INTRODUCTION TO THE THEORY OF CRITICAL POINTS
THE MOUNTAIN PASS THEOREM
EKELAND'S VARIATIONAL PRINCIPLE

J. CHABROWSKI

The University of Queensland
Department of Mathematics

Instructional Workshop in Analysis and Geometry

Canberra, 23 January - 10 February, 1995

1 INTRODUCTION

The main purpose of these lectures is to present a short introduction to the theory of critical points of C^1-functionals on a Banach space.

A number of problems in the theory of differential equations can be expressed in the form of an equation

$$Au = 0, \tag{1.1}$$

where $A : X \to Y$ is a mapping between Banach spaces X and Y. The interesting case is the situation where this equation has a variational structure, that is, there exists a functional $\phi : X \to \mathbb{R}$ such that

$$\langle A(u), v\rangle = \lim_{t\to 0} \frac{\phi(u + tv) - \phi(u)}{t},$$

where $Y = X'$, $\langle \cdot, \cdot \rangle$ is a duality pairing between X and its dual X'. In this case we can write $A = \phi'$ and equation (1.1) becomes

$$\langle \phi'(u), v\rangle = 0 \text{ for each } v \in X. \tag{1.2}$$

Equation (1.2) says that solutions of (1.1) are critical points of the functional ϕ. By writing equation (1.2) we have expressed equation (1.1) in a weak (distributional) form. The problem that we have to solve is to find critical points of ϕ. If $X = \mathbb{R}_N$, the obvious candidates for critical points are local maxima and minima of ϕ. The situation is more complicated if ϕ is a functional on an infinite - dimensional space. For example, consider a functional

$$I(u) = \int_0^\pi (\frac{1}{2}|u'|^2 - \frac{1}{4}u^4)\, dx,\ u' = \frac{du}{dx},$$

for $u \in W_\circ^{1,2}(0,\pi)$. It is easy to check that I is a (Fréchet) differentiable on $X = W_\circ^{1,2}(0,\pi)$ and has a local minimum at $u \equiv 0$. We now observe that

$$I(tu) = \int_0^\pi (\frac{t^2}{2}|u'|^2 - \frac{t^4}{4}u^4)\, dx \to -\infty \text{ as } t \to \infty$$

and also for $k \in \mathbb{N}$

$$I(\sin kx) \geq \frac{\pi}{4}k^2 - \frac{\pi}{4} \to \infty \text{ as } k \to \infty,$$

so I is neither bounded from below, nor from above. It is not clear whether $u \equiv 0$ is the only critical point of I. Therefore the problem is how to check whether I has any critical points other than $u \equiv 0$.

The method that can be used to identify the other critical points is a min - max method (the Lusternik - Schnirelman theory of critical points). In this approach we set

$$c = \inf_{A \in \mathcal{A}} \sup_{u \in A} \phi(u),$$

where $\mathcal{A}$ is a collection of subsets of X. The aim of the theory of critical points is to show that a set defined by

$$K_c = \{u \in X;\ \phi(u) = c,\ \phi'(u) = 0\}$$

is not empty. The main problem is to choose a good class of sets $\mathcal{A}$ and impose conditions on ϕ guaranteeing that $K_c \neq \emptyset$. A central result which has been extensively and successfully used to find critical points, is the mountain pass theorem. To describe it, suppose that

$$b = \inf_{\|u\|=r} \phi(u) > \max\{\phi(0), \phi(e)\},\ r > 0,$$

with $\|e\| > r$. If we interpret $\phi(u)$ as the altitude at u, then points $(0, \phi(0))$ and $(e, \phi(e))$, belonging to the graph of ϕ, are separated by a mountain range. We expect the existence of a mountain pass containing a critical point between them. This is a motivation for considering the following min - max level

$$c = \inf_{\gamma \in \Gamma} \max_{t \in [0,1]} \phi(\gamma(t)),$$

where $\Gamma = \{\gamma \in C([0,1], X);\ \gamma(0) = 0 \text{ and } \gamma(1) = e\}$. The level c can be interpreted as follows: suppose that we want to walk from $(0, \phi(0))$ to $(e, \phi(e))$ climbing as little as possible. To achieve this we must find a path crossing the mountain pass range over the lowest mountain pass. The top point of this mountain pass should be a critical point of ϕ.

In these lectures we shall discuss the mountain pass theorem and its generalizations as well as applications to the boundary value problems. In the final part of these lectures we present the Ekeland variational principle.

Throughout these lectures we use standard terminology and notations. Let X be a Banach space equipped with norm $\|\cdot\|$. By $\langle\cdot,\cdot\rangle$ we denote the duality pairing between X and X'. We denote the weak convergence in X and X' by " $\rightharpoonup$ " and the strong convergence " $\rightarrow$ ".

Let X and Y be two Banach spaces. A map $F : X \rightarrow Y$ is said to be Fréchet differentiable at $u \in X$ if there is an $F'(u) \in L(X,Y)$ such that

$$F(u+h) = F(u) + F'(u)h + \omega(u,h)$$

and $\omega(u,h) = o(\|u\|)$ as $h \rightarrow 0$. Here $\omega(u,h) = o(\|u\|)$ denotes Landau's symbol. If F is differentiable at every point $u \in X$ and $F' : X \rightarrow L(X,Y)$ is continuous, then F is said to be continuously differentiable on X. We express this by writing $F \in C^1(X,Y)$.

A functional $\phi : X \rightarrow \mathbb{R}$ is said to be Gâteaux differentiable if there is an $u^* \in X'$ such that

$$\lim_{t \rightarrow 0} t^{-1}\big(\phi(u_\circ + th) - \phi(u_\circ)\big) = \langle u^*, h\rangle = u^*(h)$$

for all $h \in X$. This means that ϕ has a directional derivative in every direction h. The functional u^* is called the Gâteaux derivative of $\phi(u_\circ)$ at $u_\circ$ and we denote it also by $\phi'(u_\circ) = u^*(h)$. We recall that if $\phi : X \rightarrow \mathbb{R}$ has a Gâteaux derivative $\phi'(u)$ at every point u in a neighbourhood of a point $u_\circ$ and $\phi'(u)$ is continuous at $u_\circ$, then ϕ is Fréchet differentiable at $u_\circ$ and the Fréchet derivative of ϕ at $u_\circ$ is equal to the Gâteaux derivative at this point.

In applications of min - max theorems we shall use functions spaces $L^p(\Omega)$, $1 \leq p \leq \infty$, and $H^1_\circ(\Omega)$, where Ω is a bounded domain in $\mathbb{R}_N$. These spaces are defined as follows. $L^p(\Omega)$, $1 \leq p < \infty$, is a space of Lebesgue measurable functions u on Ω satisfying $\int_\Omega |u|^p\,dx < \infty$ and equipped with norm

$$\|u\|_p = \left(\int_\Omega |u(x)|\,dx\right)^{\frac{1}{p}}.$$

$L^\infty(\Omega)$ is a space of Lebesgue measurable functions u on Ω which are essentially bounded and equipped with norm

$$\|u\|_\infty = \operatorname*{ess\,sup}_\Omega |u(x)|.$$

A Sobolev space $H^1_\circ(\Omega)$ is defined as the closure of $C^\infty_\circ(\Omega)$ with respect to norm

$$\|u\| = \left(\int_\Omega |Du(x)|^2\,dx\right)^{\frac{1}{2}}.$$

Spaces $L^p(\Omega)$, $1 \leq p \leq \infty$, and $H^1_\circ(\Omega)$ are Banach spaces. The dual space of $H^1_\circ(\Omega)$ is denoted by $H^{-1}(\Omega)$, that is, $H^1_\circ(\Omega)' = H^{-1}(\Omega)$.

We shall frequently refer to the following estimate: for every $u \in H^1_\circ(\Omega)$ we have

$$\|u\|_s \leq c|\Omega|^{\frac{1}{s}-\frac{1}{2^*}}\|Du\|_2$$

for $1 \leq s \leq 2^*$, where $c > 0$ is a constant depending on N and $|\Omega|$ denotes the Lebesgue measure of Ω and $2^* = \frac{2N}{N-2}$ is the so called critical Sobolev exponent. This inequality (known as the Sobolev inequality) expresses the fact that $H^1_\circ(\Omega)$ is continuously embedded into $L^s(\Omega)$, $1 \leq s \leq 2^*$. Moreover, if $1 \leq s < 2^*$, then this embedding is compact (the Sobolev compact embedding theorem).

2. Mountain pass theorem for C^2-functionals

The mountain pass theorem is true for C^1 functionals. We commence with case of C^2-functionals on a Hilbert space. In this case the proof of the mountain pass theorem is relatively easy and allows to understand difficulties that we encounter in case of C^1-functionals.

A starting point is to establish a deformation lemma. For a given functional $\phi : X \to \mathbb{R}$ we set

$$\phi^c = \{u \in X,\ \phi(u) \leq c\}.$$

Lemma 2.1. **(Deformation lemma)** *Let $\phi \in C^2(X, \mathbb{R})$, where X is a Hilbert space and let $c \in \mathbb{R}$ and $\epsilon > 0$. Suppose that*

$$\|\phi'(u)\| \geq 2\epsilon \text{ for each } u \in \phi^{-1}([c - 2\epsilon, c + 2\epsilon]), \tag{2.1}$$

then there exists $\eta \in C(X, X)$ such that

$$\begin{aligned} &(i) \qquad \eta(u) = u \text{ for each } u \notin \phi^{-1}([c - 2\epsilon, c + 2\epsilon]) \\ &(ii) \qquad \eta(\phi^{c+\epsilon}) \subset \phi^{c-\epsilon}. \end{aligned}$$

Proof. Let

$$A = \phi^{-1}([c - 2\epsilon, c + 2\epsilon]) \text{ and } B = \phi^{-1}([c - \epsilon, c + \epsilon])$$

and define a function $\psi : X \to [0, 1]$ by

$$\psi(u) = \frac{\text{dist } (u, X - A)}{\text{dist } (u, X - A) + \text{ dist } (u, B)}.$$

It is clear that ψ is locally Lipschitz and $\psi(u) = 1$ for $u \in B$ and $\psi(u) = 0$ for $u \in X - A$. We now define

$$f(u) = \begin{cases} -\frac{\psi(u)}{\|\nabla\phi(u)\|}\nabla\phi(u) & \text{for } u \in A \\ 0 & \text{for } u \in X - A \end{cases}$$

and consider the Cauchy problem

$$\sigma'(t, u) = f(\sigma(t, u))$$

$$\sigma(0,u) = u.$$

This problem has a unique solution $\sigma(t,u)$, $t \in \mathbb{R}$. Letting $\eta(u) = \sigma(1,u)$, we see that η satisfies (i). To check (ii) we first show that $\phi(\sigma(\cdot,u))$ is decreasing. Indeed, we have

$$\begin{aligned}\frac{d}{dt}\phi(\sigma(t,u)) &= \langle \nabla\phi(\sigma(t,u)), \sigma'(t,u)\rangle \\ &= -\psi(\sigma(t,u))\|\nabla\phi(\sigma(t,u))\| \leq 0.\end{aligned}$$

Let $u \in \phi^{c+\epsilon}$. If $\phi(\sigma(t,u)) < c-\epsilon$ for some $t \in [0,1)$, then $\phi(\sigma(1,u)) < c-\epsilon$, that is, $\eta(u) \in \phi^{c-\epsilon}$ (because $\phi(\sigma(t,u))$ is decreasing in t). So, it remains to consider the case

$$\sigma(t,u) \in \phi^{-1}([c-\epsilon, c+\epsilon]) \text{ for each } t \in [0,1].$$

It follows from (2.1) and the fact that $\psi(u) = 1$ on B that

$$\begin{aligned}\phi(\sigma(1,u)) &= \phi(u) + \int_0^1 \frac{d}{dt}\phi(\sigma(t,u))\,dt \\ &= \phi(u) + \int_0^1 \langle \nabla\phi(\sigma(t,u)), f(\sigma(t,u))\rangle\,dt \\ &= \phi(u) - \int_0^1 \|\nabla\phi(\sigma(t,u))\|\,dt \leq c+\epsilon-2\epsilon = c-\epsilon.\end{aligned}$$

Proposition 2.1. *Let X be a Hilbert space and let $\phi \in C^2(X,\mathbb{R})$. Suppose that there exist $r > 0$ and $e \in X$ such that $\|e\| > r$ and*

$$b = \inf_{\|u\|=r} \phi(u) > \max\big(\phi(0), \phi(e)\big). \tag{2.2}$$

Then for each $\epsilon > 0$ there exists $u \in X$ such that

$$(a) \qquad c - 2\epsilon \leq \phi(u) \leq c + 2\epsilon,$$

$$(b) \qquad \|\phi'(u)\| \leq 2\epsilon,$$

where

$$c = \inf_{\gamma\in\Gamma}\max_{t\in[0,1]} \phi(\gamma(t)), \tag{2.3}$$

and

$$\Gamma = \{\gamma \in C([0,1],X);\ \gamma(0) = 0,\ \gamma(1) = e\}.$$

Proof. It follows from (2.2) that

$$b \leq \max_{t\in[0,1]} \phi(\gamma(t)) \text{ for every } \gamma \in \Gamma$$

which implies that

$$b \leq c \leq \max_{t\in[0,1]} \phi(te).$$

Suppose that our assertion is false. Then there exists $\epsilon > 0$ such that (2.1) of Proposition 2.1 is satisfied. By (2.2) we can assume that

$$c - 2\epsilon \geq \max(\phi(0), \phi(e)). \tag{2.4}$$

It follows from the definition of c that there exists $\gamma \in \Gamma$ such that

$$\max_{t\in[0,1]} \phi(\gamma(t)) \leq c + \epsilon. \tag{2.5}$$

Let η be a deformation mapping and set $\beta = \eta \circ \gamma$. It follows from (i) of Lemma 2.1 and (2.4) that

$$\beta(0) = \eta(\gamma(0)) = \gamma(0) = 0,$$

$$\beta(1) = \eta(\gamma(1)) = \gamma(1) = e,$$

so $\beta \in \Gamma$. Relations (2.5) and (ii) of Lemma 2.1 yield that

$$c \leq \max_{t\in[0,1]} \phi(\beta(t)) \leq c - \epsilon$$

which is impossible.

This result says that to each $\epsilon > 0$ there corresponds u_ϵ satisfying (a) and (b). If we could show that $u_{\epsilon_m} \to u$ (for some sequence $\epsilon_m \to 0$), then u is a critical point of ϕ due to the continuity of ϕ'. In general, this is not true as the following example shows. Let

$$\phi(x,y) = x^2 + (1-x)^3 y^2 \text{ for } (x,y) \in \mathbb{R}_2.$$

There exists $r > 0$ such that

$$\phi(x, y) > 0 \text{ for } 0 < x^2 + y^2 < r^2$$

and there exists a point $(x_\circ, y_\circ)$, with $x_\circ^2 + y_\circ^2 > r^2$ such that $\phi(x_\circ, y_\circ) \leq 0$. However, the only critical point of ϕ is $(0, 0)$.

This means that c defined in Proposition 2.1 is not in general a critical value.

To obtain the existence of a critical point we must introduce a condition guaranteeing a compactness of the set $\{u_\epsilon, \epsilon > 0\}$ from Proposition 2.1.

Definition. Let $c \in \mathbb{R}$ and $\phi \in C^1(X, \mathbb{R})$, where X is a Banach space. We say that ϕ satisfies the Palais - Smale condition at level c ($(PS)_c$-condition for short) if each sequence $\{u_m\} \subset X$ such that $\phi(u_m) \to c$ and $\phi'(u_m) \to 0$ in X' is relatively compact in X.

Theorem 2.1. *(***Ambrosetti - Rabinowitz' mountain pass theorem***) Suppose that assumptions of Proposition 2.1 are satisfied and that the $(PS)_c$-condition holds. Then c is a critical value of ϕ.*

Proposition 2.1 can be reformulated in the following way:

Theorem 2.2. *(***the mountain pass theorem without the (PS)-condition***) Suppose that assumptions of Proposition 2.1 are satisfied. Then there exists a sequence $\{u_m\} \subset X$ such that $\phi(u_m) \to c$ and $\phi'(u_m) \to 0$ in X'.*

3. The mountain pass theorem for C^1-functionals

If ϕ is only C^1-functional then the function f introduced in the proof of the deformation lemma is only continuous and we may have difficulty with the Cauchy problem considered in its proof. To overcome this difficulty we must find a replacement for $\nabla\phi$ with better regularity.

Definition. Let $\phi \in C^1(X, \mathbb{R})$, where X is a Banach space, and let

$$\hat{X} = \{u;\ \phi'(u) \neq 0\}.$$

Let $u \in \hat{X}$, a vector $v \in X$ is called a pseudo-gradient vector for ϕ at u if

$$(i) \qquad \|v\| \leq 2\|\phi'(u)\|,$$

$$(ii) \qquad \langle \phi'(u), v\rangle \geq \|\phi'(u)\|^2.$$

A mapping $V : \hat{X} \to X$ is called a pseudo-gradient vector field for ϕ on $\hat{X}$ if V is locally Lipschitz continuous and such that

$$\|V(u)\| \leq 2\|\phi'(u)\|,$$

and

$$\langle \phi'(u), V(u)\rangle \geq \|\phi'(u)\|^2$$

for all $u \in \hat{X}$.

Example. Let X be a Hilbert space and $\phi \in C^2(X, \mathbb{R})$. The gradient of ϕ is a pseudo-gradient vector field on $\hat{X}$. Indeed, we have

$$\|\phi'(u)\|^2 = \langle \phi'(u), \phi'(u)\rangle.$$

If $\phi \in C^1(X, \mathbb{R})$, then ϕ' in general, is not a pseudo-gradient vector field.

Lemma 3.1. *Let $\phi \in C^1(X, \mathbb{R})$, where X is a Banach space, then there exists a pseudo-gradient vector field $V : \hat{X} \to X$.*

Proof. Let $\tilde{u} \in \hat{X}$. Then there exists $w \in X$ such that $\|w\| = 1$ and

$$\langle \phi'(\tilde{u}), w\rangle > \frac{2}{3}\|\phi'(\tilde{u})\|.$$

We set $v = \frac{3}{2}\|\phi'(\tilde{u})\|w$, then

$$\|v\| < 2\|\phi'(\tilde{u})\| \text{ and } \langle \phi'(\tilde{u}), v\rangle > \|\phi'(\tilde{u})\|^2. \tag{3.1}$$

The continuity of ϕ' implies the existence of an open set $N_{\tilde{u}}$, containing $\tilde{u}$, such that

$$\|v\| < 2\|\phi'(u)\| \text{ and } \langle \phi'(u), v\rangle > \|\phi'(u)\|^2$$

for all $u \in N_{\tilde{u}}$. The collection $\{N_{\tilde{u}}, \tilde{u} \in \hat{X}\}$ forms an open covering of $\hat{X}$. Since $\hat{X}$ is paracompact, as a metric space, we can find subcovering $\{M_i,\ i \in I\}$ which is locally finite refinement. This subcovering has the property: for each $i \in I$ there exists $\tilde{u} \in \hat{X}$ such that $M_i \subset N_{\tilde{u}}$. Therefore, there exists a vector $v = v_i$ such that inequalities (3.1) are satisfied on M_i. We now define

$$\rho_i(u) = \operatorname{dist}(u, X - M_i)$$

and

$$V(u) = \sum_{i \in I} \frac{\rho_i(u)}{\sum_{j \in I} \rho_j(u)} v_i.$$

V is well defined since $\{M_i,\ i \in I\}$ is locally finite. It is clear that V is locally Lipschitz pseudo-gradient vector field on $\hat{X}$ for ϕ.

From now on we shall always assume that X is a Banach space. For a given set $S \subset X$ and $\delta > 0$ we put

$$S_\delta = \{x \in X;\ \operatorname{dist}(x, S) \leq \delta\}.$$

Lemma 3.2. (**quantitative deformation lemma**) *Let $\phi \in C^1(X, \mathbb{R})$, $S \subset X$, $c \in \mathbb{R}$, $\epsilon > 0$, $\delta > 0$ be such that*

$$\|\phi'(u)\| \geq \frac{4\epsilon}{\delta} \text{ for each } u \in \phi^{-1}([c - 2\epsilon, c + 2\epsilon]) \cap S_{2\delta}. \tag{3.2}$$

Then there exists $\eta \in C([0,1] \times X, X)$ such that

$$(i) \qquad \eta(0, u) = u \text{ for each } u \in X,$$

(*ii*) $\eta(t,\cdot)$ *is a homeomorphism on* X *for each* $t\in[0,1]$,

(*iii*) $\eta(t,u)=u$ *for each* $u\notin\phi^{-1}([c-2\epsilon,c+2\epsilon])\cap S_{2\delta}$ *and each* $t\in[0,1]$,

(*iv*) $\|\eta(t,u)-u\|\le\delta$ *for* $u\in X$ *and* $t\in[0,1]$,

(*v*) $\phi(\eta(\cdot,u))$ *is decreasing for each* $u\in X$,

(*vi*) $\phi(\eta(t,u))<c$ *for each* $u\in\phi^c\cap S_\delta$ *and* $t\in(0,1]$,

(*vii*) $\eta(1,\phi^{c+\epsilon}\cap S)\subset\phi^{c-\epsilon}$,

Proof. By Lemma 3.1 there exists a pseudo-gradient vector field g for ϕ on $\hat{X}$. We set

$$A=\phi^{-1}([c-2\epsilon,c+2\epsilon])\cap S_{2\delta},$$

$$B=\phi^{-1}([c-\epsilon,c+\epsilon])\cap S_\delta,$$

and

$$\psi(u)=\frac{\operatorname{dist}(u,X-A)}{\operatorname{dist}(u,X-A)+\operatorname{dist}(u,B)}.$$

The function ψ is locally Lipschitz, $\psi(u)=1$ on B, $\psi(u)=0$ on $X-A$ and $0\le\psi(u)\le1$ on X. We put

$$f(u)=\begin{cases}-\frac{\psi(u)}{\|g(u)\|}g(u) & \text{for } u\in A\\ 0 & \text{for } u\in X-A.\end{cases}$$

For each $u\in X$ the Cauchy problem

$$\sigma'(t,u)=f(\sigma(t,u)),$$

$$\sigma(0,u)=u$$

has a solution $\sigma(\cdot,u)$ defined on $\mathbb{R}$. We define the homeomorphism η by $\eta(t,u)=\sigma(\delta t,u)$ which obviously satisfies (i), (ii) and (iii). Since

$$\|\sigma(t,u)-u\|=\|\int_0^t f(\sigma(s,u))\,ds\|\le\int_0^t\|f(\sigma(s,u))\|\,ds\le t \tag{3.3}$$

and (iv) holds. We now show that $\phi(\sigma(\cdot,u))$ is decreasing. Indeed, we have

$$\frac{d}{dt}\phi(\sigma(t,u)) = \langle \phi'(\sigma(t,)), \sigma'(t,u)\rangle = \langle \phi'(\sigma(t,u)), f(\sigma(t,u))\rangle \leq 0$$

and (v) and (vi) hold. To show (vii) let $u \in \phi^{c+\epsilon} \cap S$. If there exists $t \in [0,\delta)$ such that $\phi(\sigma(t,u)) < c-\epsilon$, then $\phi(\sigma(\delta,u)) < c-\epsilon$ which means that $\eta(1,u) \in \phi^{c-\epsilon}$. If for each $t \in [0,\delta)$ $\phi(\sigma(t,u)) \geq c-\epsilon$, then by (v) and (3.3) we have that

$$\sigma(t,u) \in \phi^{-1}([c-\epsilon, c+\epsilon]) \cap S_\delta \ \text{ for each } \ t \in [0,1].$$

Hence we obtain

$$\begin{aligned}
\phi(\sigma(\delta,u)) &= \phi(u) + \int_0^\delta \frac{d}{dt}\phi(\sigma(t,u))\,dt \\
&= \phi(u) + \int_0^\delta \langle \phi'(\sigma(t,u)), f(\sigma(t,u))\rangle\,dt \\
&= \phi(u) - \int_0^\delta \langle \phi'(\sigma(t,u)), \frac{g(\sigma(t,u))}{\|g(\sigma(t,u))\|}\rangle\,dt \\
&\leq c+\epsilon - \frac{1}{2}\int_0^\delta \|\phi'(\sigma(t,u))\|\,dt \leq c-\epsilon,
\end{aligned}$$

that is, $\eta(1,u) \in \phi^{c-\epsilon}$.

Proposition 3.1. *Let $\phi \in C^1(X,\mathbb{R})$ and suppose that there exist $r > 0$ and $e \in X$ such that $\|e\| > r$ and*

$$b = \inf_{\|u\|=r} \phi(u) > \max\big(\phi(0), \phi(e)\big). \tag{3.4}$$

Then for each $\epsilon > 0$, $\delta > 0$ and $\gamma \in \Gamma$ such that

$$\max_{t\in[0,1]} \phi(\gamma(t)) \leq c+\epsilon$$

there exists $u \in X$ such that

(a) $\quad c - 2\epsilon \leq \phi(u) \leq c + 2\epsilon$,

(b) $\quad$ *dist* $(u, \gamma([0,1])) \leq 2\delta$,

(c) $\quad \|\phi'(u)\| \leq \frac{4\epsilon}{\delta}$, *where c is given by (2.3)*

Proof. If the assertion were false, then for each u satisfying (a) and (b) we would have $\|\phi'(u)\| \geq \frac{4\epsilon}{\delta}$. Since $c \geq b$, we may assume that $\max(\phi(0), \phi(e)) < c - 2\epsilon$. We apply Lemma 3.2 with $S = \gamma([0,1])$. Let $\eta \in C([0,1] \times X, X)$ be a deformation mapping satisfying $(i) - (vii)$. We set $\beta(u) = \eta(1, \gamma(u))$. We check that $\beta(0) = 0$ and $\beta(1) = e$, so $\beta \in \Gamma$. By (vii) we see that $\max_{t\in[0,1]} \phi \circ \beta \leq c - \epsilon$, which is impossible.

Theorem 3.1. *Suppose that $\phi \in C^1(X, \mathbb{R})$ and that (3.4) holds. If ϕ satisfies the $(PS)_c$-condition then c is a critical value.*

Proposition 3.2. *Let $\phi \in C^1(X, \mathbb{R})$ be bounded from below on X. Let $\epsilon > 0$, $\delta > 0$ and $v \in X$ be such that*

$$\phi(v) \leq \inf_X \phi + \epsilon.$$

Then there exists $u \in X$ such that

$$\phi(u) \leq \inf_X \phi + 2\epsilon, \quad \|\phi'(u)\| < \frac{4\epsilon}{\delta}$$

and $\|u - v\| \leq 2\delta$.

Proof. In the contrary case for each u satisfying $\phi(u) \leq \inf_X \phi + 2\epsilon$ and $\|u - v\| \leq 2\delta$ we have $\|\phi'(u)\| > \frac{4\epsilon}{\delta}$. We then apply Lemma 3.1 with $S = \{v\}$ and $c = \inf_X \phi$ to conclude that $\eta(1, v) \in \phi^{c-\epsilon}$. This is impossible since $\phi^{c-\epsilon} = \emptyset$.

Proposition 3.2 is the Ekeland variational principle for C^1-functionals. We shall return to this problem in Section 7.

4. Applications of the mountain pass theorem

If $X = \mathbb{R}_N$ the mountain pass theorem can formulated in the following way:

Theorem 4.1. *Let $f \in C^1(\mathbb{R}_N, \mathbb{R})$ be such that $\lim_{|x|\to\infty} f(x) = \infty$. If f has two strict local minima $x_\circ$ and x_1, then it has a third critical point x_2 with $f(x_2) > \max(f(x_\circ), f(x_1))$.*

Theorem 4.1 will be used to prove the Hadamard global homeomorphism theorem.

Theorem 4.2. *Let $F \in C^1(\mathbb{R}_N, \mathbb{R}_N)$ satisfy*

$$(1) \qquad F'(x) \text{ is invertible for all } x \in \mathbb{R}_N,$$

$$(2) \qquad \|F(x)\| \to \infty \text{ as } |x| \to \infty.$$

Then F is a diffeomorphism of $\mathbb{R}_N$ onto $\mathbb{R}_N$.

Proof. By (1) and the inverse function theorem, F is an open mapping (F maps open sets into open sets). Hence $F(\mathbb{R}_N)$ is open in $\mathbb{R}_N$. Condition (2) implies that $F(\mathbb{R}_N)$ is closed in $\mathbb{R}_N$ (here we use the fact that bounded and closed sets in $\mathbb{R}_N$ are compact). Since $\mathbb{R}_N$ is connected we must have $F(\mathbb{R}_N) = \mathbb{R}_N$. To show that F is a diffeomorphism we must check that F is one - to - one. Arguing indirectly, we assume that $F(x_\circ) = F(x_1) = y$ for some $x_\circ \neq x_1$. We define a function $f(x) = \frac{1}{2}\|F(x) - y\|^2$ which is C^1. We have $f'(x) = \left(F'\right)^T(F(x) - y)$ and $\lim_{|x|\to\infty} f(x) = \infty$. Obviously, $x_\circ$ and x_1 are global minima of f. We now observe that $F(x) \neq F(x_i)$, $i = 0, 1$, for x in a neighbourhood of x_i, so $x_\circ$ and x_1 are strict local minima. By Theorem 4.1 there exists a third critical point x_2 with $f(x_2) > 0$. So we have $\|F(x_2) - y\| > 0$, that is, $F(x_2) \neq y$. Since x_2 is a critical point of f we have $(F')^T(x_2)(F(x_2) - y) = 0$, which is impossible as $F'(x_2)$ is invertible.

As a second application we consider the Dirichlet problem

$$
(4.1) \qquad \begin{cases} -\Delta u + c(x)u = f(x,u) \text{ in } \Omega, \\ u(x) = 0 \text{ on } \partial\Omega, \end{cases}
$$

where Ω is a bounded domain in $\mathbb{R}_N$.

It is assumed that

(A) $\quad c \in L^\infty(\Omega)$, $f \in C(\Omega \times \mathbb{R}, \mathbb{R})$ and there exist constants $c_1 > 0$ and $c_2 > 0$ such that

$$
|f(x,u)| \leq c_1 + c_2|u|^p \text{ on } \Omega \times \mathbb{R},
$$

where $1 < p < \frac{N+2}{N-2}$ if $N \geq 3$. If $N = 2$ then $1 < p < \infty$. (We only treat case $N \geq 3$).

(B) $\quad$ There exist $\mu > 2$ and $R > 0$ such that $0 < \mu F(x,u) \leq uf(x,u)$ for $|u| \geq R$.

Integrating, we get from (B) that

$$
(4.2) \qquad F(x,u) \geq c_3|u|^\mu - c_4
$$

for some $c_3 > 0$ and $c_4 > 0$.

A solution of (4.1) will be found as a critical point of the functional $\phi : H^1_\circ(\Omega) \to \mathbb{R}$ defined by

$$
\phi(u) = \frac{1}{2}\int_\Omega [|Du|^2 + cu^2]\,dx - \int_\Omega F(x,u)\,dx,
$$

where

$$
F(x,u) = \int_0^u f(x,s)\,ds.
$$

Lemma 4.1. *Suppose that (A) holds, then ϕ is a continuously differentiable functional on $H^1_\circ(\Omega)$.*

Proof. First we show that the directional derivative of ϕ exists. We write

$$
\phi(u) = \phi_1(u) + \phi_2(u),
$$

where

$$\phi_1(u) = \frac{1}{2}\int_\Omega [|Du|^2 + c(x)u^2]\, dx$$

and

$$\phi_2(u) = \int_\Omega F(x,u)\, dx.$$

Let $h \in H_\circ^1(\Omega$. It is easy to check that

$$\langle \phi_1'(u), h\rangle = \int_\Omega (DuDh + cuh)\, dx.$$

By the mean value theorem for each $x \in \Omega$ and $1 < |\lambda| < 1$, there exists $\theta \in (0,1)$ such that

$$\begin{aligned}\frac{1}{|\lambda|}|F(x,u(x)+\lambda h(x)) - F(x,u(x))| &= |f(x,u(x)+\theta\lambda h(x))||h(x)| \\ &\le \big(c_1 + c_2(|u(x)| + |h(x)|)^p\big)|h(x)| \\ &\le \big(c_1 + c_2 2^p(|u(x)|^p + |h(x)|^p)\big)|h(x)|\end{aligned}$$

Since $(c_1 + c_2 2^p(|u(x)|^p + |h(x)|^p))|h(x)| \in L^1(\Omega)$, it follows from the dominated convergence theorem that

$$\langle \phi_2'(u), h\rangle = \int_\Omega f(x,u(x))h(x)\, dx.$$

To show the continuity of the directional derivative we observe that the inequality

$$|\langle \phi_1'(u), h\rangle| \le \|u\|\|h\| + \|c\|_\infty\|u\|_2\|h\|_2$$

implies that

$$\|\phi_1'(u)\| \le C\|u\|,$$

which means that ϕ_1' is a continuous linear functional on $H_\circ^1(\Omega)$. By the Hölder and Sobolev inequalities we have

$$|\langle \phi_2'(u) - \phi_2'(v), h\rangle| \le \|f(\cdot,u) - f(\cdot,v)\|_{\frac{p+1}{p}}\|h\|_{p+1},$$

that is,

$$\|\phi_2'(u) - \phi_2'(v)\| \le C\|f(\cdot,u) - f(\cdot,v)\|_{\frac{p+1}{p}}.$$

Lemma 4.2. *If (A) and (B) hold, then ϕ satisfies the $(PS)_c$-condition for each $c \in \mathbb{R}$.*

Proof. Let $\{u_m\} \subset H^1_\circ(\Omega)$ be such that

$$\phi(u_m) \to c \text{ and } \phi'(u_m) \to 0 \text{ in } H^{-1}(\Omega).$$

For $\nu \in (\frac{1}{\mu}, \frac{1}{2})$ we have for m sufficiently large that

$$\begin{aligned}
c + 1 + \|u_m\| &\geq \phi(u_m) - \nu\langle\phi'(u_m), u_m\rangle \\
&= \int_\Omega \left[(\frac{1}{2} - \nu)|Du_m|^2 + (\frac{1}{2} - \nu)cu_m^2 + \nu f(x, u_m)u_m - F(x, u_m)\right] dx \\
&\geq (\frac{1}{2} - \nu)\|u_m\|^2 - (\frac{1}{2} - \nu)\|c\|_\infty\|u_m\|_2^2 + (\nu\mu - 1)\int_\Omega F(x, u_m)\, dx - d_1 \\
&\geq (\frac{1}{2} - \nu)\|u_m\|^2 - (\frac{1}{2} - \nu)\|c\|_\infty\|u_m\|_2^2 + c_3(\mu\nu - 1)\|u_m\|_\mu^\mu - d_2,
\end{aligned}$$

where $d_1 > 0$ and $d_2 > 0$ are constants independent of m. Since $\|u\|_2 \leq C(\Omega)\|u\|_\mu$, the sequence $\{u_m\}$ is bounded in $H^1_\circ(\Omega)$. Therefore we may assume that $u_m \rightharpoonup u$ in $H^1_\circ(\Omega)$. We write

(4.3)
$$\begin{aligned}
\|u_m - u\|^2 &= \langle\phi'(u_m) - \phi'(u), u_m - u\rangle \\
&\quad + \int_\Omega \left[-c(u_m - u)^2 + (f(x, u_m) - f(x, u))(u_m - u)\right] dx.
\end{aligned}$$

By assumption

$$\langle\phi'(u_m) - \phi'(u), u_m - u\rangle \to 0 \text{ as } m \to \infty. \tag{4.4}$$

By the Sobolev compact embedding theorem $u_m \to u$ in $L^{p+1}(\Omega)$. This implies that $f(x, u_m) \to f(x, u)$ in $L^{\frac{p+1}{p}}(\Omega)$. Consequently, by the Hölder inequality we get

$$\int_\Omega (f(x, u_m) - f(x, u))(u_m - u)\, dx \to 0 \text{ as } m \to \infty. \tag{4.5}$$

Finally, we observe that

$$\left|\int_\Omega c(u_m - u)^2\, dx\right| \leq \|c\|_\infty\|u_m - u\|_2^2 \to 0 \text{ as } m \to \infty. \tag{4.6}$$

The relative compactness of $\{u_m\}$ in $H^1_\circ(\Omega)$ follows from (4.3)-(4.6).

We are now in a position to establish the existence result for problem (4.1).

Theorem 4.3. *Suppose that (A) and (B) hold, $c \geq 0$ on Ω and that*

$$F(x,u) = o(|u|^2) \text{ as } |u| \to 0 \text{ uniformly in } x \in \Omega.$$

Then problem (4.1) has at least one nontrivial solution.

Proof. We apply Theorem 3.1. It follows from our assumptions that to every $\epsilon > 0$ there corresponds $C_\epsilon > 0$ such that

$$|F(x,u)| \leq \epsilon u^2 + C_\epsilon |u|^{p+1}. \tag{4.7}$$

This implies that

$$\begin{aligned}\phi(u) &\geq \int_\Omega \left(\frac{1}{2}|Du|^2 + \frac{c}{2}u^2 - \epsilon u^2 - C_\epsilon |u|^{p+1}\right) dx \\ &\geq \frac{1}{2}\|u\|^2 - \epsilon\|u\|_2^2 - C_\epsilon \|u\|_{p+1}^{p+1}.\end{aligned}$$

By the Sobolev embedding theorem there exists a constants $C_1 > 0$ and $C_2 > 0$ such that

$$\phi(u) \geq \left(\frac{1}{2} - C_1\epsilon\right)\|u\|^2 - C_\epsilon C_2 \|u\|^{p+1}.$$

If we choose ϵ so that $\left(\frac{1}{2} - C_1\epsilon\right) > 0$ and $r > 0$ sufficiently small then

$$0 < b = \inf_{\|u\|=r} \phi(u).$$

Let $u \in H_\circ^1(\Omega)$, with $u \not\equiv 0$, then by (4.2) we have

$$\phi(tu) \leq \frac{t^2}{2}\int_\Omega \left(|Du|^2 + cu^2\right) dx - t^\mu c_3 \int_\Omega |u|^\mu \, dx + c_4|\Omega|.$$

Letting $e = tu$, with $t > 0$ sufficiently large we get $\|e\| > r$ and $\phi(e) \leq 0$. Since the $(PS)_c$-condition holds for each $c \in \mathbb{R}$, by the mountain pass theorem, problem (4.1) has a nontrivial solution.

5. Linking

We discuss in this section a notion of topological linking which will be used to derive a generalization of the mountain pass theorem.

Definition. Let X be a Banach space and let $\tilde{X}$ be a closed subspace of X. Suppose that F is a closed subset of X and Q is a closed subset of $\tilde{X}$. We say that the sets F and ∂Q link if

(a) $F \cap \partial Q = \emptyset$,

(b) for each mapping $\gamma \in C(Q, X)$, with $\gamma(u) = u$ on ∂Q, we have $F \cap \gamma(Q) \neq \emptyset$.

We now introduce two important examples of linking. To describe them we need a retract of a topological space.

Definition. Let A be a subset of a topological space X. A continuous mapping $r : X \to A$ with $r(X) = A$ and $r(u) = u$ for all $u \in A$ is called a retract of X onto A.

Let

$$B^N = \{x \in \mathbb{R}_N;\ \|x\| \leq 1\} \text{ and } S^{N-1} = \{x \in \mathbb{R}_N;\ \|x\| = 1\}.$$

Theorem 5.1. *The following statements are equivalent:*

(i) *Brouwer's fixed point theorem: every continuous function $f : B^N \to B^N$ has at least one fixed point.*

(ii) *There is no retract of B^N onto S^{N-1}.*

(iii) *(Continuation theorem) Let $h \in C([0,1] \times B^N, \mathbb{R}^N)$ be such that*

(a) $h(0, u) = u$ *for all* $u \in B^N$,

(b) $h(t, u) \neq 0$ *for each* $(t, u) \in (0,1) \times S^{N-1}$.

Then there exists at least one point $u \in B^N$ such that $h(1, u) = 0$.

Proof. $(i) \Rightarrow (ii)$

If $r : B^N \to S^{N-1}$ is a retract, then a continuous function $f : B^N \to S^{N-1}$ defined by $f(u) = -r(u)$ has no fixed point.

$(ii) \Rightarrow (iii)$

Suppose that there exists a homotopy $h \in C([0,1] \times B^N, \mathbb{R}^N)$ satisfying (a) and (b) and that $h(1,u) \neq 0$ for each $u \in B^N$. We put

$$E = \left([0,1] \times S^{N-1}\right) \cup \{1\} \times B^N.$$

We define a mapping $f : B^N \to E$ in the following manner: for each $u \in B^N$ we denote by $f(u)$ the intersection point with E of a half-line emanating from $(-1,0)$ and passing through $(0,u)$. It is clear that f is continuous and we set

$$g(t,u) = \frac{h(t,u)}{\|h(t,u)\|}.$$

It is easy to check that $g \circ f$ is a retract of B^N onto S^{N-1}.

$(iii) \Rightarrow (i)$

Let $f \in C(B^N, B^N)$ and define a mapping $h : [0,1] \times B^N \to \mathbb{R}_N$ by

$$h(t,x) = x - tf(x).$$

This mapping satisfies (a) and (b) of (iii). Assuming that f has no fixed point we see that $h(1,u) \neq 0$ for each $u \in B^N$ which is impossible.

Example 5.1. *Let $X = W \bigoplus Z$ be a topological direct sum with closed subspaces W and Z and $\dim W < \infty$. We set $\tilde{X} = W$ and*

$$Q = \{w \in W;\ \|w\| \leq \rho\},\ \rho > 0.$$

Then Z and ∂Q link.

Proof. It is evident that $Z \cap \partial Q = \emptyset$. Let $\gamma \in C(Q, X)$ be such that $\gamma(u) = u$ for $u \in \partial Q$ and let $P : X \to W$ be projection of X onto W. If $\gamma(Q) \cap Z = \emptyset$, then a mapping $u \to \rho \frac{P\gamma(u)}{\|P\gamma(u)\|}$ is a retract of Q onto ∂Q, which is impossible.

Example 5.2. *Let $X = W \bigoplus Z$ be a topological direct sum with closed subspaces W and Z and $\dim W < \infty$. Let $e \in Z$, with $\|e\| = 1$, $R > 0$, $r \in (0, R)$ and $\rho > 0$. We set $\tilde{X} = W \bigoplus \mathbb{R}e$ and*

$$F = \{z \in Z;\ \|z\| = r\},$$

$$Q = \{w + te;\ w \in W, \|w\| \leq \rho, 0 \leq t \leq R\}.$$

Then F and ∂Q link.

Proof. Again it is obvious that $F \cap \partial Q = \emptyset$. Let $\gamma \in C(Q, X)$ be such that $\gamma(u) = u$ for all $u \in \partial Q$. Let P be a projection of X onto W. It is easy to construct a retract Θ of $\tilde{X} - \{re\}$ onto ∂Q. If $\gamma(Q) \cap F = \emptyset$, then a mapping

$$u \to \Theta(P\gamma(u) + \|(I - P)\gamma(u)\|e)$$

is a retract of Q on ∂Q. This is impossible because Q is homeomorphic to a finite dimensional ball.

We now establish a general linking principle.

Lemma 5.1. *Let X be a Banach space and let $F \subset X$ and $Q \subset X$. For $\phi \in C^1(X, \mathbb{R})$ we set*

$$c = \inf_{\gamma \in \Gamma} \sup_{u \in Q} \phi(\gamma(u)),$$

where

$$\Gamma = \{\gamma \in C(Q, X);\ \gamma(u) = u \text{ on } \partial Q\}.$$

If

(a) F *and* ∂Q *link,*

(b) $a = \sup_{\partial Q} \phi < b = \inf_F \phi,$

(c) $d = \sup_Q \phi < \infty,$

then $c \in [b,d]$ *and for every* $\epsilon \in \left(0, \frac{c-a}{2}\right)$, $\delta > 0$ *and* $\gamma \in \Gamma$ *such that*

$$\sup_Q \phi \circ \gamma \le c + \epsilon$$

there exists $u \in X$ *such that*

$$\begin{array}{ll} (i) & c - 2\epsilon \le \phi(u) \le c + 2\epsilon, \\ (ii) & dist\ (u, \gamma(Q)) \le 2\delta, \\ (iii) & \|\phi'(u)\| < \frac{4\epsilon}{\delta}. \end{array}$$

Proof. It follows from (a) and (b) that $b \le c$ and also by (c) we have $c \le d$. Arguing indirectly, we may assume that condition (3.2) of Lemma 3.2 holds with $S = \gamma(Q)$. According to our assumption on ϵ and γ we have

$$c - 2\epsilon > a \tag{5.1}$$

and

$$\gamma(Q) \subset \phi^{c+\epsilon}. \tag{5.2}$$

We put $\beta(u) = \eta(1, \gamma(u))$, where η is a deformation mapping from Lemma 3.2. Using (iii) of Lemma 3.2 and (5.1) we check that

$$\beta(u) = \eta(1, \gamma(u)) = \eta(1, u) = u$$

for all $u \in \partial Q$. This means that $\beta \in \Gamma$. It follows from (vii) of Lemma 3.2 and (5.2) that

$$c \le \sup_{u \in Q} \phi(\beta(u)) \le c - \epsilon$$

which is impossible.

We now list direct consequences of Lemma 5.1.

Theorem 5.1. *(* **Benci - Rabinowitz** *) Suppose that assumptions of Lemma 5.1 hold and that* ϕ *satisfies the* $(PS)_c$*-condition. Then* c *is a critical value of* ϕ*.*

Theorem 5.2. (**Rabinowitz' Saddle Point Theorem**) *Let X be a Banach space and let $X = W \bigoplus Z$ be a topological direct sum of closed subspaces W and Z and $\dim W < \infty$. Let $Q = \{u \in W : \|u\| \leq \rho\}$, $\rho > 0$. For $\phi \in C^1(X, \mathbb{R})$ we set*

$$c = \inf_{\gamma \in \Gamma} \max_{u \in Q} \phi(\gamma(u))$$

where

$$\Gamma = \{\gamma \in C(Q, X);\ \gamma(u) = u \ \text{ on } \partial Q\}.$$

If

$$(a) \qquad \max_{\partial Q} \phi < b = \inf_Z \phi,$$

$$(b) \qquad \phi \text{ satisfies the } (PS)_c\text{-condition},$$

then $c \geq b$ is a critical value of ϕ.

According to Example 5.1 Z and ∂Q link, the assertion follows from Theorem 5.1.

Theorem 5.3. (**Rabinowitz' generalization of the mountain pass theorem**) *Let X be a Banach space and let $X = W \bigoplus Z$ be a toplogical direct sum of closed subspaces W and Z with $\dim W < \infty$. Let $e \in Z$, $\|e\| = 1$, $R > 0$, $r \in (0, R)$, $\rho > 0$ and put*

$$F = \{z \in Z; \|z\| = r\},$$

$$Q = \{w + te;\ w \in W, \|w\| \leq \rho, 0 \leq t \leq R\}.$$

For $\phi \in C^1(X, \mathbb{R})$ we define a quantity c as in Theorem 5.2. If

$$(a) \qquad \max_{\partial Q} \phi < b = \inf_F \phi,$$

$$(b) \qquad \phi \text{ satisfies the } (PS)_c\text{-condition},$$

then $c \geq b$ is a critical value of ϕ.

As explained in Example 5.2 F and ∂Q link and the result follows from Theorem 5.1.

Finally, we observe that applying Theorem 5.3 with $W = \{0\}$ we deduce the mountain pass theorem (Theorem 3.1).

6. Applications of Theorems 5.2 and 5.3

We consider the Dirichlet problem

$$
(6.1) \qquad \begin{cases} -\Delta u = \lambda a(x)u + f(x,u) \text{ in } \Omega, \\ u(x) = 0 \text{ on } \partial\Omega, \end{cases}
$$

where $\Omega \subset \mathbb{R}_N$ is a bounded domain, $a \in L^\infty(\Omega)$, with $a > 0$ on Ω, and $f \in C(\Omega \times \mathbb{R}, \mathbb{R})$ is a bounded function. Here λ is a positive parameter which will be specified later.

We associate with (6.1) an eigenvalue problem

$$
(6.2) \qquad \begin{cases} -\Delta v = \mu a(x)v \text{ in } \Omega, \\ v(x) = 0 \text{ on } \partial\Omega. \end{cases}
$$

It is known that this problem possesses a sequence of eigenvalues $0 < \lambda_1 < \lambda_2 \leq \ldots \leq \lambda_j \leq \ldots$ with $\lambda_j \to \infty$ as $j \to \infty$. Here each eigenvalue is repeated according to its multiplicity.

Theorem 6.1. *Suppose that $\lambda = \lambda_k < \lambda_{k+1}$ and that*

$$
(6.3) \qquad \lim_{|u|\to\infty} F(x,u) = \lim_{|u|\to\infty} \int_0^u f(x,s)\,ds = \infty \textit{ uniformly in } x \in \Omega.
$$

Then problem (6.1) possesses a solution in $H^1_\circ(\Omega)$.

Proof. A solution to problem (6.1) will be obtained as a critical point of a functional $I : H^1_\circ(\Omega) \to \mathbb{R}$ defined by

$$
I(u) = \frac{1}{2}\int_\Omega \left(|Du|^2 - \lambda a(x)u^2\right) dx - \int_\Omega F(x,u)\,dx.
$$

Let $W =$ span $\{v_1, \ldots v_k\}$, where v_j are eigenfunctions of (6.2) corresponding to λ_j and normalized so that

$$
\int_\Omega |Dv_j|^2\,dx = 1 = \lambda_j \int_\Omega a v_j^2\,dx
$$

and let $Z = \overline{\text{span}\ \{v_j;\ j \geq k+1\}}$ so that $X = H_\circ^1(\Omega) = W \bigoplus Z$. We check that I satisfies assumptions (a) and (b) of Theorem 5.2. If $u \in Z$, then $u = \sum_{j=k+1}^{\infty} a_j v_j$ and

$$\int_\Omega (|Du|^2 - \lambda_k a u^2)\, dx = \sum_{j=k+1}^{\infty} a_j^2 (1 - \frac{\lambda_k}{\lambda_j}) \geq (1 - \frac{\lambda_k}{\lambda_{k+1}})\|u\|^2. \tag{6.4}$$

Letting $M = \sup_{(x,t)\in\bar{\Omega}\times\mathbb{R}} |f(x,t)|$ we have by the Hölder and Sobolev inequality that

$$|\int_\Omega F(x,u)\, dx| \leq M \int_\Omega |u|\, dx \leq M_1 \|u\|$$

for all $u \in H_\circ^1(\Omega)$ and some constants M and M_1. Combining the last inequalities together we see that I is bounded from below on Z. We now show that $\lim_{\|u\|\to\infty, u\in W} I(u) = -\infty$. Indeed, writing $W = W_\circ \bigoplus W_1$, with $W_\circ = \text{span}\{v_k;\ \lambda_j = \lambda_k\}$ and $W_1 = \text{span}\{v_j; \lambda_j < \lambda_k\}$ we have for $u \in W$ a decomposition $u = u_\circ + u_1$ with $u_i \in W_i$, $i = 0, 1$ and

$$I(u) = \frac{1}{2}\sum_{j<k} a_j^2 (1 - \frac{\lambda_k}{\lambda_j}) - \int_\Omega F(x,u_\circ)\, dx - \int_\Omega (F(x, u_\circ + u_1) - F(x,u_\circ))\, dx.$$

By a straightforward estimation we get

$$I(u) \leq -M_2 \|u_1\|^2 - \int_\Omega F(x,u_\circ)\, dx + M_1 \|u_1\|$$

for some $M_2 > 0$. This estimate in conjunction with (6.3) yields that

$$\lim_{\|u\|\to\infty, u\in W} I(u) = -\infty.$$

Taking as a set $Q = \{u;\ u \in W, \|u\| \leq r\}$ we see that (a) and (b) of Theorem 5.2 are satisfied provided r is sufficiently large. It remains to check the Palais - Smale condition. Towards this end it is sufficient to show that if $\{u_m\} \subset H_\circ^1(\Omega)$ is such that $\{I(u_m)\}$ is bounded and $I'(u_m) \to 0$ in $H^{-1}(\Omega)$, then $\{u_m\}$ is bounded in $H_\circ^1(\Omega)$ (see the proof of Lemma 4.2). We write $u_m = u_m^\circ + u_m^- + u_m^+$, where $u_m^\circ \in W_\circ$, $u_m^- \in W_1$ and $u_m^+ \in Z$. Since $I'(u_m) \to 0$ in $H^{-1}(\Omega)$, we have

$$|\langle I'(u_m), u_m^\pm \rangle| = |\int_\Omega (Du_m Du_m^\pm - \lambda_k a u_m u_m^\pm - f(x,u_m) u_m^\pm)\, dx| \leq \|u_m^\pm\|$$

for large m. As in (6.4) we check that

$$\|u_m^+\| \geq \big(1 - \frac{\lambda_k}{\lambda_{k+1}}\big)\|u_m^+\|^2 - M_1\|u_m^+\|,$$

which implies that $\{\|u_m^+\|\}$ is bounded. Similary, we show that $\{\|u_m^-\|\}$ is bounded. Finally, we show that $\{u_m^\circ\}$ is bounded. To verify this we observe that for some $K > 0$ and each m we have

$$\begin{aligned} K \geq |I(u_m)| = |\int_\Omega \frac{1}{2}\big(|Du_m^+|^2 + |Du_m^-|^2\big)\,dx \\ - \frac{1}{2}\lambda_k\int_\Omega a\big((u_m^+)^2 + (u_m^-)^2\big)\,dx - \int_\Omega \big(F(x,u_m) - F(x,u_m^\circ)\big)\,dx \\ - \int_\Omega F(x,u_m^\circ)\,dx. \end{aligned}$$

Since $\{\|u_m^\pm\|\}$ are bounded, there exists $K_1 > 0$ such that

$$K \geq |\int_\Omega F(x,u_m^\circ)\,dx| - K_1$$

for all m. This in conjunction with (6.3) imples that $\{\|u_m^\circ\|\}$ is bounded.

To apply Theorem 5.3 we shall use the following fact: condition (a) of this theorem is satisfied if $\phi(u) \leq 0$ on W and there exists $e \in Z$, with $\|e\| = 1$, and $\bar{R} > r$ such that $\phi(u) \leq 0$ for $u \in W \bigoplus \text{span}\{e\}$ and $\|u\| \geq \bar{R}$. Then for large R and $\rho > 0$ the set Q defined in this theorem has the property $\phi(u) \leq 0$ on ∂Q. Hence condition (a) holds if $\inf_F \phi > 0$.

As an application of Theorem 5.3 we consider problem (6.1) under slightly different set of assumptions.

Theorem 6.2. *Suppose that f satisfies (A) and (B) from Section 4 and moreover that $F(x,t) \geq 0$ on $\bar{\Omega} \times \mathbb{R}$ and $f(x,t) = o(|t|)$ as $|t| \to 0$ uniformly in Ω. Then for each $\lambda \in \mathbb{R}$ problem (6.1) possesses a nontrivial solution.*

Proof. We only consider the case $\lambda_k \leq \lambda < \lambda_{k+1}$. As in the proof of Theorem 6.1 we define $W = \text{span}\{v_1, \dots, v_k\}$ and $Z = \overline{\text{span}\{v_{k+1}, \dots\}}$, where v_j are eigenfunctions of problem (6.2). As in the proof of Theorem 6.1 we check that

$$\int_\Omega \big(|Du|^2 - \lambda a u^2\big)\,dx \geq \big(1 - \frac{\lambda}{\lambda_{k+1}}\big)\|u\|^2$$

for all $u \in Z$. Since $f(x,u) = o(|u|)$ as $|u| \to 0$ uniformly in $x \in \Omega$ we have $F(x,u) = o(|u|^2)$ as $|u| \to 0$ uniformly in Ω. Consequently, there exist constants $r > 0$ and $b > 0$ such that $I(u) \geq b$ for $\|u\| = r$ and $u \in Z$. To show that condition (a) of Theorem 5.3 is satisfied we use the remark preceding the statement of this theorem. It is clear that $I(u) \leq 0$ for $u \in W$. We choose $e \in Z$, with $\|e\| = 1$ and $\bar{R} > r$ such that $I(u) \leq 0$ for $u \in W \bigoplus$ span e and $\|u\| > \bar{R}$. Towards this end it is sufficient to notice that by assumption (B)

$$I(te) \leq \frac{t^2}{2} \int_\Omega \left(|De|^2 - \lambda a e^2\right) dx - c_3 t^\mu \int_\Omega |e|^\mu \, dx + c_4 \to -\infty$$

as $t \to \infty$ and the claim easily follows. Finally, it remains to check that I satisfies the (PS)-condition. This amounts to showing that if $|I(u_m)| \leq M$ for all m and $I'(u_m) \to 0$ in $H^{-1}(\Omega)$, then $\{u_m\}$ is relatively compact in $H^1_\circ(\Omega)$. For large m and $\beta \in \left(\frac{1}{\mu}, \frac{1}{2}\right)$ we have by (A)

(6.5)

$$\begin{aligned}
M + \|u_m\| &\geq I(u_m) - \beta\langle I'(u_m), u_m\rangle \\
&= \int_\Omega \left[\left(\frac{1}{2} - \beta\right)|Du_m|^2 - \lambda\left(\frac{1}{2} - \beta\right) a u_m^2 + \beta f(x, u_m) u_m - F(x, u_m) \right] dx \\
&\geq \left(\frac{1}{2} - \beta\right)\|u_m\|^2 - \lambda\left(\frac{1}{2} - \beta\right)\|a\|_\infty \|u_m\|_2^2 + (\beta\mu - 1)\int_\Omega F(x, u_m)\, dx - k \\
&\geq \left(\frac{1}{2} - \beta\right)\|u_m\|^2 - \lambda\left(\frac{1}{2} - \beta\right)\|a\|_\infty \|u_m\|_2^2 + (\beta\mu - 1)(c_3\|u_m\|_\mu^\mu - c_4) - k
\end{aligned}$$

for some constant $k > 0$. We now observe that by the Young inequality for each $\epsilon > 0$ there exists $K(\epsilon) > 0$ such that

$$\|u\|_2^2 \leq K(\epsilon) + \epsilon \|u\|_\mu^\mu.$$

Taking $\epsilon > 0$ so that

$$(\beta\mu - 1)c_3 - \lambda\left(\frac{1}{2} - \beta\right)\|a\|_\infty \epsilon > 0$$

we deduce from (6.5) that $\{u_m\}$ is bounded in $H^1_\circ(\Omega)$ and the relative compactness of this sequence follows as in the proof of Lemma 4.2.

7. Ekeland's variational principle

In recent years Ekeland's variational principle has been successfully used in calculus of variations.

Theorem 7.1. *Let (M, d) be a complete metric space and let $\phi : M \to \mathbb{R} \cup \{\infty\}$ be lower semicontinuous functional which is bounded from below and $\not\equiv \infty$. Then for every $\epsilon > 0$ and $\lambda > 0$ and every $u \in M$ such that*

$$\phi(u) \leq \inf_M \phi + \epsilon$$

there exists an element $v \in M$ such that

$$\phi(v) \leq \phi(u), \quad d(u, v) \leq \frac{1}{\lambda}$$

and for each $w \neq v$ in M

$$\phi(w) > \phi(v) - \epsilon\lambda d(w, v).$$

Proof. It is sufficient to prove our assertion for $\lambda = 1$. The general case is obtained by replacing d by an equivalent metric λd. We define the relation on M

$$w \prec v \iff \phi(w) + \epsilon d(v, w) \leq \phi(v).$$

It is easy to see that this relation defines a partial ordering on M. We now construct inductively a sequence $\{u_m\}$ as follows: $u_\circ = u$, assuming that u_n has been defined we set

$$S_n = \{w \in M;\ w \prec u_n\}$$

and choose $u_{n+1} \in S_n$ so that

$$\phi(u_{n+1}) \leq \inf_{S_n} \phi + \frac{1}{n+1}.$$

Since $u_{n+1} \prec u_n$, $S_{n+1} \subset S_n$ and by the lower semicontinuity of ϕ, S_n is closed. We now show that $\lim_{n\to\infty}$ diam $(S_n) = 0$. Indeed, if $w \in S_{n+1}$, then $w \prec u_{n+1} \prec u_n$ and consequently

$$\epsilon d(w, u_{n+1}) \leq \phi(u_{n+1}) - \phi(w) \leq \inf_{S_n} \phi + \frac{1}{n+1} - \inf_{S_n} \phi = \frac{1}{n+1}.$$

This estimate implies that

$$\text{diam } S_{n+1} \leq \frac{2}{\epsilon(n+1)}$$

and our claim follows. The fact that M is complete implies that

$$\bigcap_{n\geq 0} S_n = \{v\}$$

for some $v \in X$. In particular, $v \in S_\circ$, that is,

$$v \prec u_\circ = u$$

and hence

$$\phi(v) \leq \phi(u) - \epsilon d(u, v) \leq \phi(u)$$

and moreover

$$d(u, v) \leq \epsilon^{-1}\big(\phi(u) - \phi(v)\big) \leq \epsilon^{-1}\big(\inf_M \phi + \epsilon - \inf_M \phi\big) = 1.$$

To complete the proof we must show $w \prec v$ implies $w = v$. If $w \prec v$, then $w \prec u_n$ for each integer $n \geq 0$, that, is $w \in \bigcap_{n\geq 0} S_n = \{v\}$.

This result requires a very low regularity from functional ϕ. It is not clear that a bounded and lower semicontinuous functional takes on its infimum or a maximum. For example a function $f(x) = \arctan x$ on $\mathbb{R}$ neither attains its infimum nor its maximum.

The Ekeland variational principle shows that a perturbation of ϕ given by

$$\phi_{\lambda\epsilon}(w) = \lambda\epsilon d(w, v) + \phi(w)$$

has an absolute minimizer. This simple observation has very interesting consequences.

Theorem 7.2. *If X is a Banach space and $\phi \in C^1(X, \mathbb{R})$ is bounded from below, then there exists a minimizing sequence $\{u_m\}$ for ϕ such that $\phi(u_m) \to \inf_X \phi$ and $\phi'(u_m) \to 0$ in X' as $m \to \infty$.*

Proof. Let $\epsilon_m \to 0$. For each m we choose $u_m \in X$ such that

$$\phi(u_m) \leq \inf_X \phi + \epsilon_m^2.$$

Applying Theorem 7.1 with $\lambda = \frac{1}{\epsilon_m}$ and $\epsilon = \epsilon_m^2$ we find for each m en element $v_m \in X$ such that

$$\phi(v_m) \leq \phi(v_m + w) + \epsilon_m \|w\|$$

for each $w \in X$. Hence

$$\|\phi'(v_m)\| = \lim_{\delta \to 0} \sup_{\|w\| \leq \delta, \|w\| \neq 0} \frac{\phi(v_m) - \phi(v_m + w)}{\|w\|} \leq \epsilon_m \to 0$$

and the result follows.

To obtain the existence of a minimizer we need to assume that ϕ satisfies the (PS)-conndition.

Theorem 7.3. *If $\phi \in C^1(X, \mathbb{R})$ is bounded from below and satisfies the (PS)-condition, then there exists $u \in X$ such that*

$$\phi(u) = \inf_{v \in X} \phi(v).$$

It is worth mentioning that a C^1-functional ϕ bounded from below and satisfying the (PS)-condition must be coercive, that is, $\lim_{\|u\| \to \infty} \phi(u) = \infty$.

Proposition 7.1. *Let $\phi \in C^1(X, \mathbb{R})$ be bounded from below. If ϕ satisfies the (PS)-condition, then ϕ is coercive.*

Proof. If the conclusion of the theorem were not true, there would exist $c \in \mathbb{R}$ such that $\lim_{\|u\| \to \infty} \phi(u) = c$. Then for every integer $m \geq 1$ there exists u_m such that $\phi(u_m) \leq c + \frac{1}{m}$ and $\|u_m\| \geq 2m$. By virtue of Theorem 7.1, applied with $\epsilon = c + \frac{1}{m} - \inf_X \phi$ and $\lambda = \frac{1}{m}$, we get the existence of $v_m \in X$ such that

$$\phi(v_m) \leq \phi(u_m) \leq c + \frac{1}{m},\ \|v_m - u_m\| \leq m,$$

$$\|v_m\| \geq \|u_m\| - \|u_m - v_m\| \geq \|u_m\| - m \geq m,$$

and

$$\phi(w) \geq \phi(v_m) - \frac{1}{m}\big(c + \frac{1}{m} - \inf_X \phi\big)\|w - v_m\|$$

for each $w \in X$. The last inequality implies that

$$\|\phi'(v_m)\| \leq \frac{1}{m}\big(c + \frac{1}{m} - \inf_X \phi\big).$$

Since $\|v_m\| \to \infty$, $\lim_{m\to\infty} \phi(v_m) = c$. On the other hand $\phi'(v_m) \to 0$ which contradicts the (PS)-condition.

Theorem 7.1 proved to be a very powerful tool in nonlinear analysis and found a lot of interesting applications. We give here applications to the fixed point theory and nonlinear boundary value problem. In Section 8 we shall discuss an application to the optimization and control theory.

Theorem 7.4. (**Caristi**) *Let (X, d) be a complete metric space and let $f : X \to X$ be a mapping satisfying*

$$d(u, f(u)) \leq \phi(u) - \phi(f(u))$$

for each $u \in X$, where $\phi : X \to [0, \infty)$ is a prescribed lower semicontinuous function. Then f has a fixed point.

Proof. It follows from Theorem 7.1 that there exists $v \in X$ such that

$$\phi(w) \geq \phi(v) - \frac{1}{2}d(v, w)$$

for each $w \in X$. In particular, letting $w = f(v)$ we get

$$\phi(v) - \phi(f(v)) \leq \frac{1}{2}d(v, f(v)).$$

According to our assumption we have

$$d(v, f(v)) \leq \phi(v) - \phi(f(v)).$$

The last two inequalities imply that

$$d(v, f(v)) \leq \frac{1}{2} d(v, f(v)),$$

which implies that $d(v, f(v)) = 0$ and the result follows.

A remarkable feature of this theorem is that we do not require the continuity of f.

As a second application we consider the nonhomogeneous Dirichlet problem

$$\begin{cases} -\Delta u = |u|^{p-2}u + f \text{ in } \Omega \\ u(x) = 0 \text{ on } \partial\Omega, \end{cases} \tag{7.1}$$

where $\Omega \subset \mathbb{R}_N$ is a bounded domain, $f \in H^{-1}(\Omega)$, $f \not\equiv 0$, and $2 < p < \frac{2N}{N-2}$. A solution of problem (7.1) will be found as a critical point of a variational functional

$$\phi(u) = \frac{1}{2}\int_\Omega |Du|^2\, dx - \frac{1}{p}\int_\Omega |u|^p\, dx - \int_\Omega fu\, dx.$$

Theorem 7.5. *There exists a constant $M = M(N, p, |\Omega|)$ such that if $\|f\|_{H^{-1}} \leq M$, then problem (7.1) possesses a solution in $H_\circ^1(\Omega)$.*

Proof. Letting $\|u\| = t$ and using the Sobolev inequality we get

$$\phi(u) \geq t\big(\frac{t}{2} - C\frac{t^{p-1}}{p} - \|f\|_{H^{-1}}\big) = t\big(h(t) - \|f\|_{H^{-1}}\big)$$

for some constant $C > 0$. Since $p > 2$ there exists $t_\circ > 0$ such that $\sup_{t\geq 0} h(t) = h(t_\circ) > 0$. Consequently, if $\|f\|_{H^{-1}} \leq h(t_\circ)$, then $\phi(u) \geq 0$ for $\|u\| = t_\circ$. If we choose $v \in H_\circ^1(\Omega)$ such that $\int_\Omega fv\, dx > 0$, we see that $\phi(tv) < 0$ for $t > 0$ sufficiently small. Hence

$$\inf_{\|u\|\leq t_\circ} \phi(u) < 0.$$

Also, by the first step of the proof there exists $0 < \delta < t_\circ$, such that

$$\phi(u) \geq \frac{1}{2} \inf_{\|w\| \leq t_\circ} \phi(w)$$

for all $u \in \{u;\ \delta \leq \|u\| < t_\circ\}$. We apply Theorem 7.1 on a metric space $\{u;\ \|u\| \leq t_\circ\}$ equipped with the metric $d(u_1, u_2) = \|u_1 - u_2\|$. Thus there exists a sequence $\{u_m\} \subset \{u;\ \|u\| < t_\circ\}$ such that each u_m is a minimizer of

$$\inf\{\phi(u) + \delta_m \|u_m - u\|;\ \|u\| \leq t_\circ\}$$

with $\delta_m \to 0$. This implies $\phi(u_m) \to \inf_{\|u\| \leq t_\circ} \phi$ and $\phi'(u_m) \to 0$ in $H^{-1}(\Omega)$ (see the proof of Theorem 7.2). It is now a routine to show that $u_m \to u$ up to a subsequence in $H_\circ^1(\Omega)$. Obviously u is a solution of problem (7.1).

8. Application to optimization and control theory

We consider a system governed by the system of equations

$$(8.1)\qquad \begin{cases} x'(t) = f(t, x(t), u(t)) \text{ a.e.,} \\ x(0) = x_\circ, \end{cases}$$

where $x(t) \in \mathbb{R}_N$ can be interpreted as the state of the system and $u(t)$ is a control at time t and belongs to a compact metric space K. A time $T > 0$ is given and we assume that

(a) $f, \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}$ are continuous on $[0,T] \times \mathbb{R}_n \times K$,

(b) $xf(t,x,u) \leq c(1+|x|^2)$ for some constant $c > 0$ and all $(t,x,u) \in [0,T] \times \mathbb{R}_n \times K$.

Let a measurable control $u : [0,T] \to K$ be given. Condition (a) guarantees the existence of a solution $x(t)$ on $[0,\tau]$ for $\tau > 0$ small. By Gronwall's inequality we have

$$(8.2)\qquad |x(t)|^2 \leq (|x_\circ|^2 + 2cT)e^{2cT}$$

and consequently $x(t)$ is defined on $[0,T]$. Also, (8.2) implies that

$$|x'(t)| \leq \max\{f(t,x,u);\ (t,x,u) \in [0,T] \times B(0,R) \times K\},$$

where $B(0,R)$ is a ball of radius $R = \left(|x_\circ|^2 + 2cT\right)^{\frac{1}{2}} e^{cT}$. It follows from the Ascoli - Arzela theorem that the collection of all trajectories for the control system (8.1) is equicontinuous and bounded and hence relatively compact in the topology of uniform convergence.

Let $g \in C^1(\mathbb{R}_N, \mathbb{R})$. Then we wish to find a measurable control $u(t)$ such that the corresponding trajectory minimizes $g(x(T))$ among all solutions of system (8.1).

Under the above assumptions, an optimal control may not exist except in special cases. However, using the Ekeland variational principle we show the existence of "an ϵ-approximate" optimal control.

First we formulate and prove some auxiliary results. We denote by $\mathcal{U}$ the set of all measurable controls $u : [0,T] \to K$ equipped with a metric

$$\delta(u_1, u_2) = |\{t \in [0,T];\ u_1(t) \neq u_2(t)\}|.$$

We recall that for a given $A \subset \mathbb{R}_n$, $|A|$ denotes the Lebesgue measure of A.

Lemma 8.1. *$(\mathcal{U}, \delta)$ is a complete metric space.*

Proof. It is easy to check that δ is a metric. Let $\{u_m\}$ be a Cauchy sequence in $\mathcal{U}$. We select a subsequence $\{u_{m_k}\}$ such that $\delta(u_{m_k}, u_{m_{k+1}}) \leq \frac{1}{2^k}$. We show that this subsequence converges in $\mathcal{U}$. Indeed, letting

$$A_k = \bigcup_{p \geq k} \{t;\ u_{m_p}(t) \neq u_{m_{p+1}}(t)\}$$

we get

$$|A_k| = \sum_{p \geq k}^{\infty} \frac{1}{2^p} = \frac{1}{2^{k-1}} \text{ and } A_k \supset A_{k+1}.$$

We define $\bar{u} \in \mathcal{U}$ by

$$\bar{u}(t) = u_{m_k}(t) \text{ for each } t \notin A_k.$$

It is clear that the subsequence $\{u_{m_k}\}$ converges to $\bar{u}$. Since $\{u_m\}$ is a Cauchy it must converge to $\bar{u}$.

Lemma 8.2. *The mapping $F : u \in \mathcal{U} \to g(x(T))$, where $x(t)$ is a corresponding solution to u of system 8.1, is continuous on $\mathcal{U}$.*

Proof. Let $u_m \to \bar{u}$ in $\mathcal{U}$. Since the corresponding sequence of solutions $\{x_m(t)\}$ is relatively compact, we can select a uniformly convergent subsequence to $\bar{x}(t)$. It remains to show that $\bar{x}$ is a trajectory corresponding to $\bar{u}$. To show this we write

$$x_m(t) = x_\circ + \int_0^t f(s, x_m(s), u_m(s))\, ds.$$

Since the integrand $f(s, x_m(s), u_m(s))$ is uniformly bounded, it follows from the Lebesgue dominated convergence theorem that

$$\bar{x}(t) = x_\circ + \int_0^t f(s, \bar{x}(s), \bar{u}(s))\, ds.$$

Theorem 8.1. *For every $\epsilon > 0$ there exists a measurable control u_ϵ with corresponding trajectory x_ϵ such that*

$$g(x_\epsilon(T)) \leq \inf g(x(T)) + \epsilon \tag{8.3}$$

and

$$\langle f(t, x_\epsilon(t), u_\epsilon(t)), p_\epsilon \rangle \leq \min_{u \in K} \langle f(t, x_\epsilon(t), u), p_\epsilon(t) \rangle + \epsilon, \tag{8.4}$$

where p_ϵ is a solution of the following linearized system

$$\begin{cases} p'_\epsilon(t) = -{f'_x}^T(t, x_\epsilon(t), u_\epsilon(t)) p_\epsilon(t), \\ p_\epsilon(T) = g'(x_\epsilon(T)). \end{cases} \tag{8.5}$$

(If (8.3) holds with $\epsilon = 0$, so does (8.4). This means that an optimal control satisfies the Pontryagin mnimum principle).

Proof. It follows from the Ekeland variational principle that given $\epsilon > 0$ we obtain a measurable control $u_\epsilon \in \mathcal{U}$ such that

$$F(u_\epsilon) \leq \inf_{u \in \mathcal{U}} F(u) + \epsilon$$

and

$$F(u) \geq F(u_\epsilon) - \epsilon \delta(u, u_\epsilon) \tag{8.6}$$

for each $u \in \mathcal{U}$. Obviously a trajectory x_ϵ corresponding to u_ϵ satisfies

$$x'_\epsilon(t) = f(t, x_\epsilon(t), u_\epsilon(t)), \text{ a. e}$$

$$x_\epsilon(0) = x_\circ.$$

Let $t_\circ \in (0, T)$ be a point where the equality holds and let $u_\circ \in K$. We define $v_\tau \in \mathcal{U}$, $\tau \geq 0$, by

$$v_\tau(t) = \begin{cases} u_\circ & \text{for } t \in [0,T] \cap (t_\circ - \tau, t_\circ) \\ u_\epsilon(t) & \text{for } t \notin [0,T] \cap (t_\circ - \tau, t_\circ). \end{cases}$$

Let x_τ be a corresponding trajectory. We claim that

$$(8.7) \qquad \frac{d}{d\tau} g(x_\tau(T)) \mid_{\tau=0} = \langle f(t_\circ, x_\epsilon(t_\circ), u_\circ) - f(t_\circ, x_\epsilon(t_\circ), u_\epsilon(t_\circ)), p_\epsilon(t_\circ)\rangle.$$

To show this we write

$$\begin{aligned} x_\tau(t_\circ) &= x_\epsilon(t_\circ - \tau) + \int_{t_\circ - \tau}^{t_\circ} f(s, x_\tau(s), u_\circ)\, ds \\ &= x_\epsilon(t_\circ) - \tau x_\epsilon'(t_\circ) + \tau f(t_\circ, x_\epsilon(t_\circ), u_\circ) + o(\tau) \\ &= x_\epsilon(t_\circ) - \tau\big(f(t_\circ, x_\epsilon(t_\circ), u_\epsilon(t_\circ)) - f(t_\circ, x_\epsilon(t_\circ), u_\circ))\big) + o(\tau), \end{aligned}$$

and this implies that

$$(8.8) \qquad \frac{d}{d\tau} x_\tau(t_\circ) \mid_{\tau=0} = f(t_\circ, x_\epsilon(t_\circ), u_\circ) - f(t_\circ, x_\epsilon(t_\circ), u_\epsilon(t_\circ)).$$

Introducing the resolvent $R(t, t_\circ)$ of the linearized system

$$\xi'(t) = f_x'(t, x_\epsilon(t), u_\epsilon(t), t))\xi(t)$$

formula (8.8) takes the form

$$\frac{d}{d\tau} x_\tau(T) \mid_{\tau=0} = R(T, t_\circ)\left[f(t_\circ, x_\epsilon(t_\circ), u_\circ) - f(t_\circ, x_\epsilon(t_\circ), u_\epsilon(t_\circ))\right].$$

By straightforward calculations we obtain

$$\begin{aligned} \frac{d}{d\tau} g(x_\tau(T)) \mid_{\tau=0} &= \langle g'(x_\epsilon(T)), \frac{d}{d\tau} x_\tau(T) \mid_{\tau=0}\rangle \\ &= \langle g'(x_\epsilon(T)), R(T, t_\circ)[f(t_\circ, x_\epsilon(t_\circ), u_\circ) - f(t_\circ, x_\epsilon(t_\circ), u_\epsilon(t_\circ))]\rangle \\ &= \langle R(T, t_\circ)^T g'(x_\epsilon(T)), [f(t_\circ, x_\epsilon(t_\circ), u_\circ) - f(t_\circ, x_\epsilon(t_\circ), u_\epsilon(t_\circ))]\rangle. \end{aligned}$$

We now observe that

$$p_\epsilon(t) = R(T, t_\circ)^T g'(x_\epsilon(T))$$

is a solution of (8.5). Since $\delta(u_\epsilon, u_\tau) \le \tau$, inequality (8.6) implies that

$$g(x_\tau(T)) - g(x_\epsilon(T)) \ge -\epsilon\tau.$$

This combined with (8.7) gives

$$\langle f(t_\circ, x_\epsilon(t_\circ), u_\circ) - f(t_\circ, x_\epsilon(t_\circ), u_\epsilon(t_\circ)), p_\epsilon(t_\circ)\rangle \ge -\epsilon.$$

Since $u_\circ$ is arbitrary point of K and $t_\circ$ is any point of $(0, T)$, where x_ϵ satisfies our system, inequality (8.4) easily follows.

9. Mountain pass theorem and Ekeland's variational principle

The proof of the mountain pass theorem was based on the deformation lemma. It is quite interesting fact that the mountain pass theorem can be deduced from the Ekeland variational principle.

We shall use the following result the proof of which is similar to that of Lemma 3.1. In what follows X denotes a Banach space.

Lemma 9.1. *Let K be a metric space and let $f : K \to X'$ be a continuous function. Then, given $\epsilon > 0$, there exists a locally Lipschitz mapping $v : K \to X$ such that*

$$\|v(\xi)\| \leq 1$$

and

$$\langle f(\xi), v(\xi)\rangle \geq \|f(\xi)\| - \epsilon$$

for all $\xi \in X$.

Let K be a compact metric space and let K^* be a nonempty closed subset of $K \neq K$.

Let

$$\mathcal{A} = \{p \in C(K, X);\ p = p^* \text{ on } K^*\},$$

where p^* is a fixed continuous mapping on K. The set $\mathcal{A}$ equipped with a metric $d(p,q) = \max_{\xi\in K} \|p(\xi) - q(\xi)\|$ is a complete metric space.

Theorem 9.1. *For a given $\phi \in C^1(X, \mathbb{R})$ we set*

$$c = \inf_{p\in\mathcal{A}} \max_{\xi\in K} \phi(p(\xi)).$$

Suppose that for every $p \in \mathcal{A}$, $\max_{\xi\in K} \phi(p(\xi))$ is attained at some point in $K - K^*$. Then there exists a sequence $\{u_m\} \subset X$ such that

$$\phi(u_m) \to c \text{ and } \|\phi'(u_m)\| \to 0$$

in X'. If in addition ϕ satisfies the $(PS)_c$-condition, then c is a critical value. (If $K=[0,1]$, $K^*=\{0,1\}$ and $p^*(t)=t\dot{e}$, we obtain the mountain pass theorem.)

Proof. For $\xi \in K$, we set

$$d(\xi) = \min\big(\text{ dist } (\xi, K^*), 1\big)$$

and consider for every fixed $\epsilon > 0$ and $p \in \mathcal{A}$

$$G(p,\xi) = \phi(p(\xi)) + \epsilon d(\xi).$$

We set

$$\psi_\epsilon(p) = \max_{\xi \in K} G(p(\xi), \xi)$$

and let

$$c_\epsilon = \inf_{p \in \mathcal{A}} \psi_\epsilon(p).$$

We obviously have $c \le c_\epsilon \le c + \epsilon$. By the Ekeland variational principle there exists $p \in \mathcal{A}$ such that

$$\psi_\epsilon(q) - \psi_\epsilon(p) + \epsilon d(p,q) \ge 0 \tag{9.1}$$

for all $q \in \mathcal{A}$ and

$$c \le c_\epsilon \le \psi_\epsilon(p) \le c_\epsilon + \epsilon \le c + 2\epsilon.$$

According to our assumption we have

$$\psi_\epsilon(p) > \max_{\xi \in K^*} \psi(p(\xi)). \tag{9.2}$$

We put

$$B_\epsilon(p) = \{\xi \in K;\ G(p(\xi), \xi) = \psi_\epsilon(p)\}.$$

We shall prove that there exists $\xi_\circ \in B_\epsilon(p)$ such that

$$\|\phi'(p(\xi_\circ))\| \le 2\epsilon.$$

The conclusion then follows by choosing $\epsilon = \frac{1}{n}$ and $u_n = p(\xi_\circ)$. We now apply Lemma 9.1 with $f(\xi) = \phi'(p(\xi))$ to obtain a continuous mapping $v : K \to X$ such that

$$\|v(\xi)\| \leq 1 \text{ and } \langle \phi'(p(\xi)), v(\xi)\rangle \geq \|\phi'(p(\xi))\| - \epsilon$$

for all $\xi \in K$. Inequality (9.2) implies that $B_\epsilon(p) \subset K - K^*$. Hence, there exists a continuous function $\alpha : K \to [0,1]$ such that $\alpha(\xi) = 1$ on $B_\epsilon(p)$ and $\alpha(\xi) = 0$ on K^*. For small $h > 0$ we define

$$q_h(\xi) = p(\xi) - hw(\xi) = p(\xi) - h\alpha(\xi)v(\xi).$$

Since $\alpha = 0$ on K^*, $q_h \in \mathcal{A}$. We now observe that

$$\psi_\epsilon(q_h) = \max_{\xi \in K} G(q_h(\xi), \xi)$$

is attained at some point $\xi_h \in K$. Since K is compact we may assume that $\xi_{h_n} \to \xi_\circ$ for some $h_n \to 0$. Obviously, $\xi_\circ \in B_\epsilon(p)$. It follows from (9.1) with $q = q_h$ that

$$\phi(p(\xi_h) - hw(\xi_h)) + \epsilon d(\xi_h) - \psi_\epsilon(p) + \epsilon h \geq 0. \tag{9.3}$$

On the other hand we have

$$\phi(p(\xi_h) - hw(\xi_h)) = \phi(p(\xi_h)) - \langle \phi'(p(\xi_h)), hw(\xi_h)\rangle + o(h). \tag{9.4}$$

Combining (9.3) and (9.4) with the fact that

$$\phi(p(\xi_h)) + \epsilon d(\xi_h) \leq \psi_\epsilon(p)$$

we get

$$-h\langle \phi'(p(\xi_h)), w(\xi_h)\rangle + \epsilon h + o(h) \geq 0.$$

This implies that

$$\langle \phi'(p(\xi_h)), w(\xi_h)\rangle \leq \epsilon + o(1).$$

Letting $h \to 0$ we get that

$$\langle \phi'(p(\xi_\circ)), v(\xi_\circ)\rangle \leq \epsilon$$

and the result follows applying Lemma 9.1.

Bibliographical comments

Struwe's monograph [**22**] is an excellent source of information of basic methods and results in critical point theory. The mountain pass theorem appears to have been first noted in the literature by Ambrosetti and Rabiniwitz [**1**]. Their proof has been based on a deformation lemma [**19**], [**20**]. A version of the deformation lemma (Lemma 3.2) used in these lectures is taken from Willem [**24**] (see also [**22**]). A topological linking was introduced by Benci and Rabinowitz [**3**], [**2**]. The proof of Theorem 4.2 was given by Katriel [**13**]. Theorems 4.3, 6.1 and 6.2 can be found in [**7**], [**19**]. Theorem 7.1 is due to Ekeland [**7**], [**8**]. Results of Section 8 are taken from [**18**]. The fact that the Ekeland variational principle implies the mountain pass theorem has been observed many authors [**4**], [**10**]. The approach to this fact presented in Section 9 is based on paper [**4**]. There is a number of papers ([**10**], [**12**], [**16**], [**18**]) investigating the nature of critical points. A good account of these results can also be found in Ghoussoub's monograph [**11**]. One might conjucture from the geometric interpretation of the mountain pass theorem that the "mountain pass" in journey from $(0,\phi(0))$ to $(e,\phi(e))$ must be a saddle point. However, this may not occur if the mountain range surrounding $(0,\phi(0))$ everywhere has the same height. In this situation the mountain pass is a maximum. It has been proved by Pucci and Serrin [**18**] that in an infinite dimensional Banach space a set of critical points in the mountain pass theorem must contain at least one saddle point.

The Ekeland variational principle is deeply connected with the geometry of Banach spaces. The E.V.P is equivalent to the drop theorem and the petal flower theorem. To describe these striking results, let D be a convex set in a normed linear vector space X. A drop $K(D,u)$ with vertex u and basis D is defined by

$$K(D,u) = \text{convex hull of } D\cup\{u\} = \{u+t(v-u),\ v\in D, 0\le t\le 1\}.$$

A petal $P_\gamma(u,v)$ associated with $\gamma\in(0,\infty)$ and points u and v in a metric space (M,d) is the set

$$P_\gamma(u,v) = \{x\in M;\ \gamma d(x,u)+d(x,v)\le d(u,v)\}.$$

We observe that $P_\gamma(u,v) \subset P_\delta(u,v)$ if $\delta \leq \gamma$. If $B(v,r)$ is a ball with center v and radius r in a normed linear vector space X and $\gamma \leq \frac{t-r}{t+r}$, with $t = d(u,v) > r$, then by convexity $D(B(v,r),u) \subset P_\gamma(u,v)$.

Theorem A. (**the flower petal theorem**) *Let Y be a complete subset of a metric space (M,d). Let $x_\circ \in Y$ and let $u \in M - Y$ with $s = d(x_\circ,u)$ and let $r \leq d(u,Y)$. Then for every $\gamma > 0$ there exists $v \in P_\gamma(x_\circ,u) \cap Y$ (so that $d(v,x_\circ) \leq \gamma^{-1}(s-r)$) such that $P_\gamma(v,u) \cap Y = \{v\}$.*

Theorem B. (**the drop theorem**) *Let X be a Banach space, let S be a closed subset of X and let D be a closed bounded convex subset of X with $d(S,D) > 0$. If $u \in S$, then there exists a point $v \in S \cap K(D,u)$ such that $K(D,v) \cap S = \{v\}$.*

Both Theorems A and B are equivalent to the E.V.P. For proofs of these facts we refer to papers Danes̑ [**6**], Georgiev [**9**] and Penot [**15**].

Finally, another interesting observation, due to Sullivan [**23**], shows that the validity of assertion of the E.V.P on a metric space (M,d) is equivalent to the completeness of (M,d).

BIBLIOGRAPHY

1. A. Ambrosetti & P.H. Rabinowitz, *Dual variational methods in critical point theory and applications*, J. Funct. Anal. **14** (1973), 349–381.
2. P. Bartolo, V. Benci & D. Fortunato, *Abstract point theorem and applications to some nonlinear problems with "strong" resonance at infinity*, Nonlinear Analysis TMA **7**, 981–1012.
3. V. Benci & P.H. Rabinowitz, *Critical points theorems for indefinite functionals*, Invent. Math **52** (1979), 241–273.
4. H. Brezis & L. Nirenberg, *Remarks on finding critical points*, Comm. Pure Appl. Maths **44** (1991), 939–963.
5. J. Caristi, *Fixed point theorems for mappings satisfying inwardness conditions*, Trans. Am. Math. Soc. **215** (1976), 242–252.
6. J. Daneŝ, *Equivalence of some geometric and related results of nonlinear functional analysis*, Comm. Math. Univ. Carolinae **26.3** (1985), 443–454.
7. I. Ekeland, *On the variational principle*, J. Math. Anal. Appl. **47** (1974), 324–353.
8. I. Ekeland, *Nonconvex minimizations problems*, Bull. Am. Math. Soc. **1(3)** (1979), 443–474.
9. P.G. Georgiev, *The strong variational principle, the strong drop theorem and applications*, J. Math. Anal. Appl. **131** (1988), 1–21.
10. N. Ghoussoub & D. Preiss, *A general mountain pass principle for locating and classifying critical points*, Ann. Inst. H. Poincaré **6(5)** (1989), 321–330.
11. N. Ghoussoub, *Duality and perturbation methods in critical point theory*, Cambridge Univ. Press, 1993.
12. H. Hofer, *A geometric description of the neighbourhood of a critical point given by the mountain pass theorem*, J. London Math. Soc. **31** (1985), 566–570.
13. G. Katriel, *Mountain pass theorems and global homeomorphism theorems*, Ann. Inst. H. Poincaré, Analyse non linéaire **11 (2)** (1994), 189–209.

14. R. Pallu de la Barriere, *Optimal control theory*, Saunders, Philadelphia, Penn., 1967.

15. J.P. Penot, *The drop theorem, the petal theorem and Ekeland's variational principle*, Nonlinear Analysis, TMA **10(9)** (1986), 813–822.

16. P. Pucci & J. Serrin, *A mountain pass theorem*, J. Diff. Equations **60** (1985), 142–149.

17. P. Pucci & J. Serrin, *Extensions of the mountain pass theorem*, J. Funct. Analysis **59** (1984), 185–210.

18. P. Pucci & J. Serrin, *The structure of critical set in the mountain pass theorem*, Trans. Am. Math. Soc. **299(1)** (1987), 115–132.

19. P.H. Rabinowitz, *Minimax methods in critical point theory with applications to differential equations*, CBMS, Am. Math. Soc. **65** (1986).

20. P.H. Rabinowitz, *Variational methods for nonlinear eigenvalue problems*, CIME (1974).

21. M. Ramos & C. Rebelo, *A unified approach to min - max critical point theory*, Portugaliae Math. **51(4)** (1994), 489–515.

22. M. Struwe, *Variational Methods*, Springer - Verlag, Berlin Heidelberg, 1988.

23. F. Sullivan, *A characterization of complete metric spaces*, Proc. Am. Math. Soc. **83 (2)** (1981), 345–346.

24. M. Willem, *Une lemme de deformation quantitif en calcul des variations*, Institut de mathematique pure et appliquee, Université Catholique de Louvain.