit dans $E \times \mathbb{R}$ Les ensembles de niveau dans E!)

dessinons les ensembles de niveau.



Il est clair, sur cet exemple que les trois ensembles sont connexes et nous ne pouvons utiliser la caractérisation précédente: en revanche, nous voyons apparaître une autre caractéristique importante : le nombre de trous, qui doit être le même pour des ensembles de topologie équivalente. Ainsi

$$\begin{array}{ll} F^{a_1} & \text{a 1 trous} \\ F^{a_2} & \text{a 2 trous} \\ F^{a_3} & \text{a 1 trou} \\ F^{b} & \text{est vide pour } b \text{ grand}. \end{array}$$

On doit donc s'attendre à trouver une valeur critique entre a_1 et a_2 une autre entre a_2 et a_3, enfin une dernière au-dessus de a_3. C'est ce que l'on constate sur le dessin : on voit qu'il y a deux maxima, et un col.

4.2 Un peu de topologie

a) Homotopie.

Introduisons un minimum de vocabulaire pour préciser la notion de " topologies équivalentes " et aussi pour avoir des critères simples permettant de déceler des topologies dissemblables (pour des exemples élémentaires).

Considérons deux espaces topologiques (au sens du Chapitre I, c'est-à-dire muni d'une famille d'ouverts) X et Y, et l'ensemble $C^\circ(X,Y)$ des applications continues de X vers Y.

Définition 2. Soit f_1 et f_2 deux applications continues de X vers Y. On dit que f_1 est " homotope " à f_2 si et seulement si il existe une application continue F de $X \times [0,1]$ vers Y telle que

$$\begin{array}{lll} F(x,0) & = f_1(x) & \forall x \in X \\ F(x,1) & = f_2(x) & \forall x \in X \quad . \end{array}$$

F est donc un chemin continu de $C^\circ(X,Y)$ d'applications de f_1 vers f_2. On dit que F est une " déformation " de f_1.

Propriétés de l'homotopie. Il est facile de voir que la relation " être homotope à" est une relation d'équivalence. On dira donc (en raison de la symétrie) que f_1 et f_2 sont homotopes. On appelle classes d'homotopie les classes d'équivalence pour la relation " être homotope à ".

Exemple 1 : Si Y est un espace vectoriel, toute application est homotope à l'application nulle. Il suffit de prendre

$$F(x,t) = tf(x) \quad \forall x \in X$$

Exemple 2. $X = Y = S^1 = \{z \in \mathbb{C}, |z| = 1\}$. Alors l'application identité de S^1 n'est pas homotope à une application constante (voir aussi Chap.VI, Appendice).

Lorsque X et Y sont des variétés compactes de dimension finie, on montre que toute application continue est homotope à une application C^∞. [on commence par plonger Y dans un espace euclidien E de dimension finie : ensuite à l'aide des théorèmes de localisation et de régularisation on approche uniformément, une application f donnée de X vers Y, par des applications f_n de $C^\infty(X\,;E)$; on peut ensuite reprojeter ces applications sur Y]. On montre de même si f_1 et f_2 sont de classe C^1, on peut choisir la déformation F de classe C^1, (voir Appendice Chap. VI).

Exercice 1 : En utilisant la remarque précédente montrer que pour $m > k$, toute application continue de S^k vers S^m est homotope à une application constante.

b) Type d'homotopie.

Nous avons vu que la relation "être homotope à" est une relation dont les éléments sont des applications entre espaces topologiques. L'équivalence homotopique (le "type d'homotopie") s'applique aux espaces eux-mêmes.

Définition 3. Soit X et Y deux espaces topologiques. On dit que X et Y ont le même type d'homotopie, s'il existe une application f continue de X vers Y, et une application g continue de Y vers X, tels que

$$g \circ f \text{ est homotope à } Id_X$$
$$f \circ g \text{ est homotope à } Id_Y \quad .$$

En d'autre termes, cette notion couvre (à peu près) la notion de "même topologie" que nous avions employée dans l'introduction. Donnons quelques exemples.

Exemple 1. Tout espace vectoriel normé a le même type d'homotopie qu'un singleton. Soit E un tel espace. On définit

$$f : E \to \{0\} \quad , \text{par } f(x)=\{0\}$$
$$g : \{0\} \to E \quad , \text{par } g(0)=0 \quad .$$

On a donc

$$f \circ g(x) = 0 \quad \forall x \in E$$

et

$$g \circ f(0) = 0 = Id_{\{0\}}(0) \quad .$$

La propriété vient donc du fait que Id_E est homotope à l'application nulle.

De manière plus générale si X est un ensemble convexe de E (espace vectoriel), X a le type d'homotopie d'un singleton. C'est bien entendu le type d'homotopie le plus simple.

Définition 4. On dit d'un espace X qu'il est contractible si il a le type d'homotopie d'un singleton, c'est-à-dire si l'application identité de X est homotope à une application constante.

Il résulte de l'exemple 2 de la section 1 que S^1 n'est pas contractible.

Remarquons que si Y est un espace contractible alors pour tout espace X, toute application de X vers Y est homotope à une application constante, et il n'y a donc qu'une seule classe d'homotopie dans $C^0(X, Y)$.

Exemple 2 : Soit X un espace topologique et E un espace vectoriel normé. Alors $X \times E$ a le type d'homotopie de X. Prendre

$$\begin{cases} f : X \times E \to X & (x, v) \to x \quad \forall (x, v) \in X \times E \\ g : X \to X \times E & x \to (x, 0) \quad \forall x \in E \quad . \end{cases}$$

Exercice 2. On considère l'anneau $C = D^2 \backslash D^2 \left(\frac{1}{2}\right)$. Montrer que C a le type d'homotopie de S^1. Montrer de même que $\mathbb{R}^2 \backslash D^2$ a le type d'homotopie de S^1, ou encore que $\mathbb{R}^2 \backslash \{0\}$ a le type d'homotopie de S^1.

C) Invariants.

On parle d'invariants homotopiques pour des propriétés qui sont conservées lorsque deux espaces ont le même type d'homotopie. Par exemple le nombre de composantes connexes est un tel invariant, c'est-à-dire que deux espaces topologiques qui ont le même type d'homotopie ont le même nombre de composantes connexes. de même si X et Y ont le même type d'homotopie, $\forall Z$ (esp. topologique) $C^0(Z, X)$ et $C^0(Z, Y)$ ont le même nombre de classe d'homotopie... Les invariants homotopiques sont bien entendu importants pour déceler si deux espaces topologiques ont des types différents.

d) Rétractions.

Dans cette partie, on suppose $A \subset X$, et que A est muni de la topologie induite par celle de X.

Définition 5. On dit que X se rétracte, par déformation sur A s'il existe une application continue $F : X \times [0,1] \to X$ telle que

$$F(x,0) = x \quad \forall x \in X, \tag{4.2}$$
$$F(x,1) \in A \quad \forall x \in X \quad , \tag{4.3}$$
$$F(x,t) = x \quad \forall x \in A, \forall t \in [0,1] \quad . \tag{4.4}$$

Remarque: Dans la littérature, la définition que nous avons adoptée, est appelée déformation par rétraction *forte*. Il existe en effet une notion faible de déformation par rétraction (voir Spanier p.30).

Proposition 1: Si X se rétracte par déformation sur A, alors X et A ont le même type d'homotopie.

Preuve: Considérons les applications

$$\text{et} \quad \begin{cases} f_1 = F(\cdot,1) : X \to A \quad , & x \to F(x,1) \quad , \quad \forall x \in X \\ i : A \to X, & x \to x \quad , \qquad \forall x \in A \quad . \end{cases}$$

On a donc

$$\begin{aligned} & f_1 \circ i = Id_A \\ \text{et} \quad & i \circ f_1(x) = F(x,1) \quad , \quad \forall x \in X \quad . \end{aligned}$$

Il résulte donc de la définition que $i \circ f_1$ est homotope à Id_X ce qui prouve la proposition.

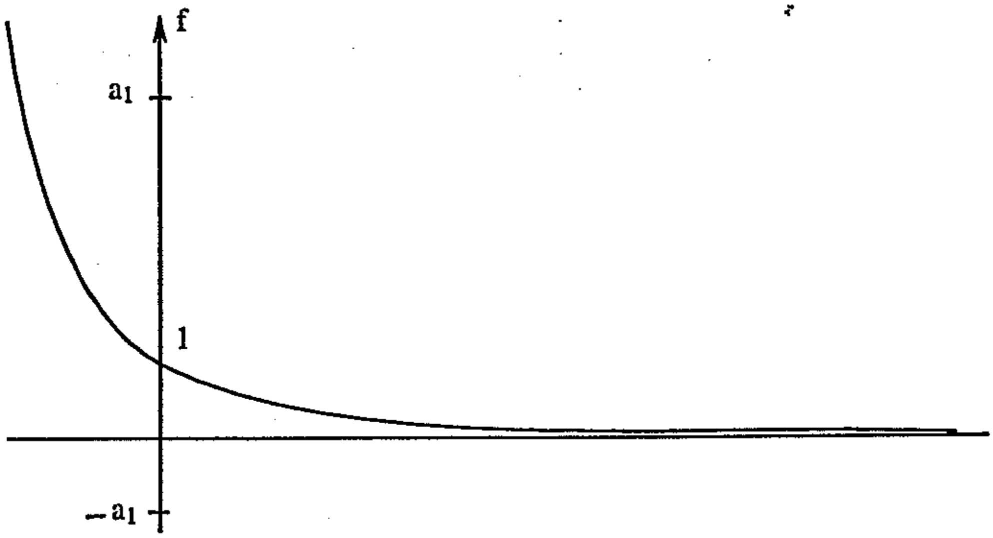
La notion de rétraction par déformation est celle qui va apparaître naturellement lorsque nous parlerons du Lemme de déformation.

Exemple 3. $\mathbb{R}^{n+1}\backslash\{0\}$ se rétracte par déformation sur S^n (sphère de dimension n). Prendre

$$F(x,t) = (1-t)x + \frac{tx}{\|x\|} \quad \forall x \in \mathbb{R}^{n+1}\backslash\{0\}\ , t \in [0,1] \quad .$$

4.3 Problèmes de Compacité

Nous avons indiqué dans l'introduction qu'une différence de topologie entre les ensembles de niveau signifiait la présence de points critiques. En fait une telle conclusion est valide, uniquement si on fait des hypothèses suffisantes de compacité. Considérons à cet effet l'exemple suivant :



Considérons $E = \mathbb{R}$, et la fonction f définie par

$$f(t) = e^{-t} \quad .$$

On a $f'(t) = -e^{-t}$, et donc f n'a pas de point critique. Si $a_1 < 0$, alors $f^{a_1} = \emptyset$, alors que pour $a_2 > 0$, on a $f^{a_2} =]-\log a_2, +\infty[$, donc a le type d'homotopie d'un point. On voit donc que f^{a_1} et f^{a_2} n'ont pas le même type d'homotopie.

Que s'est-il passé ? En fait, on constate que

$$f'(m) \to 0 \quad m \to +\infty\,, \tag{4.5}$$

et donc tout se comporte comme si le point critique qui devrait être induit par la différence de topologie était "rejeté à l'infini". De manière plus générale, on peut considérer des suites de points qui se comportent de façon similaire à (5). On introduit donc la définition suivante :

Définition 6. (suites de Palais-Smale). Soit F une fonctionnelle de classe C^1 sur un espace de Banach. On dit qu'une suite $(u_n)_{n\in\mathbb{N}}$ est une suite de Palais-Smale pour la fonctionnelle F si

$$|F(u_n)| \leq C \quad , \forall n \in \mathbb{N} \text{ où } C \in \mathbb{R}^+ \tag{4.6}$$

$$\|dF(u_n)\|_{E^*} \to 0 \quad \text{lorsque } n \to +\infty \tag{4.7}$$

[La notion de suite de Palais-Smale est, bien entendu, toujours relative à une fonctionnelle].

Les techniques du calcul des variations aboutissent bien souvent (comme dans l'exemple que nous avons donné en début de ce paragraphe) non pas directement à l'existence de points critiques mais plutôt à l'existence de suite de Palais-Smale [on peut comparer également avec les problèmes de minimisation, où on obtient des suites minimisantes]. Afin de conclure à l'existence d'un *vrai* point critique, il faut une condition supplémentaire de compacité [de même que pour trouver un vrai minimum pour un problème de minimisation on doit pouvoir en général extraire une sous-suite convergente d'une suite minimisante]. Il est alors utile d'introduire la définition suivante :

Définition 7 (condition de Palais-Smale). On dit qu'une fonctionnelle F définie sur un espace de Banach E vérifie la condition de Palais-Smale si de toute suite de Palais-Smale on peut extraire une sous-suite qui converge fortement dans F (la limite est alors un point critique).

En d'autres termes, F vérifie (P.S) (écriture abrégée pour : F vérifie la condition de Palais-Smale) si $\forall (u_n)_{n\in\mathbb{N}}$ une suite de E, telle que

$$\begin{aligned} &|F(u_n)| \leq C \\ &|dF(u_n)| \to 0 \quad \text{dans } E^* \quad , \end{aligned}$$

on peut extraire de u_n une sous-suite qui converge fortement dans E.

Dans certaines situations (voir Chapitre VI), il est utile d'introduire la définition suivante.

Définition 8. (condition $(\text{P.S})_\beta$). Soit $\beta \in \mathbb{R}$. On dit que f vérifie la condition de Palais-Smale au niveau β, (et on écrit en abrégé F vérife (ou satisfait*) la condition $(\text{P.S.})_\beta$, si pour toute suite u_n telle que

$$\begin{aligned} &F(u_n) \to \beta \\ &dF(u_n) \to 0 \quad \text{dans } E^*, \end{aligned}$$

on peut extraire une sous-suite convergente dans E.

Remarque : Il se peut évidemment qu'une telle sous-unité u_n n'existe pas auquel cas F vérifie $(\text{P.S.})_\beta$, par convention.

Exemples :

* Si $F = 0$, F ne satisfait pas (P.S) sur E.

* sur $\mathbb{R}, f(t) = e^{-t}$ ne vérifie pas la condition de Palais-Smale.

* Si F est C^1 sur $\mathbb{R}^N$ et si pour tout $a \in \mathbb{R}$, F^a est compact, alors F vérifie (P.S.).

Une dernière définition nous sera utile.

Définition 9 : On appelle valeur critique généralisée β de F (fonctionnelle C^1 sur E Banach), un nombre $\beta \in \mathbb{R}$ tel qu'il existe $(u_n)_{n\in\mathbb{N}}$ suite dans Etelle que

$$\begin{cases} F(u_n) \to \beta \\ dF(u_n) \to 0 \quad \text{lorsque } n \to +\infty. \end{cases}$$

Autrement dit dans la définition des valeurs critiques généralisée, on autorise celles qui correspondent aux " points critiques à l'infini " (terminologie introduite par A. Bahri, que nous préciserons ultérieurement).

Le lien entre cette définition et la condition de Palais-Smale est le suivant.

Proposition 2 : Si F vérifie la condition de Palais-Smale, toute valeur critique généralisée est une valeur critique.

Il est bien sûr évident que l'inverse est vrai également, c'est-à -dire que (même si F ne vérifie pas P.S.) toute valeur critique est une valeur critique généralisée.

Nous somme maintenant en mesure de présenter un des outils fondamentaux de la théorie variationnelle.

4.4 Le Lemme de déformation

Théorème 1 (lemme de déformation sans point critique). Soit F une fonctionnelle de classe C^1 sur un espace de Banach E. Soit $a < b$. On suppose que F n'a pas de valeur critique généralisée dans l'intervalle $[a-\bar{\epsilon}, b+\bar{\epsilon}]$, où $\bar{\epsilon} > 0$ (petit). Il existe alors une rétraction de F^b sur F^a, c'est-à -dire une fonction $\Phi \; : \; F^b \times [0,1] \to F^b$, continue telle que

$$\begin{aligned} \phi(v,0) &= v \, , \quad \forall v \in F^b \\ \phi(v,1) &\in F^a, \quad \forall v \in F^b \\ \phi(v,t) &= v, \quad \forall v \in F^a \quad . \end{aligned}$$

Ainsi F^b se rétracte par déformation sur F^a, s'il n'a pas par de valeur critique généralisée entre a et b, et donc F^b et F^a ont le même type d'homotopie. Dans la pratique, le résultat précédent permet de trouver des valeurs critiques généralisées lorsqu'on arrive à prouver, que deux ensembles de niveau ont des topologies différentes. C'est ainsi que l'on prouve la plupart des résultats Min Max (voir paragraphe suivant).

L'idée de la démonstration du théorème 1, lorsque E=H est un espace de Hilbert, est de "pousser" l'ensemble F^b le long des lignes de flot associées au gradient de F (qui représente, rappelons-le, la direction de plus forte de pente). En intégrant le champ de gradient on arrive donc à déformer F^b sur F^a, *sauf* si celui-ci a tendance à s'annuler: cette éventualité sera exclue grâce à l'hypothèse F n'a pas de valeur critique généralisée dans [a,b].

Afin de mettre en forme ces idées, nous allons commencer par traiter le cas où E=H est un espace de Hilbert, et F est plus régulière. Le cas général sera entrevu plus tard: nous introduirons en particulier la notion de "pseudo-gradient" (qui remplace le gradient pour un Banach).

a) Démonstration pour un espace de Hilbert, dans le cas où F est de classe C^2 et $|\nabla^2 F|$ borné.

Dans le cas d'un Hilbert H, nous pouvons identifier H et H^*: le gradient de F, au point u sera alors l'élément de H associé à $dF(u)$, élément de H^* par

$$(X, \nabla F(u)) = < dF(u), X > \quad \forall \in H \quad ,$$

où les parenthèses représentent le produit scalaire pour H et les crochets expriment la dualité entre H et H^*. On a par ailleurs

$$|\nabla F|_H = \|dF(u)\|_{H^*} \quad , \tag{4.8}$$

comme on vérifie aisément.

Rappellons que le gradient est porté par la direction de plus forte variation de F. Remarquons enfin que la définition du gradient dépend du choix du produit scalaire (et donc la direction de plus forte descente également.)

Comme nous avons supposé (dans cette partie) que F est de classe C^2 l'application

$$\begin{aligned} \nabla F \ : \ & H \to H \\ & v \to \nabla F(v) \end{aligned}$$

est de classe C^1. Considérons alors l'équation du flot associé au gradient de F (ou plutôt à son opposé!)

$$\begin{cases} \frac{d}{dt}M(t) & = -\nabla F(M(t)) \\ M(u) & = v \end{cases} \tag{4.9}$$

où v est fixé, donné dans F^b.

Comme le champ de vecteurs ∇F est lipschitzien, on peut appliquer le théorème de Cauchy-Lipschitz qui nous montre que (9) à une solution sur un petit intervalle $[0, \delta[$, où δ est un nombre positif. On vérifie alors que, pour $t \in [0, \delta[$

$$\frac{d}{dt} F(M(t)) = -\nabla F(M(t)) \cdot \frac{dM(t)}{dt} = -|\nabla F(M(t))|^2 \quad . \tag{4.10}$$

On vérifie ainsi que

$$\frac{d}{dt} F(M(t)) \leq 0,$$

F décroît donc le long des trajectoires.

a) L'intervalle d'existence de la solution est $[0, +\infty[$ (pour les temps positifs). On suppose en effet qu'il existe une constante $C > 0$ telle que

$$\|\nabla^2 F(u)\| < C, \quad \forall v \in H \, . \tag{4.11}$$

Raisonnons par l'absurde et supposons que l'intervalle maximal d'existence est $[0, \delta_0[$ ($\delta_0 > 0$). On a alors, en vertu de (11), pour tout $t \in [0, \delta_0[$,

$$\|\nabla F(M(t))\| \leq \|\nabla F(M(0))\| + C\|M(t) - M(0)\|,$$

et donc

$$\frac{d}{dt}\|M(t) - M(0)\| \leq \|\frac{d}{dt}(M(t))\| \leq C_0 + C\|M(t) - M(0)\| \, .$$

On en déduit (Lemme de Gronwall)

$$\|M(t) - M(0)\| \leq \frac{C}{C} \exp Ct \leq \frac{C_0}{C} \exp C\delta_0 \, , \quad \forall t \in [0, \delta_0[\quad ,$$

et donc $M(t)$ reste borné sur $[0, \delta_0[$, indépendamment de t. On voit alors, en revenant à (9) que $M(t)$ converge vers une limite lorsque $t \to \delta_0$. Grâce au théorème de Cauchy-Lipschitz on peut alors prolonger la solution en δ_0. L'intervalle maximal d'existence est alors $[0, \delta_0 + \mu$ [($\mu > 0$), ce qui contredit l'hypothèse de départ.

b) Il existe t_0, tel que pour tout $v \in F^b, F(M(t_0)) < a$. A cet effet, montrons qu'il existe $\eta_0 > 0$ tel que

$$|\nabla F(v)| \geq \eta_0 \quad \nabla v \in F^b \backslash F^{a-\epsilon} \tag{4.12}$$

Supposons que (12) ne soit pas vrai : alors il existerait une suite $v_n \in H$, telle que

$$F(v_n) \in [b, a - \epsilon] \tag{4.13}$$
$$\|\nabla F(v_n)\| \to 0 \quad . \tag{4.14}$$

Quitte à extraire une sous-suite, on peut toujours supposer qu'il existe $\beta \in [b, a - \epsilon]$ tel que

$$F(v_n) \to \beta$$

β serait alors une valeur critique généralisée, ce qui est contraire à l'hypothèse.

En retournant à (14), on obtient à partir de (12),

$$\frac{d}{dt}F(M(t)) = -|\nabla F(M(t))|^2 \leq -\eta_0^2 \quad .$$

Donc

$$F(M(t)) \leq F(M(0)) + \eta_0^2 t \leq b - \eta_0^2 t \quad .$$

et

$$F(M(t)) \leq a - \epsilon \quad ,$$

dès que $t \geq t_0 = \frac{b-a+\epsilon}{\eta_0^2}$.

C) Construction de la rétraction Φ. Considérons tout d'abord le " flot " $\tilde{\Phi}$ associé au champ de vecteurs $-\nabla F$. Celui-ci associe à $v \epsilon F^b$, pris compte tenu de la donnée initiale au temps zéro, la solution $M(t)$ de (9), c'est-à-dire

$$\begin{cases} \tilde{\Phi}(v,t) = M(t) \ , \ \forall v \in F^b \ , \ \forall t \in \mathbb{R}^+ \\ M(0) = v \quad . \end{cases}$$

Le théorème de Cauchy-Lipschitz nous enseigne alors que l'application $\tilde{\Phi}$ est classe C^2 sur $F^b \times \mathbb{R}^+$. De plus il résulte de b) que $\forall v \in F^b$

$$\tilde{\Phi}(v,t) \in F^a \quad \text{si } t \geq t_0 \quad .$$

Pour v donné dans F^b considérons le temps d'entrée dans F^a

$$\tilde{t}_0(v) = \inf\{t \in \mathbb{R}^+ , \ \tilde{\Phi}(v,t) \in F^a\} \quad ,$$

de sorte que

$$\tilde{t}_0(v) = 0 \ , \ \text{si } v \in F^a \quad .$$

On vérifie ensuite que l'application

$$v \to \tilde{t}_0(v)$$

est continue (en vertu de (12) et du théorème d'inversion locale, voir chap.V). Posons alors

$$\begin{aligned} \tilde{\tilde{\Phi}}(v,t) &= \tilde{\Phi}(v,t) && \text{si } t < \tilde{t}_0(v) \\ &= \tilde{\Phi}(v,\tilde{t}_0(v)) && \text{sinon} \quad . \end{aligned}$$

Il est alors clair que $\tilde{\tilde{\Phi}}$ est continue (composée de fonctions continues). Nous sommes maintenant en mesure de définir Φ par

$$\Phi(v,t) = \tilde{\tilde{\Phi}}(v, \frac{t}{t_1}) \text{ où}$$

$$t_1 = \sup\{\tilde{t}_0(v), v \in F^b\} < +\infty \text{ (cf partie b)} \quad .$$

On vérifie que Φ a les propriétés annoncées:
- Φ est continue de $F^b \times [0,1] \to F^a$
- $\forall v \in F^b, \quad \Phi(v,1) \in F^a$
- $\forall v \in F^a \quad \Phi(v,t) = v$.

Ceci termine donc la preuve du théorème dans le cas considéré.

Voyons maintenant comment adapter la preuve précédente dans le cas où F est (seulement) C^1 et E est un espace de Banach. Il est clair que dans ce cas, on ne peut plus définir la notion de gradient qui repose sur l'identification de H et H^* par l'intermédiaire du produit scalaire. On remplace alors cette notion par celle de pseudo-gradient.

Définition 10 : Soit F une fonction de classe C^1 sur E, espace de Banach. On dit qu'un champ de vecteurs $X(\cdot)$ est un pseudo-gradient pour F sur E si et seulement si

$$\text{L'application } u \to X(u) \text{ est continue de } E \text{ vers } E, \text{ localement Lipschitz} \tag{4.15}$$

$$\|X(u)\| \leq 2\min\{\|dF(u)\|_{E^*}, 1\} \tag{4.16}$$

$$< dE(u), X(u) > \geq \min\{\|dF(u)\|_{E^*}, 1\}\|dF(u)\|_{E^*} \quad . \tag{4.17}$$

Remarque: La Notion de pseudo-gradient est bien sûr toujours relative à une fonctionnelle.

Nous démontrerons plus loin le résultat suivant:

Proposition 3. On peut construire un pseudo-gradient pour toute fonction de classe C^1 sur un espace de Banach E séparable.

Démonstration du Théorème 1 dans le cas général.

Soit X un pseudo-gradient pour F. On remplace l'équation (9) par l'équation suivante, pour $v \in F^b$

$$\begin{cases} \frac{d}{dt}M(t) & = -X(u) \\ M(0) & = v \quad . \end{cases} \tag{4.18}$$

D'après le théorème de Cauchy-Lipschitz, la solution de (15) existe sur un petit intervalle $[0, \delta_0[$. En fait, d'après (16), X est uniformément borné

$$\|X(u)\| \leq 2 \quad , \tag{4.19}$$

et en raisonnant par l'absurde, comme précédemment, on montre que l'intervalle maximal d'existence est $[0, +\infty[$. Par ailleurs F décroît bien le long des courbes intégrales de X. En effet

$$\frac{d}{dt}F(M(t)) = - < dF(M(t), X(M(t)) > \leq -\min\{\|X(M(t))\|^2, 1\} \leq 0 \quad , \tag{4.20}$$

(on a utilisé (17)).

Le reste de la démonstration est ensuite pratiquement identique.

Donnons maintenant les principales idées de la preuve de la proposition 3 que nous avons laissé en suspens.

Preuve de la proposition 3. Pour vérifier (17), il est clair que l'on doit avoir $X(u) = 0$ lorsque $dF(u) = 0$. Considérons donc l'ensemble

$$K_c = \{u \in F, dF(u) = 0\}$$

des points critiques de F, et son complémentaire

$$\tilde{E} = E \backslash K_c = \{u \in F, \quad dF(u) \neq 0\} \quad .$$

Afin de mettre en lumière les idées de la preuve, considérons tout d'abord le cas simple suivant.

a) Le cas dim E $< +\infty$ et $K_c = 0$. On a alors $\tilde{E} = E$. Pour $u \in E$, considérons l'ensemble $\Lambda(u)$ des vecteurs $Y \in E$ tels que

$$\|Y\| < 2\min\{\|dF(u)\|, 1\} \tag{4.21}$$

$$< dF(u), Y >> \min\{\|dF(u)\|, 1\}\|dF(u)\| \quad . \tag{4.22}$$

On vérifie aisément que l'ensemble $\Lambda(u)$ est non vide et ouvert. Pour chaque u donné, choisissons alors un élément dans $\Lambda(u)$, que nous noterons Y_u. Comme dF est une application continue, on voit qu'il existe un voisinage $W(u)$ de u dans E, tel que $Y_u \in \Lambda(v)$ pour tout $v \in W(u)$ (c'est-à -dire que Y_u vérifie (21) et (22) avec u remplacé par v).

Considérons alors le recouvrement $\bigcup_{u \in E} W(u)$. En appliquant le théorème de Borel-Lebesque, on peut en extraire un recouvrement dénombrable $W(u_i)$ $\quad (i = 1$ à $+\infty)$ tel que

$$\bigcup_{i=1}^{\infty} W(u_i) = E$$

Pour tout compact K de $\tilde{E}$, K ne contient qu'un nombre fini de point u_i.

La connaissance des points u_i, et donc de Y_{u_i} va alors nous permettre de construire le champs $X(u)$ grâce à une partition de l'unité relative aux ouverts $W(u_i)$. Il existe des fonctions $\varphi_i \in C_c^{\infty}(W(u_i))$ telle que

$$\sum_{i=1}^{\infty} \varphi_i(u) \equiv 1 \quad \text{sur } E \quad .$$

Posons alors

$$X(u) = \sum_{i=1}^{\infty} \varphi_i(u) Y_{u_i} \quad .$$

On vérifie, comme les relations (21) et (22) peuvent être superposées par linéarité et que $X(\cdot)$ vérifie bien (16).(17). Par ailleurs il est clair que $X \in C^\infty$, et nous avons donc construit un pseudo- gradient C^∞.

b) Le cas général. On définit comme précédemment Y_u pour $u \in \tilde{E}$ (ce qui est possible). On a alors le recouvrement

$$\tilde{E} = \bigcup_{u \in \tilde{E}} W(u) \quad . \tag{4.23}$$

On ne peut évidemment pas invoquer ici le théorème de Borel-Lebesgue. En revanche, on peut montrer le résultat suivant :

Lemme 1 : V un espace métrique séparable. On peut extraire de tout recouvrement ouvert de V, (i.e $V = \bigcup_{j \in J} W_j$, où $\forall j \in J, W_j$ est un ouvert) un recouvrement localement fini $V = \bigcup_{i \in I} W_i, I \subset J$ c'est-à-dire tel que, pour tout élément $v \in V$, il existe un voisinage de v, U, tel que U n'intersecte qu'un nombre *fini* des W_i (i.e Card $\{i \in I, U \cap W_i \neq 0\} < +\infty$).

Admettons ce résultat. Considérons alors un recouvrement localement fini de $\tilde{E}$ extrait du recouvrement (23)

$$\tilde{E} = \bigcup_{i \in I} W(u_i) \quad .$$

Il nous reste maintenant à construire une partition de l'unité subordonnée au recouvrement précédent. Considérons pour $i \in I$ la fonction ρ_i définie par

$$\rho_i(u) = dist(u, E \backslash W(u_i)) \tag{4.24}$$
$$= \inf\{\|u - v\|, v \in E \backslash W(u_i)\} \quad . \tag{4.25}$$

On vérifie ρ_i est lipschitzienne, de constante de Lipschitz inférieure à 1. Par ailleurs pour u donné, il n'existe qu'un nombre fini d'indice i tels que $\rho_i(u) \neq 0$. Posons alors

$$\varphi_i(u) = \frac{\rho_i(u)}{\sum_{j \in I} \rho_j(u)} \quad .$$

(le dénominateur n'étant qu'une somme finie). On vérifie que

$$\sum \varphi_i(u) = 1 \quad , \forall u \in \tilde{E}$$
$$0 < \varphi_i \leq 1$$

et que $\qquad \varphi_i \equiv 0 \quad$ en dehors de $W(u_i) \quad .$

Comme le recouvrement est localement fini, pour tout $u \in \tilde{E}$, il existe un voisinage U de u tel que

$$Card\{j \in I, \rho_j(v) \neq 0, v \in U\} = N = C^{te} \quad ,$$

et donc φ_i est localement Lipschitzienne. On introduit ensuite

$$X(u) = \sum_{i \in I} \varphi_i(u) Y_{u_i} \quad ,$$

et on vérifie que X convient comme dans la partie a).

COMPLEMENT : Déformation globale.

Au cours du paragraphe précédent nous avons construit une déformation $\Phi \ : [0,1] \times F^b \to F^b$.

Il est parfois souhaitable de construire une déformation globale. On montre alors la variante ci-dessus du Lemme de déformation.

Théorème 2 (déformation globale). Supposons que F n'ait pas de valeur critique généralisée dans $[a-\epsilon, b+\epsilon]$. Alors il existe une déformation $\phi \ : E \times [0,1] \to E$ telle que

$$\begin{aligned}
&\phi(\cdot, 0) = Id_E \\
&\phi(F^b, t) \subset F^b \quad \forall t \in [0,1] \\
&\phi(F^b, 1) \subset F^a \\
&\phi(u, t) = u \quad \forall u \in F^a, \forall t \in [0,1] \quad .
\end{aligned}$$

Preuve : Elle est tout à fait similaire à celle du théorème 1. Il suffit de modifier la définition du champ de vecteurs. Soit φ une fonction C^∞ de $\mathbb{R} \to \mathbb{R}^*$ telle que

$$\begin{aligned}
\varphi(t) \ &= 1, \qquad \text{si } t \leq b \\
\varphi(t) \ &= 0 \qquad \text{si } t \geq b+1 \quad .
\end{aligned}$$

on considère alors le flot associé au champ de vecteurs $\varphi(F(u))X(u)$ que l'on intègre. On résout donc

$$\begin{cases} \frac{dM(t)}{dt} & = \varphi(F(M(t))X(M(t)) \\ M(u) & = v \end{cases}$$

et ceci fournit la déformation voulue.

4.5 Lemme de déformation et condition de Palais-Smale

Comme nous l'avons déjà remarqué, lorsque F^b et F^a n'ont pas le même type d'homotopie, le Théorème 1 nous fournit l'existence d'une valeur *critique généralisée* $\beta \in [a,b]$.

Bien évidemment, nous sommes en fait intéressé par les valeurs critiques (puisque nous cherchons des points critiques). Lorsque F satisfait la condition P.S. nous avons vu que {valeurs critiques}={valeurs critiques généralisées} et le Théorème 1 s'exprime alors par :

" Si F vérifie P.S. et si F^a et F^b n'ont pas le même type d'homotopie, alors il existe une valeur critique $\beta \in [a, b]$ "

Ainsi lorsqu'on traite un tel problème il est essentiel de vérifier si F satisfait (P.S.), pour pouvoir appliquer le principe précédent.

Néanmoins, si F ne satisfait pas (P.S.) la partie n'est pas perdue pour autant. Dans ce cas, on doit ***étudier les suites de Palais-Smale*** et comprendre par quels mécanismes la perte de compacité se produit : bien souvent cette perte de compacité est due à des symétries internes du problème. Ceci se produit lorsque la fonctionnelle est laissée invariante par l'action de certains groupes sur l'espace fonctionnel, et si l'orbite d'un point par le groupe est non compacte (au sens de la topologie forte de l'espace).

Une telle analyse permet parfois de déterminer les niveaux β tels que $(P.S.)_\beta$ ne soit pas vérifiées. De manière plus délicate, elle peut permettre de connaître la contribution topologique aux ensembles de niveau des suites de Palais-Smale (voir Chapitre IX).

4.6 Applications : méthodes de Min-Max

Parmi les multiples façons d'utiliser le Lemme de déformation l'une des plus utilisées et des plus faciles à mettre en oeuvre est la méthode de Min-Max. Nous en donnerons une formulation générale lorsque nous étudierons les variétés hilbertiennes (chap. VII). Néanmoins, un exemple très simple d'une telle méthode est fourni par le Lemme de Col, que nous étudierons au chapitre suivant.

Exercices

Exercice 1. Soit Ω un domaine borné régulier de $\mathbb{R}^N$. Soit V une fonction de $\Omega\times\Omega\backslash\Delta \to \mathbb{R}$ de classe C^∞, où $\Delta = \{(x, y) \in \Omega^2 \quad x = y\}$ On suppose que

$$V(x,y) \to +\infty \quad \text{lorsque } (x,y) \to \Delta$$
$$V(x,y) \to +\infty \quad \text{lorsque } x \to \partial\Omega, \quad y \to \partial\Omega \quad .$$

1) Montrer que $\inf\limits_{(x,y)\in\Omega^2\backslash\Delta} V(x,y)$ est atteint.

2) On suppose que $\theta \in \Omega$, et $B_r(0) \subset \Omega$

$$\begin{array}{rcl} S^1 & \to & \Omega \times \Omega\backslash\Delta \\ \text{expi}\,\theta & \to & (0, \text{ rexpi}\,\theta) \end{array}$$

n'est pas homotope à une application constante de $S^1 \to \Omega \times \Omega \backslash \Delta$.

En déduire qu V a au moins deux points critiques sur $\Omega \times \Omega \backslash \Delta$.

Exercice 2. - (Lemme de déformation avec point critique). Soit H un espace de Hilbert, et F une foncitonnelle C^1 sur H. On suppose que F vérifie la condition de Palais-Smale sur H. Soit $\beta \in \mathbb{R}$. On considère l'ensemble

$$K_\beta = \{x \in H, F(x) = \beta \quad , \quad dF(x) = 0\} \quad .$$

A) 1. Montrer que K_B est compact.

2. Pour $\delta > 0$, on considère l'ensemble

$$N_\delta = \{x \in H, |F(x) - \beta| \leq \delta \quad , \quad \|dF(x)\| \leq \delta\} \quad .$$

Montrer que si $\delta_1 < \delta_2, N_{\delta_1} \subset N_{\delta_2}$ et

$$\bigcap_{\delta > 0} N_\delta = K_\beta \quad .$$

3. Pour $\delta > 0$ on considère

$$U_\delta = \{x \in H, \text{ dist } (x, K_\beta) \leq \delta\} \quad .$$

Montrer que $\forall \delta > 0$, il existe $\delta_1 > 0$ tel que

$$U_{\delta_1} \subset N_\delta \quad .$$

(On utilisera le fait que K_β est compact et on pourra raisonner par l'absurde).

4. Montrer que pour tout $\delta > 0$, il existe δ_0 tel que

$$N_{\delta_0} \subset U_\delta$$

(on pourra raisonner par l'absurde) .

5. Déduire de 3) et 4) que pour tout $\delta > 0$, il existe δ_0 et $\delta > 0$ tels que

$$N_{\delta_0} \subset U_\delta \subset U_{2\delta} \subset N_\delta \quad .$$

6. Construire une fonction η continue et lipschitzienne sur H, $\delta_1 < \delta_0$ tel que

$$\eta \equiv 1 \quad \text{sur } H \backslash N_{\delta_0}$$
$$\eta \equiv 0 \quad \text{sur } N_{\delta_1}$$

et $0 \leq \eta \leq 1$, sur H.

B) . On suppose ici que F est C^2, que $|\nabla^2 F|$ est borné. Pour $u_0 \in H, \delta, \delta_0, \delta_1$ et δ comme en A) (question 5), et η construite en A6), on considère l'équation ordinaire sur H,

$$\left\{ \begin{array}{ll} \frac{d}{dt}(M(t)) & = -\eta(M(t)\frac{\nabla F(M(t)}{\|\nabla F(M(t))\|+1} \\ M(0) & = v_0 \end{array} \right. \quad \text{(I)}$$

1. Montrer que l'intervalle maximal d'existence pour I est $\mathbb{R}$ tout entier
2. Montrer que
$$\frac{d}{dt}F(M(t) = -\eta(M(t))\frac{\|\nabla F(M(t)\|^2 \leq 0}{\|\nabla F(M(t)\| + 1}$$
3. On suppose que $v_0 \notin N_\sigma$, montrer que
$$\text{mes } \{t \in \mathbb{N}^+, \quad M(t) \notin N_{\delta_0}\} \geq \rho$$
(On utilisera le fait que $\|\frac{d}{dt}M(t)\| \leq 1$, et la question A5)
4. En déduire que, si $v_0 \notin N_\delta$.
$$E(M(1)) \leq E(v_0) - \text{mes } \{t \in \mathbb{R}^+, M(t) \notin N_{\delta_0}\} \cdot -\frac{\delta_0^2}{\delta_0+1}$$
$$\leq E(v_0) - \frac{\rho\delta_0^2}{\delta_0+1} \quad .$$

C) 1. On pose $\Phi(t, v_0) = M(t)$, solution de (I) avec pour donnée initiale v_0. Montrer que $\forall \varepsilon > 0, \Phi$ est continue de $[0,1] \times F^{\beta+\varepsilon} \to F^{\beta+\varepsilon}$ et que, si on choisit ε suffisamment petit, alors
$$\Phi(1,u) \in F^{\beta-\varepsilon}, \forall u \in F^{\beta+\varepsilon} \backslash N_\delta$$
$$\text{et} \quad \Phi(1,u) \in F^{\beta-\varepsilon} \subset N_\delta \quad , \quad \forall u \in F^{\beta+\varepsilon} \quad .$$

2. De manière plus générale, soit N un voisinage quelconque de K_β. Montrer qu'il existe $\delta > 0$, tel que $N_\delta \subset N$, et en déduire qu'il existe $\varepsilon > 0$, et Φ application continue de $[0,1] \times F^{\beta+\varepsilon} \to F^{\beta+\varepsilon}$, telle que
$$\Phi(0,u) = u, \quad \forall u \in F^{\beta+\varepsilon}$$
$$F(\Phi(t,u) \quad \text{est une fonction décroissante}$$
$$\Phi(1,u) \in F^{\beta-\varepsilon} \quad, \forall u \in F^{\beta+\varepsilon} \backslash N$$
$$\Phi(1,u) \in F^{\beta-\varepsilon} \subset N \quad, \forall u \in F^{\beta+\varepsilon}$$

D) Généraliser au cas où F est seulement C^1, en utilisant un pseudo-gradient.