Mathematical Methods in Optimization of Differential Systems

Mathematics and Its Applications

Volume 310

Mathematical Methods in Optimization of Differential Systems

by

Viorel Barbu
Department of Mathematics,
University "Al. I. Cuza",
Iaşi, Romania

SPRINGER SCIENCE+BUSINESS MEDIA, B.V.

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN 978-94-010-4327-4 ISBN 978-94-011-0760-0 (eBook)
DOI 10.1007/978-94-011-0760-0

This is an updated and revised translation
of the original Romanian work
Metode Matematice in Optimazarea Sistemelor Differentiale
Editura Academiei, Bucharest © 1989

Printed on acid-free paper

Originally published by Kluwer Academic Publishers in 1994

Contents

Preface

This work is a revised and enlarged edition of a book with the same title published in Romanian by the Publishing House of the Romanian Academy in 1989. It grew out of lecture notes for a graduate course given by the author at the University if Iaşi and was initially intended for students and readers primarily interested in applications of optimal control of ordinary differential equations. In this vision the book had to contain an elementary description of the Pontryagin maximum principle and a large number of examples and applications from various fields of science.

The evolution of control science in the last decades has shown that its methods and tools are drawn from a large spectrum of mathematical results which go beyond the classical theory of ordinary differential equations and real analyses. Mathematical areas such as functional analysis, topology, partial differential equations and infinite dimensional dynamical systems, geometry, played and will continue to play an increasing role in the development of the control sciences. On the other hand, control problems is a rich source of deep mathematical problems. Any presentation of control theory which for the sake of accessibility ignores these facts is incomplete and unable to attain its goals. This is the reason we considered necessary to widen the initial perspective of the book and to include a rigorous mathematical treatment of optimal control theory of processes governed by ordinary differential equations and some typical problems from theory of distributed parameter systems.

However, this book is not a monograph and it does not intend to compete with the numerous excellent textbooks on mathematical theory of optimal control. The primary intention of this book has been to familiarize the reader with the problems, approaches and basic results of the theory with main emphasis on the mathematical ideas and tools which enriched and influenced the development of control science in the last years. For instance the convex analysis and theory of generalized gradients have provided a new setting for the formulation and the proof of Pontryagin maximum principle whilst the theory of generalized and viscosity solutions to Hamilton–Jacobi equations has contributed in significant ways to understand the nature of the dynamic programming equation and the structure of the optimal feedback controllers. The use of continuous semigroups of linear operators in Banach spaces has led to a general reformulation of linear control problems governed by partial differential equations (parameter distributed systems). The classical linear quadratic control problem on finite and infinite intervals as well as the H_∞-control problem have elegant reformulations and natural infinite dimensional generalizations into this framework. The optimal control of partial differential equations with free boundary, such as the Stefan problem describing solidification processes, owes much to new smoothing and penalty techniques for nonlinear control systems. All these topics are discussed in some details in this book. Some of the auxiliary results from convex and nonsmooth analysis are given in Chapter I.

The references are not complete and refer only to the works closely related or used in this book.

I would like to express my thanks to S. Anita and C. Popa who read the lecture notes on which this book is based and suggested improvements.

Symbols and Notations

R^n	then n-dimensional Euclidean space
R	the real line $(-\infty, +\infty)$
R^+	the half line $[0, +\infty)$
$\bar{R}$	$(-\infty, +\infty]$
$x \cdot y$	the dot product of vectors $x, y \in R^n$
$\|.\|_X$	the norm of a linear normed space X
∇f	the gradient of the map $f : X \to Y$
f_x, f_u	the partial derivatives of the map (function) $f : X \times U \to Y$
∂f	the subdifferential (generalized gradient) of the function $f : X \to \bar{R}$
$L(X, Y)$	the space of linear continuous operators from X to Y
$\|.\|_{L(X,Y)}$	the norm of $L(X, Y)$
X^*	the dual of the space X
(x, y)	the scalar product of the vectors $x, y \in H$ (a Hilbert space) If $x \in X, y \in X^*$ this is the value of y at x.
A^*	the adjoint of linear operator A
h^*	the conjugate of the function $h : X \to \bar{R}$
sgn	the signum function: sgn $r = r/\|r\|$, $r \neq 0$; sgn 0 = $[-1, 1]$
$\Omega \subset R^n$	an open subset of R^n
$C(\Omega)$	the space of real valued continuous functions on Ω
$C(\bar{\Omega})$	the space of real valued continuous functions on the closure $\bar{\Omega}$ of Ω
$C_0^\infty(\Omega)$	the space of all continuously in finitely differentiable functions with compact support in Ω
$\mathcal{D}'(\Omega)$	the space of all distributions on Ω (the dual of $C_0^\infty(\Omega)$)
$L^p(\Omega), \leq p \leq \infty$	the space of all p-summable functions on Ω
$D_{x_j}^{\alpha_j} u$	$\frac{\partial \alpha_j}{\partial_{x_j}^{\alpha_j}} u$, $j = 1, \ldots, n$
$D^\alpha u$	$D_{x_1}^{\alpha_1} \ldots D_{x_n}^{\alpha_n} u$; $\alpha = (\alpha_1, \ldots, \alpha_n)$

$W^{k,p}(\Omega)$	the Sobolev space $\{u \in L^p, D^\alpha u \in L^p(\Omega), \lvert\alpha\rvert \leq k\}$ where D^α is in the sense of distributions
$H^k(\Omega)$	the Sobolev space $W^{k,2}(\Omega)$
$H_0^1(\Omega)$	the closure of $C_0^\infty(\Omega)$ in $H^1(\Omega)$
$L^p(0,T;X)$	(X Banach space) the space of all p summable functions from $[0,T]$ to X
$y'(t)$, $\frac{dy}{dt}(t)$	the derivative of the function $y:[0,T] \to X$
$AC([0,T];X)$	the space of absolutely continuous functions from $[0,T]$ to X
$W^{1,p}([0,T];X)$	the space $\{y \in AC([0,T];y' \in L^p(0,T;X))\}$
$C([0,T];X)$	the space of all continuous functions from $[0,T]$ to X
$T_K(u)$	the tangent cone to K at u
$N_K(u)$	the normal cone to K at u

Chapter I

Generalized Gradients and Optimality

The modern theory of optimization has extended and developed the classical concept of gradient into several directions in order to treat and solve nonsmooth problems. The purpose of this chapter which has a preliminary character is two fold: to introduce the basic terminology and results of convex analysis and theory of generalized gradients and to present, albeit briefly, some implications to infinite dimensional abstract optimization.

1. Fundamentals of Convex Analysis

For a complete treatment of the subject the reader is referred to the classical texts on convex analysis (Moreau [10], Rockafellar [11]). Here we shall confine our presentation to some basic results which will be used in the subsequent chapters.

Throughout this paragraph X will be a real Banach space with the norm denoted $\|.\|$ and the dual X^*. We shall denote by $(.,.)$ the duality pairing between X and X^*.

The function $f : X \to R = (-\infty, +\infty]$ is said to be *convex* if

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) \tag{1.1}$$

for all $x, y \in X$ and $\lambda \in [0,1]$.

The function f is said to be *strictly convex* if inequality (1.1) is strict for $0 < \lambda < 1$ and $x \neq y$.

The set $D(f) = \{x \in X; f(x) < +\infty\}$ is called the *effective domain* of f and the set

$$\text{epi}(f) = \{(x,\lambda) \in X \times R; f(x) \leq \lambda\} \tag{1.2}$$

is called the *epigraph* of f. It is readily seen that the function f is convex if and only if epi(f) is a convex subset of $X \times R$. The function $f : X \to \bar{R}$ is said to be *lower semicontinuous* (l.s.c.) at x_0 if

$$\liminf_{x \to x_0} f(x) \geq f(x_0). \tag{1.3}$$

(We have used the notation

$$\liminf_{x \to x_0} f(x) = \sup_{V \in \mathcal{V} \backslash x_0} f(x)$$

where $\mathcal{V}(x_0)$ is a base of neighborhoods of x_0.) A function f which is lower semicontinuous at every point x of a subset D of X is said to be lower-semicontinuous on D. It is easily seen that f is lower semicontinuous on X if and only if its epigraph, epi(f) is closed in $X \times R$. Note also that f is lower semicontinuous on X if and only if every level set $\{x \in X; f(x) \leq \lambda\}$ is closed in X. The function $f : X \to \bar{R}$ is said to be *proper* if $f \not\equiv +\infty$; $f : X \to \bar{R}$ is said to be *concave* if $-f$ is convex. On the space X we may introduce the weak topology defined by the family of seminorms $p_{x^*}(x) = (x, x^*)$, $\forall x \in X, x^* \in X^*$.

Accordingly, a sequence $\{x_n\} \subset X$ is said to be weakly convergent to x (we denote this by the symbol $x_n \rightharpoonup x$) if $\lim_{n \to \infty}(x_n, x^*) = (x, x^*)$, $\forall x^* \in X^*$. (Recall that (x, x^*) is the value of functional $x^* \in X^*$ at $x \in X$.)

An immediate consequence of the Hahn–Banach separation theorem is that a convex set of X is closed if and only if its weakly closed (i.e., closed in the weak topology) (see e.g. Yosida [17], p. 119). Applying this result to the level set $\{x; f(x) \leq \lambda\}$ of a convex function f we deduce that

PROPOSITION 1.1. *A lower semicontinuous convex function f is weakly lower semicontinuous, i.e.,*

$$\liminf_{x_n \rightharpoonup x_0} f(x_n) \geq f(x_0) \quad \forall x_0 \in X. \tag{1.4}$$

If the space X is reflexive (i.e. $X = X^*$) then every bounded subset is sequentially weakly relatively compact and so by the Weierstrass theorem every weakly lower semicontinuous function on such a set attains its infimum. Then Proposition 1.1. leads us to the following existence result.

PROPOSITION 1.2. *Let X be a reflexive Banach space and let $f : X \to \bar{R}$ be a lower semicontinuous convex function. If K is a closed, convex and bounded subset of X then f attains its infimum on K.*

As an immediate consequence of Proposition 1.2 we get

COROLLARY 1.1. *Let X be a reflexive Banach space and let $f : X \to \bar{R}$ be a lower semicontinuous convex function such that*

$$\lim_{\|x\|\to\infty} f(x) = +\infty. \tag{1.5}$$

Then f attains its infimum on X.

Proof. By coercivity condition (1.5) we see that $\inf\{f(x); x \in X\} > -\infty$ and that for some $R > 0$ we have

$$\inf\{f(x); x \in X\} = \inf\{f(x); \|x\| \leq R\}.$$

Then one applies Proposition 1.2 to conclude that f attains its infimum on $\{x; \|x\| \leq R\}$ and therefore on X. ■

PROPOSITION 1.3. *Let $f : X \to \bar{R}$ be a lower semicontinuous, convex and proper. Then f is bounded from below by an affine function, i.e.,*

$$f(x) \geq (x, x_0^*) + \alpha \quad \forall x \in X \tag{1.6}$$

for some $x_0^* \in X^*$ and $\alpha \in R$.

Proof. Let $E =$ epi (f) and let $(x_0, \lambda_0) \in X \times R$ be such that $f(x_0) > \lambda_0$, $f(x_0) \neq \infty$. Then by the Hahn–Banach separation theorem there exists a hyperplane $H = \{(x, \lambda) \in X \times R; (x, x_1^*) + \lambda\mu_0 = 1\}$, $\mu_0 \geq 0$, which separates E and $(x_0, f(x_0))$, i.e.

$$\begin{aligned} (x, x_1^*) &+ \lambda\mu_0 > 1 \quad \forall (x, \lambda) \in E \\ (x_0, x_1^*) &+ \lambda_0\mu_0 < 1. \end{aligned} \tag{1.7}$$

Clearly, $\mu_0 > 0$, i.e., the hyperplane H is not vertical. Then by (1.7) we get (1.6) as claimed. ■

PROPOSITION 1.4. *Let $f : X \to \bar{R}$ be a lower semicontinuous, convex and proper. Then f is continuous at every interior point of $D(f)$.*

Proof. Let x_0 be an interior point of $D(f)$. Without loss of generality we may assume that $x_0 = 0$, $f(0) = 0$. Let $\lambda \in R$ be such that $f(0) < \lambda$ and set $E = \{x \in X; f(x) \leq \lambda\}$. The set E is convex, closed and $0 \in$ int $E \subset$ int $D(f)$. Since f is convex we have

$$f(x) \leq \varepsilon f(\varepsilon^{-1}x) \leq \varepsilon M \quad \forall x \in \varepsilon V$$

where V is a ball centered at 0. Hence $\lim_{x\to 0} f(x) = 0$ as claimed.

In particular, it follows by Proposition 1.4 that if f is convex, lower semicontinuous and everywhere finite on X then it is continuous on X. If the space X is finite dimensional then every convex function is continuous on the interior of the effective domain.

If C is a closed convex set of X, the function $I_C : X \to \bar{R}$,

$$I_C(x) = \begin{cases} 0 & \text{if} \quad x \in C \\ +\infty & \text{if} \quad x \bar{\in} C \end{cases}$$

is called the *indicator function* of C. Clearly I_C is convex and lower semicontinuous on X.

Given the lower semicontinuous, convex function $f : X \to \bar{R}$, the mapping $\partial f : X \to X^*$ defined by

$$\partial f(x) = \{w \in X^*; (x - y, w) \geq f(x) - f(y) \quad \forall y \in X\} \tag{1.8}$$

is called the *subdifferential* of function f. The mapping ∂f is in general multivalued and an element $w \in \partial f(x)$ is called a *subgradient* of f at x. The set $x \in X$ for which $\partial f(x) \neq \emptyset$ is called the domain of the mapping ∂f and is denoted $D(\partial f)$.

The function $f : X \to R$ is said to be *Gâteaux differentiable* at x if

$$\lim_{\lambda\to 0} \frac{f(x + \lambda y) - f(x)}{\lambda} = \eta_y$$

exists for all $y \in X$ and $y \to \eta_y$ is a linear continuous functional on X, i.e., $\eta_y = (y, \nabla f(x)), \forall y \in X$. The element $\nabla f(x) \in X^*$ is called the *gradient* of f at x.

It is easily seen that if f is convex and Gâteaux differentiable at x then $x \in D(\partial f)$ and $(\partial f)(x) = \nabla f(x)$. Indeed we have

$$\begin{aligned} -(x - y, \nabla f(x)) &= \lim_{\lambda\to 0} \lambda^{-1}(f(x - \lambda(x - y)) - f(x)) \\ &\leq f(y) - f(x) \quad \forall y \in X. \end{aligned}$$

i.e., $\nabla f(x) \in \partial f(x)$. Now let $w \in \partial f(x)$ be arbitrary but fixed. We have therefore

$$(x - y, w) \geq \lambda^{-1} (f(x) - f(x - \lambda(x - y))) \quad \forall y \in X,$$

and this implies that $(x - y, w - \nabla f(x)) \geq 0$, $\forall y \in X$, i.e., $\nabla f(x) = w$ as claimed. ■

If $f(x) = \frac{1}{2}\|x\|^2$ then ∂f is just the duality mapping of the space X, i.e.,

$$\partial f(x) = J(x) = \{w \in X^*; (x, w) = \|x\|^2; \|w\| = \|x\|\}.$$

If $f(x) = \|x\|$, then

$$f(x) = \text{ sgn } x = \begin{cases} J(x)\|x\|^{-1} & \text{if } \ x \neq 0 \\ \{y \in X^*; \|y\| \leq 1\} & \text{if } \ x = 0. \end{cases}$$

Note also that if $f = I_C$ where C is a closed convex subset of x then

$$\partial f(x) = N_C(x) = \{w \in X^*; (x - y, w) \geq 0, \forall y \in C\} \tag{1.9}$$

and $D(\partial f) = C$.

The set $N_C(x) \subset X^*$ is called the *normal cone* to C at x.

By definition of subdifferential ∂f it is also immediately seen that f attains its infimum on X at x_0 if and only if $0 \in \partial f(x_0)$. In other words, we have

$$\arg\inf(f) = (\partial f)^{-1}(0). \tag{1.10}$$

We note the following result.

PROPOSITION 1.5. *Let $f : X \to \bar{R}$ be a lower semicontinuous, convex and proper function. Then* int $D(f) \subset D(\partial f)$.

Proof. Denote by E the epigraph of f and take $x_0 \in \text{int } D(f)$. Since as seen earlier f is continuous at x_0 it follows that $(x_0, f(x_0) + \varepsilon) \in \text{int } E$ for all $\varepsilon > 0$. Since $(x_0, f(x_0))$ is a boundary point of E we infer that there is a closed tangent hyperplane to E through $(x_0, f(x_0))$. Hence there are $w \in X^*$ and $\alpha \geq 0$ such that

$$\alpha(f(x_0) - f(x)) \leq (x_0 - x, w), \quad \forall x \in D(f).$$

Since $\alpha \neq 0$ (otherwise $(w, x_0 - x) = 0, \forall x \in D(f)$, i.e., $w = 0$) we infer that $\alpha^{-1}w \in \partial f(x_0)$ as claimed. ■

The *directional derivative* of the convex function f to x at direction v is

$$f'(x, v) = \lim_{\lambda \downarrow 0}(f(x + \lambda v) - f(x))/\lambda. \tag{1.11}$$

Since the function $\lambda \to (f(x + \lambda v) - f(x))/\lambda$ is monotonically increasing on R^+, $f'(x, v)$ exists, (unambiguously a real number of $+\infty$) for all $x_0 \in D(f)$ and $v \in X$.

PROPOSITION 1.6. *Let $f : X \to \bar{R}$ be an l.s.c. convex function, $f \not\equiv +\infty$. Then for all $x_0 \in D(\partial f)$*

$$\partial f(x_0) = \{w \in X^*; f'(x_0, v) \geq (v, w) \quad \forall v \in X\}. \tag{1.12}$$

Proof. Let $w \in X^*$. Then by definition of ∂f we have

$$f(x_0) - f(x_0 + \lambda v) \leq -\lambda(v, w) \quad \forall v \in X$$

and therefore $(v, w) \leq f'(x_0, v)\ \forall v \in X$.

Conversely, since the function $\lambda \to \lambda^{-1}(f(x_0+\lambda v) - f(x_0))$ is monotonically increasing, the inequality $(v, w) \leq f'(x_0, v)$ implies that

$$(v, w) \leq \lambda^{-1}(f(x_0 + \lambda v) - f(x_0)) \quad \forall \lambda > 0$$

and therefore $w \in \partial f(x_0)$. ∎

Formula (1.12) can be taken as an equivalent definition for ∂f. If f is continuous at x_0, (1.12) can be strengthened to

$$f'(x_0, v) = \sup\{(v, w); w \in \partial f(x_0)\} \quad \forall v \in X. \tag{1.13}$$

In particular, it follows by (1.13) that if f is continuous and $\partial f(x_0)$ is simply valued then $f'(x_0, v) = (\nabla f(x_0), v)$, $\forall v \in X$, i.e., f is Gâteaux differentiable at x_0 and $\nabla f(x_0) = \partial f(x_0)$.

Given a lower semicontinuous, convex and proper function $f : X \to \bar{R}$, *the conjugate* of f denoted f^* is the function

$$f^*(p) = \sup\{(x, p) - f(x); x \in X\} \quad \forall p \in X^*. \tag{1.14}$$

As the supremum of a family of affine functions, f^* is itself convex and lower semicontinuous on X^*. Moreover, by Proposition 1.3 we see that $f^* \not\equiv +\infty$, i.e., f^* is proper.

There is a close relationship between ∂f and ∂f^* as shown in next proposition.

PROPOSITION 1.7. *Let $f : X \to \bar{R}$ be convex, l.s.c. and proper. Then the following conditions are equivalent*

$$x^* \in \partial f(x) \tag{1.15}$$

$$f(x) + f^*(x^*) = (x, x^*) \tag{1.16}$$

$$x \in \partial f^*(x^*). \tag{1.17}$$

In particular, it follows that $\partial f^* = (\partial f)^{-1}$ and

$$f(x) = \sup\{(x, x^*) - f^*(x^*); x^* \in X^*\}. \tag{1.18}$$

For the proof we refer to Barbu and Precupanu [2], p. 103.

If $f = I_C$ is the indicator function of a closed convex subset $C \subset X$, the conjugate function is the *support function* H_C of C, i.e.,

$$(I_C)^*(p) = H_C(p) = \sup\{(x, p); x \in C\}. \tag{1.19}$$

PROPOSITION 1.8. *Let C_1, C_2 be two convex, closed subsets of X. Then $C_1 \subset C_2$ if and only if*

$$H_{C_1}(w) \leq H_{C_2}(w) \quad \forall w \in X^*. \tag{1.20}$$

Proof. It is obvious that if $C_1 \subset C_2$ then (1.20) holds. Now if (1.20) holds, then it follows by (1.8) that

$$I_{C_1}(x) \geq I_{C_2}(x) \quad \forall x \in X, \quad \text{i.e.} \quad C_1 \subset C_2.$$

If C is a closed convex cone of X then the set

$$C^0 = \{x^* \in X^*; (x, x^*) \leq 0 \quad \forall x \in C\}$$

is the *polar cone* of C. If H_C is the support function of C then we have

$$C^0 = \{x^* \in X^*; H_C(x^*) = 0\}.$$

For applications some subdifferential calculus rules are necessary. In particular it is of great interest to know whether $\partial(f + g) = \partial f + \partial g$. In this context the following important result due to Rockafellar [12] is useful. ■

THEOREM 1.1. *Let f and g be two convex, lower semicontinuous proper functions on X such that* $(\text{int}\, D(f)) \cap D(g) \neq \emptyset$. *Then*

$$\partial(f + g) = \partial f + \partial g. \tag{1.21}$$

Proof. Since the inclusion $\partial f + \partial g \subset \partial(f + g)$ is obvious, we shall prove that $\partial(f + g) \subset \partial f + \partial g$. Let $x_0 \in D(\partial f) \cap D(\partial g)$ and $w \in \partial(f + g)(x_0)$ be arbitrary but fixed. We shall show that $w = w_1 + w_2$ where $w_1 \in \partial f(x_0)$ and $w_2 \in \partial g(x_0)$. Without loosing the generality we may assume that $x_0 = 0$, $w = 0$ and $f(0) = g(0) = 0$. This can be achieved by replacing f and g by

$x \to f(x+x_0)-f(x_0)-(z_1,x)$ and $x \to g(x+x_0)-g(x_0)-(z_2,x)$, respectively, where $w = z_1 + z_2$. To prove that $0 \in \partial f(0) + \partial g(0)$ we consider the sets

$$\begin{aligned} E_1 &= \{(x,\lambda) \in X \times R; f(x) \le \lambda\}, \\ E_2 &= \{(x,\lambda) \in X \times R; g(x) \le -\lambda\} \end{aligned}$$

Since $0 \in \partial(f+g)(0)$ it follows that

$$0 = (f+g)(0) = \inf\{f(x) + g(x); x \in X\}$$

and so $E_1 \cap \text{int } E_2 = \emptyset$. Then by the Hahn–Banach theorem there is a closed hyperplane in $X \times R$ which separates E_1 and E_2. In other words, there exists $(w,\alpha) \in X^* \times R$ such that

$$\begin{aligned} &(x,w) + \alpha\lambda \le 0 \qquad \forall (x,\lambda) \in E_1 \\ &(x,w) + \alpha\lambda \ge 0 \qquad \forall (x,\lambda) \in E_2. \end{aligned}$$

The hyperplane is not vertical, i.e., $\alpha \ne 0$ because otherwise it follows by the latter inequalities that the hyperplane $\{x \in X; (x,w) = 0\}$ separates $D(f)$ and $D(g)$. To be specific let us assume that $\alpha = 1$. Then we have

$$(x,w) \le -\lambda \le -f(x) \quad \forall x \in D(f)$$

and

$$(x,w) \ge -\lambda \le g(x) \quad \forall x \in D(g).$$

Hence $-w \in \partial f(0)$, $w \in \partial g(0)$ as claimed.

If the space X is finite dimensional the interiority condition of Theorem 1.1 can be weakened to

$$(\text{ri } D(F)) \cap (\text{ri } D(g)) \ne \emptyset,$$

where *ri* denotes the relative interior. ■

COROLLARY 1.2. *Let C_1, C_2 be two closed convex cones in X such that $0 \in C_1 \cap C_2$ and $(\text{int } C_1) \cap C_2 \ne \emptyset$. Then*

$$(C_1 \cap C_2)^0 = C_1^0 + C_2^0.$$

Proof. By definition of polar cone it is obvious that $C_1^0 + C_2^0 \subset (C_1 \cap C_2)^0$. To prove that $(C_1 \cap C_2)^0 \subset C_1^0 + C_2^0$ consider the support function $H_{C_1 \cap C_2}$ of $C_1 \cap C_2$ and note that in virtue of Proposition 1.7,

$$H_{C_1 \cap C_2}(p) + I_{C_1 \cap C_2}(y) = (y,p)$$

for $y \in C_1 \cap C_2$ and $p \in \partial I_{C_1 \cap C_2}(y)$.

If $p \in (C_1 \cap C_2)^0$, then $H_{c_1 \cap C_2}(p) = 0$ and so $(y, p) = 0$. On the other hand, by Theorem 1.1 $p = p_1 + p_2$ where $p_1 \in \partial I_{C_1}(y)$, $p_2 \in \partial I_{C_2}(y)$. Since $(y, p_i) \geq 0$, $i = 1, 2$ (because $0 \in C_1 \cap C_2$) we infer that $(y, p_i) = 0$, $i = 1, 2$ and so $H_{C_1}(p_1) = 0$, $H_{C_2}(p_2) = 0$. Hence $p_1 \in C_1^0$, $p_2 \in C_2^0$ as desired. ■

Now we shall consider the special case where $X = H$ is a real Hilbert space. We note first that the subdifferential ∂f of a l.s.c. convex function $f : H \to \bar{R}$ is monotone, i.e.,

$$(x_1 - x_2, y_1 - y_2) \geq 0 \quad \forall x_1, x_2 \in D(\partial f), y_i \in \partial f(x_i), \quad i = 1, 2.$$

Moreover, ∂f is *maximal monotone* in $H \times H$, i.e.,

$$R(I + \lambda \partial f) = H \quad \forall \lambda > 0. \tag{1.22}$$

(We have denoted by $R(A)$ the range of A and by I the unity operator in H.) To prove (1.22) it suffices to observe that for a given $w \in H$ the equation $x + \lambda \partial f(x) \ni w$ is equivalent with the minimization problem

$$\inf \left\{ \frac{1}{2} \|x\|^2 + \lambda f(x) - (x, w) \right\}$$

which in virtue of Proposition 1.2 (Corollary 1.1) has at least one solution x.

It is useful to mention that (1.22) is equivalent with the fact that ∂f does not admit proper monotone extensions in the space $H \times H$. For other connections between convex analysis and theory of maximal monotone operators we refer the reader to the books of Brézis [4] and Barbu [1].

We shall denote by $(\partial f)_\lambda$ the Yosida approximation of ∂f, i.e.,

$$(\partial f)_\lambda(x) = \lambda^{-1}(I - (I + \lambda \partial f)^{-1}) \quad \forall \lambda > 0 \tag{1.23}$$

and by f_λ the function

$$f_\lambda(x) = \inf \left\{ \frac{\|x - y\|^2}{2\lambda} + f(y); y \in H \right\} \quad \forall x \in X. \tag{1.24}$$

The function f_λ is a smooth convex approximation of f. More precisely we have (Brézis [3, 4])

THEOREM 1.2. *Let $f : H \to \bar{R}$ be a l.s.c., convex, proper function on Hilbert space H. Then f_λ is convex, Fréchet differentiable and $(\partial f)_\lambda = \nabla f_\lambda$, $\forall \lambda > 0$. Furthermore, we have*

$$f_\lambda(x) = \frac{\lambda}{2} \|(\partial f)_\lambda(x)\|^2 + f((I + \lambda\partial f)^{-1}x) \quad \forall x \in H \tag{1.25}$$

$$\lim_{\lambda \downarrow 0} f_\lambda(x) = f(x) \quad \forall x \in H. \tag{1.26}$$

We omit the proof which can be found in the Brézis works [3, 4] as well as in the books of Barbu [1], Barbu and Precupanu [2].

Let C be a closed convex subset of H. Applying Theorem 1.2 with $f = I_C$ it follows that

$$P_C(x) = (I + \lambda\partial I_C)^{-1}x \quad \forall x \in H \tag{1.25'}$$

$$d_C^2(x) = 2\lambda(I_C)_\lambda(x) \quad \forall x \in H \tag{1.26'}$$

where P_C is the projection operator on C and $d_C(x)$ is the distance from x to C.

In the remaining discussion of this section we shall present few results pertaining the convex integrands on the spaces $L^p(\Omega)$, $p \geq 1$ (Rockafellar [14]).

Let Ω be a measurable subset of R^N and let $L^p_m(\Omega)$ be the usual space of p-summable functions from Ω to R^m. The function $g : \Omega \times R^m \to \bar{R}$ is called a *normal convex integrand* if the following conditions hold

(i) $g(x, .) : R^m \to \bar{R}$ is convex, lower semicontinuous and proper for a.e. $x \in \Omega$,

(ii) g is $\mathcal{L} \times \mathcal{B}$ measurable on $\Omega \times R^m$, i.e., it is measurable with respect to the σ-algebra of subsets of $\Omega \times R^m$ generated by products of Lebesgue measurable subsets of Ω and Borelian subsets of R^m.

An important example of normal convex integrand is the Caratheodory integrand, i.e., a function $g : \Omega \times R^m \to \bar{R}$ which is continuous and convex in y and Lebesgue measurable in x. Indeed we have

LEMMA 1.1. *Let $g : \Omega \times R^m \to R$ be continuous in y for all $x \in \Omega$ and measurable in x on Ω for each $y \in R^m$. Then g is $\mathcal{L} \times \mathcal{B}$ measurable.*

Proof. Let $\{z_i\}$ be a dense numerable set of R^m and let $\lambda \in R$ be arbitrary but fixed. Since g is continuous in y it is clear that $g(x, y) \leq \lambda$ if and only if for every j there exists z_i such that $\|z_i - y\| \leq j^{-1}$ and $g(x, z_i) \leq \lambda + j^{-1}$. We set $\Omega_{ij} = \{x \in \Omega; g(x, z_i) \leq \alpha + j^{-1}\}$, $Y_{ij} = \{y \in R^m; \|z_i - y\| \leq j^{-1}\}$. Clearly Ω_{ij} are Lebesgue measurable and Y_{ij} are Borelian. Since

$$\{(x, y) \in \Omega \times R^m; g(x, y) \leq \lambda\} = \cap_{j=1}^{\infty} \cap_{i=1}^{\infty} \Omega_{ij} \times Y_{ij}$$