Springer Monographs in Mathematics

Viorel Barbu
Teodor Precupanu

Convexity and Optimization in Banach Spaces

Fourth Edition

Springer

S*pringer* **M***onographs* in **M***athematics*

For further volumes:
www.springer.com/series/3733

Viorel Barbu • Teodor Precupanu

Convexity and Optimization in Banach Spaces

Fourth Edition

Springer

Prof. Viorel Barbu
Department of Mathematics
University Al. I. Cuza
Iaşi
Romania
vb41@uaic.ro

Prof. Teodor Precupanu
Department of Mathematics
University Al. I. Cuza
Iaşi
Romania

ISSN 1439-7382 Springer Monographs in Mathematics
ISBN 978-94-007-2246-0 e-ISBN 978-94-007-2247-7
DOI 10.1007/978-94-007-2247-7
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2011942142

Mathematics Subject Classification (2010): 46A55, 46N10, 49J20, 49K25

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

This is the fourth English edition of the book *Convexity and Optimization in Banach Spaces*. With respect to the previous edition published by Kluwer in 1986, this book contains new results pertaining to new concepts of subdifferential for convex functions and new duality results in convex programming. The last chapter of the book, concerned with convex control problems, was rewritten for this edition and completed with new results concerning boundary control systems, the dynamic programming equations in optimal control theory, periodic optimal control problems. Also, the bibliographical list and bibliographical comments were updated. The contents, as well as the structure of the book, were modified in order to include a few fundamental results and progress in the theory of infinite-dimensional convex analysis which were obtained in the last 25 years.

Iaşi, Romania

Viorel Barbu
Teodor Precupanu

Contents

Acronyms

A_∞	The asymptotic cone of the set A
$\operatorname{aff} A$	The affine hull of the set A
$\varGamma$	The field of scalars of a linear space ($\mathbb{R}$ or $\mathbb{C}$)
$\operatorname{hypo} f$	The hypograph of the function f
$\operatorname{span} A$	The linear subspace generated by the set A
$\operatorname{cone} A$	The cone generated by the set A
$\ker f$	The kernel of the function f
A^0	The polar of the set A
cA	The complement of the set A
p_A	The Minkowski functional of the set A
$\mathbb{C}$	The set of all complex numbers
$\mathbb{N}$	The set of all natural numbers
$\mathbb{N}^*$	The set of all nonzero natural numbers
$\mathbb{R}$	The real line $(-\infty, \infty)$
$\mathbb{R}^n$	The n-dimensional Euclidean space
$\mathbb{R}^+$	$= (0, +\infty)$
$\mathbb{R}^-$	$= (-\infty, 0)$
$\overline{\mathbb{R}}$	$= [-\infty, +\infty]$
$\overline{\mathbb{R}}^*$	$=]-\infty, +\infty]$
$\mathbb{R}^n_+$	$= \{(x_1, \ldots, x_n); x_n > 0\}$
$\varOmega$	An open subset of $\mathbb{R}^n$
$\partial\varOmega$	The boundary of $\varOmega$
Q	$= \varOmega \times (0, T)$
$\varSigma$	$= \partial\varOmega \times (0, T)$, where $0 < T < \infty$
$\|\cdot\|_X$	The norm of the linear normed space X
X^*	The dual of space X
$(\cdot, \cdot)_X$	The scalar product of the Hilbert space X
$x \cdot y$	The scalar product of the vectors $x, y \in \mathbb{R}^n$
$L(X, Y)$	The space of linear continuous operators from X to Y
∇f	The gradient of the function f
∂f	The subdifferential of the function f

f^*	The conjugate of the function f
B^*	The adjoint of the operator B
$\overline{C}$	The closure of the set C
$\operatorname{int} C$	The interior of the set C
$\operatorname{conv} C$	The convex hull of the set C
$\operatorname{ri} C$	The relative interior of the set C
$\operatorname{cl} f$	The closure of the function f
$\operatorname{Dom}(f)$	The effective domain of the function f
$D(A)$	The domain of the operator A
$R(A)$	The range of the operator A
I_C	The indicator function of the set C
$\operatorname{epi} f$	The epigraph of the function f
sign	The signum function on X : $\operatorname{sign} x = x/\\|x\\|_X$ if $x \neq 0$ $\operatorname{sign} 0 = \{x;\ \\|x\\| \leq 1\}$
$C^k(\Omega)$	The space of real-valued functions on Ω that are continuously differentiable up to order k, $0 \leq k \leq \infty$
$C_0^k(\Omega)$	The subspace of functions in $C^k(\Omega)$ with compact support in Ω
$\mathscr{D}(\Omega)$	The space $C_0^\infty(\Omega)$
$\frac{\mathrm{d}^k u}{\mathrm{d}t^k},\ u^{(k)}$	The derivative of order k of $u : [a,b] \to X$
$\mathscr{D}'(\Omega)$	The dual of $\mathscr{D}(\Omega)$ (i.e., the space of distributions on Ω)
$C(\overline{\Omega})$	The space of continuous functions on $\overline{\Omega}$
$L^p(\Omega)$	The space of p-summable functions $u : \Omega \to \mathbb{R}$ endowed with the norm $\\|u\\|_p = (\int_\Omega \lvert u(x)\rvert^p \,\mathrm{d}x)^{1/p}$, $1 \leq p < \infty$, $\\|u\\|_\infty = \operatorname{ess\,sup}_{x\in\Omega} \lvert u(x)\rvert$ for $p = \infty$
$L_m^p(\Omega)$	The space of p-summable functions $u : \Omega \to \mathbb{R}^m$
$W^{m,p}(\Omega)$	The Sobolev space $\{u \in L^p(\Omega);\ D^\alpha u \in L^p(\Omega),\ \lvert\alpha\rvert \leq m,\ 1 \leq p \leq \infty\}$
$W_0^{m,p}(\Omega)$	The closure of $C_0^\infty(\Omega)$ in the norm of $W^{m,p}(\Omega)$
$W^{-m,q}(\Omega)$	The dual of $W_0^{m,p}(\Omega)$; $\frac{1}{p} + \frac{1}{q} = 1$, $p < \infty$, $q > 1$
$H^k(\Omega)$, $H_0^k(\Omega)$	The spaces $W^{k,2}(\Omega)$ and $W_0^{k,2}(\Omega)$, respectively
$L^p(a,b;X)$	The space of p-summable functions from (a,b) to X (Banach space) $1 \leq p \leq \infty$, $-\infty \leq a < b \leq \infty$
$AC([a,b];X)$	The space of absolutely continuous functions from $[a,b]$ to X
$BV([a,b];X)$	The space of functions with bounded variation on $[a,b]$
$BV(\Omega)$	The space of functions with bounded variation on Ω
$W^{1,p}([a,b];X)$	The space $\{u \in AC([a,b];X);\ \mathrm{d}u/\mathrm{d}t \in L^p([a,b];X)\}$
$\frac{\partial u}{\partial \nu}$	The normal derivative of the function u
A^{ac}	The algebraic closure of the set A
A^{i}	The algebraic interior of the set A
A^{ri}	The algebraic relative interior of the set A

Chapter 1
Fundamentals of Functional Analysis

The purpose of this preliminary chapter is to introduce the basic terminology and results of functional and convex analysis which are used in the sequel.

1.1 Convexity in Topological Linear Spaces

In this section, we concentrate on basic definitions and properties of convex sets in linear infinite-dimensional spaces.

1.1.1 Classes of Topological Linear Spaces

The general framework for functional analysis is the structure of the *topological linear space*, which is a linear space endowed with a topology for which the operations of addition and scalar multiplication are continuous. In this case, we say that the *topology is compatible* with the algebraic structure of the linear space or that the *topology is linear*. In the following, we recall some basic properties of topological linear spaces, most of them being immediate consequences of the definition.

We denote by X a linear space over a field of scalars $\varGamma$. (In our discussion, the field $\varGamma$ will always be the real field $\mathbb{R}$ or the complex field $\mathbb{C}$.)

Theorem 1.1 *The mappings $x \to x + x_0$ and $x \to \lambda x$, where $\lambda \neq 0$, $\lambda \in \varGamma$, are homeomorphisms of X onto itself.*

In particular, a linear topology can be defined if we know a base of neighborhoods at the origin because by translation we can obtain a base of neighborhoods for every other point $x \in X$; each neighborhood of a point x is of the form $x + V$, where V is a neighborhood of the origin. Consequently, we easily as a result find that *a linear mapping between two topological linear spaces is continuous if and only if it is continuous at the origin.*

V. Barbu, T. Precupanu, *Convexity and Optimization in Banach Spaces*,
Springer Monographs in Mathematics,
DOI 10.1007/978-94-007-2247-7_1,

As concerns the continuity of linear functionals, we can prove the following characterization theorem.

Theorem 1.2 *If f is a linear functional on a topological linear space, then the following statements are equivalent*:

(i) *f is continuous*
(ii) *The kernel of f*, $\ker f = \{x;\ f(x) = 0\}$, *is closed*
(iii) *There is a neighborhood of the origin on which f is bounded.*

A linear space of finite-dimension possesses a unique separated linear topology. Therefore, *every separated topological linear space of dimensions $n \in \mathbb{N}$ is isomorphic with Γ^n.*

Definition 1.3 A mapping $p : X \to \mathbb{R}$ is called a *seminorm* on X if it has the following properties:

(i) $p(x) = |\lambda| p(x)$, for every $x \in X$ and $\lambda \in \Gamma$
(ii) $p(x + y) \le p(x) + p(y)$, for every $x, y \in X$.

From conditions (i) and (ii), as a result we find that

(iii) $p(x) \ge 0$, for every $x \in X$.

If p has the stronger property

(iii)′ $p(x) > 0$, for every $x \in X \setminus \{0\}$,

then p is called a *norm* on X.

A particular class of topological linear spaces with richer properties is the class of *locally convex spaces*; these are topological linear spaces with the property that *for every element there exists a base of neighborhoods consisting of convex sets.* It is well known that any locally convex topology on a linear space may be generated by a family of seminorms.

Let $\mathscr{P} = \{p_i;\ i \in I\}$ be a family of seminorms on the linear space X. Consider for every $x \in X$ the family of subsets of X

$$\mathscr{V}(x) = \left\{V_{i_1,i_2,\ldots,i_k,\varepsilon}(x);\ k \in \mathbb{N}^*,\ i_1,\ldots,i_k \in I,\ \varepsilon > 0\right\}, \quad x \in X, \tag{1.1}$$

where

$$V_{i_1,i_2,\ldots,i_k,\varepsilon}(x) = \left\{u \in X;\ p_{i_j}(u - x) < \varepsilon,\ \forall j = 1, 2, \ldots, k\right\}. \tag{1.2}$$

We can easily see that $\mathscr{V}(x),\ x \in X$, is a *base of neighborhoods for a locally convex topology $\tau_{\mathscr{P}}$* on X. The topological properties for $\tau_{\mathscr{P}}$ can be characterized analytically by means of the seminorms of $\mathscr{P}$.

Theorem 1.4 *The locally convex topology $\tau_{\mathscr{P}}$ is the coarsest linear topology on X for which all seminorms of the family $\mathscr{P}$ are continuous.*

We recall that a seminorm $p : X \to \mathbb{R}$ is continuous for $\tau_{\mathscr{P}}$ if and only if there are $k > 0$ and $p_1, p_2, \ldots, p_n \in \mathscr{P}$ such that

$$p(x) \leq k \max_{1 \leq i \leq n} p_i(x), \quad \forall x \in X. \tag{1.3}$$

This implies that a linear mapping $T : X \to Y$, where X and Y are locally convex spaces endowed with the topologies $\tau_{\mathscr{P}}$ and $\tau_{\mathscr{Q}}$, respectively, is *continuous* if and only if for each $q \in \mathscr{Q}$ there are $k_q > 0$ and $p_1, \ldots, p_n \in \mathscr{P}$ such that

$$q(Tx) \leq k_q \max_{1 \leq i \leq n} p_i(x), \quad \forall x \in X. \tag{1.4}$$

A sequence $\{x_n\}$ of points from X is $\tau_{\mathscr{P}}$*-convergent* to $x_0 \in X$ if and only if the numerical sequence

$$p(x_1 - x_0), p(x_2 - x_0), \ldots, p(x_n - x_0), \ldots \tag{1.5}$$

is convergent to zero for every $p \in \mathscr{P}$.

A set $M \subset X$ is $\tau_{\mathscr{P}}$*-bounded* if and only if every seminorm belonging to the family $\mathscr{P}$ is bounded on M, that is, if and only if, for every $p \in \mathscr{P}$, there exists a constant $k_p > 0$ such that $p(x) \leq k_p$, $\forall x \in M$.

The locally convex topology $\tau_{\mathscr{P}}$ is *separated* if and only if the family of seminorms $\mathscr{P}$ possesses the following property:

$$\text{for every } x \in X \setminus \{0\} \text{ there is } p \in \mathscr{P} \text{ such that } p(x) \neq 0. \tag{1.6}$$

A linear space X endowed with a norm $\|\cdot\|$ is called a *linear normed space*.

In particular, we can obtain the topology of a linear normed space if we take $\mathscr{P} = \{\|\cdot\|\}$. On the other hand, the topology of a linear normed space can be generated by the distance defined by $d(x, y) = \|x - y\|$, $\forall x, y \in X$. In this way, for linear normed spaces, the metric properties interweave with the topological properties of a locally convex space. Generally speaking, a locally convex topology is *metrizable* (in other words, there exists a distance which generates its topology), if and only if this topology is separated and can be generated by a countable family of seminorms. The importance of the metrizability consists of the fact that all the topological properties can be characterized by sequences.

Let X, Y be linear normed spaces of the same nature. A linear operator $T : X \to Y$ is continuous if and only if it is *bounded*, that is, $T(M)$ is bounded for every bounded set $M \subset X$. In other words, there exists $K > 0$ such that

$$\|Tx\| \leq K\|x\|, \quad \forall x \in X. \tag{1.7}$$

The set $L(X, Y)$ of all linear continuous operators defined on X with values in Y becomes a linear normed space by

$$\|T\| = \sup\big\{\|Tx\|;\ \|x\| \leq 1\big\} = \inf\big\{K;\ \|Tx\| \leq K\|x\|,\ \forall x \in X\big\}. \tag{1.8}$$

If $X = Y$, we shortly denote $L(X) = L(X, X)$.

A complete linear normed space is called a *Banach space*. If Y is a Banach space, then $L(X, Y)$ is also a Banach space.

In particular, if $Y = \Gamma$, we find that $X^* = L(X, \Gamma)$, called the *dual* of X, is a Banach space relative to the norm of the functionals, which by (1.8) becomes

$$\|x^*\| = \sup\big\{\big|x^*(x)\big|;\ \|x\| \le 1\big\}. \tag{1.9}$$

If X is a real linear normed space, then we also have

$$\|x^*\| = \sup\big\{x^*(x);\ \|x\| \le 1\big\}, \tag{1.9$'$}$$

while, if X is a complex linear normed space, we have

$$\|x^*\| = \sup\big\{\mathrm{Re}\ x^*(x);\ \|x\| \le 1\big\}.$$

Consequently, the following fundamental inequality holds:

$$\big|x^*(x)\big| \le \|x^*\|\,\|x\|, \quad \text{for all } x \in X,\ x^* \in X^*. \tag{1.10}$$

A family $\mathscr{A} \subset L(X, Y)$ is called *uniformly bounded* if it is bounded in the norm (1.8). Hence, regarding $\mathscr{A}$ as a family of functions on X, it is uniformly bounded on the closed unit ball of X. On the other hand, we say that the family $\mathscr{A} \subset L(X, Y)$ is *pointwise bounded* if $\mathscr{A}_x = \{Tx;\ T \in \mathscr{A}\}$ is a bounded set of Y for every $x \in X$. It is clear that every uniformly bounded family is pointwise bounded; the converse is not true. But it is well known that for Banach spaces we have the very useful *principle of uniform boundedness*.

Theorem 1.5 *If X is a Banach space, then every pointwise bounded family of linear continuous operators from $L(X, Y)$ is uniformly bounded.*

In the special case of sequences of operators, this result leads to the *Banach–Steinhauss Theorem*.

Theorem 1.6 (Banach–Steinhauss) *Let $\{A_n\}_{n\in N}$ be a sequence from $L(X, Y)$ pointwise convergent to an operator A. If X is a Banach space, then $A \in L(X, Y)$ and*

$$\|A\| \le \liminf_{n\to\infty} \|A_n\|.$$

Definition 1.7 A mapping $\langle\cdot,\cdot\rangle : X \times X \to \Gamma$ is said to be an *inner product* if it has the following properties:

(i) $\langle x, x\rangle \ge 0,\ \forall x \in X$, and $\langle x, x\rangle = 0$ implies $x = 0$
(ii) $\langle x, y\rangle = \overline{\langle y, x\rangle},\ \forall x, y \in X$
(iii) $\langle ax + by, z\rangle = a\langle x, z\rangle + b\langle y, z\rangle,\ \forall a, b \in \Gamma,\ \forall x, y \in X$.

From condition (i), we find as a result that, for every $x_1, x_2 \in X$, the Hermitian form

$$f(\lambda_1, \lambda_2) = \langle \lambda_1 x_1 + \lambda_2 x_2, \lambda_1 x_1 + \lambda_2 x_2 \rangle, \quad \lambda_1, \lambda_2 \in \Gamma,$$

is nonnegative. Conditions (ii) and (iii) give the well-known *Cauchy–Schwarz inequality*,

$$\left|\langle x, y\rangle\right|^2 \le \langle x, x\rangle \langle y, y\rangle, \quad \forall x, y \in X. \tag{1.11}$$

The mapping $\|\cdot\| : X \to \mathbb{R}$, defined by

$$\|x\| = \langle x, x\rangle^{\frac{1}{2}}, \quad \forall x \in X, \tag{1.12}$$

is a norm on X. Inequality (1.11) becomes

$$\left|\langle x, y\rangle\right| \le \|x\| \, \|y\|, \quad \forall x, y \in X. \tag{1.13}$$

A linear space endowed with an inner product is called a *pre-Hilbert space*. A pre-Hilbert space is also considered as a linear normed space by the norm defined by (1.12).

Two elements x and y in a pre-Hilbert space are said to be *orthogonal* if $\langle x, y\rangle = 0$; we denote this by $x \perp y$. We remark that the orthogonality relation is linear and symmetric.

We also mention the following consequence of property (i) from Definition 1.7:

$$\text{if } x \perp y, \ \forall y \in X, \quad \text{then } x = 0. \tag{1.14}$$

Proposition 1.8 *The elements x, y are orthogonal if and only if*

$$\|x + \lambda y\| \ge \|x\|, \quad \forall \lambda \in \Gamma.$$

If a pre-Hilbert space is complete in the norm associated to the given inner product, then it is called a *Hilbert space*.

The general form of the continuous linear functionals is expressed more precisely by the so-called *Riesz Representation Theorem*.

Theorem 1.9 (Riesz) *If f is a continuous linear functional on the Hilbert space X, then there exists a unique element $a \in X$ such that*

$$f(x) = \langle x, a\rangle, \quad \forall x \in X, \tag{1.15}$$

$$\|f\| = \|a\|. \tag{1.16}$$

On the other hand, using the Cauchy–Schwarz inequality (1.13), we easily observe that, for every $a \in X$, the linear functional $f_a : X \to \Gamma$ defined by

$$f_a(x) = \langle x, a\rangle, \quad \forall x \in X,$$

is continuous, hence $f_a \in X^*$, and at the same time $\|f_a\| = \|a\|$, $\forall a \in X$.

If we define the mapping $J : X \to X^*$ by

$$Ja = f_a, \quad \forall a \in X, \tag{1.17}$$

then the Riesz representation Theorem 1.9 says that J is an isometric bijection (anti-linear). With the aid of J, called the *canonical isomorphism* of the Hilbert space X, we can convey some properties from X to X^*. In fact, we observe that the natural norm on X^*, given by relation (1.9), is associated to the inner product on X^*, defined by

$$\langle f_1, f_2 \rangle = \left\langle J^{-1} f_2, J^{-1} f_1 \right\rangle, \quad \forall f_1, f_2 \in X^*. \tag{1.18}$$

Therefore, the dual of a Hilbert space is also a Hilbert space.

Let X_1, X_2 be two Hilbert spaces of the same nature and let $T : X_1 \to X_2$ be a linear continuous operator. For every $y \in X_2$, the function $f_y : X_1 \to \Gamma$ defined by

$$f_y(x) = \langle Tx, y \rangle, \quad \forall x \in X_1, \tag{1.19}$$

is a continuous linear functional on X_1 and so, according to the Riesz Representation Theorem 1.9, there exists a unique element $\bar{y} \in X_1$ such that

$$f_y(x) = \langle x, \bar{y} \rangle, \quad \forall x \in X_1. \tag{1.20}$$

If we define the mapping $T^* : X_2 \to X_1$ by $T^* y = \bar{y}$, for every $y \in X_2$, it is easy to see that T^* is the unique operator from X_2 into X_1 which satisfies the relation

$$\langle Tx, y \rangle = \langle x, T^* y \rangle, \quad \forall x \in X_1,\ \forall y \in X_2. \tag{1.21}$$

This operator T^* (which satisfies relation (1.21)) is called the *adjoint* of the continuous linear operator T. The adjoint is also a continuous linear operator and we have

$$\|T^*\| = \|T\|, \quad \forall T \in L(X_1, X_2). \tag{1.22}$$

In particular, if $X_1 = X_2 = X$ and $T \in L(X)$, then $T^* \in L(X)$. We say that the operator $T \in L(X)$ is *self-adjoint* if $T^* = T$. Therefore, the operator $T \in L(X)$ is self-adjoint if and only if we have

$$\langle Tx, y \rangle = \langle x, Ty \rangle, \quad \forall x, y \in X. \tag{1.23}$$

1.1.2 Convex Sets

Let X be a real linear space.

Definition 1.10 A subset of the linear space X is said to be *convex* if whenever it contains x_1 and x_2, it also contains $\lambda_1 x_1 + \lambda_2 x_2$, where $\lambda_1 \geq 0$, $\lambda_2 \geq 0$ and $\lambda_1 + \lambda_1 = 1$.

We denote

$$[x_1, x_2] = \{\lambda_1 x_1 + \lambda_2 x_2;\ \lambda \geq 0,\ \lambda_2 \geq 0,\ \lambda_1 + \lambda_2 = 1\},$$

called the *segment* generated by elements x_1, x_2.

Definition 1.11 A subset of the linear space X is said to be an *affine set* if, whenever it contains x_1 and x_2, it also contains $\lambda_1 x_1 + \lambda_2 x_2$ for arbitrary $\lambda_1, \lambda_2 \in \mathbb{R}$ so that $\lambda_1 + \lambda_1 = 1$. If $x_1, x_2, \dots, x_n$ are finitely many elements of X, every element of the form $\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n$, where $\lambda_i \in \mathbb{R}$ and $\lambda_1 + \lambda_2 + \cdots + \lambda_n = 1$ is called an *affine combination* of $x_1, x_2, \dots, x_n$. Moreover, if $\lambda_i \geq 0$, the affine combination is called a *convex combination*.

Proposition 1.12 *Any convex (affine) set contains all the convex (affine) combinations formed with its elements.*

Proof Applying mathematical induction, we observe that a convex (affine) combination of n elements $x_1, x_2, \dots, x_n$ is a convex (affine) combination of $n-1$ elements,

$$\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n = \lambda_1 x_1 + \cdots + \lambda_{n-2} x_{n-2} + (\lambda_{n-1} + \lambda_n)\tilde{x}_{n-1},$$

where

$$\tilde{x}_{n-1} = \frac{\lambda_{n-1}}{\lambda_{n-1} + \lambda_n} x_{n-1} + \frac{\lambda_n}{\lambda_{n-1} + \lambda_n} x_n$$

is also a convex (affine) combination of two elements whenever $\lambda_{n-1} + \lambda_n \neq 0$. (The case $\lambda_i + \lambda_j = 0$ for all $i, j = 1, 2, \dots, n$, $i \neq j$, is not possible.) Hence, according to Definitions 1.10 and 1.11, it is an element of the same convex (affine) set. □

From the above definitions, it follows immediately that:

(i) *The intersection of many arbitrary convex (affine) sets is again a convex (affine) set.*
(ii) *The union of a directed by inclusion family of convex (affine) sets is a convex (affine) set.*
(iii) *If $A_1, A_2, \dots, A_n$ are convex (affine) sets and $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$, then $\lambda_1 A_1 + \cdots + \lambda_n A_n$ is a convex (affine) set.*
(iv) *The linear image and the linear inverse image of a convex (affine) set are again convex (affine) sets.*
(v) *If X is a linear topological space, then the closure and the interior of a convex (affine) set is a set of the same kind.*

The property of stability under intersection leads to the introduction of the *convex (affine) hull*, denoted by $\operatorname{conv} A(\operatorname{aff} A)$, of an arbitrary set A as the intersection of all convex (affine) sets containing it. In other words, $\operatorname{conv} A$ ($\operatorname{aff} A$) is the smallest convex (affine) set which contains the set A. Using Proposition 1.12, we can easily show that the elements of the convex (affine) hull can be represented only with the elements of the given set.

Theorem 1.13 *The convex (affine) hull of a set A of X coincides with the set of all convex (affine) combinations of elements belonging to A, that is,*

$$\operatorname{conv} A = \left\{ \sum_{i=1}^{n} \lambda_i x_i;\ n \in \mathbb{N}^*, \lambda_i \geq 0,\ x_i \in A, \sum_{i=1}^{n} \lambda_i = 1 \right\}, \tag{1.24}$$

$$\operatorname{aff} A = \left\{ \sum_{i=1}^{n} \lambda_i x_i;\ n \in \mathbb{N}^*, \lambda_i \in \mathbb{R},\ x_i \in A, \sum_{i=1}^{n} \lambda_i = 1 \right\}. \tag{1.25}$$

We remark that an affine set is a linear subspace if and only if it contains the origin. In general, we have the following proposition.

Proposition 1.14 *In a real linear space, a set is affine if and only if it is a translation of a linear subspace.*

Definition 1.15 A point $x_0 \in X$ is said to be *algebraic relative interior* of $A \subset X$ if, for every straight line through x_0 which lies in $\operatorname{aff} A$, there exists an open segment contained in A which contains x_0. If $\operatorname{aff} A = X$, the point x_0 is called the *algebraic interior* of A. The set of all the algebraic (relative) interior points of A is called the *algebraic (relative) interior* of the set A and we denote it by $(A^{\mathrm{ri}})A^{\mathrm{i}}$.

Definition 1.16 If X is a topological vector space, then a point $x_0 \in X$ is said to be a *relative interior* of $A \subset X$ if it is an interior point (in a topological sense) of A with respect to the topology induced on $\operatorname{aff} A$. The set of all relative interior points of A is called the *relative interior* of A, and we denote it by $\operatorname{ri} A$. Also, we denote the interior of A by $\operatorname{int} A$.

Thus, $\operatorname{ri} A = \operatorname{int} A$ if $\operatorname{aff} A = X$. On the other hand, it is clear that $\operatorname{aff} A = X$ if $\operatorname{int} A \neq \emptyset$ or $A^{\mathrm{i}} \neq \emptyset$.

Similarly, we can define A^{ac}, the *algebraic closure* of a convex set A, as the set of all points $x \in X$ for which there exists $u \in A$ such that $[u, x[\subset A$, where $[u, x[$ is the segment joining u and x, including u and excluding x.

Now, we define the *Minkowski functional* associated to a set A which contains the origin by

$$p_A(x) = \inf \left\{ \lambda \in \overline{\mathbb{R}}_+;\ \frac{1}{\lambda} x \in A \right\}, \quad \forall x \in X, \tag{1.26}$$

where $\overline{\mathbb{R}}_+ =]0, +\infty]$ and we admit that $\frac{1}{+\infty} = 0$ and $0 \cdot \infty = 0$.

We denote $\operatorname{Dom} p_A = \{y \in Y;\ p_A(x) < \infty\}$.

The Minkowski functional has the following properties:

(i) $p_A(x) \geq 0,\ \forall x \in X$ and $p_A(0) = 0$
(ii) $p_A(\lambda x) = \lambda p_A(x),\ \forall \lambda \geq 0,\ \forall x \in X$
(iii) $A \subset \{x \in X;\ p_A(x) \leq 1\}$
(iv) $p_{A_1}(x) \leq p_{A_2}(x),\ \forall x \in X$, if $A_1 \supset A_2$.

Generally, any functional having property (ii) is called a *positive-homogeneous functional*.

It is easy to show that $p_A(x) < \infty$, $\forall x \in X$, if and only if A is an *absorbent* set; that is, *for every* $x \in X$ *there is a* $\lambda > 0$ *such that* $\lambda x \in A$.

If we mean by the *radial boundary* of a set A, denoted by A^{rb}, the set of all elements $x \in X$ for which $[u, x] \cap A \neq \emptyset$ for every $u \in]0, x[$, and $\lambda x \notin A$ for every $\lambda > 1$, then we see that

$$p_A = p_{\{0\} \cup A^{\mathrm{rb}}}.$$

From this result, we remark that the Minkowski functional is perfectly determined only by the radial boundary. Also, we have

$$A^{\mathrm{rb}} = \left\{x \in X;\ p_A(x) = 1\right\}.$$

Moreover, *if the set* A *is convex, then* p_A *is subadditive*; hence, p_A is a *sublinear functional* (positive-homogeneous and subadditive). In this case, property (iii) can be completed by

(iii)$'$ $\{x \in X;\, p_A(x) < 1\} \subset A \subset \{x \in X;\, p_A(x) \leq 1\}$.

To establish some sufficient conditions under which we have equality in the right-hand side or in the left-hand side of (iii)$'$, we shall use the algebraic notions given in Definition 1.15.

However, these notions simultaneously have a topological role to play: thus, in finite topological linear spaces, they coincide with the notions contained in Definition 1.16 (see Proposition 1.17). We also note that if the set A is convex, then a point $x_0 \in A$ is algebraic relative interior to A if and only if for every $x \in \operatorname{aff} A$ there is $\rho > 0$ such that $x_0 + \rho(x - x_0) \in A$. In other words, this is the case if and only if $A - x_0$ is absorbent in the linear subspace generated by $A - x_0$. In particular, if A is absorbent, then the origin is an algebraic interior point.

Proposition 1.17 *Let* X *be a finite-dimensional separated topological linear space and let* A *be a convex set of* X. *A point* $x_0 \in A$ *is algebraic interior of* A *if and only if* x_0 *is an interior point* (*in the topological sense*) *of* A.

Proof Let $\mathscr{V}$ be the family of all symmetric, absorbent and convex sets of X. It is well known that $\mathscr{V}$ generates a separated linear topology on X. Then we observe that a point is algebraic interior to a convex set if and only if it is interior with respect to this linear topology generated by $\mathscr{V}$. On the other hand, on a finite-dimensional linear space there exists a unique separated linear topology, from which the following statement results. □

Corollary 1.18 *A point $x_0 \in A$, where A is a convex set from a finite-dimensional separated topological linear space, is an algebraic relative interior point of A if and only if it is a relative interior point of A.*

If X is a separated topological linear space, it can easily be shown that every (relative) interior point of a set is again an algebraic (relative) interior point of this set, that is,

$$\operatorname{int} A \subset A^{\mathrm{i}} \quad \text{and} \quad \operatorname{ri} A \subset A^{\mathrm{ri}}. \tag{1.27}$$

If A is convex, this result can be completed by the following.

Proposition 1.19 *If A is a convex set for which the origin is an algebraic relative interior point, then*

$$A^{\mathrm{ri}} = \big\{x \in X;\ p_A(x) < 1\big\} \quad \textit{and} \quad A^{\mathrm{ac}} = \big\{x \in X;\ p_A(x) \le 1\big\}.$$

Proof Since $0 \in A$, we find as a result that lin A is a linear subspace and then $x_0 \in A$ is an algebraic relative interior if and only if for every $x \in \operatorname{aff} A$ there is $\rho > 0$ such that $x_0 + \rho x \in A$. Thus, if $x_0 \in A^{\mathrm{ri}}$, there is $\rho_0 > 0$ such that $x_0 + \rho_0 x_0 \in A$; hence, $p_A(x_0) \le \frac{1}{1+\rho_0} < 1$.

Conversely, if $p_A(x_0) < 1$ and $x \in \operatorname{aff} A$, there exists $\rho > 0$ such that $p_A(x_0) + \rho p_A(x) < 1$. From this, we have $p_A(x_0 + \rho x) < 1$, which implies $x_0 + \rho x \in A$. By a similar argument, we can obtain the other equality. □

Corollary 1.20 *The interior of a convex set is either an empty set or it coincides with its algebraic interior.*

Proof If $\operatorname{int} A \ne \emptyset$, we can assume without loss of generality that $0 \in \operatorname{int} A$. Thus, there exists a neighborhood V of the origin such that $V \subset A$. Hence, $p_A(x) \le 1$, $\forall x \in V$. Let x_0 be an algebraic interior point of A. According to Proposition 1.19, there exists $\varepsilon_0 > 0$ such that $p_A(x_0)+\varepsilon_0 < 1$. Since $p_A(x_0+\varepsilon_0 x) \le p_A(x_0)+\varepsilon_0 p_A(x) < 1$, $\forall x \in V$, as a result of property (iii)′ of convex sets we see that $x_0 + \varepsilon V \subset A$, that is, $x_0 \in \operatorname{int} A$, which together with relation (1.27) implies $\operatorname{int} A = A^{\mathrm{i}}$. □

Corollary 1.21 *The Minkowski functional of a convex, absorbent set A of a topological linear space is continuous if and only if $\operatorname{int} A \ne \emptyset$. In this case, we have*

$$\operatorname{int} A = A^{\mathrm{i}}, \qquad \overline{A} = A^{\mathrm{ac}}, \qquad \operatorname{Fr} A = A^{\mathrm{rb}},$$

where $\operatorname{Fr} A = \overline{A} \cap \overline{cA}$.

Finally, let us see what happens when a functional on a linear space coincides with the Minkowski functional of a convex set.

Proposition 1.22 *If $p : X \to\ ]-\infty, \infty]$ is a proper sublinear, nonnegative function, then*:

(i) $A_1 = \{x \in X;\ p(x) < 1\}$ *has only relative algebraic interior points*
(ii) $A_2 = \{x \in X;\ p(x) \le 1\}$ *coincides with its algebraic closure*
(iii) $p_A = p$, *if* $A_1 \subset A \subset A_2$.

Proof First, we see that $0 \in A_1 \cap A_2$, $\operatorname{aff} A_1 = \operatorname{aff} A_2$, which are proper linear subspaces. As we have seen in the second half of the proof of Proposition 1.19, the origin is an algebraic relative interior point to A_1; hence it is of the same type for A_2. Consequently, it is sufficient to prove only statement (iii).

But for every $\lambda > 0$ such that $p(x) < \lambda$, we find as a result that $\frac{1}{\lambda}x \in A_1$ and so $p_{A_1}(x) \le \lambda$, that is, $p_{A_1}(x) \le p(x)$.

If $p(x) = 0$, then $\lambda x \in A_1 \cap A_2$, for every $\lambda > 0$, and we obtain $p_{A_1}(x) = p_{A_2}(x) = 0$.

If $p(x) \ne 0$, for every $\lambda \in \mathbb{R}$ with $0 < \lambda < p(x)$, we have $\frac{1}{\lambda}x \notin A_2$; hence, $\lambda \le p_{A_2}(x)$, that is, $p(x) \le p_{A_2}(x)$. Since $A_1 \subset A_2$, we find as a result that $p_{A_2} \le p_{A_1}$.

Now, using the two inequalities established above, we obtain $p_{A_1} = p_{A_2} = p$, which implies statement (iii). □

Corollary 1.23 *Let X be a topological linear space and let A be a convex set of X containing the origin. Then*

(i) $A = \{x \in X;\ p_A(x) \le 1\}$ *if A is closed*
(ii) $A = \{x \in X;\ p_A(x) < 1\}$ *if A is open.*

Remark 1.24 An important problem is to specify the conditions under which the relative interior of a set is nonempty. For instance, we can show that *every nonempty convex set of* $\mathbb{R}^n$ *has a nonempty relative interior.* On the other hand, *in a Banach space, for a closed convex set its interior is the same as its algebraic interior*, because every Banach space is of the second category.

Many special properties of convex sets in a linear space may be found in the books of Eggleston [13] and Valentine [28].

Definition 1.25 A maximal affine set is called a *hyperplane*. We say that the hyperplane is *homogeneous* (*nonhomogeneous*) if it contains (does not contain) the origin.

Since every affine set is the translation of a linear subspace, as a result we find that a set is a hyperplane if and only if it is the translation of a maximal linear subspace. In particular, the homogeneous hyperplanes coincide with the maximal linear subspaces.

Proposition 1.26 *In a real topological linear space X, any homogeneous hyperplane is either closed or dense in X.*

Proof If H is a homogeneous hyperplane, then $\overline{H}$ is evidently a linear subspace; hence, from the maximality of H, since $H \subset \overline{H}$, we find as a result that $\overline{H} = H$ or $\overline{H} = X$, as claimed. □

The next theorem, concerning the connection between hyperplanes and linear functionals, represents a fundamental result in the theory of hyperplanes.

Theorem 1.27 *The kernel of a nontrivial linear functional is a homogeneous hyperplane. Conversely, for every homogeneous hyperplane H there exists a functional, uniquely determined up to a nonzero multiplicative constant, with the kernel H.*

Proof If it is a nontrivial linear functional, it can be observed that its *kernel*, $\ker f = f^{-1}(\{0\})$, is a linear subspace. Let Y be a linear subspace, which strictly contains the kernel, that is, there exists $y_0 \in Y$ such that $f(y_0) \neq 0$. For every $x \in X$ we have

$$x = \left(x - \frac{f(x)}{f(y_0)}\, y_0\right) + \frac{f(x)}{f(y_0)}\, y_0 = u + \lambda y_0 \in Y,$$

because $u = x - \frac{f(x)}{f(y_0)}\, y_0 \in \ker f \subset Y$. Thus, $X = Y$. Hence, $\ker f$ is maximal, that is, it represents a homogeneous hyperplane. Now, let H be a homogeneous hyperplane and let $z_0 \in X \setminus H$. Since the linear subspace spanned by $H \cup \{z_0\}$ strictly contains H, it must coincide with X. Therefore, every $x \in X$ can be represented uniquely as $x = u + \lambda z_0$, $u \in H$, $\lambda \in \mathbb{R}$. We define the functional f on X by $f(x) = \lambda$, if $x = u + \lambda z_0$. One can easily verify that f is linear and $\ker f = H$.

Let f_1 and f_2 be two nontrivial linear functionals such that $\ker f_1 = \ker f_2$. If $x_0 \overline{\in} \ker f_1$, we have as a result that $x - \frac{f_1(x)}{f_1(x_0)}\, x_0 \in \ker f_1$ for every $x \in X$. Consequently, $f_2(x - \frac{f_1(x)}{f_1(x_0)}\, x_0) = 0$, that is, $f_2(x) = k f_1(x)$ for every $x \in X$, where $k = \frac{f_2(x_0)}{f_1(x_0)}$ is a real constant. Hence, the theorem ensures the uniqueness of the functional up to a nonzero multiplicative constant. □

Corollary 1.28 *If f is a nontrivial linear functional on the linear space X, then $\{x \in X;\, f(x) = k\}$ is a hyperplane of X, for every $k \in \mathbb{R}$. Conversely, for every hyperplane H, there exists a linear functional f and $k \in \mathbb{R}$, such that $H = \{x \in X;\, f(x) = k\}$.*

Corollary 1.29 *A hyperplane is closed if and only if it is determined by a nonidentically zero continuous linear functional.*

Proof Use Theorems 1.2 and 1.27. □

From the above considerations, it follows that every hyperplane verifies an equation of the form

$$f(x) = k, \quad k \in \mathbb{R}.$$

For $k \neq 0$, we can put this equation in the form

$$f(x) = 1.$$

In this form, the linear functional f is uniquely determined by the nonhomogeneous hyperplane.

If the hyperplane is homogeneous, then f is nonunique. More precisely, if $\ker f_1 = \ker f_2$, then $f_1 = af_2$, where a is a nonzero constant. Along these lines, we inductively obtain a more general result.

Theorem 1.30 *If $f, f_1, f_2, \ldots, f_n$ are $n+1$ linear functionals, such that $f(x) = 0$, whenever $f_1(x) = f_2(x) = \cdots = f_n(x) = 0$, then f is a linear combination of $f_1, f_2, \ldots, f_n$.*

Proof Applying mathematical induction, we observe that in $\ker f_n$ we have n linear functionals $f, f_1, f_2, \ldots, f_{n-1}$ having the property that $f(x) = 0$, whenever $f_1(x) = f_2(x) = \cdots = f_{n-1}(x) = 0$ if $x \in \ker f_n$. Consequently, there exists $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$, such that $f(x) = \sum_{i=1}^{n-1} \lambda_i f_i(x)$, $x \in \ker f_n$.

Now, we observe that f_n and $f - \sum_{i=1}^{n-1} \lambda_i f_i$ are two linear functionals having the same kernels. Thus, by Theorem 1.27, there exists $\lambda_n \in \mathbb{R}$, such that $f - \sum_{i=1}^{n-1} f_i = \lambda_n f_n$, as claimed. $\square$

1.1.3 Separation of Convex Sets

If $f(x) = k$, $k \in \mathbb{R}$, is the equation of a hyperplane in a real linear space X, we have two *open half-spaces* $\{x \in X;\ f(x) < k\}$, $\{x \in X;\ f(x) > k\}$ and two *closed half-spaces* $\{x \in X;\ f(x) \leq k\}$, $\{x \in X;\ f(x) \geq k\}$. It is clear that the algebraic boundary of each of the four half-spaces is just the hyperplane $f(x) = k$. It should be emphasized that a *convex set which contains no point of a hyperplane is contained in one of the two open half-spaces determined by that hyperplane.* Indeed, if $f(x_1) > k$ and $f(x_2) < k$, there exists $\lambda \in\ ]0, 1[$, such that $f(\lambda x_1 + (1-\lambda)x_2) = k$, hence x_1 and x_2 cannot be contained in a convex set which is disjoint from the hyperplane $f(x) = k$.

Remark 1.31 An open half-space has only algebraic interior points and each closed half-space coincides with its algebraic closure. If X is a topological linear space, then *the open half-spaces are open sets and the closed half-spaces are closed sets if and only if the linear functional f which generated them is continuous, or, equivalently, the hyperplane $\{x \in X;\ f(x) = k\}$ is closed* (cf. Corollary 1.29).

In the following, we deal with sufficient conditions which ensure that two convex sets can be separated by a hyperplane. Such results are immediate consequences of the *Hahn–Banach Theorem*. For this purpose, we define the concept of a convex function which generalizes that of a subadditive functional.

Definition 1.32 A function $p : X \to\ ]-\infty, +\infty]$ is called *convex* if

$$p(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 p(x_1) + \lambda_2 p(x_2), \tag{1.28}$$

for all $x_1, x_2 \in X$ and $\lambda_1 \geq 0$, $\lambda_2 \geq 0$, with $\lambda_1 + \lambda_2 = 1$. If inequality (1.28) is strict for $x_1 \neq x_2$ in $\mathrm{Dom}(p)$ and $\lambda_1 > 0$, $\lambda_2 > 0$, then the function p is called *strictly convex*.

It easily follows that inequality (1.28) is equivalent to the property

$$(a_1 + a_2)\,p\left(\frac{a_1 x_1 + a_2 x_2}{a_1 + a_2}\right) \le a_1 p(x_1) + a_2 p(x_2), \tag{1.29}$$

for all $x_1, x_2 \in X$ and $a_1 > 0$, $a_2 > 0$.

The Minkowski functionals of convex sets which contain the origin are examples of convex functions.

Theorem 1.33 (Hahn–Banach) *Let X be a real linear space, let p be a real convex function on X and let Y be a linear subspace of X. If a linear functional f_0 defined on Y satisfies*

$$f_0(y) \le p(y), \quad \forall y \in Y, \tag{1.30}$$

then f_0 can be extended to a linear functional f defined on all of X, satisfying

$$f(x) \le p(x), \quad \forall x \in X. \tag{1.31}$$

Proof If $u, v \in Y$, $x_0 \in X \setminus Y$ and $\alpha > 0$, $\beta < 0$, according to relations (1.29) and (1.30), it follows that

$$\begin{aligned}
\alpha f_0(u) - \beta f_0(v) &= f_0(\alpha u - \beta v) = (\alpha - \beta) f_0\left[\frac{\alpha(u + \frac{1}{\alpha}x_0)}{\alpha - \beta} + \frac{-\beta(v + \frac{1}{\beta}x_0)}{\alpha - \beta}\right] \\
&\le (\alpha - \beta)\,p\left[\frac{\alpha(u + \frac{1}{\alpha}x_0)}{\alpha - \beta} + \frac{-\beta(v + \frac{1}{\beta}x_0)}{\alpha - \beta}\right] \\
&\le \alpha p\left(u + \frac{1}{\alpha}x_0\right) - \beta p\left(v + \frac{1}{\beta}x_0\right).
\end{aligned}$$

Thus, there exists $c \in \mathbb{R}$ such that

$$\begin{aligned}
&\sup\left\{\beta\left[p\left(v + \frac{1}{\beta}x_0\right) - f_0(v)\right];\ v \in Y,\ \beta < 0\right\} \\
&\quad \le c \le \inf\left\{\alpha\left[p\left(u + \frac{1}{\alpha}x_0\right) - f_0(u)\right];\ u \in Y,\ \alpha > 0\right\}.
\end{aligned} \tag{1.32}$$

First, we prove that f_0 can be extended to the linear subspace $X_1 = \operatorname{span}(Y \cup \{x_0\})$ preserving the linearity and the boundedness condition (1.30). We observe that each element $x_1 \in X_1$ has the form $x_1 = y + \lambda x_0$, with $y \in Y$ and $\lambda \in \mathbb{R}$ uniquely determined. We define the functional f_1 on X_1 by $f_1(x_1) = f_0(y) + \lambda c$ if $x_1 = y + \lambda x_0$, with $y \in Y$ and $\lambda \in \mathbb{R}$. It can easily be seen that f_1 is linear on X_1 and $f|_Y = f_0$. To prove the boundedness property (1.30) on X_1, we consider two cases: $\lambda < 0$ and $\lambda > 0$ (the case $\lambda = 0$ is obvious). Thus, if $x_1 = y + \lambda x_0$, $y \in Y$ and $\lambda \ne 0$, then we have

$$f_1(x_1) = f_0(y) + \lambda c \le f_0(y) + \lambda\left[\frac{1}{\lambda}\,p(y + \lambda x_0) - \frac{1}{\lambda}\,f_0(y)\right] = p(x_1)$$

as follows from the left-hand side of relation (1.32) for $v = y$ and $\beta = \frac{1}{\lambda}$ if $\lambda < 0$ or from the right-hand side of (1.32) for $u = y$ and $\alpha = \frac{1}{\lambda}$ if $\lambda > 0$. Using the *Zorn Lemma* (*any nonvoid ordered set has at least one maximal element*), it is clear that every maximal element of the set of all the linear functionals which extend f_0 and preserve the boundedness property (1.30) on the linear subspaces on which they are defined is again a linear functional defined on the whole of X. Indeed, otherwise, according to the above, a strict extension would exist, which would contradict the maximality property. Thus, any maximal element has all the required properties of the theorem. □

Remark 1.34 From the above proof, we see that the theorem remains valid if the convex function p takes also infinite values but has the following property:

if $x \in X \setminus Y$ and $p(y + kx) = \infty$, for all $y \in Y$ and $k > 0$, then $p(y + kx) = \infty$ for all $y \in Y$ and $k < 0$. Particularly, it suffices that

$$Y \cap \left(\mathrm{Dom}(p)\right)^{\mathrm{ri}} \neq \emptyset.$$

It can be easily seen that the Hahn–Banach Theorem 1.33 may be, equivalently, reformulated in the following form.

Theorem 1.35 *If A is a convex set with $A^{\mathrm{ri}} \neq \emptyset$ and M is an affine set such that $M \cap A^{\mathrm{ri}} = \emptyset$, then there exists a hyperplane containing M, which is disjoint from A^{ri}.*

Proof We may suppose, without loss of generality, that $0 \in A^{\mathrm{ri}}$. Hence, $0 \overline{\in} M$ and M is a maximal affine set in $\mathrm{span}\, M$. According to Corollary 1.28, there exists a linear functional f_0 on $\mathrm{span}\, M$ such that $M = \{y \in \mathrm{span}\, M;\ f_0(y) = 1\}$. On the other hand, by Proposition 1.19 we have $A^{\mathrm{ri}} = \{x \in X;\ p_A(x) < 1\}$. If $f_0(y) > 0$, it follows that $\frac{y}{f_0(y)} \in M$, hence $\frac{y}{f_0(y)} \overline{\in} A^{\mathrm{ri}}$ and this implies $p_A(\frac{y}{f_0(y)}) \geq 1$ or $f_0(y) \leq p_A(y)$. This inequality is obvious in the case $f_0(y) \leq 0$. Thus, we have $f_0(y) \leq p_A(y)$ for all $y \in \mathrm{span}\, M$. By the Hahn–Banach Theorem 1.33 (see also the last part of Remark 1.34), there exists a linear extension f of f_0 on the whole of X such that $f(x) \leq p_A(x)$, $x \in X$. If $u \in A^{\mathrm{ri}}$, we have $p_A(u) < 1$, so that $f(u) < 1$. Hence, A^{ri} is disjoint from the nonhomogeneous hyperplane $f(x) = 1$, which contains M.

If X is finite-dimensional, according to Remark 1.24, the hypothesis $A^{\mathrm{ri}} \neq \emptyset$ is fulfilled for any nonvoid convex set.

The result of the algebraic type established by Theorem 1.35 may be improved if X becomes a linear topological space. In this context, we have the well-known *geometric form of the Hahn–Banach Theorem.* □

Theorem 1.36 *If A is a convex set with a nonempty interior and if M is an affine set which contains no interior point of A, then there exists a closed hyperplane which contains M and which again contains no interior point of A.*

Proof In our hypothesis, the interior of A coincides with its algebraic interior (cf. Corollary 1.20). Thus, it is sufficient to prove that the linear functional found in the proof of the previous theorem is continuous. Continuity holds since the Minkowski functional associated to A is continuous (cf. Corollary 1.21). □

Corollary 1.37 *On a topological linear space there exist nontrivial continuous linear functionals* (*or closed hyperplanes*) *if and only if there exist proper convex sets with nonempty interior. On any proper locally convex space there exist nontrivial continuous functionals and closed hyperplanes.*

A hyperplane H is called a *supporting hyperplane* of a set A if H contains at least one point of A and A lies in one of the two closed half-spaces determined by H. In the analytic form, this fact ensures the existence of a nontrivial linear functional f and an element $x_0 \in A$, such that

$$\sup\{f(x);\ x \in A\} = f(x_0).$$

In a linear topological space, *any supporting hyperplane of a set with a nonempty interior is closed*. A point of A through which a supporting hyperplane passes is called a *support point* of A. It is clear that an algebraic interior point cannot be a support point. Hence, any support point is necessarily an algebraic boundary point. Now, we give a simple condition under which a boundary point is a support point.

Theorem 1.38 *If the interior of a convex set is nonempty*, *then all the boundary points are support points.*

Proof Apply Theorem 1.36 for $M = \{x_0\}$, where x_0 is an arbitrary boundary point. □

Remark 1.39 The uniqueness of the supporting hyperplane passing through a support point depends on the differentiability property of the Minkowski functional associated with that set (see Sect. 2.2.2). We restrict our attention to the case in which two convex sets may be separated by a hyperplane, that is, they are contained in different half-spaces. If they are contained even in different open half-spaces, the sets are said to be *strictly separated* by that hyperplane.

Theorem 1.40 *If A_1 and A_2 are two nonempty convex sets and if at least one of them has a nonempty interior and is disjoint from the other set*, *then there exists a separating hyperplane. Moreover*, *if A_1 and A_2 are open*, *the separation is strict.*

Proof Suppose that $\operatorname{int} A_2 \neq \emptyset$ and $A_1 \cap \operatorname{int} A_2 = \emptyset$. The set $A = A_1 - \operatorname{int} A_2$ is open convex and does not contain the origin. For $M = \{0\}$, from Theorem 1.36 there exists a closed hyperplane H_1 such that $H_1 \cap A = \emptyset$ and $0 \in H_1$. Therefore, H_1 is a homogeneous hyperplane for which A lies in one of the open half-spaces. According to Theorem 1.27 and Corollary 1.29, there exists a continuous linear functional f which has as kernel H_1. Hence, f keeps a constant sign of A. Suppose that $f(x) > 0$, $\forall x \in A$, that is, $f(u) > f(v)$, $\forall u \in A_1$, $\forall v \in A_2$, which implies

$$\inf\{f(u);\ u \in A_1\} \geq \sup\{f(x);\ v \in A_2\}. \tag{1.33}$$

Since A_2 being convex, $A_2 \subset \overline{\operatorname{int} A_2}$. Clearly, the hyperplane $f(x) = k$, where $k = \inf_{u \in A_1} f(u)$, is closed and separates A_1 and A_2. If A_1 and A_2 are open, we have $f(u) > k > f(v)$, $\forall u \in A_1$, $\forall v \in A_2$, since A_1 and A_2 have no boundary points. □

Corollary 1.41 *If A_1 and A_2 are two nonempty disjoint convex sets of $\mathbb{R}^n$, there exists a nonzero element $c = (c_1, c_2, \ldots, c_n) \in \mathbb{R}^n \setminus \{0\}$, such that*

$$\sum_{i=1}^{n} c_i u_i \le \sum_{i=1}^{n} c_i v_i, \quad \forall u = (u_i) \in A_1, \ \forall v = (v_i) \in A_2.$$

Proof Let us write $A = A_1 - A_2$. Since $A^{\mathrm{ri}} \neq \emptyset$ (see Remark 1.24) and $0 \overline{\in} A$, we can apply Theorem 1.35 for $M = \{0\}$. Taking into account the form of nontrivial linear functionals on $\mathbb{R}^n$, we find $c = (c_1, c_2, \ldots, c_n) \in \mathbb{R}^n \setminus \{0\}$, such that $\sum_{i=1}^{n} c_i a_i < 0$ for all $(a_1, a_2, \ldots, a_n) \in A^{\mathrm{ri}}$ and so $\sum_{i=1}^{n} c_i a_i \le 0$ for all $(a_1, a_2, \ldots, a_n) \in A = A_1 - A_2$, because the interior of any segment joining a point of A and a point of A^{ri} contains only points of A^{ri}. □

Remark 1.42 From Theorem 1.35, we can obtain a result, being of algebraic type, similar to that obtained using Theorem 1.40.

Remark 1.43 If we drop the condition that A_1 and A_2 are open, Theorem 1.40 is no longer true, as can be shown by counter-examples that can readily be constructed. Thus, the disjoint convex sets

$$A_1 = \{(x_1, x_2);\ x_1 \le 0\} \quad \text{and} \quad A_2 = \{(x_1, x_2);\ x_1 x_2 \ge 1,\ x_1 \ge 0,\ x_2 \ge 0\}$$

have nonempty interiors in $\mathbb{R}^2$ but cannot be strictly separated; the single separation hyperplane is $x_1 = 0$.

Theorem 1.44 *If F_1 and F_2 are two disjoint nonempty closed convex sets in a separated locally convex space such that at least one of them is compact, then there exists a hyperplane strictly separating F_1 and F_2. Moreover, there exists a continuous linear functional f such that*

$$\sup\{f(x);\ x \in F_1\} < \inf\{f(x);\ x \in F_2\}. \tag{1.34}$$

Proof Suppose that F_2 is compact. Since any separated locally convex space is regular, there exists an open symmetric and convex neighborhood of the origin V such that $F_1 \cap (F_2 + V) = \emptyset$.

Thus, we may apply Theorem 1.40 to conclude that there exists a nontrivial continuous functional f such that

$$\sup\{f(x);\ x \in F_1\} \le \inf\{f(y);\ y \in F_2 + V\}.$$

But $\inf\{f(v); v \in V\} < 0$, since V is absorbent and f is a nontrivial linear functional. Hence, relation (1.34) holds. It is clear that $\{x \in X; f(x) = k\}$ is a strict separation hyperplane for any $k \in]\sup\{f(x); x \in F_1\}, \inf\{f(x); x \in F_2\}[$. □

Corollary 1.45 *If $x_0 \overline{\in} F$, where F is a nonempty closed convex set of a separated locally convex space, then there exists a closed hyperplane strictly separating F and x_0, that is, there is a nontrivial continuous linear functional such that*

$$\sup\big\{f(x);\ x \in F\big\} < f(x_0).$$

Remark 1.46 Generally, if inequality (1.34) is fulfilled, we say that the sets F_1 and F_2 are *strongly separated* by the hyperplane $f(x) = k$, where $k \in]\sup\{f(x); x \in F_1\}, \inf\{f(x); x \in F_2\}[$. We observe that *two convex sets A, B can be (strongly) separated if and only if the origin can be (strongly) separated from $A - B$.*

Remark 1.47 If the set F from Corollary 1.45 is a closed linear subspace, then f must be null on F. Therefore, *a linear subspace is dense in X if and only if every continuous linear functional which is null on it, is null on X.*

Now, as a consequence of the separation theorems, we obtain the following remarkable theorem concerning the characterization of closed convex sets.

Theorem 1.48 *A proper convex set of a separated locally convex space is closed if and only if it coincides with an intersection of closed half-spaces.*

Proof The sufficiency is obvious. To prove the necessity, we consider the set $\{S_i; i \in I\}$ of all closed half-spaces which contain the proper convex closed set F.

For every $x_0 \overline{\in} F$, taking f as in Corollary 1.45, we have $F \subset \{x;\ f(x) \le k_0\}$ and $f(x_0) > k_0$, where $k_0 = \sup\{f(x); x \in F\}$. Therefore, there exists $i_0 \in I$ such that $S_{i_0} = \{x;\ f(x) \le k_0\}$, hence $CF \subset \bigcup_{i\in I} CS_i$, that is, $F \supset \bigcap_{i\in I} S_i$. On the other hand, it is clear that $F \subset \bigcap_{i\in I} S_i$, hence $F = \bigcap_{i\in I} S_i$ and the theorem is completely proven. □

Corollary 1.49 *A closed convex set with nonempty interior of a separated locally convex space coincides with the intersection of all half-spaces generated by its supporting hyperplanes.*

In the following, we consider the special case of linear normed spaces. The Hahn–Banach Theorem 1.33 ensures in this case the existence of continuous linear extensions which preserve the norm. We recall that the dual X^* of a normed linear space X is the set of all continuous linear functionals on X; it is again a normed space.

Theorem 1.50 *Let f_0 be a continuous linear functional on a linear subspace Y of a linear normed space X. Then there exists a continuous linear functional f on the whole of X, that is, $f \in X^*$, such that*

(i) $f|_Y = f_0$
(ii) $\|f\| = \|f_0\|$.

Proof Since f_0 is continuous on Y, by (1.10) we have $f_0(y) \le \|f_0\|\,\|y\|$, $\forall y \in Y$. □

Now, we can apply the Hahn–Banach Theorem (Theorem 1.33) for f_0 and for the convex function $p(x) = \|f_0\|\,\|x\|$.

A specialization of this theorem yields a whole class of existence results. In this context, we present a general and classical theorem concerning the existence of continuous linear functionals with important consequences in the duality theory of linear normed spaces.

Theorem 1.51 *Let m be a nonnegative number and let $h : A \to \mathbb{R}$ be a given real function, where A is a nonempty set of the linear normed space X. Then h has a continuous linear extension f on all of X such that $\|f\| \le m$ if and only if the following condition holds*:

$$\left|\sum_{i=1}^{n} \lambda_i h(a_i)\right| \le m \left\|\sum_{i=1}^{n} \lambda_i a_i\right\|, \quad \forall n \in \mathbb{N}^*,\ \lambda_i \in \mathbb{R},\ a_i \in A. \tag{1.35}$$

Proof From relations (1.8) and (1.9) it is clear that condition (1.35) is necessary. To prove sufficiency, we consider $Y = \operatorname{span} A$ and we define f_0 on Y by

$$f_0(y) = \sum_{i=1}^{n} \lambda_i h(a_i), \quad \text{if } y = \sum_{i=1}^{n} \lambda_i a_i \in Y,\ a_i \in A.$$

First, using condition (1.35), we observe that f_0 is well defined on Y. Moreover, from condition (1.35), continuity of f_0 on Y follows and $\|f_0\| \le m$. Thus, any extension given under Theorem 1.50 has all the required properties. □

Theorem 1.52 *For any linear subspace Y of a normed linear space X and $x \in X$ there exists $f \in X^*$ with the following properties*:

(i) $f|_Y = 0$
(ii) $f(x) = d^2(x; Y)$
(iii) $\|f\| = d(x; Y)$.

Proof We take $A = Y \cup \{x\}$ and $h : A \to \mathbb{R}$ defined by $h(y) = 0$, $y \in Y$, and $h(x) = d^2(x; Y)$. We observe that, for any $\lambda \ne 0$, we have

$$\begin{aligned}\left|\lambda h(x) + \sum_{i=1}^{n} \lambda_i h(a_i)\right| &= \left|\lambda h(x)\right| = |\lambda| d^2(x; Y)\\ &\le |\lambda| d(x; Y)\left\|x + \sum_{i=1}^{n} \frac{\lambda_i}{\lambda} a_i\right\| = d(x; Y)\left\|\lambda x + \sum_{i=1}^{n} \lambda_i a_i\right\|,\end{aligned}$$

which is just inequality (1.35). The desired result then follows by applying the previous theorem. Indeed, we have properties (i) and (ii) and $\|f\| \le d(x; Y)$ since

$m = d(x; Y)$. On the other hand, if we consider a sequence $\{y_n\} \subset Y$ such that $\|x + y_n\| \to d(x; Y)$, we obtain

$$\|f\| \geq f\left(\frac{x + y_n}{\|x + y_n\|}\right) = \frac{f(x)}{\|x + y_n\|} = \frac{d^2(x; Y)}{\|x + y_n\|} \to d(x; Y),$$

which implies $\|f\| \geq d(x; Y)$. Hence, property (iii) also holds. □

Corollary 1.53 *In a linear normed space X, for every $x \in X$ there exists a continuous linear functional $f \in X^*$ such that*

(i) $f(x) = \|x\|^2$
(ii) $\|f\| = \|x\|$.

Moreover, if $x \neq 0$, there exists $g \in X^*$ such that

(i)′ $g(x) = \|x\|$
(ii)′ $\|g\| = 1$.

Proof Theorem 1.52 can be applied where $Y = \{0\}$, hence $d(x; Y) = \|x\|$. □

In particular, relations (1.9) and (1.10) and the second part of Corollary 1.53 yield the following corollary.

Corollary 1.54 *For any $x \in X$, we have*

$$\|x\| = \max_{\|x^*\| \leq 1} \left|(x, x^*)\right| = \max_{\|x^*\| \leq 1} (x, x^*). \tag{1.36}$$

This formula is known as the *dual formula of the norm* in a linear normed space.

1.1.4 Closedness of the Sum of Two Sets

It is well known that, generally, in a linear topological space, the sum of two closed sets is not a closed set. But if one of the two sets is compact, then the sum is also closed (the property can be immediately verified using the nets). Furthermore, Klee [18] showed that in a Banach space the sum of two bounded closed convex sets is always closed if and only if the space is reflexive (see, also, Köthe [19], p. 322).

Because the compactness (or boundedness and reflexivity) hypotheses are too strong, in practice we use suitable sufficient conditions which ensure the closedness property of the sum of certain pairs of sets. Such a result was established by Dieudonné [10], replacing the compactness condition by a local condition. Next, Dieudonné's criterion is extended in several directions. Thus, we must remark that the closedness problem of the sum of two sets may be regarded in the framework of a more general problem, namely, of the *closedness of the image of a set by a multi-valued function*. In the following, we present a general result due to Dedieu [8]. This result extends to nonconvex case the generalizations of Dieudonné's criterion [10] established by Gwinner [14] in the convex case.

If $\Lambda \subset \Gamma$ and $A \subset X$, we denote

$$\Lambda A = \{\lambda x;\ \lambda \in \Lambda,\ x \in A\}.$$

Definition 1.55 For a given set A, the set A_∞ defined by

$$A_\infty = \bigcap_{\varepsilon > 0} \overline{[0, \varepsilon]A} \tag{1.37}$$

is called the *asymptotic cone* of A.

Remark 1.56 The asymptotic cone A_∞ is a closed cone with the vertex in the origin consisting of all the limits of the convergent nets of the form $(\lambda_i x_i)_{i\in I}$, where $(\lambda_i)_{i\in I}$ is a net of positive numbers convergent to zero and $(x_i)_{i\in I}$ is a net of elements of A. Also, $A_\infty = A$ if and only if A is a closed cone.

If the set A is convex and closed, then the asymptotic cone can be equivalently defined by

$$A_\infty = \bigcap_{\varepsilon > 0} \varepsilon(A - a), \tag{1.38}$$

where a is an arbitrary fixed element of A. In this case, A_∞ is also convex.

Definition 1.57 A set A is called *asymptotically compact* if there exists $\varepsilon_0 > 0$ and a neighborhood V_0 of the origin such that $([0, \varepsilon_0]A) \cap V_0$ is relatively compact.

It is easy to observe that a closed convex set or a closed cone is asymptotically compact if and only if it is locally compact. Generally, for nonconvex sets these concepts are different. If A is closed and asymptotically compact, then A and A_∞ are locally compact, but the converse is false.

Theorem 1.58 *Let E_1, E_2 be two separated linear topological spaces and let A be an asymptotically compact and closed subset of E_1. If $F : E_1 \to E_2$ is a closed multi-valued mapping, that is,* Graph $F = \{(x, y) \in E_1 \times E_2;\ x \in E_1,\ y \in F(x)\}$ *is closed in $E_1 \times E_2$, satisfying the condition*

$$\big(A_\infty \times \{0_{E_2}\}\big) \cap (\mathrm{Graph}\, F)_\infty = \{0_{E_1\times E_2}\}, \tag{1.39}$$

then $F(A)$ is closed in E_2.

Proof Let $\{y_i\}_{i\in I}$ be a net of elements of $F(A)$ convergent to an element $y_0 \in E_2$ and let us prove that $y_0 \in F(A)$. Take $\{x_i\}_{i\in I}$, a net of elements of A such that $y_i \in F(x_i)$, for every $i \in I$. We can suppose that $x_i \neq 0$ for all $i \in I$ (taking a subnet) since if there exists a subnet $x_{i'} = 0$ for all $i' \in I'$, then $y_0 \in F(0)$, as claimed. Let ε_0 and V_0 be as in Definition 1.57. Since V_0 is absorbent, for each $i \in I$ there exist positive numbers $\lambda > 0$ such that $\lambda x_i \in V_0$. We suppose that V_0 is an open circled neighborhood of the origin and denote $\lambda_i = \min(\{\varepsilon_0\} \cup \{\lambda;\ \lambda > 0 \text{ and } 2\lambda x_i \overline{\in} V_0\})$. Since $\lambda_i \in]0, \varepsilon_0]$ and $\lambda_i x_i \in V_0$, the net $\{\lambda_i x_i\}_{i \in I}$ contains a subnet convergent to an element $x_0 \in E_1$. Without loss of generality, we may assume that even $\{\lambda_i x_i\}_{i \in I}$ is convergent. Also, since $\{\lambda_i\}_{i \in I}$ is bounded, we can suppose that it is convergent to an element $\lambda_0 \geq 0$. If $\lambda_0 \neq 0$, the net $\{x_i\}_{i \in I}$ is also convergent to $x_0' = \frac{1}{\lambda_0} x_0 \in A$ since A is closed. Hence, $\{(x_i, y_i)\}_{i \in I} \subset \operatorname{Graph} F$ is convergent to $(x_0', y_0) \in \operatorname{Graph} F$, since F is closed, that is, $y_0 \in F(x_0') \subset F(A)$ and the proof is finished. Now, we prove that it is not possible to have $\lambda_0 = 0$. Indeed, if $\lambda_0 = 0$, there exists $i_0 \in I$ such that $\lambda_i < \varepsilon_0$ for all $i > i_0$, and so, from the definition of λ_i, we have

$$2\lambda_i x_i \overline{\in} V_0, \quad \text{for all } i > i_0. \tag{1.40}$$

On the other hand, according to Remark 1.56 it follows that $x_0 \in A_\infty$ and therefore the net $\{(\lambda_i x_i, \lambda_i y_i)\}_{i \in I}$ is convergent to $(x_0, 0) \in A_\infty \times \{0_{E_2}\}$. But $(\lambda_i x_i, \lambda_i y_i) = \lambda_i (x_i, y_i) \in \lambda_i \operatorname{Graph} F$ and, by virtue of the same Remark 1.56, its limit belongs to $(\operatorname{Graph} F)_\infty$. By condition (1.39), we have $x_0 = 0$, contradicting property (1.40). Thus, the proof is complete. □

Theorem 1.59 *Let E_1, E_2 be two linear separated topological spaces and let T be a positive-homogeneous operator defined on a cone of E_1 into E_2. Let $A \subset E_1$, $B \subset E_2$ be two closed sets such that $\overline{\operatorname{cone} A} \subset \operatorname{dom} T$ and such that the restriction of T to $\overline{\operatorname{cone} A}$ is continuous. If*

$$A_\infty \cap T^{-1}(B_\infty) = \{0\} \tag{1.41}$$

and A is asymptotically compact, then $T(A) - B$ is closed.

Proof It is readily seen that for the multi-valued function F defined by $Fx = Tx - B$ for $x \in \overline{\operatorname{cone} A}$ and $Fx = \emptyset$ otherwise, the asymptotic separation property (1.39) is just condition (1.41). Also, $\operatorname{Graph} F$ is closed in $E_1 \times E_2$. Indeed, if $\{(x_i, Tx_i - b_i)\}_{i \in I}$ is convergent to (x_0, y_0), then $x_0 \in A$ and $(b_i)_{i \in I}$ is convergent to $Tx_0 - y_0 \in B$, and so $y_0 \in F(x_0)$. But A being asymptotically compact and closed, according to Theorem 1.58, it follows that $F(A) = T(A) - B$ is closed. □

Taking $B = \{0\}$ and $T = I$ (the identity operator), we obtain the following two special results.

Corollary 1.60 *Let T and A be as in Theorem 1.59. If*

$$A_\infty \cap \ker T = \{0\}, \tag{1.42}$$

then $T(A)$ is closed.

Corollary 1.61 (Dieudonné) *Let A, B be two closed sets such that*

$$A_\infty \cap B_\infty = \{0\}. \tag{1.43}$$

If A or B is asymptotically compact, then their difference is also closed.

If the sets A and B are also convex, then Corollary 1.61 is just the well-known *Dieudonné criterion* (in the convex case, the asymptotical compactness becomes even local compactness). For closed cones, the preceding results can be improved by taking into account the following simple result on the characterization of asymptotical compactness.

Theorem 1.62 *A cone is locally compact* (*or closed asymptotically compact*) *if and only if it can be generated by a compact set which does not contain the origin.*

Proof Let A be an asymptotically closed cone and let ε_0 and V_0 be as in Definition 1.57. We can suppose that V_0 is closed and circled. Since A is a cone, we have $]0, \varepsilon_0]A = A$ and so $A \cap V_0$ is compact. Also, $K = A \cap (V_0 \setminus \operatorname{int} \frac{1}{2} V_0)$ is compact and does not contain the origin. Conversely, if $A = \operatorname{cone} K$, where K is a compact set and $0 \overline{\in} K$, then there exists a closed circled neighborhood of origin V_0 such that $V_0 \cap K = \emptyset$. But $V_0 \cap A \subset [0, 1]K$. Indeed, if $x \in V_0 \cap A$, there exists $\lambda > 0$ such that $\lambda x \in K$. If $\lambda \geq 1$, we have $x \in \frac{1}{\lambda} K \subset [0, 1]K$. If $\lambda < 1$, it follows that $\lambda x \in V_0$ since V_0 is circled and thus $V_0 \cap K \neq \emptyset$, which is a contradiction. Therefore, $V_0 \cap A$ is compact, that is, A is locally compact. □

Corollary 1.63 *A locally compact cone is necessarily closed.*

Corollary 1.64 *If A is a locally compact cone and T is a linear continuous operator such that*

$$A \cap \ker T = \{0\}, \tag{1.44}$$

then $T(A)$ is also a locally compact cone.

Proof Let K be a compact set such that $A = \operatorname{cone} K$ and $0 \overline{\in} K$. Then $T(A) = \operatorname{cone} T(K)$. The set $T(K)$ is compact and, according to condition (1.44), it does not contain the origin. Therefore, $T(A)$ is locally compact. □

1.2 Duality in Linear Normed Spaces

We now briefly survey the basic concepts and results related to dual pairs of linear topological spaces and weak topologies on linear normed spaces.

1.2.1 The Dual Systems of Linear Spaces

Two linear spaces X and Y over the same scalar field Γ define a dual system if a fixed bilinear functional on their product is given:

$$(\cdot, \cdot) : X \times Y \to \Gamma. \tag{1.45}$$

The bilinear functional is sometimes omitted. The dual system is called *separated* if the following two properties hold:

(i) *For every* $x \in X \setminus \{0\}$ *there is* $y \in Y$ *such that* $(x, y) \neq 0$
(ii) *For every* $y \in Y \setminus \{0\}$ *there is* $x \in X$ *such that* $(x, y) \neq 0$.

In other words, X separates points in Y and Y separates points in X.

In the following, we consider only *separated dual systems*.

For each $x \in X$, we define the application $f_x : Y \to \Gamma$ by

$$f_x(y) = (x, y), \quad \forall y \in Y. \tag{1.46}$$

We observe that f_x is a linear functional on Y and the mapping

$$x \to f_x, \quad \forall x \in X, \tag{1.47}$$

is linear and injective, as can be seen from condition (i). Hence, the correspondence (1.47) is an embedding. Thus, the elements of X can be identified with the linear functionals on Y. In a similar way, the elements of Y can be considered as linear functionals of X, identifying an element $y \in Y$ with $g_y : X \to \Gamma$, defined by

$$g_y(x) = (x, y), \quad \forall x \in X. \tag{1.48}$$

Therefore, each dual system of linear spaces defines a mapping from either of the two linear spaces into the space of linear functionals on the other.

We set

$$p_y(x) = \big|(x, y)\big| = \big|g_y(x)\big|, \quad \forall x \in X, \tag{1.49}$$

$$q_x(y) = \big|(x, y)\big| = \big|f_x(y)\big|, \quad \forall y \in Y, \tag{1.50}$$

and we observe that $\mathscr{P} = \{p_y; y \in Y\}$ is a family of seminorms on X and $\mathscr{Q} = \{q_x; x \in X\}$ is a family of seminorms on Y. The locally convex topology defined by $\mathscr{P}$ on X is called the *weak topology* or Y*-topology of* X induced by the duality (X, Y), and we denote it by $\sigma(X, Y)$. Similarly, the *weak topology* or X*-topology of* Y, denoted by $\sigma(Y, X)$, is the locally convex topology on Y generated by $\mathscr{Q}$. Clearly, the roles of X and Y are interchangeable here, since there is a natural duality between Y and X which determines a dual system (Y, X). Thus, it is sufficient to establish the properties only for the linear space X. According to the well-known results concerning the locally convex topologies generated by families of seminorms, we immediately obtain the following result.

Proposition 1.65

(i) *$\sigma(X,Y)$ is the weakest topology on X which makes the linear functionals g_y, defined by (1.48), continuous for any $y \in Y$.*
(ii) *$\sigma(X,Y)$ is separated.*
(iii) *The family of all the sets of the form*

$$V_{y_1,y_2,\dots,y_n;\varepsilon}(x) = \left\{u \in X;\ \left|(u-x;y_i)\right| < \varepsilon,\ i = 1,2,\dots,n\right\}, \tag{1.51}$$

where $n \in N^*$, $y_1, y_2, \dots, y_n \in Y$, $\varepsilon > 0$, is a fundamental neighborhood system of the element $x \in X$ for $\sigma(X,Y)$.
(iv) *A sequence $\{x_n\} \subset X$ is $\sigma(X,Y)$-convergent to $x_0 \in X$ if and only if $\{(x_n,y)\}$ converges to (x_0,y) in Γ, for each $y \in Y$.*
(v) *If Z is a locally convex space with the topology generated by a family $\mathscr{P}$ of seminorms, then a linear operator $T : X \to Z$ is $\sigma(X,Y)$-continuous if and only if for any $p \in \mathscr{P}$ there are $k_p > 0$ and $y_1, y_2, \dots, y_n \in Y$ such that*

$$p(Tx) \le k_p \max_{1 \le i \le n} \left|(x,y_i)\right|, \quad \forall x \in X. \tag{1.52}$$

From assertion (i), we find as a result that g_y is a $\sigma(X,Y)$-continuous linear functional on X for every $y \in Y$. It is natural to investigate if the set of linear functionals of this type coincides with the dual of locally convex space $(X, \sigma(X,Y))$. The answer is affirmative. Thus, in view of embedding $y \to g_y$, it is possible to regard Y as the dual of X endowed with weak topology.

Definition 1.66 A linear topology τ on X is called *compatible* with the duality (X,Y) if $(X,\tau)^* = Y$. Similarly, a linear topology μ on Y is called *compatible* with the duality (X,Y) if $(Y,\mu)^* = X$.

Hence, $\sigma(X,Y)$ and $\sigma(Y,X)$ are the weakest compatible topologies.

Other properties of the weak topologies are consequences of the fact that each of these topologies may be considered as a relativized product topology. In this connection, let us recall some basic results on product topology.

Let $\{X_\alpha; \alpha \in A\}$ be a family of topological spaces. Consider their *product space*

$$X = \prod_{\alpha \in A} X_\alpha = \left\{x : A \to \bigcup_{\alpha \in A} X_\alpha;\ x(\alpha) \in X_\alpha,\ \forall \alpha \in A\right\}. \tag{1.53}$$

Write $x(\alpha) = x_\alpha$, $\forall \alpha \in A$, and $x = (x_\alpha)_{\alpha \in A}$. For each $\alpha \in A$, consider the *projections* $P_\alpha : X \to X_\alpha$ defined by $P_\alpha x = x_\alpha$, $\forall x \in X$. The space X endowed with the weakest topology which makes each projection continuous is called the *topological product space* of the topological spaces X_α, $\alpha \in A$. Thus, a basis for the product topology is given by the sets of the form $\prod_{\alpha \in A} D_\alpha$, where D_α is an open set in X_α, $\forall \alpha \in A$, and $D_\alpha = X_\alpha$, $\forall \alpha \in A \setminus F$, with F a finite subset of A. Also, a fundamental neighborhood system of an element $x = (x_\alpha)_{\alpha \in A} \in X$ is given by the sets having the form

$$V_{F,\{V_\alpha;\alpha\in F\}}(x) = \{u = (u_\alpha) \in X;\ u_\alpha \in V_\alpha(x_\alpha),\ \forall\alpha \in F\}, \tag{1.54}$$

where F is a finite subset of A and for each α, $V_\alpha(x_\alpha)$ runs through a fundamental neighborhood system of $x_\alpha \in X_\alpha$. In particular, a topological product space is separated if and only if each factor space is separated. A remarkable result, with various consequences, is the well-known *Tychonoff Theorem*.

Theorem 1.67 (Tychonoff) *A topological product is compact if and only if each coordinate space is compact.*

Corollary 1.68 *A subset $M \subset \prod_{\alpha\in A} X_\alpha$ is compact if and only if it is closed and $P_\alpha(M)$ is relatively compact in X_α for each $\alpha \in A$.*

Proof It suffices to observe that $M \subset \prod_{\alpha\in A} P_\alpha(M)$. □

Proposition 1.69 *$\sigma(X,Y)$ coincides with the topology induced on X by the topological product Γ^Y.*

Proof We recall that Γ^Y is the set of all applications defined on Y with values in Γ. Thus, Γ^Y may be regarded as a topological product. On the other hand, using the embedding (1.47), the linear space X may be considered as a subspace of Γ^Y identifying $x \in X$ with the functional $f_x \in \Gamma^Y$. According to our convention, we identify x with $(x_y)_{y\in Y}$, where $x_y = (x,y)$. Using formula (1.54), a neighborhood at $x = (x_y)_{y\in Y}$ in the topological product Γ^Y has the form $V = \{u = (u_y)_{y\in Y} \in \Gamma^Y;\ |u_{y_i} - x_{y_i}| < \varepsilon, \forall i = 1,2,\dots,n\}$, where $\varepsilon > 0$ and $\{y_1, y_2, \dots, y_n\}$ is a finite subset of Y. It is clear that $V \cap X$ is a neighborhood of the form (1.51) since, in this case, the element $u = (u_y)_{y\in Y} \in \Gamma^Y$ corresponds to an element $u \in X$ such that $u_y = (u,y)$, $\forall y \in Y$. □

Corollary 1.70 *A set $M \subset X$ is $\sigma(X,Y)$-compact if and only if it is closed in Γ^Y and if, for every $y \in Y$, there exists $k_y > 0$ such that $|(x,y)| \le k_y$, $\forall x \in M$.*

Proof Apply Corollary 1.68 and recall that in Γ a set is relatively compact if and only if it is bounded. □

1.2.2 Weak Topologies on Linear Normed Spaces

Let X be a real normed space and let X^* be its dual, that is, the space of all real continuous linear functionals on X. We recall that $X^* = L(X,\Gamma)$ is a Banach space. As is well known, the norm of an element $x^* \in X^*$ is defined by

$$\|x^*\| = \sup_{\|x\|\le 1} |x^*(x)|. \tag{1.55}$$

There is a natural duality between X and X^* determined by the bilinear functional $(\cdot,\cdot) : X \times X^* \to \Gamma$, defined by

$$(x, x^*) = x^*(x), \quad \forall x \in X,\ x^* \in X^*. \tag{1.56}$$

In the preceding section, we have generated the weak topologies $\sigma(X, X^*)$ and $\sigma(X^*, X)$. The properties of these topologies which do not depend on topological structures on X and X^* are similar. However, different properties still exist because the two normed spaces X and X^* do not play symmetric roles; in general, X is not the dual of X^* as linear normed space. It can be observed that the previous property characterizes a special class of normed spaces called *reflexive*. In fact, X^* as a normed space generates the weak topologies $\sigma(X^*, X^{**})$ and $\sigma(X^{**}, X^*)$, where X^{**} is the dual space of X^*, called the *bidual* of X. In general, the topologies $\sigma(X^*, X)$ and $\sigma(X^*, X^{**})$ are different.

Denote $\sigma(X, X^*) = w$ and $\sigma(X^*, X) = w^*$, preserving the name of weak topology only for w. The topology w^* will be called the *weak-star topology* on X^*, being in general different from the *weak topology* on X^*, which is $\sigma(X^*, X^{**})$.

In contrast to these topologies, the initial topologies on X and X^* generated by the usual norms will be called the *strong topology* on X and the *strong topology* on X^*, respectively.

We denote by $\to$, $\xrightarrow{w}$ and $\xrightarrow{w^*}$ the *strong convergence* and the *weak*, *weak-star convergence* in X and X^*, respectively.

As follows from Proposition 1.65(iii), a neighborhood base at x_0 for the topology w on X is formed by the sets

$$V_{x_1^*, x_2^*, \ldots, x_n^*; \varepsilon}(x_0) = \left\{ x \in X;\ \left| x_i^*(x - x_0) \right| < \varepsilon,\ \forall i = 1, 2, \ldots, n \right\}, \tag{1.57}$$

where $\varepsilon > 0$, $n \in \mathbb{N}^*$ and $x_1^*, x_2^*, \ldots, x_n^* \in X^*$.

Similarly, the sets of the form

$$V_{x_1, x_2, \ldots, x_n; \varepsilon}(x_0^*) = \left\{ x^* \in X^*;\ \left| (x^* - x_0^*)(x_i) \right| < \varepsilon,\ \forall i = 1, 2, \ldots, n \right\}, \tag{1.58}$$

where $\varepsilon > 0,\ n \in \mathbb{N}^*,\ x_1, x_2, \ldots, x_n \in X$, constitute a neighborhood base at x_0^* for the topology w^* on X^*.

Proposition 1.71 below sums up some elementary properties of the weak (weak-star) topology.

Proposition 1.71 *If X is a linear normed space, then*

(i) *$w(w^*)$ is a separated locally convex topology on $X(X^*)$.*
(ii) *$w(w^*)$ is the coarsest linear topology on $X(X^*)$ for which $X^*(X)$ is the dual space.*
(iii) *The original norm topology on $X(X^*)$ is always finer than the weak topology w (weak-star topology w^*). The equality holds if and only if X is finite-dimensional.*

Proof According to Corollary 1.53, the sets of seminorms which generate the topologies w and w^* are sufficient, that is, satisfy property (1.6). Hence, assertion (i) holds. To obtain assertion (ii), we make use of Proposition 1.65(i), since the elements of $X^*(X)$ may be viewed as linear functionals on $X(X^*)$. To prove assertion (iii), we observe that the norm is weakly continuous if w coincides with the norm topology on X. According to Proposition 1.65(v), there are $x_1^*, x_2^*, \ldots, x_n^* \in X^*$ and $k > 0$ such that

$$\|x\| \leq k \max_{1 \leq i \leq n} \left| x_i^*(x) \right|, \quad \forall x \in X.$$

Thus, we have

$$\bigcap_{i=1}^{n} \ker x_i^* = \{0\} \subset \ker x^*, \quad \forall x^* \in X^*,$$

which implies that x^* is a linear combination of x_i^*, $i = 1, 2, \ldots, n$ (cf. Theorem 1.30), that is, X^* is a finite-dimensional space. However, this holds if and only if X is finite-dimensional. Similarly, if w^* coincides with the norm topology of X^*, it follows that X is finite-dimensional.

In the following, we prove certain properties of w which, in general, are not true for w^*. □

Proposition 1.72 *A linear functional of normed space X is continuous if and only if it is weakly continuous, that is, $(X, w)^* = X^*$.*

Proof We apply Proposition 1.71(ii). □

Proposition 1.73 *A convex set is closed if and only if it is weakly closed.*

Proof Using Proposition 1.71(iii), it follows that any w-closed set is closed. On the other hand, from Theorem 1.48 we find as a result that a closed convex set is an intersection of closed half-spaces. However, it is clear that a closed half-space is also w-closed (Proposition 1.72 and Theorem 1.2). □

Corollary 1.74 *The closure of a convex set coincides with its weak closure.*

Corollary 1.75 *The closed unit ball of a normed space is weakly closed.*

From these two results, we obtain the following useful statements.

Corollary 1.76 *If $x_n \xrightarrow{w} x_0$, then there exists a sequence of convex combinations of $\{x_n\}$ which converges strongly to x_0. Moreover, we have*

$$\|x_0\| \leq \liminf_{n \to \infty} \|x_n\|. \tag{1.59}$$

Corollary 1.77 *If a sequence is weakly convergent and norm fundamental, then it is strongly convergent.*

Remark 1.78 Weak topologies can be similarly defined in a separated locally convex space (X, τ). It is obvious that $w \leq \tau$ and τ is a compatible topology sequence, and all nonmetric results are also true if they only depend on X^*. For instance, Propositions 1.72, 1.73, and Corollary 1.74 hold in every locally convex space.

Now, we consider the special case of Hilbert spaces. In view of the Riesz Theorem 1.9, a sequence $\{x_n\}$ in a Hilbert space X is weakly convergent to $x_0 \in X$ if and only if

$$\lim_{n\to\infty} \langle x_n, a\rangle = \langle x_0, a\rangle, \tag{1.60}$$

for every $a \in X$. (See Proposition 1.65(iv).)

Proposition 1.79 *If* $x_n \xrightarrow{w} x_0$, *where* $\{x_n\}$ *is a sequence in a Hilbert space, and* $\|x_n\| \to \|x_0\|$, *then* $x_n \to x_0$.

Proof We use the identity

$$\|x_n - x_0\|^2 = \|x_n\|^2 + \|x_0\|^2 - 2\,\mathrm{Re}\langle x_n, x_0\rangle.$$ □

Proposition 1.80 *In a normed space* X *the bounded and weakly bounded sets are the same. If* X *is a Banach space, then the bounded and weak-star bounded sets of* X^* *are the same.*

Proof Let $M \subset X$ be a weakly bounded set. Define $f_x : X^* \to \mathbb{R}$ by $f_x(x^*) = x^*(x)$, $x^* \in X^*$, for every $x \in M$. We observe that $\{f_x;\ x \in M\}$ is a pointwise bounded family of continuous linear functionals on the Banach space X^*. According to the *principle of uniform boundedness* (see Theorem 1.5), it follows that this family is uniformly bounded, that is, $\|f_x\| \leq k$, $\forall x \in M$, for some $k > 0$. By virtue of relations (1.10) and (1.36), we obtain

$$\|f_x\| = \sup_{\|x^*\|\leq 1} f_x(x^*) = \sup_{\|x^*\|\leq 1} x^*(x) = \|x\|.$$

Therefore, $\|x\| \leq k$, $\forall x \in M$, which shows that M is bounded. On the other hand, it is obvious that any bounded set is weakly bounded (see Proposition 1.71(iii)). The second part of the assertion follows in a similar way.

In particular, Propositions 1.72, 1.73 and 1.80 show that, for the weak topology and the norm topology, the continuity of the linear functionals, the boundedness of the sets, and the closure of the convex sets are identical. □

Theorem 1.81 *The closed unit ball of the dual of a normed space is weak-star compact.*

Proof Let X be a normed space. In view of Corollary 1.70, we need only to prove that the closed unit ball of X^* is closed in Γ^X. Let $f_0 : X \to \Gamma$ be an adherent point of $S^* = \{x^* \in X^*;\ \|x^*\| \le 1\}$ in the product topology of Γ^X. First, we prove that f_0 is linear. For $\varepsilon > 0$, $x, y \in X$ and $a \in \Gamma$, we consider the neighborhood $V(f_0)$ having the form (1.54) with $F = \{x, y, ax + y\}$. By hypothesis, there exists $x_0^* \in V(f_0) \cap S^*$, that is, $|x_0^*(x) - f_0(x)| < \varepsilon$, $|x_0^*(y) - f_0(y)| < \varepsilon$ and $|x_0^*(ax+y) - f_0(ax + y)| < \varepsilon$. From these relations, we obtain $|f_0(ax + y) - af_0(x) - f_0(y)| \le \varepsilon(2 + |a|)$, $\forall \varepsilon > 0$. Hence, f_0 is a linear functional on X. Also, we have

$$\left|f_0(x)\right| \le \left|x_0^*(x)\right| + \left|x_0^*(x) - f_0(x)\right| \le \|x\| + \varepsilon, \quad \forall \varepsilon > 0,$$

which implies $\|f_0\| \le 1$; hence $f_0 \in S^*$, that is, S^* is closed in Γ^X. □

Corollary 1.82 *The closed unit ball of the dual of a normed space is weak-star closed.*

Corollary 1.83 *If $\{x_n^*\} \subset X^*$ is weak-star convergent to $x_0^* \in X^*$, then*

$$\|x_0^*\| \le \liminf_{n\to\infty} \|x_n^*\|. \tag{1.61}$$

If, in addition, $\{x_n^*\}$ is norm fundamental, it is strongly convergent to x_0^*.

Now, we prove some properties of linear operators related to the duality theory.

Proposition 1.84 *Any linear continuous operator is weakly continuous.*

Proof Let $T : X \to Y$ be a linear continuous operator, where X, Y are two linear normed spaces. It is clear that $y^* \circ T$ is a linear continuous functional on X for every $y^* \in Y^*$. According to Proposition 1.72, $y^* \circ T$ is a weakly continuous functional and, therefore, T is weakly continuous (Proposition 1.65(v)).

Let T be a linear operator defined on a linear subspace $D(T)$ of X, with values in Y. We observe that $y^* \circ T$ is a linear functional on $D(T)$ for every $y^* \in Y^*$. The problem is to find the conditions which should ensure that there exists a unique element $x^* \in X^*$ such that $x^*|_{D(T)} = y^* \circ T$. First, it is necessary that $y^* \circ T$ is bounded on $D(T)$. Moreover, $y^* \circ T$ should admit a unique extension on the whole space X, that is, $D(T)$ should be dense in X. Furthermore, the linear operator T must be densely defined ($\overline{D(T)} = X$). In this case, we denote by $D(T^*)$ the set of all elements $y^* \in Y^*$, which have the property that $y^* \circ T$ is bounded on $D(T)$, that is,

$$\begin{aligned} D(T^*) = \big\{y^* \in Y;\ &\text{there is } k > 0 \text{ such that } (x, y^* \circ T) \le k\|x\|, \\ &\text{for all } x \in D(T)\big\}. \end{aligned} \tag{1.62}$$

Thus, for every $y^* \in D(T^*)$, there is a unique element $x^* \in X^*$ such that $(x, x^*) = (x, y^* \circ T)$ for any $x \in D(T)$.

Define the operator $T^* : D(T^*) \to X^*$ by $T^* y^* = x^*$, called the adjoint of T. In other words, T^* is well defined by the relation

$$(Tx, y^*) = (x, T^* y^*), \quad \forall x \in D(T) \text{ and } y^* \in D(T^*). \tag{1.63}$$

□

Proposition 1.85 *The adjoint of a densely defined linear operator is a closed linear operator.*

Proof We recall that an operator is said to be *closed* if its graph is closed. Let $T : D(T) \subset X \to Y$, $\overline{D(T)} = X$, be a linear operator and let

$$G(T^*) = \big\{(y^*, T^* y^*);\ y^* \in D(T^*)\big\} \subset Y^* \times X^* \tag{1.64}$$

be the graph of T^*. Clearly, T^* is linear. If $(t_0^*, x_0^*) \in \overline{G(T^*)}$, then there exist $\{y_n^*\} \subset D(T^*)$ such that $y_n^* \to y_0^*$ and $T^* y_n^* \to x_0^*$. On the other hand, by relation (1.63) we have $(Tx, y_n^*) = (x, T^* y_n^*)$, $\forall x \in D(T)$. For $n \to \infty$, we obtain $(Tx, y_0^*) = (x, x_0^*)$; hence $y_0^* \in D(T^*)$. Moreover, $(Tx, y_0^*) = (x, T^* y_0^*)$, $\forall x \in D(T)$, that is, $x_0^* = T^* y_0^*$, because $\overline{D(T)} = X$. Therefore, $(y_0^*, x_0^*) = (y_0^*, T^* y_0^*) \in G(T^*)$. Hence, $G(T^*)$ is closed in $Y^* \times X^*$, thus proving Proposition 1.85. □

Remark 1.86 If $T \in L(X, Y)$, that is, $D(T) = X$ and T is a continuous linear operator, then $D(T^*) = Y^*$ and $T^* \in L(Y^*, X^*)$. Also, $T^{**}|_X = T$.

In the special case of a Hilbert space, the adjoint T^* defined here differs from the adjoint T' defined in Sect. 1.1.1. But one easily verifies that

$$T' = J_1^{-1} \circ T^* \circ J_2, \tag{1.65}$$

where J_1 and J_2 are canonical isomorphisms of the Hilbert spaces X and Y generated by (1.88). It should be observed that $T' \in L(Y, X)$, while $T^* \in L(Y^*, X^*)$.

1.2.3 Reflexive Banach Spaces

It is clear that, for each $x \in X$, the functional $f_x : X \to \Gamma$, defined by

$$f_x(x^*) = x^*(x), \quad \forall x \in X, \tag{1.66}$$

is linear, and the weak-star topology on X^* is the coarsest topology on X^* for which f_x is continuous for every $x \in X$. Also, we observe that f_x is strongly continuous on X^* since

$$\big|f_x(x^*)\big| = \big|x^*(x)\big| \le \|x^*\|\,\|x\|, \quad \forall x^* \in X^*.$$

Furthermore, f_x is an element of the dual space of X^*, that is, $f_x \in X^{**}$. The mapping $\Phi : X \to X^{**}$, defined by $\Phi(x) = f_x$, $\forall x \in X$, is a linear isometric injection (cf. equality (1.36)). Hence, Φ is an imbedding of linear normed spaces, called the *natural imbedding of X into* X^{**}.

Definition 1.87 A linear normed space is called *reflexive* if the natural imbedding is surjective, that is, it can be identified under Φ with its bidual.

However, a linear normed space X is reflexive if and only if for every continuous linear functional F on X^* there is an element $x \in X$ such that

$$F(x^*) = x^*(x), \quad \forall x^* \in X^*.$$

Proposition 1.88 *A Hilbert space is reflexive.*

Proof If J is a canonical isomorphism of the Hilbert space X defined by (1.17), that is, $(x, x^*) = \langle x, J^{-1}x^*\rangle$, $x \in X$, $x^* \in X^*$, according to (1.17) and (1.18), we have

$$(x^*, \Phi x) = (x, x^*) = \left\langle x, J^{-1}x^*\right\rangle = \langle x^*, Jx\rangle = \left(x^*, \tilde{J}Jx\right), \quad \forall x \in X,$$

where $\tilde{J}$ is the canonical isomorphism of X^*. Therefore, $\Phi = \tilde{J}J$. Hence, Φ is onto X^{**}, since J and $\tilde{J}$ are onto X^* and X^{**}, respectively. □

Proposition 1.89 *A linear normed space X is reflexive if and only if one of the following three equivalent statements is satisfied*:

(i) *Each continuous linear functional on* X^* *is weak-star continuous.*
(ii) *The norm topology on* X^* *is compatible with the natural duality between X and* X^*.
(iii) *Each closed convex set of* X^* *is weak-star closed.*

Proposition 1.90 *A Banach space is reflexive if and only if its dual is reflexive.*

Proof If X is reflexive, it is clear that X^* is also reflexive. Conversely, let X^* be reflexive and suppose that X is not reflexive. Thus, by natural imbedding, X can be considered as being a proper closed linear subspace of X^{**} because X is complete.

According to Theorem 1.52, there is a nonidentically zero continuous linear functional $x_0''' \in X^{***}$ which is null on X, that is,

$$\left(f_x, x_0'''\right) = 0, \quad \forall x \in X.$$

On the other hand, since X^* is reflexive, there is $x_0' \in X^*$ such that

$$\left(x'', x_0'''\right) = \left(x_0', x''\right), \quad \forall x'' \in X^{**}.$$

From the latter two relations, we obtain $(x_0', f_x) = 0$, that is, $(x, x_0') = 0$, $\forall x \in X$. Hence, x_0' is the trivial functional on X. By natural imbedding, it follows that x_0'''

is also the null element of X^{***}. The contradiction we arrived at concludes the proof. $\square$

The condition that X is a Banach space cannot be dropped in Proposition 1.90. This condition is quite natural because a reflexive normed space is always complete.

Theorem 1.91 *The closed unit ball of a linear normed space is dense in the closed unit ball of the bidual in the weak-star topology $\sigma(X^{**}, X^*)$ (under natural imbedding).*

Proof For simplicity, we suppose that X is a real linear normed space. Denote $S = \{x \in X; \|x\| \le 1\}$ and $S^{**} = \{x'' \in X^{**}; \|x''\| \le 1\}$. It is clear that $\Phi(S) \subset S^{**}$. From Theorem 1.81 (for X^*), it follows that S^{**} is a $\sigma(X^{**}, X^*)$-compact set; hence it is a $\sigma(X^{**}, X^*)$-closed set. Let $\widetilde{\Phi(S)}$ be the closure of $\Phi(S)$ with respect to the topology $\sigma(X^{**}, X^*)$, hence $\widetilde{\Phi(S)} \subset S^{**}$. Suppose that there is an element $x_0'' \in S^{**} \setminus \widetilde{\Phi(S)}$. By Corollary 1.45, we find a $\sigma(X^{**}, X^*)$-continuous linear functional on X^{**}, that is, an element $x_0' \in X^*$ (cf. Proposition 1.71(ii)) such that

$$\sup\{x''(x_0');\ x'' \in \widetilde{\Phi(S)}\} < x_0''(x_0'),$$

which implies

$$\sup\{x_0'(x);\ x \in S\} < x_0''(x_0') \le \|x_0''\|\,\|x_0'\| \le \|x_0'\|.$$

This contradicts the norm definition of x_0', given by relation (1.55); hence $\widetilde{\Phi(S)} = X^{**}$. $\square$

Corollary 1.92 *Each linear normed space X is $\sigma(X^{**}, X^*)$-dense in its bidual. In particular, X is reflexive if and only if it is $\sigma(X^{**}, X^*)$-closed (equivalently, if and only if the closed unit ball is weak-star closed in its bidual).*

Remark 1.93 The topology $\sigma(X^{**}, X^*)$ is the weak-star topology of the bidual, and it coincides with the weak topology $\sigma(X^{**}, X^{***})$ of the bidual if and only if X^* is reflexive.

Theorem 1.94 *A linear normed space is reflexive if and only if its closed unit ball is weakly compact.*

Proof From Theorem 1.81 it follows that S^{**} is $\sigma(X^{**}, X^*)$-compact. If X is reflexive, $\Phi(S)$ is $\sigma(X^{**}, X^*)$-compact since $\Phi(S) = S^{**}$, that is, S is weakly compact. Conversely, let X be a normed space with its unit closed ball weakly compact. Since, by natural imbedding, the relativization of the topology $\sigma(X^{**}, X^*)$ with respect to X is $\sigma(X, X^*)$, it follows that $\Phi(S)$ is also $\sigma(X^{**}, X^*)$-closed. According to Theorem 1.91, we obtain $S^{**} = \Phi(S)$ and, therefore, X is reflexive. $\square$

Corollary 1.95 *A linear normed space is reflexive if and only if each bounded subset is a relatively weakly compact set.*

We observe that weak compactness may be replaced by *weak-sequential compactness or even by weak-countable compactness* (see, for instance, [9, 16]). Thus, we obtain the well-known *Eberlein Theorem*.

Theorem 1.96 (Eberlein) *A linear normed space is reflexive if and only if its closed unit ball is sequentially weakly compact.*

Corollary 1.97 *A linear normed space is reflexive if and only if each bounded sequence contains a weakly convergent subsequence or, equivalently, if and only if each bounded set is a weakly sequentially relatively compact set.*

Corollary 1.98 *A linear normed space is reflexive if and only if each separable and closed subspace is reflexive.*

1.2.4 Duality Mapping

Let X be a real linear normed space and let X^* be its dual.

Definition 1.99 The operator $F : X \to \mathscr{P}(X^*)$ defined by

$$Fx = \left\{x^* \in X^*;\ (x, x^*) = \|x\|^2 = \|x^*\|^2\right\} \tag{1.67}$$

is called the *duality mapping* of X.

If X is a real Hilbert space, it is clear that the duality mapping is even the canonical isomorphism given by the Riesz Theorem.

From Corollary 1.53, it follows that $Fx \neq \emptyset$, $\forall x \in X$, hence F is well defined. Also, it is clear that Fx is a convex set of X^* for every $x \in X$. Moreover, Fx is bounded and w^*-closed, thus it is w^*-compact (cf. Corollary 1.70).

According to relation (1.67), we observe that $x^* \in Fx$, with $x^* \neq 0$, if and only if the element x maximizes x^* on the closed ball $\overline{S}(0, \|x\|)$, or equivalently if and only if $x^*(u) = \|x\|^2$, $u \in X$, is the equation of a closed supporting hyperplane to $\overline{S}(0, \|x\|)$. This condition can be translated to the closed unit ball $\overline{S}(0, 1)$ by replacing x by $x\|x\|^{-1}$. Thus, it is natural to consider the linear normed spaces which have the following property: *at most one closed support hyperplane passes through every boundary point of the closed unit ball.* A linear normed space which has this property is called *smooth*.

However, from Theorem 1.38 there exists a closed supporting hyperplane through each boundary point of the closed unit ball. Consequently, a linear normed space is smooth if and only if there is exactly one supporting hyperplane through each boundary point of the closed unit ball. Thus, a supporting hyperplane is at the same time a *tangent hyperplane*.

Remark 1.100 It is clear that the duality mapping is single-valued if and only if the normed space is smooth.

A property which is the dual of that given above is the following. *Any nonidentically zero continuous linear functional takes a maximum value on the closed unit ball at most at one point.* If a linear normed space enjoys this property, it is called a *strictly convex space*. In terms of supporting hyperplanes, this property may be expressed as: *distinct boundary points of the closed unit ball have distinct supporting hyperplanes.*

Theorem 1.101 *If X^* is smooth* (*strictly convex*), *then X is strictly convex* (*smooth*).

Proof Let Σ and Σ^* be the boundaries of the closed unit balls of X and X^*. If X is not strictly convex there exist $x_0^* \in \Sigma^*$ and $x_1, x_2 \in \Sigma$, with $x_1 \neq x_2$, such that $(x_1, x_0^*) = (x_2, x_0^*) = 1$. Thus, two distinct supporting hyperplanes pass through $x_0^* \in \Sigma^* : (x_1, u^*) = 1$, $(x_2, u^*) = 1$, $u^* \in X^*$. Therefore, X^* is not smooth. If X is not smooth, there exist $x_0 \in \Sigma$ and $x_1^*, x_2^* \in \Sigma^*$, with $x_1^* \neq x_2^*$, such that $x_1^*(x_0) = x_2^*(x_0) = 1$; that is, x_0 determines a continuous linear functional on X^* which takes the maximum value on the closed unit ball of X in two distinct points x_1^*, x_2^*. Hence, X^* is not strictly convex. □

Complete duality clearly holds in the reflexive case, namely, we have the following.

Corollary 1.102 *A reflexive normed space is smooth* (*strictly convex*) *if and only if its dual is strictly convex* (*smooth*).

Proposition 1.103 *A linear normed space is strictly convex if and only if one of the following equivalent properties holds*:

(i) *If* $\|x + y\| = \|x\| + \|y\|$ *and* $x \neq 0$, *there is* $t \geq 0$ *such that* $y = tx$.
(ii) *If* $\|x\| = \|y\| = 1$ *and* $x \neq y$, *then* $\|\lambda x + (1 - \lambda)y\| < 1$ *for all* $\lambda \in]0, 1[$.
(iii) *If* $\|x\| = \|y\| = 1$ *and* $x \neq y$, *then* $\|\frac{1}{2}(x + y)\| < 1$.
(iv) *The function* $x \to \|x\|^2$, $x \in X$, *is strictly convex.*

Proof Let X be strictly convex and let $x, y \in X \setminus \{0\}$ be such that $\|x + y\| = \|x\| + \|y\|$. From Corollary 1.53 there exists x^* such that $(x + y, x^*) = \|x + y\|$, $\|x^*\| = 1$. Since $(x, x^*) \leq \|x\|$, $(y, x^*) \leq \|y\|$, we must have $(x, x^*) = \|x\|$ and $(y, x^*) = \|y\|$, that is, $(\frac{x}{\|x\|}, x^*) = (\frac{y}{\|y\|}, x^*) = 1$. Because X is strictly convex, it follows that $\frac{x}{\|x\|} = \frac{y}{\|y\|}$, hence property (i) holds with $t = \frac{\|y\|}{\|x\|}$.

To prove that (i)→(ii), we assume by contradiction that there exists $x \neq y$ such that $\|x\| = \|y\| = 1$ and $\|\lambda x + (1 - \lambda)y\| = 1$, where $\lambda \in]0, 1[$. Therefore, we have $\|\lambda x + (1 - \lambda)y\| = \|\lambda x\| + \|(1 - \lambda)y\|$. According to property (i), there exists $t \geq 0$ such that $\lambda x = t(1 - \lambda)y$. Since $\|x\| = \|y\|$, we obtain $\lambda = t(1 - \lambda)$ and so $x = y$, which is a contradiction.

The implications (ii)→(iii) and (iv)→(ii) are obvious.

Now, we assume that X is not strictly convex. Therefore, there exist $x_0^* \in X^*$ and $x_1, x_2 \in X$ with $\|x_0^*\| = 1$, $\|x_1\| = \|x_2\| = 1$, $x_1 \neq x_2$, such that $(x_1, x_0^*) = (x_2, x_0^*) = 1$, hence $(\frac{1}{2}(x_1 + x_2), x_0^*) = 1$. Thus,

$$\left\| \frac{1}{2}(x_1 + x_2) \right\| = \sup_{\|x^*\| \leq 1} \left(\frac{1}{2}(x_1 + x_2), x^* \right) \geq \left(\frac{1}{2}(x_1 + x_2), x_0^* \right) = 1,$$

contradicting property (iii). Hence, property (iii) implies the strict convexity of X. Now, from the equality

$$\lambda \|x\|^2 + (1 - \lambda)\|y\|^2 = \big(\lambda \|x\| + (1 - \lambda)\|y\|\big)^2 + \lambda(1 - \lambda)\big(\|x\| - \|y\|\big)^2,$$

it follows that

$$\big\|\lambda x + (1 - \lambda) y\big\|^2 \leq \big(\lambda \|x\| + (1 - \lambda)\|y\|\big)^2 < \lambda \|x\|^2 + (1 - \lambda)\|y\|^2,$$

for all $x, y \in X$ with $\|x\| \neq \|y\|$ and $\lambda \in]0, 1[$. If $\|x\| = \|y\|$, we obtain the strict convexity of the function $x \to \|x\|^2$, $x \in X$, from (ii). Thus, the implication (ii)→(iv) is established and the proof is complete. □

Corollary 1.104 *A normed space is strictly convex if and only if each two-dimensional linear subspace is strictly convex.*

This problem which arises naturally is whether equivalent norms exist which are simultaneously strictly convex and smooth. A remarkable result of this type is the *renorming theorem* due to Asplund [1, 2], which will be frequently used in the following work.

Theorem 1.105 (Asplund) *Let X be a reflexive Banach space. Then there exists an equivalent norm on X, such that, under this new norm, X and X^* are strictly convex, that is, X and X^* are simultaneously smooth and strictly convex.*

The proof is omitted since it involves some special considerations. In the following, other special properties of the duality mapping will be examined.

Theorem 1.106 *If X is smooth, then the duality mapping is continuous from X with strong topology into X^* with weak-star topology, that is, F is demicontinuous.*

Proof It is clear that, under our hypothesis, the duality mapping is single-valued (see Remark 1.100).

Let $\{x_n\}$ be any sequence of X convergent to $x_0 \in X$. The sequence $\{Fx_n\}$ is bounded in X^* and hence, using Theorem 1.81, it has a w^*-adherent point x_0^* and $\|x_0^*\| \leq \|x_0\|$. Thus, for every $\varepsilon > 0$ and $n \in \mathbb{N}$, there exists $k_n \in \mathbb{N}$ with $k_n \geq n$, such that $|(x_0^* - Fx_n, x_0)| < \varepsilon$. But we have

$$\begin{aligned}\left|(x_0^*, x_0) - \|x_{k_n}\|^2\right| &= \left|(x_0^*, x_0) - (Fx_{k_n}, x_{k_n})\right| \\ &\le \left|(x_0^* - Fx_{k_n}, x_0)\right| + \left|(Fx_{k_n}, x_{k_n} - x_0)\right| \\ &< \varepsilon + \|Fx_{k_n}\| \, \|x_{k_n} - x_0\|,\end{aligned}$$

from which, for $n \to \infty$, we obtain $(x_0^*, x_0) = \|x_0\|^2$. This also implies $\|x_0\| \le \|x_0^*\|$. Hence, $\|x_0^*\| = \|x_0\|$, that is, $x_0^* = Fx_0$. Thus, the sequence $\{Fx_n\}$, which clearly is w^*-compact, has a unique w^*-adherent point x_0^*. Therefore, $\{Fx_n\}$ is w^*-convergent to Fx_0, as claimed. □

A stronger property than strict convexity is uniform convexity.

Definition 1.107 A linear normed space is called *uniformly convex* if, for each $\varepsilon \in]0, 2[$, there exists a $\delta(\varepsilon) > 0$, for which $\|x\| \le 1$, $\|y\| \le 1$ and $\|x - y\| \ge \varepsilon$ imply

$$\left\| \frac{1}{2}(x + y) \right\| \le 1 - \delta(\varepsilon). \tag{1.68}$$

A function $\varepsilon \to \delta(\varepsilon)$, $\forall \varepsilon \in]0, 2[$, with the above property is called a *modulus of convexity* of X.

The following characterization is obvious.

Proposition 1.108 *A linear normed space X is uniformly convex if and only if* $\lim_{n\to\infty} \|x_n - y_n\| = 0$ *whenever* $\|x_n\| \le 1$, $\|y_n\| \le 1$ *and* $\lim_{n\to\infty} \|\frac{1}{2}(x_n + y_n)\| = 1$.

As examples of uniformly convex spaces, we have the Hilbert spaces and the Banach spaces $L^p[a, b]$, ℓ^p for $p \in]1, \infty[$ (see, for instance, [19]).

In the next section, we show that the property of weak convergence established for Hilbert spaces in Proposition 1.79 can also be extended to uniformly convex spaces.

Theorem 1.109 *If $x_n \xrightarrow{w} x_0$, where $\{x_n\}$ is a sequence in a uniformly convex space, and $\|x_n\| \to \|x_0\|$, then $x_n \to x_0$.*

Proof If $x_0 = 0$, the statement is obvious. Suppose $x_0 \ne 0$; we can consider $\|x_n\| > 0$ for all $n \in \mathbb{N}$. Let us put $x_n' = \|x_n\|^{-1} x_n$, $n \in \mathbb{N}$, $x_0' = \|x_0\|^{-1} x_0$ and so $\|x_n'\| = \|x_0'\| = 1$. Since $\{x_n' + x_0'\}_{n\in\mathbb{N}}$ converges weakly to $2x_0'$, from the second part of Corollary 1.76 we have

$$2 = \left\|2x_0'\right\| \le \liminf_{n\to\infty} \left\|x_n' + x_0'\right\| \le \limsup_{n\to\infty} \left\|x_n' + x_0'\right\| \le \left\|x_0'\right\| + \lim_{n\to\infty} \left\|x_n'\right\| = 2.$$

By virtue of Proposition 1.108, we obtain

$$\lim_{n\to\infty} \|x_n' - x_0'\| = 0,$$

that is, $\{x_n\}$ converges in the norm to x_0. □

Remark 1.110 The above statement is also true for the nets, so we find that in a uniformly convex space the weak and the strong (norm) topologies coincide on the boundary of a closed ball.

Theorem 1.111 *Every uniformly convex Banach space is reflexive.*

Proof Let x^{**} be an arbitrary element of X^{**} such that $\|x^{**}\| = 1$. According to Theorem 1.91, there exists a net $\{x_i\}_{i\in I} \subset \overline{S}(0,1)$, $\sigma(X^{**}, X^*)$-convergent to x^{**}. Since the net $\{\frac{1}{2}(x_i + x_j)\}_{(i,j)\in I\times I}$ is also $\sigma(X^{**}, X^*)$-convergent to x^{**}, and $\|x^{**}\| = 1$, applying Corollary 1.83, we have

$$1 \le \liminf_{i,j\in I} \left\| \frac{1}{2}(x_i + x_j) \right\| \le \limsup_{i,j\in I} \left\| \frac{1}{2}(x_i + x_j) \right\| \le \frac{1}{2} \limsup_{i,j\in I} \big(\|x_i\| + \|x_j\|\big) \le 1,$$

which says that

$$\lim_{i,j\in I} \left\| \frac{1}{2}(x_i + x_j) \right\| = 1.$$

According to the uniform convexity, it follows that $\{x_i\}$ is a Cauchy net. Thus, $\{x_i\}$ converges to an element $x \in X$. Therefore, $x^{**} = x \in X$ and the proof is complete. □

The notion dual to uniform convexity is the notion of uniform smoothness.

Definition 1.112 A normed linear space is said to be *uniformly smooth* if for each $\varepsilon > 0$ there exists an $\eta(\varepsilon) > 0$ for which

$$\|x\| = 1 \quad \text{and} \quad \|y\| \le \eta(\varepsilon) \quad \text{always implies } \|x+y\| + \|x-y\| < 2 + \varepsilon\|y\|. \tag{1.69}$$

Proposition 1.113 *A linear normed space X is uniformly smooth if and only if for each $\varepsilon > 0$ there exists an $\eta'(\varepsilon) > 0$ such that*

$$\|x\| \ge 1, \qquad \|y\| \ge 1 \quad \textit{and} \quad \|x-y\| \le \eta'(\varepsilon)$$
$$\textit{imply } \|x+y\| \ge \|x\| + \|y\| - \varepsilon\|x-y\|. \tag{1.70}$$

Proof If X is uniformly smooth, it is easy to establish property (1.70) for $\eta'(\varepsilon) = 2\eta(2\varepsilon)(1+\eta(2\varepsilon))^{-1}$, $\varepsilon > 0$.

Conversely, taking $\eta(\varepsilon) = \eta'(\frac{\varepsilon}{2})(2 + \eta'(\frac{\varepsilon}{2}))^{-1}$, $\varepsilon > 0$, property (1.69) follows from (1.70). □

Theorem 1.114 *A Banach space X is uniformly convex (smooth) if and only if its dual space X^* is uniformly smooth (convex).*

Proof Let us assume that X is a uniformly convex Banach space. We claim that (1.69) is verified for X^* with $\eta(\varepsilon) < \delta(\varepsilon)(2-\delta(\varepsilon))^{-1}$, where δ is a modulus of convexity of X. Indeed, if $x^*, y^* \in X^*$ and $\|x^*\| = 1$, $\|y^*\| \le \eta(\varepsilon)$, since X is reflexive (Theorem 1.111), there exist $x, y \in X$ with $\|x\| = 1$, $\|y\| = 1$, such that $(x^* + y^*, x) = \|x^* + y^*\|$ and $(x^* - y^*, y) = \|x^* - y^*\|$.

Thus, we have

$$\begin{aligned}\|x^* + y^*\| + \|x^* - y^*\| &= (x^*, x + y) + (y^*, x - y)\\ &\le \|x + y\| + \|x - y\|\,\|y^*\| \le 2 + \|x - y\|\,\|y^*\|. \quad (1.71)\end{aligned}$$

On the other hand, we have

$$\begin{aligned}\frac{1}{2}\|x + y\| &\ge \frac{1}{2}(x^* + y^*, x + y)\big(1 + \eta(\varepsilon)\big)^{-1}\\ &\ge \frac{1}{2}\big[(x^* + y^*, x) + (x^* - y^*, y) + 2(y^*, y)\big]\big(1 + \eta(\varepsilon)\big)^{-1}\\ &\ge \frac{1}{2}\big[\|x^* + y^*\| + \|x^* - y^*\| - 2\eta(\varepsilon)\big]\big(1 + \eta(\varepsilon)\big)^{-1}\\ &\ge \frac{1}{2}\big[\|x^* + y^* + x^* - y^*\| - 2\eta(\varepsilon)\big]\big(1 + \eta(\varepsilon)\big)^{-1}\\ &> \big(1 - \eta(\varepsilon)\big)\big(1 + \eta(\varepsilon)\big)^{-1} > 1 - \delta(\varepsilon),\end{aligned}$$

from which, according to property (1.68), we obtain $\|x - y\| < \varepsilon$. Therefore, using (1.71), (1.69) follows for X^*, that is, X^* is uniformly smooth.

Now, let us assume that X is uniformly smooth. To prove that X^* is uniformly convex, we establish property (1.68) with $\delta(\varepsilon) = \frac{\varepsilon}{12}\eta(\frac{\varepsilon}{2})$. For $\|x^*\| \le 1$, $\|y^*\| \le 1$, $\|x^* - y^*\| \ge \varepsilon$, there exists $x_\varepsilon \in X$ with $\|x_\varepsilon\| = \eta(\frac{\varepsilon}{2})$, such that

$$(x^* - y^*, x_\varepsilon) > \frac{2\varepsilon}{3}\eta\left(\frac{\varepsilon}{2}\right).$$

If $\|x\| = 1$, we have

$$\begin{aligned}(x, x^* + y^*) &= (x + x_\varepsilon, x^*) + (x - x_\varepsilon, y^*) - (x_\varepsilon, x^* - y^*)\\ &\le \|x + x_\varepsilon\| + \|x - x_\varepsilon\| - \frac{2\varepsilon}{3}\eta\left(\frac{\varepsilon}{2}\right)\\ &< 2 + \frac{\varepsilon}{2}\|x_\varepsilon\| - \frac{2\varepsilon}{3}\eta\left(\frac{\varepsilon}{2}\right) = 2 - \frac{\varepsilon}{6}\eta\left(\frac{\varepsilon}{2}\right)\end{aligned}$$

and therefore $\|x^* + y^*\| < 2(1 - \delta(\varepsilon))$, that is, X^* is uniformly convex. By virtue of Proposition 1.90, Theorem 1.111, the proof is complete. □

Corollary 1.115 *A uniformly smooth Banach space is reflexive.*

A renorming result analogous to Theorem 1.105 can be stated as follows.

If a Banach space X is endowed with two equivalent norms $\|\cdot\|_1$, $\|\cdot\|_2$, such that $(X, \|\cdot\|_1)$ is uniformly convex and $(X, \|\cdot\|_2)$ is uniformly smooth, then there exists a third equivalent norm $\|\cdot\|_3$ which is both uniformly convex and uniformly smooth.

A detailed study of the *uniform convexifiability problem* was given by James [17] (see also Diestel [9] and van Dulst [29]). For the special properties of duality and convexity in Banach spaces we refer the reader to the books of Day [7], Köthe [19] and Holmes [16].

Theorem 1.116 *If X is uniformly smooth, then the duality mapping is uniformly continuous on every bounded set M of X.*

Proof First, we remark that the duality mapping is single-valued (see Remark 1.100). Without any loss of generality, we may consider $M = \{x \in X; \|x\| = 1\}$. Let $x, y \in M$, $\|x - y\| < 2\delta(\varepsilon)$, where δ is a modulus of convexity of X^*. We have

$$\begin{aligned}\|Fx + Fy\| &\geq (x, Fx + Fy) = (x, Fx) + (y, Fy) + (x - y, Fy)\\ &\geq \|x\|^2 + \|y\|^2 - \|x - y\|\,\|Fy\| > 2\bigl(1 - \delta(\varepsilon)\bigr).\end{aligned}$$

Therefore, using property (1.68) for X^*, we obtain $\|Fx - Fy\| < \varepsilon$, that is, the duality mapping is uniformly continuous on M.

Finally, we describe some basic properties of the duality mapping. □

Proposition 1.117 *The duality mapping of a real Banach space X has the following properties*:

(i) *It is homogeneous*
(ii) *It is additive if and only if X is a Hilbert space*
(iii) *It is single-valued if and only if X is smooth*
(iv) *It is surjective if and only if X is reflexive*
(v) *It is injective or strictly monotone if and only if X is strictly convex*
(vi) *It is single-valued and uniformly continuous if and only if $X(X^*)$ is uniformly smooth* (*convex*).

Remark 1.118 The duality mapping can be replaced in property (vi) by one of its selections.

1.3 Vector-Valued Functions and Distributions

This section presents the notation, definitions and other necessary background information on vector-valued functions required for the following treatment. Most of the terminology and basic results used here are well known and will be used without further comment.

1.3.1 The Bochner Integral

Let X be a real (or complex) Banach space and let $[a, b]$ be a real (closed) interval. A vector-valued function x, defined on $[a, b]$ with values in X, is said to be finitely valued if it is a constant vector $\neq 0$ on each of a finite number of disjoint measurable sets $A_k \subset [a, b]$ and equal to 0 on $[a, b] \setminus \bigcup_k A_k$. The function x is said to be strongly measurable on $[a, b]$ if there is a sequence $\{x_n\}$ of finite-valued functions which converges strongly almost everywhere on $[a, b]$ to x.

A function x on $[a, b]$ to X is said to be Bochner integrable if there exists a sequence $\{x_n\}$ of a finitely valued function on $[a, b]$ to X, which converges strongly almost everywhere to x in such a way that

$$\lim_{n\to\infty} \int_a^b \left\| x(t) - x_n(t) \right\| \mathrm{d}t = 0.$$

A necessary and sufficient condition that x on $[a, b]$ to X is Bochner integrable is that x is strongly measurable and that $\int_a^b \|x(t)\| \,\mathrm{d}t < +\infty$. More generally, the space of all (classes of) strongly measurable functions x on $[a, b]$ to X, such that $\int_a^b \|x(t)\|^p \,\mathrm{d}t < +\infty$, for $1 \le p < \infty$, and $\operatorname{ess\,sup}_{t\in[a,b]} \|x(t)\| < +\infty$, $p = \infty$, is a Banach space $L^p(a, b; X)$ with the norm

$$\|x\|_p = \left(\int_a^b \left\| x(t) \right\|^p \mathrm{d}t \right)^{\frac{1}{p}}, \quad 1 \le p < \infty, \tag{1.72}$$

with the usual modification in the case $p = \infty$.

If X is reflexive and $1 \le p < \infty$, the dual of $L^p(a, b; X)$ is $L^q(a, b; X^*)$, where $\frac{1}{p} + \frac{1}{q} = 1$. More precisely, we have the following theorem.

Theorem 1.119 *Let X be a reflexive Banach space. Then to every $f \in (L^p(a, b; X))^*$ there corresponds a unique element $y_f \in L^q(a, b; X^*)$, $1 \le p < +\infty$, $\frac{1}{p} + \frac{1}{q} = 1$, such that*

$$\langle x, f \rangle = \int_a^b \big(x(t), y_f(t) \big) \,\mathrm{d}t, \quad x \in L^p(a, b; X) \tag{1.73}$$

and $\|f\| = \|y_f\|_q$.

Conversely, any $y \in L^q(a, b; X^*)$ defines a functional $f_y \in (L^p(a, b; X))^*$ such that $\|f_y\| = \|y\|_q$ and

$$\langle x, f_y \rangle = \int_a^b \big(x(t), y(t) \big) \,\mathrm{d}t, \quad \forall x \in L^p(a, b; X). \tag{1.74}$$

Remark 1.120 In the special case $1 < p < +\infty$, Theorem 1.119 is due to Philips and we refer the reader to Edward's book [12] for the proof. It should be noticed that in this case Theorem 1.119 remains true if X^* is separable.

The classical result of Dunford and Pettis asserts that a subset $\mathscr{A}$ of $L^1(\Omega)$, where $\Omega \subset \mathbb{R}^n$, is weakly sequentially compact if and only if it is bounded and the integrals $\int u \, \mathrm{d}t$ are uniformly absolutely continuous. This criterion still applies in the space $L^1(a, b; X)$ as shown in the next theorem.

Theorem 1.121 *Let X be a reflexive Banach space or a separable dual space. In order for a subset $\mathscr{A} \subset L^1(a, b; X)$ to be weakly sequentially compact, it is necessary and sufficient that the following two conditions be fulfilled:*

(a)

$$\sup\left\{\int_a^b \|x(t)\| \, \mathrm{d}t;\ x \in \mathscr{A}\right\} < +\infty.$$

(b) *Given $\varepsilon > 0$, there exists a number $\delta(\varepsilon) > 0$, such that*

$$\int_E \|x(t)\| \, \mathrm{d}t \leq \varepsilon, \quad \forall x \in \mathscr{A}, \tag{1.75}$$

provided that $E \subset [a, b]$ is measurable and Lebesgue measure $\mu(E)$ of E is $\leq \delta(\varepsilon)$.

For the proof, see Brooks and Dinculeanu [5].

We denote by $L^p_{\mathrm{loc}}(\mathbb{R}^+; X)$ the space of all strongly measurable functions $x : \mathbb{R}^+ \to X$ such that $x \in L^p(0, T; X)$ for all $T > 0$.

1.3.2 Bounded Variation Vector Functions

If X is a Banach space with norm $\|\cdot\|$ and $x : [a, b] \to X$ is a given function on $[a, b]$, then the total variation of x is defined by

$$\operatorname{Var}\bigl(x; [a, b]\bigr) = \sup \sum_{i=1}^{n} \bigl\|x(t_{i-1}) - x(t_i)\bigr\|, \tag{1.76}$$

where the supremum is taken over all partitions $\Delta = \{a = t_0 < t_1 < \cdots < t_n = b\}$ of $[a, b]$. If $\operatorname{Var}(x; [a, b]) < +\infty$, then the function x is said to be of *bounded variation* over $[a, b]$. We denote by $BV([a, b]; X)$ the space of all such functions.

Proposition 1.122 *Let $x : [a, b] \to X$ be a function of bounded variation. Then x is bounded and strongly measurable over $[a, b]$ and $x(t \pm 0)$ exists at all $t \in [a, b[$ and $t \in]a, b]$, respectively. Moreover, x is continuous apart from a countable set of points and the following inequality holds:*

$$\int_a^{b-h} \bigl\|x(t + h) - x(t)\bigr\| \, \mathrm{d}t \leq h \operatorname{Var}\bigl(x; [a, b]\bigr), \tag{1.77}$$

for all positive h such that $a \leq b - h$.

Proof Let $V_x(\cdot)$ be the real-valued function defined by

$$V_x(t) = \mathrm{Var}\big(x; [a,t]\big), \quad a \le t \le b.$$

Observe that $t \to V_x(t)$ is nondecreasing on $[a,b]$ and

$$\big\|x(t_1) - x(t_2)\big\| \le V_x(t_2) - V_x(t_1), \quad a \le t_1 \le t_2 \le b. \tag{1.78}$$

In particular, this implies that

$$x(t_0 + 0) = \lim_{\substack{t \to t_0 \\ t > t_0}} x(t)$$

and

$$x(t_0 - 0) = \lim_{\substack{t \to t_0 \\ t < t_0}} x(t)$$

exist at every $a \le t_0 < b$ (respectively, at every $a < t_0 \le b$). The same inequality shows that $x(t_0 - 0) = x(t_0) = x(t_0 + 0)$ apart from a countable set of discontinuities t_0. Thus, x is measurable over $[a,b]$. The remaining part of Proposition 1.122 is a straightforward consequence of inequality (1.78). □

Contrary to the case of numerical functions, the X-valued function of bounded variation does not necessarily need to be almost everywhere differentiable. However, if the space X is reflexive, we have the following theorem, due to Y. Komura (see, e.g., Barbu [3]).

Theorem 1.123 *Let X be a reflexive Banach space and let $x \in BV([a,b]; X)$. Then*

$$\big(\dot{x}(t), x^*\big) = \lim_{h \to 0} \left(\frac{x(t+h) - x(t)}{h}, x^* \right) \quad \text{for all } x^* \in X^* \tag{1.79}$$

exists, a.e. on $]a,b[$. Moreover, $\dot{x} \in L^1(a,b; X)$ and

$$\int_a^b \big\|\dot{x}(t)\big\| \,\mathrm{d}t \le \mathrm{Var}\big(x; [a,b]\big). \tag{1.80}$$

An X-valued function x defined over $[a,b]$ is said to be *absolutely continuous* on $[a,b]$ if, for each $\varepsilon > 0$, there exists $\delta(\varepsilon) > 0$, such that

$$\begin{aligned} &\sum_{n=1}^{N} \big\|x(t_n) - x(s_n)\big\| \le \varepsilon \quad \text{whenever} \sum_{n=1}^{N} |t_n - s_n| \le \delta(\varepsilon) \quad \text{and} \\ &]t_n, s_n[\, \cap \,]t_m, s_m[\, = \emptyset \quad \text{for } m \ne n. \end{aligned} \tag{1.81}$$

Theorem 1.124 *Let X be a reflexive Banach space. Then every X-valued absolutely continuous function x over $[a, b]$ is a.e. differentiable on $[a, b]$ and can be represented as*

$$x(t) = x(a) + \int_a^t \left(\frac{\mathrm{d}x}{\mathrm{d}s}\right)(s)\,\mathrm{d}s, \quad a \le t \le b, \tag{1.82}$$

where $\frac{\mathrm{d}x}{\mathrm{d}s} \in L^1(a, b; X)$ is the strong derivative of x.

Proof Using Theorem 1.123, the weak derivative $\dot{x}$ of x exists and belongs to $L^1(a, b; X)$. Let $\tilde{x} : [a, b] \to X$ be defined by

$$\tilde{x}(t) = x(a) + \int_a^t \dot{x}(s)\,\mathrm{d}s \quad \text{for } t \in [a, b].$$

Obviously, $(\tilde{x}(t), x^*) = (x(t), x^*)$ for all $t \in [a, b]$ and all x^* in X^*. Hence, $\tilde{x} = x$. On the other hand, $\tilde{x}$ is almost everywhere strongly differentiable on $]a, b[$ and $(\frac{\mathrm{d}\tilde{x}}{\mathrm{d}t})(t) = \dot{x}(t)$, a.e. on $]a, b[$ because $\dot{x} \in L^1(a, b; X)$. This completes the proof. □

Basic properties concerning the theory of real functions and vector measures can be found in the books of Dinculeanu [11], Edwards [12] and Precupanu [25].

1.3.3 Vector Measures and Distributions on Real Intervals

Let I be an interval of real axis and let X be a Banach space with the norm $\|\cdot\|$. Let $\mathscr{K}(I)$ be the space of all continuous scalar (real or complex) functions on I with compact supports in I. Given a compact subset K of I, we denote by $\mathscr{K}_K(I)$ the linear subspace of $\mathscr{K}(I)$ consisting of all continuous functions with support in K; the space $\mathscr{K}_K(I)$ is a Banach space with the norm $|||\varphi||| = \sup\{|\varphi(x)|; x \in K\}$. By definition, a *measure* (*Radon measure*) μ on I with values in X is a linear operator from $\mathscr{K}(I)$ to X whose restrictions to every $\mathscr{K}_K(I)$ are continuous. In other words, the linear operator $\mu : \mathscr{K}(I) \to X$ is a measure on I if and only if to each compact subset K of I there corresponds a number $m_K > 0$ such that

$$\left\|\mu(\varphi)\right\| \le m_k |||\varphi|||, \quad \forall \varphi \in \mathscr{K}_K(I). \tag{1.83}$$

If the constant m_K which occurs in (1.83) can be chosen independent of K, then the measure μ is said to be a *bounded measure* on I. The measure $\mu : \mathscr{K}(I) \to X$ is said to be majorized if there exists a scalar positive measure $\nu : \mathscr{K}(I) \to \mathbb{R}$ such that

$$\left\|\mu(\varphi)\right\| \le \nu\big(|\varphi|\big), \quad \forall \varphi \in \mathscr{K}(I).$$

If μ is a *majorized measure*, then there exists a smallest measure $\nu : \mathscr{K}(I) \to \mathbb{R}^+$ which majorizes μ. This positive scalar measure is called the *absolute value* of μ and will be denoted by $|\mu|$.

If I is an open interval of real axis, then we denote by $\mathscr{D}(I)$ the space of all infinitely differentiable real-valued functions on I with compact support in I. The space $\mathscr{D}(I)$ is topologized as a strict inductive limit of $\mathscr{D}_K(I)$ where K ranges over all compact subsets of I and $\mathscr{D}_K(I) = \{\varphi \in \mathscr{D}(I); \text{support}\,\varphi \subset K\}$. We denote by $\mathscr{D}'(I; X)$ the space of all linear continuous operators from $\mathscr{D}(I)$ to X. An element $u \in \mathscr{D}'(I; X)$ is called an *X-valued distribution* on I. If $u \in \mathscr{D}'(I; X)$ and j is a positive integer, then the relation

$$u^{(j)}(\varphi) = (-1)^j u\big(\varphi^{(j)}\big), \quad \forall \varphi \in \mathscr{D}(I) \tag{1.84}$$

defines another distribution $u^{(j)}$ called the *derivative of order j* of u.

Now, let I be an arbitrary interval of the real axis and let $w : I \to X$ be a function which is of bounded variation on every compact subinterval of I. We associate to w the scalar-valued function $V_w : I \to \mathbb{R}^+$ defined by

$$V_w(t) - V_w(s) = \mathrm{Var}\big(w; [s,t]\big) \quad \text{for } s, t \in I.$$

It is well known that w defines an X-valued measure on I (the Lebesgue–Stieltjes measure associated with w). In the sequel, we briefly recall the construction of this measure.

Let $d : t_0 < t_1 < \cdots < t_n$ be a finite partition of the interval I. We say that partition d' is finer than d and we write this as $d \le d'$ if every point of d is a point of d'. The family P of all finite partitions d of I is a directed set with this order relation. For every $\varphi \in \mathscr{K}(I)$ and $d \in P$, consider the *Riemann–Stieltjes sum*

$$S(d, \varphi) = \sum_{i=1}^{n} \varphi(t_i)\big(w(t_i) - w(t_{i-1})\big).$$

It turns out that there exists the limit of $S(d, \varphi)$ through the directed set P, which will be denoted by $\int \varphi\, \mathrm{d}w$ (the *Riemann–Stieltjes integral* of φ with respect to w). If $[a, b]$ is a compact interval of I containing the support of φ, we have

$$\left\| \int \varphi\, \mathrm{d}w \right\| \le |||\varphi|||\, \mathrm{Var}\big(w; [a,b]\big). \tag{1.85}$$

Hence, the map $\varphi \to \int \varphi\, \mathrm{d}w$ is a measure on I. Furthermore, it follows by (1.85) that the measure $\varphi \to \int \varphi\, \mathrm{d}w$, which we simply denote by $\mathrm{d}w$, is majorizable on I and, therefore, its absolute value $\mathrm{d}|w|$ exists. More precisely, one has

$$\left\| \int \varphi\, \mathrm{d}w \right\| \le \int \varphi\, \mathrm{d}V_w, \quad \forall \varphi \in \mathscr{K}(I),\ \varphi \ge 0.$$

As a matter of fact, the integral $\int \varphi\, \mathrm{d}w$ can be defined by the formula

$$\left(\int \varphi\, \mathrm{d}w, x^* \right) = \int \varphi\, \mathrm{d}(w, x^*), \quad \forall x^* \in X^*,$$

where $\int \varphi \, \mathrm{d}(w, x^*)$ is the classical Riemann–Stieltjes integral. In particular, the above inequality shows that every scalar valued $\mathrm{d}V_w$ integrable function f on I is integrable with respect to $\mathrm{d}w$ and

$$\left\| \int f \, \mathrm{d}w \right\| \leq \int |f| \, \mathrm{d}V_w.$$

If I' is a subinterval of I, then, by definition, the dw-measure of I' is

$$\mathrm{d}w(I') = \int \chi_{I'} \, \mathrm{d}w,$$

where $\chi_{I'}$ is the characteristic function of I'. A little calculation reveals that

$$\begin{aligned}
\mathrm{d}w\big([a,b]\big) &= w(b+0) - w(a-0),\\
\mathrm{d}w\big([a,b[\big) &= w(b-0) - w(a-0),\\
\mathrm{d}w\big(]a,b]\big) &= w(b+0) - w(a+0),\\
\mathrm{d}w\big(]a,b[\big) &= w(b-0) - w(a+0),
\end{aligned} \tag{1.86}$$

where $a \leq b$. Here, we used the usual convention $w(a-0) = w(a)$ if I is of the form $[a, t_2]$, and $w(b+0) = w(b)$ if I is of the form $[t_1, b]$.

If $f \in C([a,b])$, we denote by $\int_{[a,b]} f \, \mathrm{d}w$ the integral $\int f \chi_{[a,b]} \, \mathrm{d}w$, and by $\int_a^b f \, \mathrm{d}w$ the Riemann–Stieltjes integral on $[a,b]$. We have

$$\int_{[a,b]} f \, \mathrm{d}w - \int_a^b f \, \mathrm{d}w = f(b)\big(w(b+0) - w(b)\big) - f(a)\big(w(a-0) - w(a)\big).$$

Moreover, if $f : I \to \mathbb{R}$ is a continuously differentiable function on I, then, for every interval $[a,b] \in I$, one has the following formula for integrating by parts:

$$\int_{[a,b]} f \, \mathrm{d}w + \int_{[a,b]} f' w \, \mathrm{d}t = w(b+0) f(b) - w(a-0) f(a), \tag{1.87}$$

where $\mathrm{d}t$ is the Lebesgue measure on $\mathbb{R}$.

In particular, we see by (1.87) that the measure $\mathrm{d}w$ on $]a,b[$ is just the derivative w' of w in the sense of X-valued distributions on $]a,b[$.

Let us now assume that the space X is reflexive. Then, by Theorem 1.123, the function w is a.e. weakly differentiable on $[a,b]$ and we may write

$$w(t) = \int_a^t \dot{w}(s) \, \mathrm{d}s + w_s(t), \quad a \leq t \leq b, \tag{1.88}$$

where $\dot{w} \in L^1(a,b;X)$ is the weak derivative of w. The function $w_s \in BV([a,b];X)$ will be called the *singular part* of w. In accordance with (1.88), we have the Lebesgue decomposition of the measure $\mathrm{d}w$ in two parts: $\dot{w} \, \mathrm{d}t, \mathrm{d}w_s$; $\mathrm{d}w = \dot{w} \, \mathrm{d}t + \mathrm{d}w_s$.

Here, the measure $\dot{w}\,\mathrm{d}t$ defined by $\varphi \to \int_a^b \varphi\dot{w}\,\mathrm{d}t$ is the *absolutely continuous part* of $\mathrm{d}w$ with respect to the Lebesgue measure and $\mathrm{d}w_s$ is the *singular part* of $\mathrm{d}w$. In other words, there exists a closed subset $\Omega \subset [a,b]$ with the Lebesgue measure zero, such that $\mathrm{d}w_s = 0$, on $[a,b] \setminus \Omega$.

Given a compact interval $[a,b] \subset \mathbb{R}$, a continuous function $f : [a,b] \to X^*$, and $w \in BV([a,b]; X)$, we denote by $\int_a^b (\mathrm{d}w, f)$ the Riemann–Stieltjes integral:

$$\int_a^b (\mathrm{d}w, f) = \lim_d \sum_{i=1}^n \big(f(t_i), w(t_i) - w(t_{i-1})\big),$$

where the limit is taken through the directed set P of all the finite partitions $d : a = t_1 < t_1 < \cdots < t_n = b$. (Here, $(\cdot,\cdot)$ denotes the pairing between X and X^*.) By a classical device, it follows that this limit exists and the following estimate is satisfied:

$$\left|\int_a^b (\mathrm{d}w, f)\right| \le \mathrm{Var}\big(w; [a,b]\big) \sup\big\{\|f(t)\|;\ t \in [a,b]\big\}. \tag{1.89}$$

Now, let Y be a reflexive Banach space and let $F : [a,b] \to L(X,Y)$ be such that the function $F^* : [a,b] \to L(Y^*, X^*)$ is strongly continuous on $[a,b]$ (F^* is the adjoint operator). By definition, the Riemann–Stieltjes integral of F with respect to w is the element $\int_a^b F\,\mathrm{d}w \in Y$ given by

$$\left(\int_a^b F\,\mathrm{d}w, y^*\right) = \int_a^b \big(\mathrm{d}w, (F^* y^*)\big), \quad \forall y^* \in Y^*. \tag{1.90}$$

Given a Banach space Z, we denote by $C([a,b]; Z)$ the Banach space of all continuous functions from $[a,b]$ to Z endowed with the supp norm.

Proposition 1.125 *Let $w \in BV([a,b]; X)$, $f \in C([a,b]; Y^*)$ and $U : [a,b] \times [a,b] \to L(X,Y)$ be given such that $U(t,s)$ and $U^*(t,s)$ are strongly continuous on $[a,b] \times [a,b]$. Further, assume that the space Y is reflexive. Then we have*

$$\int_a^b \left(f(t), \int_t^b U(t,s)\,\mathrm{d}w(s)\right)\mathrm{d}t = \int_a^b \left(\mathrm{d}w(s), \int_a^s U^*(t,s) f(t)\,\mathrm{d}t\right). \tag{1.91}$$

For the proof, which is classical, we refer the reader to Höenig's book [15].

Now, we give a weak form of the classical Helly Theorem in Banach spaces.

Theorem 1.126 *Let X be a reflexive separable Banach space with separable dual X^*. Let $\{w_n\} \subset BV([a,b]; X)$ be such that $\|w_n(t)\| \le C$ for $t \in [a,b]$ and* $\mathrm{Var}(w_n; [a,b]) \le C$ *for all n.*

Then there exists a subsequence $\{w_{n_k}\} \subset \{w_n\}$ and a function $w \in BV([a,b]; X)$ such that, for $k \to \infty$,

$$w_{n_k}(t) \to w(t), \quad \textit{weakly in } X \textit{ for } t \in [a,b], \tag{1.92}$$

$$\int_a^b \varphi \, \mathrm{d}w_{n_k} \to \int_a^b \varphi \, \mathrm{d}w, \quad \textit{weakly in } X, \ \forall \varphi \in C\big([a,b]\big). \tag{1.93}$$

Proof Let $x^* \in X^*$ be arbitrary, but fixed. By the classical theorem of Helly, there exists a subsequence again denoted by $\{w_n\}$ and $f_{x^*} \in BV([a,b])$ such that

$$\big(w_n(t), x^*\big) \to f_{x^*}(t), \quad t \in [a,b].$$

Let S be a countable and dense subset of X^*. Then, applying the diagonal process, we obtain a subsequence $\{w_{n_k}\} \subset \{w_n\}$ having the property that

$$\lim_{n_k \to \infty} \big(w_{n_k}(t), x^*\big) = f_{x^*}(t), \quad \forall x^* \in S, \ t \in [a,b].$$

The density of S entails

$$\lim_{n_k \to \infty} \big(w_{n_k}(t), x^*\big) = f_{x^*}(t), \quad \forall x^* \in X^*, \ t \in [a,b]$$

and

$$\lim_{n_k \to \infty} \int_a^b \varphi \, \mathrm{d}(w_{n_k}, x^*) = \int_a^b \varphi \, \mathrm{d}f_{x^*}, \quad \forall \varphi \in C\big([a,b]\big).$$

By the uniform boundedness principle, it follows that there exists $w : [a,b] \to X$, such that

$$\big(w(t), x^*\big) = f_{x^*}(t), \quad \forall t \in [a,b], \ x^* \in X^*.$$

□

Next, since by the assumption $\mathrm{Var}(f_{x^*}; [a,b]) \le C\|x^*\|$, where C is independent of $x^* \in X^*$, we conclude that $w \in BV([a,b]; X)$.

Remark 1.127 If X is a general Banach space and for each $t \in [a,b]$ the family $\{w_n(t)\}$ is compact in X, then the sequence $\{w_{n_k}\}$ can be chosen strongly convergent on $[a,b]$. This strong version of the Helly Theorem is due to Foiaş and can be found in the book of Nicolescu [23].

Corollary 1.128 *Let $f \in C([a,b]; X^*)$ be given. Then under the conditions of Theorem 1.126 we have*

$$\lim_{n_k \to \infty} \int_a^b (\mathrm{d}w_{n_k}, f) = \int_a^b (\mathrm{d}w, f). \tag{1.94}$$

Proof By Theorem 1.126, relation (1.94) is satisfied for all $f \in C([a,b]; X^*)$ of the form

$$f(t) = \sum_{i=1}^N a_i \varphi_i(t), \quad t \in [a,b], \tag{1.95}$$

where $a_i \in X^*$ and $\varphi_i \in C([a,b])$, $i = 1, \ldots, N$. Since the space of functions of the form (1.95) is dense in $C([a,b]; X^*)$ and $\mathrm{Var}(w_{n_k}; [a,b]) \le C$, we infer by (1.89) that relation (1.94) holds for every $f \in C([a,b]; X^*)$. $\square$

Theorem 1.129 below gives a useful characterization of functions of bounded variation. For the proof, we refer the reader to Brezis' book [4].

Theorem 1.129 *Let X be a Banach space and let $x \in L^1(a,b;X)$ be given. Let C be a positive constant. Then the following two conditions are equivalent*:

(i) *There exists a function $y \in BV([a,b]; X)$ such that $y(t) = x(t)$, a.e. $t \in]a,b[$ and* $\mathrm{Var}(y; [a,b]) \le C$.
(ii) $|\int_a^b (x(t), \frac{d\varphi}{dt})\, dt| \le C \|\varphi\|_{C([a,b];X^*)}, \ \forall \varphi \in \mathscr{D}(a,b;X^*)$.

Here, $\mathscr{D}(a,b;X^*)$ is the space of all infinitely differentiable X^*-valued functions with compact support in $]a,b[$.

We denote by $W^{1,p}([a,b];X)$, $1 \le p \le \infty$, the space of all vector distributions $u \in \mathscr{D}'(a,b;X)$ having the property that $u, u' \in L^p(a,b;X)$, where u' is the distributional derivative of u.

Let $A^{1,p}([a,b];X)$ be the space of all absolutely continuous functions $u : [a,b] \to X$, whose strong derivatives $\frac{du}{dt}(t)$ exist a.e. on $]a,b[$, belong to $L^p(a,b;X)$, and which can be represented by

$$u(t) = u(a) + \int_a^t g(s)\, ds, \quad \text{for } a \le t \le b,\ g \in L^p(a,b;X).$$

Theorem 1.130 *Let X be a Banach space and let $u \in L^p(a,b;X)$, $1 \le p < \infty$, be given. The following conditions are equivalent*:

(j) $u \in W^{1,p}([a,b];X)$
(jj) *There exists $u_1 \in A^{1,p}([a,b];X)$ such that $u(t) = u_1(t)$, a.e. $t \in]0,T[$*
(jjj) $\lim_{h\to 0} \int_0^{-h} \|h^{-1}(u(t+h) - u(t)) - u'(t)\|^p\, dt = 0$.

For the proof, we refer to Brezis [4].

In particular, it follows by Theorem 1.130 that the space $W^{1,p}([a,b];X)$ can be identified with the space $A^{1,p}([a,b];X)$ endowed with the norm

$$\|u\|_{1,p} = \left(\|u(a)\|^p + \int_a^b \|u'(t)\|^p\, dt \right)^{\frac{1}{p}}$$

and, for every $u \in W^{1,p}([a,b];X)$, the distributional derivative u' coincides with the strong derivative $\frac{du}{dt}$.

Let V and H be a pair of real Hilbert spaces such that $V \subset H \subset V'$ in the algebraic and topological sense. The norms in V and H will be denoted by $\|\cdot\|$ and $|\cdot|$, respectively. V' is the dual space of V, and H is identified with its own dual. We denote by (v_1, v_2) the inner product of $v_1 \in V$ and $v_2 \in V'$; if $v_1, v_2 \in H$, this is the ordinary inner product in H. We set

$$W(0,T) = \left\{u \in L^2(0,T;V);\ u' \in L^2(0,T;V')\right\},$$

where u' is the distributional derivative of u. By virtue of Theorem 1.133, every $u \in W(0,T)$ may be identified with a V'-valued absolutely continuous function on $[0,T]$, and u' is the strong derivative of $u : [0,T] \to V'$.

Proposition 1.131 *Any function $u \in W(0,T)$ coincides almost everywhere on $[0,T]$ with a continuous function from $[0,T]$ to H. Moreover, if $u, v \in W(0,T)$, then the function $t \to (u(t), v(t))$ is absolutely continuous on $[0,T]$ and*

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(u(t), v(t)\big) = \left(\frac{\mathrm{d}u}{\mathrm{d}t}(t), v(t)\right) + \left(u(t), \frac{\mathrm{d}v}{\mathrm{d}t}(t)\right) \quad \text{a.e. } t \in \,]0,T[. \tag{1.96}$$

Proof Let u and v be two elements of $W(0,T)$. Define $\psi(t) = (u(t), v(t))$. An easy calculation, using Theorem 1.130(jjj), reveals that

$$\lim_{h\to 0}\int_0^{T-h}\left|h^{-1}\big(\psi(t+h) - \psi(t)\big) - \left(\frac{\mathrm{d}u}{\mathrm{d}t}(t), v(t)\right) - \left(u(t), \frac{\mathrm{d}v}{\mathrm{d}t}(t)\right)\right| \mathrm{d}t = 0.$$

□

Hence, $\psi \in W^{1,1}([0,T];\mathbb{R})$ and $\frac{\mathrm{d}\psi}{\mathrm{d}t}(t) = (\frac{\mathrm{d}u}{\mathrm{d}t}(t), v(t)) + (u(t), \frac{\mathrm{d}v}{\mathrm{d}t}(t))$ a.e. on $]0,T[$, as claimed.

Now, in (1.96), we put $v = u$ and integrate on $[s,t]$, to get

$$\frac{1}{2}\left(\left|u(t)\right|^2 - \left|u(s)\right|^2\right) = \int_s^t \left(\frac{\mathrm{d}u}{\mathrm{d}\tau}, u\right) \mathrm{d}\tau.$$

Thus, the function $t \to |u(t)|^2$ is absolutely continuous on $[0,T]$. Since the function u is continuous from $[0,T]$ to V' (more precisely, it coincides a.e. with a continuous function), we infer that, for every $v \in V$, $t \to (u(t), v)$ is continuous on $[0,T]$. As the space V is dense in H, we may conclude that it is weakly continuous on H. Inasmuch as u is continuous on $[0,T]$, this implies that $|u(t)|$ is strongly continuous from $[0,T]$ to H.

Remark 1.132 Proposition 1.131 extends to a pair of reflexive Banach spaces V, V' which are in duality and $V \subset H \subset V'$ algebraically and topologically; $u \in L^{p_1}(0,T;V)$, $v \in W^{1,p_1'}([0,T];V')$. The details are omitted.

1.3.4 Sobolev Spaces

We assume familiarity with the concepts and fundamental results in the theory of scalar distributions. However, we recall for easy reference the basic notation and definitions.

Let Ω be an open subset of $\mathbb{R}^n$. Denote by $\mathscr{D}(\Omega)$ the space of all real-valued infinitely differentiable functions with compact support in Ω, and by $\mathscr{D}'(\Omega)$ the space of all scalar distributions defined on Ω, that is, the dual space of $\mathscr{D}(\Omega)$. We use the multi-index notation

$$D^\alpha u(x) = D_1^{\alpha_1} D_2^{\alpha_2} \cdots D_n^{\alpha_n} u(x), \quad x \in \Omega,\ \alpha = (\alpha_1, \ldots, \alpha_n),$$

where $D_i = \frac{\partial}{\partial x_i}$, $i = 1, \ldots, n$. The distribution $D^\alpha u$ defined by

$$D^\alpha u(\varphi) = (-1)^{|\alpha|} u\big(D^\alpha \varphi\big), \quad \forall \varphi \in \mathscr{D}(\Omega),\ |\alpha| = \alpha + 1 + \alpha_2 + \cdots + \alpha_n,$$

is called the derivative of order α of $u \in \mathscr{D}'(\Omega)$.

Let $L^p(\Omega)$, $1 \le p \le \infty$, denote the usual Banach space of Lebesgue measurable functions of equivalence classes from Ω to $\mathbb{R}$ under the norm

$$\|u\|_p = \left(\int_\Omega \big|u(x)\big|^p \,\mathrm{d}x\right)^{\frac{1}{p}} \quad \text{if } 1 \le p < +\infty.$$

For $1 \le p < \infty$ and $k \ge 1$, we denote by $W^{k,p}(\Omega)$ the set of all functions u defined in Ω, such that u and all its derivatives $D^\alpha u$ up to order k belong to $L^p(\Omega)$. $W^{k,p}(\Omega)$ is a Banach space under the norm

$$\|u\|_{k,p} = \sum_{|\alpha| \le k} \big\| D^\alpha u \big\|_p. \tag{1.97}$$

Let $C_0^k(\Omega)$ denote the space of all functions $u \in C^k(\Omega)$ with compact support in Ω. The completion of the space $C_0^k(\Omega)$, normed by (1.97), will be denoted by $W_o^{k,p}(\Omega)$. For simplicity, we write $W^{k,2}(\Omega) = H^k(\Omega)$. The space $W_0^{k,2}(\Omega)$ will, similarly, be denoted by $H_0^k(\Omega)$. Finally, denote by $W^{-k,p'}(\Omega)$, $1 \le p' < +\infty$, the set of all $u \in \mathscr{D}'(\Omega)$ which can be represented as

$$u = \sum_{|\alpha| \le k} D^\alpha f_\alpha, \quad f_\alpha \in L^{p'}(\Omega).$$

Now, we state without proof some important theorems concerning the Sobolev spaces (for a proof, see Lions and Magenes [22], Chap. 1).

Theorem 1.133 *The dual space of $H_0^k(\Omega)$ coincides with the space $H^{-k}(\Omega) = W^{-k,2}(\Omega)$.*

For every $s \ge 0$, define $H^s(\mathbb{R}^n) = \{u \in L^2(\mathbb{R}^n);\ (1 + |\xi|^2)^{\frac{s}{2}} \hat{u}(\xi) \in L^2(\mathbb{R}^n)\}$, where $\hat{u}$ denotes the Fourier transform of u. If $k = s$ is a positive integer, then, by Parseval's formula, one can easily deduce that $H^s(\mathbb{R}^n) = H_0^s(\Omega)$, where $\Omega = \mathbb{R}^n$.

Now we suppose that Ω is a bounded and open subset of $\mathbb{R}^n$ with sufficiently smooth boundary. More precisely, it will be assumed that the boundary Γ is an

$n-1$-dimensional manifold of class C^∞. This allows us to define the Sobolev space $H^s(\Gamma)$ for any real s. Let $L^2(\Gamma)$ be the space of all square integrable functions on Γ with respect to measure $d\sigma$. Let $\mathscr{N}$ be a family of local charts on Γ and let $\{\alpha_i\}$ be a finite partition of unity subordinated to it. For any $u \in L^2(\Gamma)$, we have $u \in \sum_i \alpha_i u$.

Let $u_i = \alpha_i u$. We say that $u \in H^s(\Gamma)$ if $u_i \in H^s(\mathbb{R}^{n-1})$ for all i. $H^s(\Gamma)$ is a Hilbert space with the norm defined in an obvious manner. If $s < 0$, we set $H^s(\Gamma) = (H^{-s}(\Gamma))^*$.

Let $C^\infty(\overline{\Omega})$ be the space of all infinitely differentiable functions on $\overline{\Omega}$. For any $u \in C^\infty(\overline{\Omega})$ we can define the derivatives of order j outward normal to Γ : $\frac{\partial^j u}{\partial \nu^j}$. It turns out that the mapping $u \to \{u, \frac{\partial u}{\partial \nu}, \ldots\}$ can be extended by continuity to all u in $H^k(\Omega)$.

Theorem 1.134 *The mapping* $u \to \{\frac{\partial^j u}{\partial \nu^j}; j = 0, 1, \ldots, \mu\}$ *from* $C^\infty(\overline{\Omega})$ *to* $(C^\infty(\Gamma))^{\mu+1}$ *extends to a linear continuous operator* $u \to \{\frac{\partial^j u}{\partial \nu^j}; j = 0, 1, \ldots, \mu\}$ *from* $H^k(\Omega)$ *onto* $\prod_{j=0}^{\mu} H^{k-j-\frac{1}{2}}(\Gamma)$, *where* μ *is the largest integer such that* $\mu \le k - \frac{1}{2}$.

In particular, the above theorem shows that, for each $u \in H^k(\Omega)$, the $\frac{\partial^j u}{\partial \nu^j}$, $0 \le j \le \mu$, are well defined and belong to $H^{k-j-\frac{1}{2}}(\Gamma)$. The space $H_0^k(\Omega)$ can, equivalently, be defined as $H_0^k(\Omega) = \{u \in H^k(\Omega);\ \frac{\partial^j u}{\partial \nu^j} = 0,\ j = 0, 1, \ldots, k-1\}$.

1.4 Maximal Monotone Operators and Evolution Systems in Banach Spaces

This section summarizes some significant results on maximal monotone operators and linear differential equations in Banach spaces. The generality is confined to the context needed as a prerequisite for Chaps. 2 and 4. For a general approach to the theory of nonlinear monotone operators, the reader is referred to the survey of Browder [6] and to the books of Lions [21], Brezis [4], and Barbu [3]. As regards the linear evolution equations, we refer the reader to the books of Yosida [30], Pazy [24], and Krein [20].

1.4.1 Definitions and Fundamental Results

If X and Y are linear spaces, then $X \times Y$ will denote their Cartesian product space. An element of the product space $X \times Y$ will be written in the form $[x, y]$ for $x \in X$ and $y \in Y$.

A multi-valued operator A from X to Y will be viewed as a subset of $X \times Y$.

If $A \subset X \times Y$, we define

$$Ax = \{y \in Y;\ [x, y] \in A\}, \qquad D(A) = \{x \in X;\ Ax \neq \emptyset\},$$
$$R(A) = \bigcup \{Ax;\ x \in D(A)\}, \qquad A^{-1} = \{[y, x];\ [x, y] \in A\}. \tag{1.98}$$

If $A, B \subset X \times Y$, and λ is real, we set

$$\lambda A = \{[x, \lambda y];\ [x, y] \in A\},$$
$$A + B = \{[x, y + z];\ [x, y] \in A,\ [x, z] \in B\}. \tag{1.99}$$

In the following, the operators from X to Y will not be distinguished from their graphs in $X \times Y$. If A is single-valued, Ax will denote either a value of A at x, or the set defined in formula (1.98).

Throughout this chapter, X will be a real Banach space and X^* will denote its dual space. The notation for norms, convergence, duality mapping, and scalar product will be the same as in Sect. 1.1. In particular, the value of $x^* \in X^*$ at $x \in X$ will be denoted by either (x, x^*) or (x^*, x).

Definition 1.135 A subset $A \subset X \times X^*$ is called *monotone* if

$$(x_1 - x_2, y_1 - y_2) \geq 0, \tag{1.100}$$

for any $[x_i, y_i] \in A$, $i = 1, 2$. A monotone subset of $X \times X^*$ is said to be *maximal monotone* if it is not properly contained in any other monotone subset of $X \times X^*$.

If A is a single-valued operator from X to X^*, then the monotonicity condition (1.99) becomes

$$(x_1 - x_2, Ax_1 - Ax_2) \geq 0 \quad \text{for all } x_1, x_2 \in D(A). \tag{1.101}$$

It must be noticed that if $A \subset X \times X^*$ is maximal monotone, then, for each $x \in X$, *Ax is a closed convex subset of X^**. This is easily seen from the obvious formula

$$Ax = \{x^* \in X^*;\ (x^* - v, x - u) \geq 0 \quad \text{for all } [u, v] \in A\}.$$

Definition 1.136 A subset $A \subset X \times X^*$ is said to be *locally bounded at $x_0 \in X$* if there exists a neighborhood V of x_0 such that $A(V) = \bigcup\{Ax;\ x \in D(A) \cap V\}$ is a bounded subset of X^*. The operator $A : X \to X^*$ is said to be *bounded* if it maps every bounded subset of X into a bounded set of X^*.

Definition 1.137 Let A be a single-valued operator defined from X into X^*. A is said to be *demicontinuous* if it is strongly–weak-star continuous from X to X^*, that is,

$$\lim_{n \to \infty} Ax_n = Ax_0 \quad \text{weak-star in } X^*, \tag{1.102}$$

for any sequence $\{x_n\} \subset D(A)$ strongly convergent to x_0 in X.

Definition 1.138 The (multi-valued) operator $A : X \to X^*$ is called *coercive* if

$$\lim_{n\to\infty} \frac{(x_n^*, x_n - x_0)}{\|x_n\|} = +\infty, \tag{1.103}$$

for some $x_0 \in X$ and all $[x_n, x_n^*] \in A$ such that $\lim_{n\to\infty} \|x_n\| = +\infty$.

We begin with a rather technical result; the proof may be found in the first author's book [3].

Theorem 1.139 *Let X be a reflexive Banach space and let A and B be two monotone subsets of $X \times X^*$ such that $0 \in D(A)$ and $B : X \to X^*$ is demicontinuous and bounded, and*

$$\lim_{n\to\infty} \frac{(Bx_n, x_n)}{\|x_n\|} = +\infty, \tag{1.104}$$

for every sequence $\{x_n\}$ such that $\lim_{n\to\infty} \|x_n\| = +\infty$.

Then there exists $x \in \overline{\operatorname{conv} D(A)}$ such that

$$(u - x, Bx + v) \geq 0 \quad \textit{for all } [u, v] \in A. \tag{1.105}$$

With this tool in hand, the proof of the following basic theorem on maximal monotone operators is straightforward.

Corollary 1.140 (Minty) *Let X be a reflexive Banach space and let B be a monotone, demicontinuous and bounded operator from X to X^* satisfying condition (1.104). Let A be a maximal monotone subset of $X \times X^*$. Then*

$$R(A + B) = X^*.$$

Proof Let y_0 be arbitrary but fixed in X^*. By Theorem 1.139, there exists $x_0 \in X$ such that $(u - x_0, Bx_0 - y_0 + v) \geq 0$, for all $[u, v] \in A$. Hence, $x_0 \in D(A)$ and $y_0 - Bx_0 \in Ax_0$, because A is maximal monotone. We have, therefore, proved that $y_0 \in R(A + B)$, as claimed. □

Theorem 1.141 *Let X be reflexive and let $F : X \to X^*$ be the duality mapping of X. Let A be any monotone subset of $X \times X^*$. Then A is maximal monotone in $X \times X^*$ if and only if, for any $\lambda > 0$ (equivalently, for some $\lambda > 0$), $R(A + \lambda F)$ is all of X^*.*

Proof *If* part. Assume that $R(A + \lambda F) = X^*$, for some $\lambda > 0$. We have to show that A is maximal monotone. Suppose that this is not the case, and that there exists $[x_0, y_0] \in X \times X^*$ such that $[x_0, y_0] \overline{\in} A$ and

$$(x - x_0, y - y_0) \geq 0, \quad \text{for all } [x, y] \in A. \tag{1.106}$$

By hypothesis, there exists $[x_1, y_1] \in A$ such that $\lambda F x_1 + y_1 = \lambda F x_0 + y_0$.

Substituting $[x_1, y_1]$ for $[x, y]$ in inequality (1.106), we obtain

$$\big(x_1 - x_0, F(x_1) - F(x_0)\big) \le 0.$$

Since, by Theorem 1.105, the spaces X and X^* can be chosen strictly convex, the above inequality implies that $x_1 = x_0$, so that we have $[x_0, y_0] = [x_1, y_1] \in A$, which is a contradiction.

Only if part. Renorming the spaces X and X^* (see Theorem 1.105), we may assume without loss of generality that X as well as X^* are strictly convex. Then the duality mapping $F : X \to X^*$ is monotone, single-valued, demicontinuous and bounded, and it satisfies condition (1.104). Then we may apply Corollary 1.140, where $B = \lambda F$, to conclude the proof. □

Theorem 1.141 is due to Rockafellar [27]. When specialized to the case when X is a Hilbert space, this theorem yields the classical theorem of Minty.

Corollary 1.142 *Let X be reflexive and let B be monotone, everywhere defined and demicontinuous from X to X^*. Then B is maximal monotone.*

Proof Suppose that B is not maximal monotone. Then we may find x_0 in X and y_0 in X^* such that $y_0 \neq Bx_0$, and

$$(Bx - y_0, x - x_0) \ge 0 \quad \text{for every } x \in X = D(B).$$

For each $t \in [0, 1]$, we set $x_t = tx_0 + (1 - t)u$, where u is arbitrary in X. Then

$$(Bx_t - y_0, x_0 - u) \le 0 \quad \text{for all } t \in [0, 1].$$

In particular, it follows from the demicontinuity of B that

$$(Bx_0 - y_0, x_0 - u) \le 0.$$

Thus, the arbitrariness of $u \in X$ implies that $y_0 = Bx_0$, which contradicts the assumption. □

Theorem 1.143 *Let X be reflexive and let A be a maximal monotone and coercive operator from X to X^*. Then $R(A) = X^*$.*

Proof Let x_0^* be any fixed element of X^*. Using the renorming theorem, we may assume that X and X^* are strictly convex. Then it follows by Theorem 1.141 that, for every $\lambda > 0$, the equation

$$\lambda F x_\lambda + A x_\lambda \ni x_0^* \tag{1.107}$$

has at least one solution $x_\lambda \in X$. Using the monotonicity of A, we see, after multiplying equation (1.107) by $(x_\lambda - x_0)$ (x_0 is the element arising in condition (1.103)), that

$$\lambda \|x_\lambda\|^2 + (Ax_\lambda, x_\lambda - x_0) \le \|x_0^*\| \, \|x_\lambda - x_0\| + \lambda \|x_\lambda\| \, \|x_0\|.$$

Since A is coercive, this implies that $\{\|x_\lambda\|\}$ is bounded for $\lambda \to 0$. We may, therefore, assume that $\{x_\lambda\}$ converges weakly to x_0 in X and $\{Ax_\lambda\}$ converges strongly to x_0^* in X^* as $\lambda \to 0$. Thus, by the monotonicity of A, we find that

$$(x_0^* - y, x_0 - x) \geq 0 \quad \text{for all } [x, y] \in A,$$

and, since A is maximal monotone, we may infer that $Ax_0 \ni x_0^*$. Thus, we have shown that the range of A is all of X^*. □

Theorem 1.144 *Let X be a real Banach space and let A be any monotone subset of $X \times X^*$. If $x_0 \in D(A)$ is an interior point of $D(A)$, then A is locally bounded at x_0.*

Proof Assuming that A is not locally bounded at x_0, we derive a contradiction. Let $\{x_n\} \subset X$ and $y_n \in Ax_n$ be such that $\|y_n\| = \lambda_n \to +\infty$ and $x_n \to x_0$ as $n \to \infty$. Define

$$\alpha_n = \max\left\{\lambda_n^{-1},\ \|x_n - x_0\|^{\frac{1}{2}}\right\}.$$

Obviously, $\alpha_n \to 0$, $\alpha_n\lambda_n \geq 1$ and $\alpha_n \geq \alpha_n^{-1}\|x_n - x_0\| \to 0$ for $n \to \infty$. Let z be any element of X and let $u_n = x_0 + \alpha_n z$. Then, for n large enough, $u_n \in D(A)$. Let $v_n \in Au_n$. First, we show that $\{v_n\}$ is bounded in X^*. Let $\rho > 0$ be such that $x_0 + \rho z \in D(A)$ and let $w_0 \in A(x_0 + \rho z)$. Since A is monotone in $X \times X^*$, we have

$$(v_n - w_0, z)(\alpha_n - \rho) \geq 0.$$

If $\alpha_n < \rho$, this implies $(v_n, z) \leq (w_0, z)$. Hence $\{(v_n, z)\}$ is bounded. Now, let $x_0 + t_0 x \in D(A)$ and $w \in A(x_0 + t_0 x)$, where x is an arbitrary element of X. Once again, using the monotonicity of A, we obtain $0 \leq (v_n - w, \alpha_n z - t_0 x) = \alpha_n(v_n, z) - t_0(v_n, x) - (w, \alpha_n z - t_0 x)$ and, therefore,

$$\limsup_{n\to\infty}(v_n, x) \leq (w, x). \tag{1.108}$$

Using the uniform boundedness Theorem 1.5, relation (1.108) implies that $\{v_n\}$ is bounded in X^*. Next, the inequality $(y_n - v_n, x_n - u_n) \geq 0$ implies

$$\begin{aligned}(y_n, z) &\leq \left(\frac{x_n - x_0}{\alpha_n}, y_n\right) - \left(\frac{x_n - x_0}{\alpha_n} - z, v_n\right) \\ &\leq \|y_n\|\frac{\|x_n - x_0\|}{\alpha_n} + M\left(\frac{\|x_n - x_0\|}{\alpha_n} + \|z\|\right) \leq \lambda_n\alpha_n + M\big(\alpha_n + \|z\|\big),\end{aligned}$$

where M is a positive constant independent of n and z. Therefore,

$$\limsup_{n\to\infty}\left(\frac{y_n}{\alpha_n\lambda_n}, z\right) \leq 1 + M\|z\| < \infty \quad \text{for every } z \in X,$$

which contradicts the uniform boundedness theorem (see Theorem 1.5), since

$$\frac{\|y_n\|}{\lambda_n\alpha_n} = \frac{1}{\alpha_n} \to \infty \quad \text{as } n \to \infty.$$

□

Remark 1.145 Theorem 1.139 is due to Rockafellar [26].

We assume now that the space X is reflexive and strictly convex along with the dual X^*. Let A be a maximal monotone subset of $X \times X^*$. Since the duality mapping $F : X \to X^*$ is demicontinuous, it follows by Corollary 1.140 that, for every $x \in X$ and $\lambda > 0$, the equation

$$F(x_\lambda - x) + \lambda A x_\lambda \ni 0 \tag{1.109}$$

has at least one solution x_λ. Since X is strictly convex, F^{-1} is single-valued and along with the monotonicity of F and A this implies the uniqueness of x_λ.

We set

$$x_\lambda = J_\lambda x; \quad A_\lambda x = \lambda^{-1} F(x - x_\lambda). \tag{1.110}$$

In Proposition 1.146, we collect for later use some properties of the operators J_λ and A_λ defined above.

First, we notice that the maximality of A implies that, for every $x \in D(A)$, Ax is a closed and convex subset of X^*. Hence, if the space X^* is strictly convex, there exists a unique element of minimum norm in Ax, which will be denoted $A^0 x$. In other words,

$$\|A^0 x\| = |Ax| = \inf\{\|y\|;\ y \in Ax\}.$$

Proposition 1.146 *For every $\lambda > 0$, we have the following*:

(i) A_λ *is monotone, bounded on bounded subsets and demicontinuous from X to X^*. If X^* is uniformly convex, then A_λ is continuous.*
(ii) $\|A_\lambda x\| \le |Ax|$ *for all* $x \in D(A)$.
(iii) $\lim_{\lambda \to 0} J_\lambda x = x$ *for all* $x \in \overline{\mathrm{conv}\, D(A)}$.
(iv) *For every $x \in D(A)$, $A_\lambda x \to A^0 x$ weakly in X^* for $\lambda \to 0$. If X^* is uniformly convex, then the convergence is strong.*
(v) *If, for some sequence $\lambda_n \to 0$, $x_n \to x$ strongly and $A_{\lambda_n} x_n \to y$ weakly, then* $[x, y] \in A$.
(vi) *If $X = H$ is a Hilbert space, then $J_\lambda = (I + \lambda A)^{-1}$ is a contraction on H, and A_λ is Lipschitzian with constant λ^{-1}.*

Proof The proof of the monotonicity, boundedness, demicontinuity as well as of (ii), (iii), (v), and (vi) can be found in the first author's book [3], p. 42, so it will be omitted. Here, we prove continuity of A_λ (under the assumption that X^* is uniformly convex) and property (iv). Let $\{x_n\} \subset X$ be strongly convergent to x and let $u_n = J_\lambda x_n$, $v_n = A_\lambda x_n \in A u_n$ and $y_n = u_n - x_n$. We have

$$F y_n + \lambda v_n = 0.$$

Since A and F are monotone, the latter yields

$$(y_n - y_m, F y_n - F y_m) \le C \|x_n - x_m\|, \tag{1.111}$$

because $\{v_n\}$ is bounded. On the other hand, since A_λ is demicontinuous, we have

$$Fy_n \to F(J_\lambda x - x) \quad \text{weakly in } X^*,$$
$$y_n \to J_\lambda x - x \quad \text{weakly in } X.$$

Then, by (4.14) and Lemma 1.3, p. 42, in the book cited above, it follows that

$$\lim_{n\to\infty} \left(\|Fy_n\|^2 = (Fy_n, y_n)\right) = \left\|F(J_\lambda x - x)\right\|^2.$$

Since the space X^* is uniformly convex, the latter implies via Proposition 1.79 that $\lim_{n\to\infty} Fy_n = F(J_\lambda x - x)$. Hence, $\lim_{n\to\infty} A_\lambda x_n = A_\lambda x$, as claimed.

(iv) Let $x \in D(A)$ be fixed. Since $\{A_\lambda x\}$ is bounded in X^*, there exists $\xi \in X^*$ such that $A_{\lambda_n} x \to \xi$ weakly in X^* on some sequence $\lambda_n \to 0$. Since $x_{\lambda_n} \to x$ and $A_{\lambda_n} x \in Ax_{\lambda_n}$, we may infer that $\xi \in Ax$ (A is strongly–weakly closed in $X \times X^*$ as consequence of the maximal monotonicity).

Next, by (ii) we see that $\|\xi\| \le \|A^0 x\|$ and hence $\xi = A^0 x$. We have, therefore, proved that $A_\lambda x \to A^0 x$ weakly in X^* and $\|A_\lambda x\| \to \|A^0 x\|$ for $\lambda \to 0$. If X^* is uniformly convex, the latter implies via Proposition 1.85

$$A_\lambda x \to A^0 x \quad \text{strongly in } X^*. \tag{1.112}$$

□

1.4.2 Linear Evolution Equations in Banach Spaces

Let X be a (real or complex) Banach space with norm $\|\cdot\|$ and dual X^*. By $L(X)$, we denote in the sequel the algebra of linear continuous operators on X.

Consider the Cauchy problem

$$\begin{aligned} &x'(t) = A(t)x(t) + f(t), \quad 0 \le t \le T, \\ &x(0) = x_0, \end{aligned} \tag{1.113}$$

where $f \in L^1(0, T; X)$ and $x_0 \in X$ are given and $\{A(t);\ 0 \le t \le T\}$ is a family of closed and densely defined linear operators with domain and range both in X.

Consider also the homogeneous equation

$$x'(t) = A(t)x(t), \quad 0 \le t \le T. \tag{1.114}$$

Definition 1.147 The Cauchy problem for (1.114) is said to be *well posed* if there exists a function $U : \Delta = \{(s,t);\ 0 \le s \le t \le T\} \to L(X)$ having the following properties:

(i) For each $x_0 \in X$, the function $(t,s) \to U(t,s)x_0$ is continuous on Δ (that is, U is strongly continuous on Δ).

(ii) $U(s,s) = I$ (the identity operator) for every $s \in [0,T]$.
(iii) $U(t,s)U(s,\tau) = U(t,\tau)$ for $0 \le \tau \le s \le t \le T$.
(iv) For each $s \in [0,T]$, there exists a densely linear subspace E_s of X such that, for each $x_0 \in E_s$, the function $t \to U(t,s)x_0$ is continuously differentiable on $[s,T]$ and

$$\frac{\partial}{\partial t} U(t,s)x_0 = A(t)U(t,s)x_0, \quad s \le t \le T. \tag{1.115}$$

There is $C > 0$ such that

$$\|U(t,s)\|_{L(X)} \le C \quad \text{for } (t,s) \in \Delta.$$

If the conditions of Definition 1.147 are satisfied, we say that the family $\{A(t); 0 \le t \le T\}$ generates the *evolution operator* $U(t,s)$.

If the Cauchy problem (1.113) is well posed, then by the solution to the nonhomogeneous equation (1.113) we mean the continuous function $x : [0,T] \to X$ given by the formula

$$x(t) = U(t,0)x_0 + \int_0^t U(t,s)f(s)\,\mathrm{d}s, \quad 0 \le t \le T. \tag{1.116}$$

This is the so-called "*mild*" *solution* to (1.114). By a *strong solution* to (1.113), we mean an absolutely continuous function x on $[0,T]$ which is almost everywhere differentiable on $]0,T[$ and satisfies (1.113) a.e. on $]0,T[$.

It is well known that every strong solution to (1.113) can be written in this form; (1.113) itself may not be satisfied by x given by the variation of the constant formula (1.116), however.

Now, we point out some standard circumstances when the Cauchy problem for (1.114) is well posed.

Time-Independent Equations If $A(t) = A$ is independent of t, then the conditions of Definition 1.147 are satisfied if and only if A is the infinitesimal generator of a strongly continuous semigroup $\{S(t); t \ge 0\}$ of linear continuous operators on X (*semigroup of class* C_0). By the classical Hille–Yosida Theorem (see, for instance, Yosida [30], p. 246) this happens if and only if A is closed, densely defined and there is $\omega \in \mathbb{R}$, such that

$$\|(\lambda I - A)^{-n}\|_{L(X)} \le M(\operatorname{Re}\lambda - \omega)^{-n} \quad \text{for } \operatorname{Re}\lambda > \omega,\ n = 1,2,\ldots.$$

In this case, the evolution operator associated to A exists for all $0 \le s \le t < \infty$ and has the form: $U(t,s) = S(t-s)$. For each $x_0 \in D(A)$ (the domain of A), the function $x(t) = S(t)x_0$ is continuously differentiable and satisfies (1.114) on $\mathbb{R}^+$. The operator A is called the infinitesimal generator of the semigroup $S(t)$. If A is *dissipative*, that is, $\operatorname{Re}(Ax, Fx) \le 0$, $\forall x \in D(A)$, where $F : X \to X^*$ is the duality mapping of X, then the semigroup $S(t)$ is contractant, that is,

$$\|S(t)\|_{L(X)} \le 1, \quad \forall t \ge 0.$$

If the operator A is dissipative and $R(I - A) = X$ (that is, A is m-dissipative), then A generates a *contraction semigroup* on X. Sometimes we denote by e^{At} the semigroup $S(t)$ generated by A.

Finally, we notice that if A satisfies the condition

$$\left\|(\lambda I - A)^{-1}\right\|_{L(X)} \le \frac{C}{|\lambda|} \quad \text{for } |\arg\lambda| > \theta > \frac{\pi}{2}, \tag{1.117}$$

then A generates a semigroup $S(t)$ which is *analytic* in t on $\mathbb{R}$. (See, e.g., [20, 24, 30].) It is worth noting that in this case, for each $x_0 \in X$, $x(t) = S(t)x_0$ is continuously differentiable on $\mathbb{R}^+$ and satisfies (1.114) on all of $\mathbb{R}^+$. Moreover, we have the following proposition.

Proposition 1.148 *Let X be a Hilbert space and let A satisfy condition* (1.117). *Then, for each $f \in L^2(0, T; X)$ and $x_0 \in D(A)$, problem* (1.113) *has a unique strong solution $x \in W^{1,2}([0, T]; X)$ satisfying*

$$\int_0^T \left\|x'(t)\right\|^2 \mathrm{d}t \le C\left(\|Ax_0\|^2 + \int_0^T \left\|f(t)\right\|^2 \mathrm{d}t\right), \tag{1.118}$$

where C is independent of x_0 and f.

Proof To prove the proposition, it suffices to verify estimate (1.118) for any strong solution x to (1.113). If x is such a solution, let us denote by $\tilde{x}$ the solution to

$$\begin{aligned} &\tilde{x}' = A\tilde{x} + f_T, \quad t \ge 0, \\ &\tilde{x}(0) = x_0, \end{aligned} \tag{1.119}$$

where $f_T(t) = f(t)$ for $0 \le t \le T$ and $f_T(t) = 0$ for $T \le t < \infty$. Let $\alpha > 0$ be such that $\|\mathrm{e}^{At}\|_{L(X)} \le M\mathrm{e}^{\frac{\alpha t}{2}}$ for all $t \ge 0$ (such a constant α always exists). For $\lambda = -\alpha + i\xi$, $\xi \in \mathbb{R}$, we set

$$\begin{aligned} \hat{x}(\lambda) &= \int_0^\infty \mathrm{e}^{\lambda t}\tilde{x}(t)\,\mathrm{d}t, \\ \hat{f}_T(t) &= \int_0^\infty \mathrm{e}^{\lambda t} f_T(t)\,\mathrm{d}t. \end{aligned}$$

By (1.119), it follows that

$$(\lambda I + A)\hat{x}(\lambda) = -x_0 - \hat{f}_T(t)$$

and, therefore,

$$\lambda\hat{x}(\lambda) + x_0 = (\lambda I + A)^{-1}Ax_0 + \lambda(\lambda I + A)^{-1}\hat{f}_T(t) - \hat{f}_T(t).$$

Thus, by (1.117),

$$\left\|\lambda\hat{x}(\lambda) + x_0\right\| \le C\|Ax_0\|\,|\alpha - i\xi|^{-1} + C\left\|\hat{f}_T(\lambda)\right\|, \quad \xi \in \mathbb{R}. \tag{1.120}$$

Recalling that

$$\lambda\hat{x}(\lambda) + x_0 = -\int_0^\infty e^{-\alpha t} e^{i\xi t}\tilde{x}'(t)\,dt, \quad \xi \in \mathbb{R},$$

it follows, by (1.120) and the Parseval formula, that

$$\int_0^\infty e^{-2\alpha t}\left\|\tilde{x}'(t)\right\|^2 dt \le C_1\left(\|Ax_0\|^2 + \int_0^\infty e^{-2\alpha t}\left\|f_T(t)\right\|^2 dt\right).$$

Since $\tilde{x}(t) = x(t)$ for $t \in [0, T]$, the latter yields (1.118). □

Time-Dependent "Parabolic" Equations, Hilbert Theory Let H be a real Hilbert space and let V be another real Hilbert space with the dual V' such that $V \subset H \subset V'$ and such that the inclusion mapping of V into H is continuous and densely defined. The norms in V and H will be denoted by $\|\cdot\|$ and $|\cdot|$, respectively. Denote by $\|\cdot\|_*$ the norm (dual) of V' and by (v_1, v_2) the value of $v_1 \in V'$ in $v_2 \in V$; if $v_1, v_2 \in H$, this is the ordinary inner product in H. We are given a family of linear operators $A(t) : V \to V'$, $0 \le t \le T$, which are assumed to satisfy the following conditions:

(j) *For every $u \in V$, the function $t \to A(t)u$ is strongly measurable on* $[0, T]$.
(jj) *For every $t \in [0, T]$, $A(t)$ is continuous from V to V' and there exists $C > 0$ such that $\|A(t)\|_{L(V,V')} \le C$, a.e. $t \in [0, T]$.*
(jjj) *There are $\omega > 0$ and $\alpha \in \mathbb{R}$, such that*

$$\big(A(t)y, y\big) + \omega\|y\|^2 \le \alpha|y|^2 \quad \textit{for all } y \in V, \text{ a.e. } t \in\,]0, T[. \tag{1.121}$$

Proposition 1.149 *Let $x_0 \in H$ and $f \in L^2(0, T; V')$ be given. Then, under assumptions* (j)–(jjj), *there exists one and only one function $x \in W(0, T)$ satisfying*

$$\begin{aligned} &x'(t) = A(t)x(t) + f(t) \quad \textit{a.e. } t \in\,]0, T[,\\ &x(0) = x_0. \end{aligned} \tag{1.122}$$

If $A(t) = A$ is independent of t, $Ax_0 \in H$, $f \in L^2(0, T; H)$, then $x \in W^{1,2}([0, T]; H)$ and

$$\int_0^T \left|x'(t)\right|^2 dt \le C\left(|Ax_0|^2 + \int_0^T \left|f(t)\right|^2 dt\right), \tag{1.123}$$

where C is independent on x_0 and f.

The proof can be found in Lions and Magenes [22], Chap. 4. The first part of the proposition remains valid for nonlinear hemicontinuous monotone operators $A(t) : V \to V'$ (Lions [21], Barbu [3]). The second part follows by Proposition 1.148, since, as is easily seen, if $A(t) \equiv A$ is independent on t and satisfies (1.121), then A is the infinitesimal generator of an analytic semigroup on H.

In applications, X is usually a space of functions defined on a domain Ω of the Euclidean space $\mathbb{R}^n$, and $A(t)$ is a linear partial differential operator on Ω with null boundary conditions. Another motivation for considering infinite-dimensional systems of the form (1.113) comes from differential functional systems. Here, we present briefly some important examples of this type.

Parabolic Equations with Dirichlet Conditions Let Ω be a bounded and open domain of $\mathbb{R}^n$ with a sufficient smooth boundary Γ. We consider on Ω the second order differential operator defined by

$$Ly = -\sum_{i,j=1}^{n} (a_{ij} y_{x_j}) x_i + ay,$$

where $a_{ij}, a_i \in C^1(\overline{\Omega})$, $a \in L^\infty(\Omega)$, $a \geq 0$ in Ω, and

$$\sum_{i,j=1}^{n} a_{ij}\xi_i\xi_j \geq \omega|\xi|^2, \quad \text{a.e. } x \in \Omega,\ \forall \xi \in \mathbb{R}^n, \tag{1.124}$$

where $\omega > 0$. (Here, the subscript x_i denotes partial differentiation with respect to x_i.) According to a classical result due to Agmon and Nirenberg, the operator A defined by $Ay = -Ly$ for $y \in D(A) = W_0^{1,p}(\Omega) \cap W^{2,p}(\Omega)$ is the *infinitesimal generator* of a contraction semigroup of class C_0 on $L^p(\Omega)$, $1 \leq p \leq \infty$ [24].

Parabolic Equations with Homogeneous Neumann Conditions In the space $L^2(\Omega)$, we define the linear operator $Ay = -Ly$ with the domain

$$D(A) = \left\{ y \in H^2(\Omega);\ \alpha y + \frac{\partial y}{\partial \nu} = 0 \ \text{ on } \Gamma \right\},$$

where α is a nonnegative constant, and where $\frac{\partial}{\partial \nu}$ denotes the outward normal derivative corresponding to L. The operator A is m-dissipative on $L^2(\Omega)$ and, therefore, it generates a contraction semigroup e^{At} on $L^2(\Omega)$. Moreover, since A is self-adjoint, the semigroup e^{At} is analytic.

Parabolic Equations with Mixed Boundary Value Conditions Let Ω be a bounded and open domain of $\mathbb{R}^n$, whose boundary Γ consists of two disjoint smooth parts Γ_1 and Γ_2. Assume that conditions (1.124) hold. We set $V = \{y \in H^1(\Omega);\ y = 0 \text{ in } \Gamma_1\}$ and define the operator $A \in L(V, V')$

$$(Ay, z) = \int_\Omega a_{ij} y_{x_j} z_{x_i}\, dx + \int_\Omega ayz\, dx \quad \text{for all } z \in H^1(\Omega).$$

Assumptions (j)–(jjj) are satisfied and therefore the operator $-A$ with the domain $D(A) = \{y \in V;\ Ay \in L^2(\Omega)\}$ generates a C_0-semigroup on $L^2(\Omega)$. On the other hand, by Green's formula we see that $\frac{\partial y}{\partial \nu} = 0$ in Γ_2 for all $y \in D(A)$. (Since

$y \in H^1(\Omega)$ and $Ay \in L^2(\Omega)$, $\frac{\partial y}{\partial \nu}$ may be defined as an element of $H^{-\frac{1}{2}}(\Gamma)$.) In other words, $D(A)$ may be regarded as the set of all functions $y \in H^1(\Omega)$ which satisfy the boundary value conditions $y = 0$ in Γ_1, $\frac{\partial y}{\partial \nu} = 0$ in Γ_2.

1.5 Problems

1.1 Let A, B be two nonvoid disjoint convex sets in a linear space X. Show that there exist two disjoint convex sets A_0, B_0 such that $A \subset A_0$, $B \subset B_0$, $A_0 \cup B_0 = X$.

Hint. Consider a maximal element of the family $\mathscr{F}$ of all pairs (U, V) of disjoint convex sets such that $A \subset U$, $B \subset V$, endowed with the order relation given by inclusion using the Zorn Lemma.

1.2 Let A, B be two convex sets. Show that $\operatorname{co}(A \cup B) = \bigcup_{\substack{x \in A \\ y \in B}} [x, y]$.

Hint. Use formula (1.23).

1.3 Find p_A, $\operatorname{Dom}(p_A)$, A^{ri}, A^{ac}, A^{rb} of the following sets in R^2:

(i) $A = \{(x_1, x_2);\, x_1^2 + x_2^2 \le 1 \text{ if } x_1 \le 0 \text{ and } |x_2| \le 1 \text{ if } x_1 > 0\}$
(ii) $A = \{(x_1, x_2);\, x_1 x_2 \le 1\}$
(iii) $A = \{(x_1, x_2);\, x_1 + |x_2| \le 1 \text{ and } x_1 \ge 0\}$.

Hint. (i) $p_A(x_1, x_2) = (x_1^2 + x_2^2)^{\frac{1}{2}}$ if $x_1 \le 0$ and $p_A(x_1, x_2) = |x_2|$ if $x_1 > 0$, $\operatorname{Dom}(p_A) = \mathbb{R}$.

(ii) $p_A(x_1, x_2) = (x_1, x_2)^{\frac{1}{2}}$ if $x_1 x_2 \ge 0$, and $p_A(x_1, x_2) = \infty$ if $x_1 x_2 < 0$, $\operatorname{Dom}(p_A) = \{(x_1, x_2);\ x_1 x_2 \ge 0\}$.

(iii) $p_A(x_1, x_2) = x_1 + |x_2|$ if $x_1 \ge 0$ and $p_A(x_1, x_2) = \infty$ if $x_1 < 0$.

1.4 Are the equalities of Propositions 1.19 and 1.22 and Corollaries 1.21 and 1.23 true for the sets in Problem 1.3?

1.5 Let $A_1, A_2, \ldots, A_n$ be convex sets which contain the origin. Find p_A if $A = \bigcap_{i=1}^n A_i$ in terms of p_{A_i}, $i = \overline{1, n}$.

Hint. $p_A = \max_{1 \le i \le n} p_{A_i}$.

1.6 Let $\mathbb{R}^\infty$ be the linear space of all sequences of real numbers having a finite number of elements different to zero endowed with the Euclidean norm $\|x\| = (\sum_{i=1}^\infty x_i^2)^{\frac{1}{2}}$, $x = (x_i)_{i \in \mathbb{N}} \in \mathbb{R}^\infty$. Show that the set $S = \{m^{-1}(\delta_{nm})_{n \in \mathbb{N}^*};\ m \in \mathbb{N}^*\} \cup \{0\}$ is a compact set although its convex hull if not compact.

Hint. We denote $e_m = (\delta_{nm})_{n \in \mathbb{N}^*}$, $m \in \mathbb{N}^*$. The sequence $(m^{-1} e_m)_{m \in \mathbb{N}^*} \to 0$ in $\mathbb{R}^\infty$ since $\|m^{-1} e_m\| = m^{-1}$. Hence, S is compact. Now, let us consider a sequence

$(x_n)_{n\in\mathbb{N}^*}$ in conv S. By (1.23), we have $x_n = \sum_{m=1}^{\alpha_n} \lambda_{nm} m^{-1} e_m$, where $\lambda_{nm} \geq 0$ and $\sum_{m=1}^{\alpha_n} \lambda_{nm} = 1$. Take $\lambda_{nm} = \frac{2^n}{2^m(2^n-1)}$, $n, m \in \mathbb{N}^*$, $\alpha_n = n$, $n \in \mathbb{N}^*$. Then the sequence $(\lambda_{nm})_{n\in\mathbb{N}^*} \to \frac{1}{2^n}$ for each $n \in \mathbb{N}^*$. Thus, $(x_n)_{n\in\mathbb{N}}$ converges in ℓ_2 to an element which is not in $\mathbb{R}^\infty$. Hence, any subsequence of the sequence $(x_n)_{n\in\,\mathbb{N}^*}$ is not convergent in conv S.

1.7 Show that the norm in a linear normed space X is generated by a *semi-inner product*, which is an application $\langle\cdot,\cdot\rangle : X \times X \to \mathbb{R}$ with the following properties:

(i) $\langle x, x\rangle \geq 0$ for all $x \in X$ and $\langle x, x\rangle = 0$ if and only if $x = 0$
(ii) $\langle a_1x_1 + a_2x_2, y\rangle = a_1\langle x_1, y\rangle + a_2\langle x_2, y\rangle$, for all $a_1, a_2 \in \mathbb{R}$, $x_1, x_2, y \in X$
(iii) $\langle x, y\rangle^2 \leq \langle x, x\rangle\langle y, y\rangle$, for all $x, y \in X$.

Hint. Define $\langle x, y\rangle = (fy)(x)$, $x, y \in X$, where f is a selection of the duality mapping F (see Definition 1.99), that is, $fy \in Fy$ for every $y \in X$.

1.8 Prove that the Dieudonné criterion (Corollary 1.61) is also in order if $A_\infty \cap B_\infty$ is a closed linear subspace.

Hint. It suffices to consider the quotient space with respect to the linear subspace $A_\infty \cap B_\infty$.

1.9 Let A be a closed convex set. Show that $x \in A_\infty$ if and only if $A + tx \subset A$ for all $t \geq 0$. In particular, $A + x \subset A$ for all $x \in A_\infty$ and so any semi-straight line starting from an element in A which has the direction of an element in A_∞ is contained in A.

Hint. Use formula (1.37).

1.10 Show that in $C[0, 1]$ any weakly convergent sequence is pointwise convergent.

Hint. Consider the Dirac functional defined by $\delta_t(x) = x(t)$ for all $x \in C[0, 1]$, where $t \in [0, 1]$. Obviously, $\delta_t \in (C[0, 1])^*$ for every $t \in [0, 1]$. Now apply Proposition 1.65(iv).

References

1. Asplund E (1967) Averaged norms. Isr J Math 5:227–233
2. Asplund E (1968) Fréchet-differentiability of convex functions. Acta Math 121:31–47
3. Barbu V (1976) Nonlinear semigroups and evolution equations in Banach spaces. Noordhoff International Publishing, Leyden. Ed Acad, Bucureşti, Romania
4. Brezis H (1973) Opérateurs maximaux monotones et semigroupes de contractions dans les espaces de Hilbert. Math Studies, vol 5. North Holland, Amsterdam
5. Brooks JK, Dinculeanu N (1977) Weak compactness in spaces of Bochner integrable functionsand applications. Adv Math 24:172–188

6. Browder F (1968) Nonlinear operators and nonlinear equations of evolution in Banach spaces. In: Proc Amer math soc symposium on nonlinear functional analysis, Chicago
7. Day M (1958) Normed linear spaces. Springer, Berlin
8. Dedieu J-P (1978) Critères de femeture pour l'image d'un fermé non convexe par une multiplication. C R Acad Sci Paris 287:941–943
9. Diestel J (1975) Geometry of Banach spaces. Selected topics. Lecture notes in mathematics. Springer, Berlin
10. Dieudonné J (1966) Sur la séparation des ensembles convexes. Math Ann 163:1–3
11. Dinculeanu N (1967) Vector measures. Pergamon, London
12. Edwards RE (1965) Functional analysis. Holt, Reinhart and Wiston, New York
13. Eggleston H (1958) Convexity. Cambridge University Press, Cambridge
14. Gwinner J (1977) Closed images of convex multivalued mappings in linear topological spaces with applications. J Math Anal Appl 60:75–86
15. Höenig ChS (1975) Volterra Stieltjes integral equations. Mathematical studies. North Holland, Amsterdam
16. Holmes RB (1975) Geometric functional analysis and its applications. Springer, Berlin
17. James RC (1964) Characterization of reflexivity. Stud Math 23:205–216
18. Klee V (1969) Separation and support properties of convex sets. In: Balakrishnan AV (ed) Control theory and the calculus of variations. Academic Press, New York, pp 235–303
19. Köthe G (1969) Topological vector spaces. I. Springer, Berlin
20. Krein SG (1967) Linear differential equations in Banach spaces. Nauka, Moscow (Russian)
21. Lions JL (1969) Quelques méthodes de résolution des problèmes aux limites nonlinéaires. Dunod, Gauthier–Villars, Paris
22. Lions JL, Magenes E (1970) Problèmes aux limites non homogènes et applications. Dunod, Gauthier–Villars, Paris
23. Nicolescu M (1958/1960) Mathematical analysis I, II. Editura Tehnică, Bucureşti (Romanian)
24. Pazy A (1983) Semigroups of linear operators and applications to partial differential equations. Springer, New York
25. Precupanu A (1976) Mathematical analysis. Real functions. Editura Didactică, Bucureşti (Romanian)
26. Rockafellar RT (1969) Local boundedness of nonlinear monotone operators. Mich Math J 16:397–407
27. Rockafellar RT (1970) On the maximality of sums of nonlinear operators. Trans Am Math Soc 149:75–88
28. Valentine FA (1964) Convex sets. McGraw-Hill, New York
29. van Dulst D (1978) Reflexive and superreflexive Banach spaces. Mathematical centre tracts, vol 102. Mathematisch Centrum, Amsterdam
30. Yosida K (1980) Functional analysis. Springer, Berlin

Chapter 2
Convex Functions

In this chapter, the basic concepts and the properties of extended real-valued convex functions defined on a real Banach space are described. The main topic, however, is the concept of *subdifferential* and its relationship to maximal monotone operators. In addition, concave–convex functions are examined because of their importance in the duality theory of minimization problems as well as in min-max problems.

2.1 General Properties of Convex Functions

We develop here the basic concepts and results on convex functions which were briefly presented in Chap. 1.

2.1.1 Definitions and Basic Properties

In Chap. 1, we have already become familiar with convex functions (see Definition 1.32) and their relationship to convex sets. In this section, the concept of convex function on a real linear space X will be extended to include functions with values in $\overline{\mathbb{R}} = [-\infty, +\infty]$ (*extended real-valued functions*).

Definition 2.1 The function $f : X \to \overline{\mathbb{R}}$ is called *convex* if the inequality

$$f\bigl(\lambda x + (1-\lambda)y\bigr) \le \lambda f(x) + (1-\lambda) f(y) \tag{2.1}$$

holds for every $\lambda \in [0, 1]$ and all $x, y \in X$ such that the right-hand side is well defined. The function f is called *strictly convex* if an inequality strictly holds in inequality (2.1) for every $\lambda \in]0, 1[$ and for all pairs of distinct points x, y in X with $f(x) < \infty$ and $f(y) < \infty$.

V. Barbu, T. Precupanu, *Convexity and Optimization in Banach Spaces*, Springer Monographs in Mathematics,
DOI 10.1007/978-94-007-2247-7_2,

The function $g : X \to \overline{\mathbb{R}}$ is said to be (*strictly*) *concave* if the function $-g$ is (strictly) convex. It should be observed that if f is convex, then the inequality

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \le \sum_{i=1}^{n} \lambda_i f(x_i), \quad \lambda_i \ge 0, \ \sum_{i=1}^{n} \lambda_i = 1$$

holds for all $x_1, \ldots, x_n$ in X, for which the right-hand side makes sense.

Another consequence of convexity of $f : X \to \overline{\mathbb{R}}$ is the convexity of the *level sets*,

$$\{x \in X; \ f(x) \le \lambda\},$$

where $\lambda \in \overline{\mathbb{R}}$. However, as is readily seen, the functions endowed with this property are not necessarily convex. Such functions are called *quasi-convex*.

The function f is called *proper convex* if $f(x) > -\infty$ for every $x \in X$, and if f is not the constant function $+\infty$ (that is, $f \not\equiv +\infty$). Given any convex function $f : X \to \overline{\mathbb{R}}$, we denote by $\operatorname{Dom}(f)$ (sometimes $\operatorname{dom} f$) the convex set

$$\operatorname{Dom}(f) = \{x \in X; \ f(x) < +\infty\}. \tag{2.2}$$

Such a set $\operatorname{Dom}(f)$ is called the *effective domain* of f. If f is proper, then $\operatorname{Dom}(f)$ is the finiteness domain of f. Conversely, if A is a nonempty convex subset of X and if f is a finite and convex function on A, then one can obtain a proper convex function on X by setting $f(x) = +\infty$ if $x \in X \setminus A$. Using all this, we are able to introduce an important example of convex function. Given any nonempty subset A of X, the function I_A on X, defined by

$$I_A(x) = \begin{cases} 0, & \text{if } x \in A, \\ +\infty, & \text{if } x \overline{\in} A, \end{cases} \tag{2.3}$$

is called the *indicator function* of A.

The characterization of convexity follows.

Proposition 2.2 *The subset A of X is convex if and only if its indicator function I_A is convex.*

Let $f : X \to \overline{\mathbb{R}}$ be any extended real-valued function on X. The set

$$\operatorname{epi} f = \{(x, \alpha); \ x \in X, \ \alpha \in \mathbb{R}, \ f(x) \le \alpha\} \tag{2.4}$$

is called the *epigraph* of f. The set

$$\operatorname{hypo} f = \{(x, \alpha); \ x \in X, \ \alpha \in \mathbb{R}, \ f(x) \ge \alpha\} \tag{2.5}$$

is called the *hypograph* of f.

Proposition 2.3, which follows, demonstrates that the above-mentioned theory of convex functions and that of convex sets overlap considerably.

Proposition 2.3 *A function $f : X \to \overline{\mathbb{R}}$ is convex if and only if its epigraph is a convex subset of $X \times \mathbb{R}$.*

Proof Sufficiency. Suppose that f is convex and $(x, \alpha), (y, \beta) \in \operatorname{epi} f$ and $\lambda \in [0, 1]$. We set $w = (1 - \lambda)x + \lambda y$ and $t = (1 - \lambda)\alpha + \lambda\beta$. From the inequality

$$f(w) \leq (1 - \lambda) f(x) + \lambda f(y) \leq t$$

it follows that $(w, t) \in \operatorname{epi} f$. This proves that $\operatorname{epi} f$ is a convex set of $X \times \mathbb{R}$.

Necessity. Suppose that $\operatorname{epi} f$ is convex, but for some $x, y \in X$ and some $\lambda \in [0, 1]$ the inequality

$$f(w) = f\big((1 - \lambda)x + \lambda y\big) > (1 - \lambda) f(x) + \lambda f(y)$$

holds. In particular, the latter shows that $0 < \lambda < 1$ and that neither $f(x)$ nor $f(y)$ can be $+\infty$. Thus, there exist real numbers α, β such that (x, α) and (y, β) belong to $\operatorname{epi} f$. Thus, for each x, y and λ, one has

$$\inf\big\{(1 - \lambda)\alpha + \lambda\beta;\ (x, \alpha), (y, \beta) \in \operatorname{epi} f\big\} = (1 - \lambda) f(x) + \lambda f(y).$$

Since the epigraph of f is convex, we have

$$f(w) = \inf\big\{t;\ (w, t) \in \operatorname{epi} f\big\} \leq (1 - \lambda) f(x) + \lambda f(y) < f(w).$$

The contradiction we arrived at concludes the proof. □

A similar characterization of concave function can be given in terms of its hypograph.

2.1.2 Lower-Semicontinuous Functions

Let X be a topological space.

Definition 2.4 The function $f : X \to \overline{\mathbb{R}}$ is called *lower-semicontinuous* (*upper-semicontinuous*) *at* x_0 if

$$f(x_0) = \liminf_{x \to x_0} f(x) \quad \Big(f(x_0) = \limsup_{x \to x_0} f(x)\Big). \tag{2.6}$$

We recall that

$$\liminf_{x \to x_0} f(x) = \sup_{V \in \mathscr{V}(x_0)} \inf_{s \in V} f(s) \tag{2.7}$$

and

$$\limsup_{x \to x_0} f(x) = \inf_{V \in \mathscr{V}(x_0)} \sup_{s \in V} f(s), \tag{2.8}$$

where $\mathscr{V}(x_0)$ is a base of neighborhoods of x_0 in X.

A function which is lower-semicontinuous at each point of X is called *lower-semicontinuous* on X.

Let us denote by τ_ℓ the topology on $\overline{\mathbb{R}}$ defined by the following basis of open sets:

$$\tau_\ell = \left\{]a, +\infty[;\ a \in [-\infty, +\infty[\right\} \cup \left\{ \emptyset, \overline{\mathbb{R}} \right\}.$$

It is readily seen that the function $f : X \to \overline{\mathbb{R}}$ is lower-semicontinuous (l.s.c.) at x_0 if and only if $f : X \to (\overline{\mathbb{R}}, \tau_\ell)$ is continuous at x_0. The topology τ_ℓ is called the *lower-topology* of $\overline{\mathbb{R}}$. The *upper-semicontinuity* is similarly defined.

Since a function f is upper-semicontinuous if and only if $-f$ is lower-semicontinuous, the following considerations will be restricted to the basic properies of lower-semicontinuous functions as required for the purpose of the next section.

Proposition 2.5 *Let X be a topological space and let $f : X \to \overline{\mathbb{R}}$ be any extended real-valued function on X. Then, the following conditions are equivalent*:

(i) *f is lower-semicontinuous on X.*
(ii) *The level sets $\{x \in X;\ f(x) \le \lambda\}$, $\lambda \in \mathbb{R}$, are closed.*
(iii) *The epigraph of the function f is closed in $X \times \mathbb{R}$.*

Proof It is well known that a function is continuous if and only if the inverse image of every closed subset is closed. Since $\{x \in X;\ f(x) \le \lambda\} = f^{-1}([-\infty, \lambda])$ and (i) is equivalent to the continuity of $f : X \to (\overline{\mathbb{R}}, \tau_\ell)$, we may conclude that conditions (i) and (ii) are equivalent.

We define

$$\varphi(x, t) = f(x) - t \quad \text{for } x \in X \text{ and } t \in \mathbb{R}$$

and observe that f is lower-semicontinuous on X if and only if $\varphi : X \times \mathbb{R} \to \overline{\mathbb{R}}$ is lower-semicontinuous on the product space $X \times \mathbb{R}$. Furthermore, the equivalence of conditions (i) and (ii) for φ implies that (ii) and (iii) are also equivalent, since

$$\operatorname{epi} f - (0, \lambda) = \left\{ (x, t) \in X \times \mathbb{R};\ \varphi(x, t) \le \lambda \right\},$$

that is, the level sets of the function φ are translates of epi f. Proposition 2.5 has now been proved. □

Corollary 2.6 *The upper-envelope of a family of lower-semicontinuous functions is also a lower-semicontinuous function.*

Proof It suffices to apply Proposition 2.5, condition (ii), and to observe that

$$\left\{ x \in X;\ \sup_{i \in I} f_i(x) \le \lambda \right\} = \bigcap_{i \in I} \left\{ x \in X;\ f_i(x) \le \lambda \right\}.$$

□

Corollary 2.7 *A subset A of X is closed if and only if its indicator function I_A is lower-semicontinuous.*

An important property of lower-semicontinuous functions is given by the following well-known *Weierstrass theorem.*

Theorem 2.8 (Weierstrass) *A lower-semicontinuous function f on a compact topological space X takes a minimum value on X. Moreover, if it takes only finite values, it is bounded from below.*

Proof Since, by Proposition 2.5, every level subset of f is closed, using the nonempty ones among them we form a filter base on the compact space X. This filter base has at least one adherent point x_0 which clearly lies in all the nonempty level subsets. Thus, $f(x_0) \le f(x)$ for all x in X, thereby proving Theorem 2.8. □

2.1.3 Lower-Semicontinuous Convex Functions

Throughout this section, X is a topological linear space over a real field. It may be seen that, if a convex function f takes the value $-\infty$, then the set of all points where f is finite is quite "rare". If f is actually convex and lower-semicontinuous on X, then f is nowhere finite on X. Namely, one has the following proposition.

Proposition 2.9 *Let $f : X \to \overline{\mathbb{R}}$ be a convex and lower-semicontinuous function. Assume that there exists $x_0 \in X$ such that $f(x_0) = -\infty$. Then f is nowhere finite on X.*

Proof If there was a $y_0 \in X$ such that $-\infty < f(y_0) < +\infty$, then the convexity of f would imply that $f(\lambda x_0 + (1-\lambda)y_0) = -\infty$, for each $\lambda \in]0, 1]$.

Inasmuch as f is lower-semicontinuous, letting λ approach to zero, $f(y_0) = -\infty$ would hold, which contradicts the assumption. The proof is now complete.

Let $f : X \to \overline{\mathbb{R}}$ be any convex function on X. The *closure* of the function f, denoted by $\operatorname{cl} f$, is by definition the lower-semicontinuous hull of f, that is, $\operatorname{cl} f = \liminf_{y\to x} f(y)$ for all $x \in X$ if $\liminf_{y\to x'} f(y) > -\infty$ for every $x' \in X$ or $\operatorname{cl} f() = -\infty$ for all $x \in X$ if $\liminf_{y\to x'} f(y) = -\infty$ for some $x' \in X$. The convex function f is said to be *closed* if $\operatorname{cl} f = f$. Particularly, *a proper convex function is closed if and only if it is lower-semicontinuous.*

For every proper closed convex function one has

$$(\operatorname{cl} f)(x) = \liminf_{y\to x} f(y), \quad \forall x \in X. \tag{2.9}$$

As a consequence of equality (2.9), one obtains

$$\operatorname{epi}(\operatorname{cl} f) = \overline{\operatorname{epi} f}, \tag{2.10}$$

or, more specifically,

$$\{x \in X;\ (\mathrm{cl}\, f)(x) \le \alpha\} = \bigcap_{\lambda > \alpha} \overline{\{x \in X;\ f(x) \le \lambda\}}$$

for every $\alpha \in \mathbb{R}$. In particular, it follows from (2.7) that

$$\inf\{f(x);\ x \in X\} = \inf\{(\mathrm{cl}\, f)(x);\ x \in X\}. \tag{2.11}$$

Likewise, it should be observed that in general the closure of the convex function f is the greatest closed convex function majorized by f (namely, the pointwise supremum of the collection of all closed convex functions g, such that $g(x) \le f(x)$, for every $x \in X$). □

Furthermore, we give some simple results pertaining to lower-semicontinuous convex functions.

Proposition 2.10 *Let X be a locally convex space. A proper convex function $f : X \to]-\infty, +\infty]$ is lower-semicontinuous on X if and only if it is lower-semicontinuous with respect to the weak topology on X.*

Proof We have already seen in Chap. 1 (Proposition 1.73 and Remark 1.78) that a convex subset is (strongly) closed if and only if it is closed in the corresponding weak topology on X. In particular, we may infer that epi f is (strongly) closed if it is weakly closed. This establishes Proposition 2.10. □

Theorem 2.11 *Let f be a lower-semicontinuous, proper and convex function on a reflexive Banach space X. Then f takes a minimum value on every bounded, convex and closed subset M of X. In other words, $x_0 \in M$ exists such that*

$$f(x_0) = \inf\{f(x);\ x \in M\}.$$

Proof We apply Theorem 2.8 to the space X endowed with weak topology. (According to Corollary 1.95, every closed and bounded subset of a reflexive Banach space is weakly compact.) □

Remark 2.12 If in Theorem 2.11 we further suppose that f is strictly convex, then the minimum point x_0 is unique.

Remark 2.13 In Theorem 2.11, the condition that M is bounded may be replaced by the *coercivity condition*

$$\lim_{\substack{\|x\| \to +\infty \\ x \in M}} f(x) = +\infty. \tag{2.12}$$

In fact, let $x_1 \in \mathrm{Dom}(f)$ and $k > 0$ be such that

$$f(x) > f(x_1) \quad \text{for } \|x\| > k,\ x \in M.$$

Obviously,

$$\inf\{f(x);\ x \in M\} = \inf\{f(x);\ x \in M \cap \overline{S(0,k)}\},$$

where $\overline{S(0,k)} = \{x \in X;\ \|x\| \leq k\}$. Thus, we may apply the preceding theorem where M is replaced by $M \cap \overline{S(0,k)}$.

Now, we divert our attention to the continuity properties of the convex functions. The main result is contained in the following theorem.

Theorem 2.14 *Let X be a topological linear space and let $f : X \to]-\infty, +\infty]$ be a proper convex function on X. Then, the function f is continuous on* $\operatorname{int}\operatorname{Dom}(f)$ *if and only if f is bounded from above on a neighborhood of an interior point of* $\operatorname{Dom}(f)$.

Proof Since the *necessity* is obvious, we restrict ourselves to proving the *sufficiency*. To this end, consider any point x_0 which is interior to the effective domain $\operatorname{Dom}(f)$. Let $V \in \mathscr{V}(x_0)$ be a circled neighborhood of x_0 such that $f(x) \leq k$ for all $x \in V$. Since X is a linear topological space, the function f is continuous at $x = x_0$ if and only if the function $x \to f(x + x_0) - f(x_0)$ is continuous at $x = 0$. Thus, without any loss of generality, we may assume that $x_0 = 0$ and $f(x_0) = 0$. Furthermore, we may assume that V is a circled neighborhood of 0. Since f is convex, we have

$$f(x) = f\left(\varepsilon\, \frac{x}{\varepsilon} + (1-\varepsilon)0\right) \leq \varepsilon f\left(\frac{x}{\varepsilon}\right) \leq \varepsilon k,$$

for all $x \in \varepsilon V$, where $\varepsilon \in]0,1[$. On the other hand,

$$0 = f(0) \leq \frac{1}{2}\big(f(x) + f(-x)\big)$$

and therefore

$$-f(x) \leq f(-x) \leq \varepsilon k \quad \text{for every } x \in -\varepsilon V = \varepsilon V.$$

Thus, we have shown that $|f(x)| \leq \varepsilon k$ for each $x \in \varepsilon V$. In other words, the function f is continuous at the origin. Now, we prove that f is continuous on $\operatorname{int}\operatorname{Dom}(f)$. Let z be any point in $\operatorname{int}\operatorname{Dom}(f)$ and let $\rho > 1$ be such that $z_0 = \rho z \in \operatorname{Dom}(f)$. According to the first part of the proof, it suffices to show that f is bounded from above on a neighborhood of z. Let V be the neighborhood of the origin given above, and let $V(z)$ be a neighborhood of z defined by

$$V(z) = z + \left(1 - \frac{1}{\rho}\right)V.$$

Once again, making use of the convexity of f, we obtain

$$f(u) = f\left(\frac{1}{\rho} z_0 + \left(1 - \frac{1}{\rho}\right)x\right) \le \frac{1}{\rho} f(z_0) + \left(1 - \frac{1}{\rho}\right) f(x)$$
$$\le \frac{1}{\rho} f(z_0) + \left(1 - \frac{1}{\rho}\right)k \quad \text{for all } u \in V(z).$$

Hence, f is bounded above on $V(z)$, as claimed. This completes the proof. □

As a consequence, we obtain the next corollary.

Corollary 2.15 *If a proper convex function $f : X \to]-\infty, +\infty]$ is upper-semicontinuous at a point which is interior to its effective domain* Dom(f), *then f is continuous on* int Dom(f).

For a lower-semicontinuous convex function, this result may be clarified as follows.

Proposition 2.16 *Let X be a real Banach space and let $f : X \to]-\infty, +\infty]$ be a lower-semicontinuous proper convex function. Then f is continuous at every algebraic interior point of its effective domain* Dom(f).

Proof Without any loss of generality, we may restrict ourselves again to the case in which the origin in an algebraic interior to the effective domain Dom(f). We choose any real number α such that $\alpha > f(0)$ and set $A = \{x \in X;\ f(x) \le \alpha\}$. The level set A is convex, closed and contained in the effective domain of f. Let us observe that the origin is an algebraic interior point of A. Indeed, for every $x \in X$, there corresponds $\rho > 0$ such that $x_0 = \rho x \in \mathrm{Dom}(f)$. Here, we have used the fact that the origin is an algebraic interior point of Dom(f). Since f is convex, we have

$$f(\lambda\rho x) = f\big(\lambda x_0 + (1 - \lambda)0\big) \le \lambda\big(f(x_0) - f(0)\big) + f(0),$$

for every $\lambda \in [0, 1]$. Therefore, there exists $\delta > 0$ such that $f(\lambda\rho x) \le \alpha$ for every $\lambda \in [0, \delta]$. This shows that the origin is an algebraic interior point of A. According to Remark 1.24, this fact implies that the origin is an interior point of the closed convex set A. In other words, we have shown that f is bounded from above by α on the neighborhood A of the origin. Applying Theorem 2.14, we may infer that f is continuous on this neighborhood, thereby proving Proposition 2.16. □

If X is a finite-dimensional space, Proposition 2.16 can be considerably strengthened. More precisely, we have the next proposition.

Proposition 2.17 *Every proper convex function f on a finite-dimensional separated topological liner space X is continuous on the interior of its effective domain.*

Proof We suppose again that the origin belongs to the interior of the effective domain $\mathrm{Dom}(f)$ of the function f. Let $\{e_i; i = 1, 2, \ldots, n\}$ be a basis of the n-dimensional space X, and let a be a sufficiently small positive number such that

$$U = \left\{ x \in X;\ x = \sum_{i=1}^{n} x_i e_i,\ 0 < x_i < \frac{a}{n},\ i = 1, 2, \ldots, n \right\} \subset \mathrm{Dom}(f).$$

Using the convexity of f, since

$$x = \sum_{i=1}^{n} x_i e_i = \sum_{i=1}^{n} \frac{x_i}{a}\, a e_i + \left(1 - \sum_{i=1}^{n} \frac{x_i}{a} \right) \cdot 0,$$

we obtain the inequality

$$f(x) \le \sum_{i=1}^{n} \frac{x_i}{a} f(a e_i) + \left(1 - \sum_{i=1}^{n} \frac{x_i}{a} \right) f(0) \le \frac{1}{n} \sum_{i=1}^{n} \left| f(a e_i) \right| + \left| f(0) \right|$$

for every $x \in U$.

Thus, the function f is bounded from above on $U \subset \mathrm{Dom}(f)$. But it is obvious that U is open. This implies, according to Theorem 2.14, that f is continuous on $\operatorname{int}\mathrm{Dom}(f)$, which completes the proof. □

Concerning the continuity of proper convex functions, the results are similar to those obtained for linear functionals: the continuity at a point implies the continuity everywhere and this is equivalent to the boundedness on a certain neighborhood. However, for convex functions these facts are restricted to the interior of effective domain. In this context, our attention has to be restricted to those points of $\mathrm{Dom}(f)$ which do not belong to $\operatorname{int}\mathrm{Dom}(f)$. In addition to the continuity of f on X, we introduce the concept of continuity on $\mathrm{Dom}(f)$. These two concepts are clearly equivalent on $\operatorname{int}\mathrm{Dom}(f)$, but not necessarily on $\mathrm{Dom}(f)$. Also, we notice for later use that

$$\operatorname{int}(\operatorname{epi} f) = \left\{ (x, \alpha) \in X \times \mathbb{R};\ x \in \operatorname{int}\mathrm{Dom}(f),\ f(x) < \alpha \right\}. \tag{2.13}$$

2.1.4 Conjugate Functions

Let X be a real linear locally convex space and let X^* be its conjugate space. Consider any function $f : X \to \overline{\mathbb{R}}$. The function $f^* : X^* \to \overline{\mathbb{R}}$ defined by

$$f^*(x^*) = \sup\left\{ (x, x^*) - f(x);\ x \in X \right\}, \quad x^* \in X^* \tag{2.14}$$

is called the *conjugate function* of f. The conjugate of f^*, that is, the function f^{**} on X defined by

$$f^{**}(x) = \sup\left\{ (x, x^*) - f^*(x^*);\ x^* \in X^* \right\}, \quad x \in X, \tag{2.15}$$

is called the *biconjugate* of f (with respect to the natural dual system given by X and X^*). The conjugate of order n, denoted by $f^{(n)*}$, of the function f is similarly defined.

We pause briefly to observe that relations (2.14) and (2.15) yield

$$f(x) + f^*(x^*) \geq (x, x^*) \tag{2.16}$$

and

$$f^*(x^*) + f^{**}(x) \geq (x, x^*), \tag{2.17}$$

for all $x \in X$ and $x^* \in X^*$. Inequality (2.16) is known as the *Young inequality*. Observe also that if f is proper, then "sup" in relation (2.14) may be restricted to the points x which belong to $\mathrm{Dom}(f)$.

Example 2.18 The conjugate of the indicator function I_A of a subset A of X is given by

$$I_A^*(x^*) = \sup\big\{(x, x^*);\ x \in A\big\}. \tag{2.18}$$

The function I_A^*, usually denoted by s_A, is called the *support functional* of A. It should be observed that A is contained in a closed half-space, $\{x \in X; (x, x^*) \leq \alpha\}$ if and only if $\alpha \geq I_A^*(x^*)$. Thus, $I_A^*(x^*)$ may be determined by the minimal half-space containing A. In other words, if the linear function $x \to (x, x^*)$ reaches its maximum on A, then $(x, x^*) = I_A^*(x^*)$ represents the equation of a supporting hyperplane of A.

Let A° be the *polar* of A, that is,

$$A^\circ = \big\{x^* \in X^*;\ (x, x^*) \leq 1,\ \forall x \in A\big\}. \tag{2.19}$$

In terms of I_A^* defined above, the polar of A may be expressed as

$$A^\circ = \big\{x^* \in X^*;\ I_A^*(x^*) \leq 1\big\}. \tag{2.20}$$

We observe that, if $A = C$ is a cone with vertex in 0, then the polar set C° is again a cone with vertex in 0, which is given by

$$C^\circ = \big\{x^* \in X^*;\ (x, x^*) \leq 0,\ \forall x \in C\big\} \tag{2.21}$$

and is called the *dual cone* of C.

If $A = Y$ is a linear subspace of X, then

$$Y^\circ = \big\{x^* \in X^*;\ (x, x^*) = 0,\ \forall x \in Y\big\} \tag{2.22}$$

is also a linear subspace, called the *orthogonal of the space* Y, sometimes denoted by $Y^\perp$.

As is readily seen, the polar A° of a subset A is a closed convex subset which contains the origin. If we take into account (2.20) and Corollary 1.23, the question arises whether I_A^* is a Minkowski functional associated with the subset A°. In general, the answer is negative. However, we have

$$p_{A^\circ}(x^*) = \max\{I_A^*(x^*), 0\}, \quad \forall x^* \in X^*. \tag{2.23}$$

Therefore, if $0 \in A$, then $I_A^*(x^*) \geq 0$ and

$$p_{A^\circ} = I_A^*. \tag{2.24}$$

Furthermore,

$$p_A^* = I_{A^\circ} \quad \text{for every } A \subset X \text{ with } 0 \in A. \tag{2.25}$$

Indeed, if $x^* \overline{\in} A^\circ$, then there exists $\bar{x} \in A$ such that $(x^*, \bar{x}) > 1$. This implies that

$$\begin{aligned} p_A^*(x^*) &= \sup_{x \in X}\{(x, x^*) - p_A(x)\} \geq \lambda(\bar{x}, x^*) - p_A(\lambda\bar{x}) \\ &= \lambda\big[(\bar{x}, x^*) - p_A(\bar{x})\big] \geq \lambda\big[(\bar{x}, x^*) - 1\big], \quad \forall \lambda > 0. \end{aligned}$$

Hence, $p_A^*(x^*) = +\infty$ for every $x^* \overline{\in} A^\circ$. Now, if $x^* \in A^\circ$, since for every $x \in \mathrm{Dom}(p_A)$, $x \in (p_A(x) + \varepsilon)A$, for all $\varepsilon > 0$, we have

$$\begin{aligned} p_A^*(x^*) &= \sup\{(x, x^*) - p_A(x);\ x \in \mathrm{Dom}(p_A)\} \\ &\leq \sup_{a \in A} \sup_{x \in \mathrm{Dom}(p_A)} \{(p_A(x) + \varepsilon)(a, x^*) - p_A(x)\} \leq \varepsilon, \quad \forall \varepsilon > 0. \end{aligned}$$

Hence, $p_A^*(x^*) \leq 0$. Because $0 \in \mathrm{Dom}(p_A)$, we may infer that $p_A^*(x^*) \geq 0$, which completely proves relation (2.25).

Proposition 2.19 contains some elementary facts concerning conjugacy relations.

Proposition 2.19 *Let $f : X \to \overline{\mathbb{R}}$ be any function on X. Then*

(i) *The functions f^* and f^{**} are always convex and lower-semicontinuous in the weak-star topology of X^* and in the weak topology of X, respectively.*
(ii) $f^{**} \leq f$.
(iii) *$f^{(n)*} = f^*$ or $f^{(n)} = f^{**}$ depending on whether n is odd or even.*
(iv) *$f_1 \leq f_2$ implies that $f_1^* \geq f_2^*$.*

Proof We observe that f^* is the supremum of a family of convex and weak-star continuous functions on X^*. Similarly, relation (2.15) shows that f^{**} is the supremum of a family of convex and weakly continuous functions on X. Thus, we obtain part (i) as an immediate consequence of Corollary 2.6.

As already mentioned, it follows from relation (2.14) that

$$(x, x^*) - f^*(x^*) \leq f(x) \quad \text{for all } x \in X,\ x^* \in X^*,$$

which clearly implies that $f^{**} \leq f$, as claimed. Part (iv) is immediate, and therefore its proof is omitted. To prove part (iii), it suffices to show that $f^{***} = f^*$. In fact, it follows from part (ii) that $f^{***} \le f^*$, while part (iv) implies that $f^* \le f^{***}$, as claimed.

We observe from the definition of f^* that, if the function f is not proper, that is, if f takes on $-\infty$ or it is identically $+\infty$, then its conjugate is also not proper. Furthermore, the conjugate f^* may not be proper on X^* though f is proper on X. This is the reason for saying that *a function admits conjugate if its conjugate is proper*. In particular, it follows from Proposition 2.19 that, if f admits a conjugate, then it admits conjugate of every order. We shall see later that a lower-semicontinuous convex function is proper if and only if it admits conjugate. This assertion will follow from the Proposition 2.20 below. $\square$

Proposition 2.20 *Any convex, proper and lower-semicontinuous function is bounded from below by an affine function.*

Proof Let $f : X \to]-\infty, +\infty]$ be any convex and lower-semicontinuous function on X, $f \not\equiv +\infty$. As already seen, the epigraph $\operatorname{epi} f$ of f is a proper convex and closed subset of product space $X \times \mathbb{R}$. If $x_0 \in \operatorname{Dom}(f)$, then $(x_0, f(x_0) - \varepsilon) \overline{\in} \operatorname{epi} f$ for every $\varepsilon > 0$. Thus, using the Hahn–Banach theorem (see Corollary 1.45), there exists $u \in (X \times \mathbb{R})^*$ such that

$$\sup_{(x,t)\in \operatorname{epi} f} u(x,t) < u\big(x_0, f(x_0) - \varepsilon\big).$$

Identifying the dual space $(X \times \mathbb{R})^*$ with $X^* \times \mathbb{R}$, we may infer that there exist $x_0^* \in X^*$ and $\alpha \in \mathbb{R}$, not both zero, such that

$$\sup_{(x,t)\in \operatorname{epi} f} \big\{x_0^*(x) + t\alpha\big\} < x_0^*(x_0) + \alpha\big(f(x_0) - \varepsilon\big).$$

We observe that $\alpha \ne 0$ and must be negative, since $(x_0, f(x_0) + n) \in \operatorname{epi} f$ for every $n \in \mathbb{N}$. On the other hand, $(x, f(x)) \in \operatorname{epi} f$ for every $x \in \operatorname{Dom}(f)$. Thus,

$$x_0^*(x) + \alpha f(x) \le x_0^*(x_0) + \alpha f(x_0), \quad \forall x \in \operatorname{Dom}(f),$$

or

$$f(x) \ge -\frac{1}{\alpha} x_0^*(x) + \frac{1}{\alpha} x_0^*(x_0) + f(x_0), \quad \forall x \in \operatorname{Dom}(f),$$

but the function in the right-hand side is affine, as claimed. $\square$

Corollary 2.21 *A lower-semicontinuous convex function is proper if and only if its conjugate is proper.*

Proof If the function $f : X \to]-\infty, +\infty]$ is convex lower-semicontinuous and nonidentically $+\infty$, then relation (2.14) and Proposition 2.20 show that $f^* \not\equiv +\infty$ and $f^*(x^*) > -\infty$ for every $x^* \in X^*$. Next, we assume that f^* is proper on X^*. Then, inequality (2.16) implies that f is nowhere $-\infty$ on X while relation (2.14) shows that f must be nonidentically $+\infty$. $\square$

Now, we establish a central result of Convex Analysis which is known in the literature as the *biconjugate theorem*.

Theorem 2.22 *Let $f : X \to]-\infty, +\infty]$ be any function nonidentically $+\infty$. Then $f^{**} = f$ if and only if f is convex and lower-semicontinuous on X.*

Proof If $f = f^{**}$, then Proposition 2.19 implies that f is convex and lower-semicontinuous. Now, we assume that f is proper, convex and lower-semicontinuous on X. Since the conjugate f^* of f is proper, using Corollary 2.21, we may infer that $f^{**} > -\infty$ everywhere on X. Moreover, Proposition 2.19(ii) implies that $f^{**}(x) \le f(x)$, for every $x \in X$. Suppose that there exists $x_0 \in X$ such that $f^{**}(x_0) < f(x_0)$ and we argue from this to a contradiction. Thus, $(x_0, f^{**}(x_0)) \overline{\in}$ epi f, so that, using the same reasoning as in the proof of Proposition 2.20, we may conclude that there exist $x_0^* \in X^*$ and $\alpha \in \mathbb{R}$ such that

$$x_0^*(x_0) + \alpha f^{**}(x_0) > \sup\{x_0^*(x) + \alpha t;\ (x,t) \in \operatorname{epi} f\}. \tag{2.26}$$

Since $(x, t+n) \in \operatorname{epi} f$ for every $n \in \mathbb{N}$ and $(x,t) \in \operatorname{epi} f$, relation (2.26) implies that $\alpha \le 0$. Furthermore, α must be negative. Indeed, otherwise (that is, $\alpha = 0$), inequality (2.26) implies that

$$x_0^*(x_0) > \sup\{x_0^*(x);\ x \in \operatorname{Dom}(f)\}. \tag{2.27}$$

Let $h > 0$ and $y_0^* \in \operatorname{Dom}(f^*)$ be arbitrarily chosen. (We recall that $\operatorname{Dom}(f^*) \neq \emptyset$ because f^* is proper.) One obtains

$$\begin{aligned} f^*(y_0^* + hx_0^*) &= \sup\{(x, y_0^*) + h(x, x_0^*) - f(x);\ x \in \operatorname{Dom}(f)\} \\ &\le \sup\{(x, y_0^*) - f(x);\ x \in \operatorname{Dom}(f)\} \\ &\quad + h \sup\{(x_0^*, x);\ x \in \operatorname{Dom}(f)\} \\ &= f^*(y_0^*) + h \sup\{(x_0^*, x);\ x \in \operatorname{Dom}(f)\}. \end{aligned}$$

On the other hand, a simple calculation involving the latter expression and inequality (2.17) yields

$$\begin{aligned} f^{**}(x_0) &\ge (y_0^* + hx_0^*, x_0) - f^*(y_0^* + hx_0^*) \\ &\ge (y_0^*, x_0) - f^*(y_0^*) + h\big[(x_0^*, x_0) - \sup\{(x_0^*, x);\ x \in \operatorname{Dom}(f)\}\big]. \end{aligned}$$

Comparing this inequality with (2.27) and letting $h \to +\infty$, we obtain $f^{**}(x_0) = +\infty$, which is absurd. Therefore, α is necessarily negative. Thus, we may divide inequality (2.26) by $-\alpha$ to obtain

$$\begin{aligned} x_0^*\left(-\frac{x_0}{\alpha}\right) - f^{**}(x_0) &> \sup\left\{x_0^*\left(-\frac{x}{\alpha}\right) - t;\ (x,t) \in \operatorname{epi} f\right\} \\ &= \sup\left\{\left(-\frac{1}{\alpha}x_0^*, x\right) - f(x);\ x \in \operatorname{Dom}(f)\right\} = f^*\left(-\frac{1}{\alpha}x_0^*\right). \end{aligned}$$

But this inequality obviously contradicts inequality (2.17). Hence, $f^{**}(x_0) = f(x_0)$ for every $x_0 \in \mathrm{Dom}(f^{**})$. Since $f^{**}(x) = f(x)$, for all $x \overline{\in} \mathrm{Dom}(f^{**})$, it results that $f^{**}(x) = f(x)$ for all $x \in X$. Thus, the proof is complete. □

More generally, if f is not lower-semicontinuous, then $f^{**} = \mathrm{cl}\, f$. Thus, we obtain the following corollary.

Corollary 2.23 *The biconjugate of a convex function f coincides with its closure, that is*, $f^{**} = \mathrm{cl}\, f$.

Proof It is clear that $\mathrm{cl}\, f$ is lower-semicontinuous if it is proper and, therefore, $(\mathrm{cl}\, f)^{**} = \mathrm{cl}\, f$ as a consequence of Theorem 2.22. But as has already been mentioned, $f^* = (\mathrm{cl}\, f)^*$, which shows that $f^{**} = \mathrm{cl}\, f$, as claimed. If $\mathrm{cl}\, f$ is not proper, the result is immediately clear, since $f^{**} = (\mathrm{cl}\, f)^{**} = \mathrm{cl}\, f \equiv -\infty$. □

Corollary 2.24 *A proper function f is convex and lower-semicontinuous on X if and only if it is the supremum of a family of affine continuous functions.*

Proof If f is a proper convex and lower-semicontinuous, then $f(x) = f^{**}(x) = \sup\{(x, x^*) - f^*(x^*);\ x^* \in D(f^*)\}$ for every $x \in X$, and $x \to (x, x^*) - f^*(x^*)$ is an affine continuous function for each $x^* \in \mathrm{Dom}(f^*)$, as claimed. The converse is obvious (see Corollary 2.6). □

There is a close connection between the effective domain $\mathrm{Dom}(f)$ of a lower-semicontinuous convex function $f : X \to \overline{\mathbb{R}}^*$ and the growth properties of its conjugate $f^* : X^* \to \overline{\mathbb{R}}^*$.

Proposition 2.25 *Assume that X is a reflexive Banach space. Then the following two conditions are equivalent*:

(i) $\mathrm{int}\, \mathrm{Dom}(f) \neq \emptyset$.
(ii) *There are $\rho > 0$ and $C > 0$ such that*

$$f^*(p) \geq \rho \|p\|_{X^*} - C, \quad \forall p \in X. \tag{2.28}$$

Moreover, $\mathrm{Dom}(f) = X$ *if and only if*

$$\lim_{\|p\| \to \infty} \frac{f^*(p)}{\|p\|} = +\infty. \tag{2.29}$$

Proof If $\mathrm{int}\, \mathrm{Dom}(f) \neq \emptyset$, then there is a ball $B(x_0, \rho) \subset \mathrm{int}\, \mathrm{Dom}(f)$ and by Theorem 2.14, f is bounded on $B(x_0, \rho)$. Then, by the duality formula (2.14), we have (for simplicity, assume $x_0 = 0$)

$$f^*(p) \geq \rho \|p\|_{X^*} - f\left(\rho \frac{x}{\|x\|_X}\right) \geq \rho \|p\|_{X^*} - C, \quad \forall p \in X^*,$$

as claimed.

If (ii) holds, then by (2.15) we see that

$$f(x) = f^{**}(x) \le \sup_x \big\{ (x, x^*) - \rho \|x^*\|_{X^*} - C \big\} \le \infty \quad \text{for } \|x\|_X \le \rho$$

and therefore $B(0,\rho) \subset \mathrm{Dom}(f)$, as claimed.

Now, if $\mathrm{Dom}(f) = X$, then by the above argument it follows that (2.28) holds for all $\rho > 0$, that is, for all $\rho > 0$,

$$f^*(p) \ge \rho \|p\|_{X^*} - C_\rho, \quad \forall p \in X^*,$$

which implies that (2.29) holds. Conversely, if (2.29) holds, then, by (2.15), we see that $\mathrm{Dom}(f) = X$, as claimed. □

Theorem 2.22 and Corollary 2.23, in particular, yield a simple proof for the well-known *bipolar theorem* (Theorem 2.26 below), which plays an important role in the duality theory.

Theorem 2.26 *The bipolar $A^{\circ\circ}$ of a subset A of X is the closed convex hull of the origin and of A, that is,*

$$A^{\circ\circ} = \overline{\mathrm{conv}\big(A \cup \{0\}\big)}. \tag{2.30}$$

Proof Inasmuch as the polar is convex, weakly closed and contains the origin, it suffices to show that $A^{\circ\circ} = A$ for every convex, closed subset of X, which contains the origin. In this case, relations (2.24) and (2.25) imply that

$$I_{A^{\circ\circ}} = p^*_{A^\circ} = I^{**}_A = I_A,$$

because I_A is convex and lower-semicontinuous. Hence, $A = A^{\infty}$, as claimed. □

Remark 2.27 We notice that the conjugate correspondence $f \to f^*$ is one-to-one between convex and lower-semicontinuous convex functions on X and weak-star lower-semicontinuous convex functions on X^*. In this context, the concept of conjugate defined above seems to be more suitable for convex functions.

For concave functions, it is more natural to introduce a concept of conjugate which preserves the concavity and upper-semicontinuity. Given any function $g : x \to \overline{\mathbb{R}}$, the function $g^* : X^* \to \overline{\mathbb{R}}$ defined by

$$g^*(x) = \inf\big\{ (x, x^*) - g(x);\ x \in X \big\}, \tag{2.31}$$

is called the *concave conjugate function* of g. We observe that the concave conjugate g^* of a function g can be equivalently expressed with the aid of convex conjugate defined by relation (2.14) as it follows that

$$g^*(x^*) = -(-g)^*(-x^*) \quad \text{for every } x^* \in X^*,$$

where the conjugate in the right-hand side is in the convex sense.

In general, facts and definitions for concave conjugate functions are obtained from those above by interchanging $\leq$ with $\geq$, $+\infty$ with $-\infty$ and infimum with supremum wherever these occur. Typically, we consider the concave conjugate for concave functions and the conjugate for convex functions.

Remark 2.28 Let f be a convex function on a linear normed space X and let $f^* : X^* \to \overline{\mathbb{R}}$ be the conjugate function of f. Let $(f^*)^*; X^{**} \to \overline{\mathbb{R}}$ be the conjugate of f^* defined on the bidual X^{**} of X. It is natural also to call $(f^*)^*$ the biconjugate of f and, if X is reflexive, obviously $(f^*)^*$ coincides with f^{**}. In general, the restriction of $(f^*)^*$ to X (when X is regarded in the canonical way as the linear subspace of X^{**}) coincides with f^{**}.

Remark 2.29 The theory of conjugate functions can be developed in a context more general than that of the linear locally convex space. Specifically, let X and Y be arbitrary real linear spaces paired by a bilinear functional $(\cdot, \cdot)$ and let X and Y be endowed with compatible topologies with respect to this pairing. Let $f : X \to \overline{\mathbb{R}}$ be any extended real-valued function on X. Then the function f^* on Y defined by

$$f^*(y) = \sup\big\{(x, y) - f(x);\ x \in X\big\}, \quad y \in Y, \tag{2.32}$$

is called the *conjugate* of f (with respect to the given pairing). A closer examination of the proofs shows that the above results on conjugate functions are still valid in this general framework.

2.2 The Subdifferential of a Convex Function

The subdifferential of a convex is a basic concept for convex analysis and it will be developed in detail in this section.

2.2.1 Definition and Fundamental Results

Throughout this section, X denote a real Banach space with dual X^* and norm $\|\cdot\|$. As usually, $(\cdot, \cdot)$ denote the canonical pairing between X and X^*.

Definition 2.30 Given the proper convex function $f : X \to]-\infty, +\infty]$, the *subdifferential* of such a function is the (generally multivalued) mapping $\partial f : X \to X^*$ defined by

$$\partial f(x) = \big\{x^* \in X^*;\ f(x) - f(u) \leq (x - u, x^*),\ \forall u \in X\big\}. \tag{2.33}$$

The elements $x^* \in \partial f(x)$ are called *subgradients* of f at x.

It is clear from relation (2.33) that $\partial f(x)$ is always a closed convex subset of X^*. The set $\partial f(x)$ may well be empty as happens, e.g., if $f(x) = +\infty$ and $f \not\equiv +\infty$.

The set of those x for which $\partial f(x) \neq \emptyset$ is called the *domain* of ∂f and is denoted by $D(\partial f)$. Clearly, if f is not the constant $+\infty$, $D(\partial f)$ is a subset of $\mathrm{Dom}(f)$. The function f is said to be *subdifferentiable* at x, if $x \in D(\partial f)$.

Example 2.31 Let K be a closed convex subset of X. The *normal cone* $N_K(x)$ to K at a point $x \in K$ consists, by definition, of all the normal vectors to half-spaces that support K at x, that is,

$$N_K(x) = \left\{x^* \in X^*;\ (x^*, x - u) \geq 0 \text{ for all } u \in K\right\}.$$

This is a closed convex cone containing the origin and, in terms of the indicator function I_K of K, we can write it as

$$N_K(x) = \partial I_K(x), \quad x \in K.$$

Clearly, $D(\partial I_K) = K$ and $\partial I_K(x) = \{0\}$ when $x \in \operatorname{int} K$. In particular, if K is a linear subspace of X, then $\partial I_K(x) = K^\perp$ for all $x \in K$ ($K^\perp$ is the subspace of X^* orthogonal to K).

Example 2.32 Let $f(x) = \frac{1}{2}\|x\|^2$. Then, f is a convex continuous function on X. Furthermore, f is everywhere subdifferentiable on X and the subdifferential ∂f coincides with the duality mapping $F : X \to X^*$ (see Definition 1.99). Indeed, if $x^* \in F(x)$, then, by the definition of F, one has

$$\begin{aligned}(x - u, x^*) &= \|x\|^2 - (u, x^*) \geq \|x\|^2 - \|u\|\,\|x\| \\ &\geq \frac{1}{2}\big(\|x\|^2 - \|u\|^2\big), \quad \text{for every } u \in X.\end{aligned}$$

In other words, $x^* \in \partial f(x)$. Conversely, suppose that $x^* \in \partial f(x)$. Hence,

$$(x - u, x^*) \geq \frac{1}{2}\big(\|x\|^2 - \|u\|^2\big), \quad \forall u \in X.$$

Taking in the latter inequality $u = x + \lambda v$, where $\lambda \in \mathbb{R}^+$ and $v \in X$, we see that

$$-\lambda(v, x^*) \geq -\frac{1}{2}\big(2\lambda\|x\|\,\|v\| + \lambda^2\|v\|^2\big).$$

Therefore

$$\big|(v, x^*)\big| \leq \|v\|\,\|x\|, \quad \forall v \in X.$$

Furthermore, we take $u = (1 - \lambda)x$, divide by λ and let $\lambda \searrow 0$; we get

$$(x, x^*) \geq \|x\|^2.$$

Combining these inequalities, we obtain

$$(x, x^*) = \|x\|^2 = \|x^*\|^2.$$

Thus, we have shown that $x^* \in F(x)$, as claimed.

In the general theory of convex optimization, the following trivial consequence of Definition 2.30 plays an important role.

If f is a proper convex function on X, then the minimum (global) of f over X is attained at the point $x \in X$ if and only if $0 \in \partial f(x)$.

It must be observed that, if f is strictly convex, then for every $x^* \in X^*$ the function $f(x) - (x, x^*)$ attains its minimum in at most one point $x = (\partial f)^{-1}(x^*)$. Hence, in this case, the map $(\partial f)^{-1}$ is single valued.

To make use of this minimum (necessary and sufficient condition), it is necessary to calculate the subdifferentials of certain convex functions; this can be easy or difficult, depending on the nature and the complexity of the given function. It is found as a result that, if f is lower-semicontinuous, the subdifferential ∂f^* of the conjugate function f^* coincides with $(\partial f)^{-1}$. More precisely, one has the following proposition.

Proposition 2.33 *Let $f : X \to]-\infty, +\infty]$ be a proper convex function. Then, the following three properties are equivalent:*

(i) $x^* \in \partial f(x)$.
(ii) $f(x) + f^*(x^*) \le (x, x^*)$.
(iii) $f(x) + f^*(x^*) = (x, x^*)$.

If, in addition, f is lower-semicontinuous, then all of these properties are equivalent to the following one.

(iv) $x \in \partial f^*(x^*)$.

Proof The Young inequality (relation (2.16)) shows that (i) and (iii) are equivalent. If statement (iii) holds, then, using again the Young inequality, we find that

$$f(u) - f(x) \ge (u - x, x^*), \quad \forall u \in X,$$

that is, $x^* \in \partial f(x)$. Using a similar argument, it follows that (i) implies (iii). Thus, we have shown that (i), (ii) and (iii) are equivalent. Now, we assume that f is a lower-semicontinuous, proper convex function on X. Since statements (i) and (iii) are equivalent for f^*, relation (iv) can be equivalently expressed as

$$f^*(x^*) + (f^*)^*(x) = (x, x^*), \tag{2.34}$$

where $(f^*)^* : X^{**} \to]-\infty, +\infty]$ is the conjugate function of f^*. As mentioned in Sect. 2.1.4, the restriction of $(f^*)^*$ to X (which, from the canonical viewpoint, is regarded as a subspace of X^{**}) is f^{**} and the latter coincides with f (see Theorem 2.22). Thus, (iii) and (iv) are equivalent. This completes the proof of Proposition 2.33. $\square$

Remark 2.34 Since the set of all minimum points of the function f coincides with the set of solutions x of the equation $0 \in \partial f(x)$, Proposition 2.33 implies that in the lower-semicontinuous case, a function f attains its infimum on X if and only if its conjugate function f^* is subdifferentiable at the origin, that is, $\partial f^*(0) \cap X^* \ne \emptyset$.

Remark 2.35 If the space X is reflexive, then it follows from Proposition 2.33 that $\partial f^* : X^* \to X^{**} = X$ is just the inverse of ∂f, in other words,

$$x \in \partial f^*(x^*) \quad \Longleftrightarrow \quad x^* \in \partial f(x). \tag{2.35}$$

If X is not reflexive, ∂f^* is a (multivalued) mapping from X^* to the bidual X^{**}, which strictly contains X, and the relation between ∂f and ∂f^* is more complicated (see, for example, Rockafellar [59]).

Proposition 2.36 *If the convex function $f : X \to]-\infty, +\infty]$ is (finite and) continuous at x_0, then f is subdifferentiable at this point, that is, $x_0 \in D(\partial f)$.*

Proof Let us denote by H the epigraph of the function f, that is,

$$H = \{(x, \lambda) \in X \times \mathbb{R};\ f(x) \le \lambda\}.$$

H is a convex subset of $X \times \mathbb{R}$ and $(x_0, f(x_0) + \varepsilon) \in \operatorname{int} H$ for every $\varepsilon > 0$, because f is continuous at x_0. We denote by $\overline{H}$ the closure of H and observe that $(x_0, f(x_0))$ is a boundary point of $\overline{H}$. Thus, there exists a closed supporting hyperplane of H which passes through $(x_0, f(x_0))$ (see Theorem 1.38). In other words, there exist $x_0^* \in X^*$ and $\alpha_0 \in \mathbb{R}^+$, such that

$$\alpha_0\big(f(x_0) - f(x)\big) \le (x_0 - x, x_0^*) \quad \text{for every } x \in \operatorname{Dom}(f). \tag{2.36}$$

It should be observed that $\alpha_0 \ne 0$ (that is, the hyperplane is not vertical) because, otherwise, $(x_0 - x, x_0^*) = 0$ for all x in $\operatorname{Dom}(f)$, which is a neighborhood of x_0. But this would imply that $x_0^* = 0$, which is not possible. However, inequality (2.36) shows that $\frac{x_0^*}{\alpha_0}$ is a subgradient of f at x_0, thereby proving Proposition 2.36. □

Remark 2.37 From the above proof, it follows that a proper convex function f is subdifferentiable in an element $x_0 \in \operatorname{Dom}(f)$ if and only if there exists a nonvertical closed support hyperplane of the epigraph passing through $(x_0, f(x_0))$.

Corollary 2.38 *Let f be a lower-semicontinuous proper convex function on a Banach space X. Then*

$$\operatorname{int} \operatorname{Dom}(f) \subset D(\partial f). \tag{2.37}$$

Proof We have seen in Sect. 2.1.3 (Proposition 2.16) that f is continuous at every interior point of its effective domain $\operatorname{Dom}(f)$. Thus, relation (2.37) is an immediate consequence of Proposition 2.36.

The question of when a convex function is subdifferentiable at a given point is connected with the properties of the directional derivative at this point. Also, we shall see later that the subdifferential of a convex function is closely related to other classical concepts, such as the *Gâteaux* (or *Fréchet*) *derivative*.

First, we review the definition and some basic facts about directional and weak derivatives.

Let f be an proper convex function on X. If f is finite at the point x, then, for every $h \in X$, the difference quotient $\lambda \to \lambda^{-1}(f(x+\lambda h) - f(x))$ is monotonically increasing on $]0, \infty[$. Thus, the *directional derivative at x in the direction h*

$$f'(x,h) = \lim_{\lambda \downarrow 0} \lambda^{-1}\big(f(x+\lambda h) - f(x)\big) = \inf_{\lambda > 0} \lambda^{-1}\big(f(x+\lambda h) - f(x)\big) \tag{2.38}$$

exists for every $h \in X$. The function $h \to f'(x,h)$ is called the *directional differential* of f at x. It is immediate from the definition that for fixed $x \in \mathrm{Dom}(f)$, $f'(x,h)$ is a positively homogeneous subadditive function on X. The function f is said to be *weakly* or *Gâteaux differentiable* at x if $h \to f'(x,h)$ is a linear continuous function on X. In particular, this implies that

$$-f'(x,-h) = f'(x,h) = \lim_{\lambda \to 0} \lambda^{-1}\big(f(x+\lambda h) - f(x)\big)$$

for every $h \in X$. If f is weakly differentiable at x, then we denote by $\nabla f(x)$ or $\mathrm{grad}\, f(x)$ (the *gradient of f at x*) the element of X^* defined by

$$f'(x,h) = \big(h, \mathrm{grad}\, f(x)\big) \quad \text{for every } h \in X.$$

The function f is said to be *Fréchet differentiable* at x if the difference quotients in (2.38) as a function of h converges uniformly on every bounded set. □

Proposition 2.39 *Let $f : X \to\ ]-\infty, +\infty]$ be a proper convex function. If f is finite and continuous at x_0, then*

$$f'(x_0,h) = \sup\big\{(h,x^*);\ x^* \in \partial f(x_0)\big\} \tag{2.39}$$

and, in general, one has

$$\partial f(x_0) = \big\{x^* \in X;\ (h,x^*) \le f'(x_0,h),\ \forall h \in X\big\}. \tag{2.40}$$

Proof Since (2.40) is immediate from the definition of ∂f and (2.38), we confine ourselves to prove (2.39). For the sake of simplicity, we denote by f_0 the function $f_0(h) = f'(x_0,h)$, $\forall h \in X$. Inasmuch as f is continuous at x_0, the inequality

$$(h,w) \le f_0(h) \le f(x_0+h) - f(x_0), \quad \forall w \in \partial f(x_0)$$

implies that f_0 is everywhere finite and continuous on X. Furthermore, a simple calculation involving the definition of conjugate (see relation (2.14)) shows that the conjugate of the function $x \to \lambda^{-1}(f(x_0+\lambda x) - f(x_0))$ is just the function $x^* \to \lambda^{-1}(f^*(x^*) + f(x_0) - (x_0,x^*))$. Therefore,

$$f_0^*(x^*) = \sup_{\lambda > 0} \lambda^{-1}\big(f(x_0) + f^*(x^*) - (x_0,x^*)\big),$$

because

$$f_0(h) = \inf_{\lambda > 0} \lambda^{-1}\big(f(x_0+\lambda h) - f(x_0)\big).$$

According to Proposition 2.33, one has

$$\partial f(x_0) = \big\{x^* \in X^*;\ f(x_0) + f^*(x^*) - (x_0, x^*) = 0\big\}$$

and, therefore,

$$f_0^*(x^*) = \begin{cases} 0, & \text{if } x^* \in \partial f(x_0), \\ +\infty, & \text{otherwise.} \end{cases}$$

Thus, $f_0^{**} = f_0$ is the support functional of the closed convex set $\partial f(x_0) \subset X^*$. This, clearly, implies relation (2.39), thereby proving Proposition 2.39.

If $\partial f(x_0)$ happens to consist of a single element, Proposition 2.39 says that $f'(x_0, h)$ can be written as

$$f'(x_0, h) = \big(h, \partial f(x_0)\big) \quad \text{for every } h \in X.$$

In particular, this implies that f is Gâteaux differentiable at x_0 and $\operatorname{grad} f(x_0) = \partial f(x_0)$. It follows that the converse result is also true. □

Namely,

Proposition 2.40 *If the convex function f is Gâteaux differentiable at x_0, then $\partial f(x_0)$ consists of a single element $x_0^* = \operatorname{grad} f(x_0)$. Conversely, if f is continuous at x_0 and if $\partial f(x_0)$ contains a single element, then f is Gâteaux differentiable at x_0 and* $\operatorname{grad} f(x_0) = \partial f(x_0)$.

Proof Suppose that f is Gâteaux differentiable at x_0, that is,

$$\big(h, \operatorname{grad} f(x_0)\big) = \lim_{\lambda \to 0} \lambda^{-1}\big(f(x_0 + \lambda h) - f(x_0)\big), \quad \forall h \in X.$$

However,

$$\lambda^{-1}\big(f(x_0 + \lambda h) - f(x_0)\big) \le f(x_0 + h) - f(x_0) \quad \text{for } \lambda \in]0, 1[$$

because f is convex. This implies that

$$f(x_0) - f(x_0 + h) \le -\big(h, \operatorname{grad} f(x_0)\big) \quad \text{for all } h \in X,$$

that is, $\operatorname{grad} f(x_0) \in \partial f(x_0)$. Now, let x_0^* be any element of $\partial f(x_0)$. We have

$$f(x_0) - f(u) \le (x_0 - u, x_0^*), \quad \forall u \in X,$$

and, therefore,

$$\lambda^{-1}\big(f(x_0 + \lambda h) - f(x_0)\big) \ge (h, x_0^*) \quad \text{for every } \lambda > 0.$$

This show that $(\operatorname{grad} f(x_0) - x_0^*, h) \ge 0$ for all $h \in X$, that is, $x_0^* = \operatorname{grad} f(x_0)$. We conclude the proof by noting that the second part of Proposition 2.40 has already been proven by the above remarks. □

Remark 2.41 Let f be a continuous convex function on X. If f^* is strictly convex, then, as noticed earlier, $(\partial f^*)^{-1} = \partial f$ is single valued. Then, by Proposition 2.40, f is Gâteaux differentiable. In particular, if $f(x) = \frac{1}{2}\|x\|^2$, this fact leads to a well-known result in the metric theory of normed spaces. (See Theorem 1.101.) Namely, *if the dual X^* of X is strictly convex, then X is itself smooth.*

Remark 2.42 If g is a concave function on X, then, by definition its subdifferential is $\partial g = -\partial(-g)$. In other words, $x^* \in \partial g(x)$ if and only if

$$g(x) - g(u) \geq (x - u, x^*) \quad \text{for every } u \in X.$$

2.2.2 Further Properties of Subdifferential Mappings

It is apparent from Definition 2.30 that every subdifferential mapping $\partial f : X \to X^*$ is monotone in $X \times X^*$. In other words,

$$(x_1 - x_2, x_1^* - x_2^*) \geq 0 \quad \text{for } x_i^* \in \partial f(x_i),\ i = 1, 2. \tag{2.41}$$

The theorem below ensures us that any subdifferential mapping is maximal monotone.

Theorem 2.43 (Rockafellar) *Let X be a real Banach space and let f be a lower-semicontinuous proper convex function on X. Then, ∂f is a maximal monotone operator from X to X^*.*

Proof In order to avoid making the treatment too ponderous, we confine ourselves to proving the theorem in the case in which X is reflexive. We refer the reader to Rockafellar's work [59] for the proof in a general context. Then, using the renorming theorem, we may assume without any loss of generality that X and X^* are strictly convex Banach spaces. Using Theorem 1.141, the maximal monotonicity of ∂f is equivalent to $R(F + \partial f) = X^*$, where, as usual, $F : X \to X^*$ stands for the duality mapping of X. Let x_0^* be any fixed element of X^*. We must show that the equation

$$F(x) + \partial f(x) \ni x_0^*,$$

has at least one solution $x_0 \in D(\partial f)$. To this end, we define

$$f_1(x) = \frac{\|x\|^2}{2} + f(x) - (x, x_0^*) \quad \text{for every } x \in X.$$

Clearly, $f_1 : X \to]-\infty, +\infty]$ is convex and lower-semicontinuous on X. Moreover, since f is bounded from below by an affine function, we may infer that

$$\lim_{\|x\| \to +\infty} f_1(x) = +\infty.$$

Thus, using Theorem 2.11 (see Remark 2.13), the infimum of f_1 on X is attained. In other words, there is $x_0 \in \mathrm{Dom}(f)$ such that

$$f_1(x_0) \le f_1(x) \quad \text{for every } x \in X.$$

We write this inequality in the form

$$f(x_0) - f(x) \le (x_0 - x, x_0^*) + \big(x - x_0, F(x)\big) \quad \text{for every } x \in X$$

and set $x = tx_0 + (1-t)u$, where $t \in [0,1]$, and u is any element of X. Since the function f is convex, one obtains

$$f(x_0) - f(u) \le (x_0 - u, x_0^*) + \big(u - x_0, F\big(tx_0 + (1-t)u\big)\big).$$

Passing to limit $t \to 1$, we obtain

$$f(x_0) - f(u) \le (x_0 - u, x_0^*) + \big(u - x_0, F(x_0)\big)$$

because F is demicontinuous from X to X^* (see Theorem 1.106). Since u was arbitrary, we may conclude that

$$x_0^* - F(x_0) \in \partial f(x_0),$$

as we wanted to prove. □

Corollary 2.44 *Let $f : X \to]-\infty, +\infty]$ be a lower-semicontinuous proper and convex function on X. Then $D(\partial f)$ is a dense subset of* $\mathrm{Dom}(f)$.

Proof For simplicity, we assume that X is reflexive. Let x be any element of $\mathrm{Dom}(f)$. Then, Theorem 1.141 and Corollary 1.140 imply that, for every $\lambda > 0$, the equation

$$F(x_\lambda - x) + \lambda \partial f(x_\lambda) \ni 0 \tag{2.42}$$

has a unique solution $x_\lambda \in D(\partial f)$. By the definition of ∂f, we see that, multiplying equation (2.42) by $x_\lambda - x$, we obtain

$$\|x_\lambda - x\|^2 + \lambda f(x_\lambda) \le \lambda f(x)$$

and therefore

$$\lim_{\lambda \to 0} \|x_\lambda - x\| = 0,$$

because f is bounded from below by an affine function. Therefore, $x \in \overline{D(\partial f)}$ and the corollary has been proved. □

It is well known that not every monotone operator arises from a convex function. For instance (see Proposition 2.51 below), a positive linear operator acting in a real Hilbert space is the subdifferential of a proper convex function on H if and only if it is self-adjoint. Thus, we should look for properties which should characterize the maximal monotone operators which are subdifferentials.

Definition 2.45 The operator (multivalued) $A : X \to X^*$ is said to be *cyclically monotone* if

$$(x_0 - x_1, x_0^*) + \cdots + (x_{n-1} - x_n, x_{n-1}^*) + (x_n - x_0, x_n^*) \geq 0, \tag{2.43}$$

for every finite set of points in the graph of A, that is, $x_i^* \in Ax_i$ for $i = 0, 1, \ldots, n$. The operator A is said to be *maximal cyclically monotone* if it is cyclically monotone and has no cyclically monotone extension in $X \times X^*$.

Obviously, every cyclically monotone operator is also monotone. If f is a proper convex function on X, then a simple calculation involving the definition of ∂f shows that the operator ∂f is cyclically monotone. Moreover, it follows from Theorem 2.43 that, if f is in addition lower-semicontinuous on X, then its subdifferential ∂f is cyclically maximal monotone. Surprisingly, it turns out that condition (2.43) is both necessary and sufficient for an operator A to be the subdifferential of some proper convex function. The next theorem is more precise.

Theorem 2.46 *Let X be a real Banach space and let A be an operator from X to X^*. In order that a lower-semicontinuous proper convex function f on X exists such that $A = \partial f$, it is necessary and sufficient that A be a maximal cyclically monotone operator. Moreover, in this case, A determines f uniquely up to an additive constant.*

Proof The *necessity* of the condition was proved in the above remarks. To prove the *sufficiency*, we suppose therefore that A is maximal cyclically monotone in $X \times X^*$. We fix $[x_0, x_0^*]$ in A. For every $x \in X$, let

$$f(x) = \sup\big\{(x - x_n, x_n^*) + \cdots + (x_1 - x_0, x_0^*)\big\},$$

where $x_i^* \in Ax_i$ for $i = 1, \ldots, n$ and the supremum is taken over all possible finite sets of pairs $[x_i, x_i^*] \in A$. We shall prove that $A = \partial f$. Clearly, $f(x) > -\infty$ for all $x \in X$. Note also that f is convex and lower-semicontinuous on X. Furthermore, $f(x_0) = 0$ because A is cyclically monotone. Hence, $f \not\equiv +\infty$. Now, choose any $\tilde{x}$ and $\tilde{x}^*$ with $\tilde{x}^* \in A\tilde{x}$. To prove that $[\tilde{x}, \tilde{x}^*] \in \partial f$, it suffices to show that, for every $\lambda < f(\tilde{x})$, we have

$$f(x) \geq \lambda + (x - \tilde{x}, \tilde{x}^*) \quad \text{for all } x \in X. \tag{2.44}$$

Let $\lambda < f(\tilde{x})$. Then, by the definition of f there exist the pairs $[x_i, x_i^*] \in A$, $i = 1, \ldots, m$, such that

$$\lambda < (\tilde{x} - x_m, x_m^*) + \cdots + (x_1 - x_0, x_0^*).$$

Let $x_{m+1} = \tilde{x}$ and $x_{m+1}^* = \tilde{x}^*$. Then, again by the definition of f, one has

$$f(x) \geq (x - x_{m+1}, x_{m+1}^*) + (x_{m+1} - x_m, x_m^*) + \cdots + (x_1 - x_0, x_0^*),$$

for all $x \in X$, which implies inequality (2.44).

By the arbitrariness of $[\tilde{x}, \tilde{x}^*] \in A$, we conclude that $A \subset \partial f$. Since A is maximal in the class of cyclical sets of $X \times X^*$, it follows that $A = \partial f$, as claimed. It remains to be shown that f is uniquely determined up to an additive constant. This fact will be shown later (see Corollary 2.60 below). □

As mentioned earlier (see Theorem 1.143 and Corollary 1.140), if a maximal monotone operator $A : X \to X^*$ is coercive, then its range is all of X^*. We would like to know more about A^{-1} in the case in which A is cyclically maximal monotone. This information is contained in the following proposition.

Proposition 2.47 *Let X be reflexive and $A = \partial f$, where $f : X \to]-\infty, +\infty]$ is a lower-semicontinuous proper convex function. Then, the following conditions are equivalent.*

$$\lim_{\|x\| \to +\infty} \frac{f(x)}{\|x\|} = +\infty, \tag{2.45}$$

$$R(A) = X^* \quad \text{and} \quad A^{-1} \text{ is bounded on bounded subsets.} \tag{2.46}$$

Proof 1°. (2.45)⇒(2.46). Let x_0 be arbitrary, but fixed in $D(A)$. By the definition of ∂f, one has

$$\big(\partial f(x), x - x_0\big) \geq f(x) - f(x_0) \quad \text{for any } x \in D(A)$$

and therefore

$$\lim_{\substack{\|x\| \to \infty \\ [x,y] \in A}} \frac{(x - x_0, y)}{\|x\|} = +\infty.$$

Thus, Corollary 1.140 quoted above implies that $R(A) = X^*$. Moreover, it is readily seen that the operator A^{-1} is bounded on every bounded subset of X^*.

2°. (2.46)⇒(2.45). Inasmuch as f is bounded from below by an affine function, no loss of generality results in assuming that $f \geq 0$ on X. Let $r > 0$. Then, for every $z \in X^*$, $\|z\| \leq r$, $v \in D(A)$ and $C > 0$ such that

$$z \in Av, \quad \|v\| \leq C.$$

Next, by

$$f(u) - f(v) \geq (u - v, z) \quad \text{for all } u \text{ in } X,$$

it follows that $(u, z) \leq f(u) + Cr$ for any $u \in \mathrm{Dom}(f)$ and z in X with $\|z\| \leq r$. Hence,

$$f(u) + Cr \geq r\|u\|,$$

or

$$\frac{f(u)}{\|u\|} \geq r - \frac{Cr}{\|u\|} \quad \text{for all } u \in X.$$

This shows that condition (2.45) is satisfied, thereby completing the proof. □

Remark 2.48 A convex function f satisfying condition (2.45) is called *cofinite* on X. Recalling that $(\partial f)^{-1}$ is just the subdifferential ∂f^* of the conjugate function f^* (see Proposition 2.33). Proposition 2.47 says that a lower-semicontinuous proper convex function f is cofinite on X if and only if its conjugate f^* is everywhere finite and ∂f^* is bounded on every bounded subset of X^*. In particular, if $X = \mathbb{R}$, then condition (2.46) and $\mathrm{Dom}(f^*) = \mathbb{R}$ are equivalent. Thus, in this case, a lower-semicontinuous convex function f is *cofinite* if and only if $f^* \neq +\infty$ everywhere on X^*.

We conclude this section with some examples of cyclically monotone operators.

Example 2.49 (Maximal monotone graphs in $\mathbb{R} \times \mathbb{R}$) *Every maximal monotone graph in* $\mathbb{R}^2$ *is cyclically monotone*. Indeed, let β be a maximal monotone graph in $\mathbb{R} \times \mathbb{R}$. We prove that *there exists a lower-semicontinuous convex function* $j : R \to]-\infty, +\infty]$ such that $\partial j = \beta$. Indeed, there exist $-\infty \leq a \leq b \leq +\infty$ such that $]a, b[\subset \mathrm{Dom}(\beta) \subset [a, b]$. Let $\beta^\circ : \mathrm{Dom}(\beta) \to \mathbb{R}$ be the minimal section of β, that is, $|\beta^\circ(r)| = \inf\{|w|;\ \ w \in \beta(r)\}$ (see Sect. 1.4.1). Clearly, the function β° is single valued, monotonically increasing and, for each $r \in]a, b[$, $\beta(r) = [\beta^\circ(r-0), \beta^\circ(r+0)]$ while $\beta(a) =]-\infty, \beta^\circ(a+0)]$ if $a \in \mathrm{Dom}(\beta)$ and $\beta(b) = [\beta^\circ(b-0), +\infty[$ if $b \in \mathrm{Dom}(\beta)$ (this is an immediate consequence of the maximality).

Now, let r_0 be fixed in $\mathrm{Dom}(\beta)$ and define the function $j : \mathbb{R} \to]-\infty, +\infty]$

$$j(r) = \begin{cases} \int_{r_0}^{t} \beta^\circ(s)\,\mathrm{d}s, & \text{if } r \in [a, b], \\ +\infty, & \text{if } r \,\overline{\in}\, [a, b]. \end{cases}$$

Then, we have

$$j(r) - j(t) \leq \int_t^r \beta^\circ(s)\,\mathrm{d}s \leq \xi(r-t),$$

for all $r \in \mathrm{Dom}(\beta)$, $t \in \mathbb{R}$ and $\xi \in \beta(r)$. Hence, $\beta(r) \in \partial j(r)$ for all $r \in \mathrm{Dom}(\beta)$. We have therefore proved that $\beta = \partial j$.

By Corollary 2.60 below, the function j is uniquely defined up to an additive constant.

Example 2.50 (Self-adjoint operators in Hilbert spaces) Let H be a real Hilbert space whose norm and inner product are denoted $|\cdot|$ and $(\cdot, \cdot)$, respectively. Let A be a single-valued, linear and densely defined maximal monotone operator in H.

Proposition 2.51 *A is cyclically maximal monotone if and only if it is self-adjoint. Moreover, in this case, $A = \partial f$, where*

$$f(x) = \begin{cases} \frac{1}{2}|A^{\frac{1}{2}}x|^2, & \text{if } x \in D(A^{\frac{1}{2}}), \\ +\infty, & \text{otherwise}. \end{cases} \tag{2.47}$$

Proof First, suppose that A is self-adjoint. Then, f defined by (2.47) ($A^{\frac{1}{2}}$ denotes the square-root of the operator A) is convex and lower-semicontinuous on H (because $A^{\frac{1}{2}}$ is closed). Let $x \in D(A)$. We have

$$\frac{1}{2}\left|A^{\frac{1}{2}}x\right|^2 - \frac{1}{2}\left|A^{\frac{1}{2}}u\right|^2 \le (Ax, x-u), \quad \text{for all } u \in D\bigl(A^{\frac{1}{2}}\bigr),$$

because $(Ax, u) = (A^{\frac{1}{2}}x, A^{\frac{1}{2}}u)$ for all x in $D(A)$ and $u \in D(A^{\frac{1}{2}})$. Hence, $A \subset \partial f$.

On the other hand, it follows by a standard device that A is maximal, that is, $R(I + A) = H$. (One proves that $R(I + A)$ is simultaneously closed and dense in H.) We may conclude, therefore, that $A = \partial f$.

Suppose now that A is cyclically maximal monotone. According to Theorem 2.46, there exists $f : H \to]-\infty, +\infty]$ convex and lower-semicontinuous, such that $A = \partial f$. Inasmuch as $A0 = 0$, we may choose the function f such that $f(0) = 0$. Let $g(t)$ be the real-valued function on [0, 1] defined by

$$g(t) - f(tu),$$

where $u \in D(A)$. By the definition of the subgradient, we have

$$g(t) - g(s) \le (t - s)t(Au, u) \quad \text{for } t, s \in [0, 1].$$

The last inequality shows that g is absolutely continuous on [0, 1] and $\frac{\mathrm{d}}{\mathrm{d}t} g(t) = t(Au, u)$ almost everywhere on this interval. By integrating the above relation on [0, 1], we obtain

$$f(u) = \frac{1}{2}(Au, u) \quad \text{for every } u \in D(A)$$

and, therefore,

$$\partial f(u) = \frac{1}{2}(Au + A^*u) \quad \text{for every } u \in D(A) \cap D(A^*).$$

This, clearly, implies that $A = A^*$, as claimed. □

Example 2.52 (Convex integrands and integral functionals) Let Ω be a Lebesgue measurable subset of $\mathbb{R}^n$ and let $L^p_m(\Omega)$, $1 \le p < \infty$, be the usual Banach space of p-summable functions $y : \Omega \to \mathbb{R}^m$.

A function $g : \Omega \times \mathbb{R}^m \to \overline{\mathbb{R}}^* =]-\infty, +\infty]$ is said to be a *normal convex integrand* on $\Omega \times \mathbb{R}^m$ if the following conditions are satisfied:

(i) $g(x, \cdot) : \mathbb{R}^m \to \overline{\mathbb{R}}^*$ *is convex, lower-semicontinuous and* $\not\equiv +\infty$, *a.e.* $x \in \Omega$.
(ii) g *is measurable with respect to* σ*-field of subsets of* $\Omega \times \mathbb{R}^m$ *generated by products of Lebesgue sets in* Ω *and Borel sets in* $\mathbb{R}^m$.

It is easy to see that, if g is a normal convex integrand on $\Omega \times \mathbb{R}^m$, then for every measurable function $y : \Omega \to \mathbb{R}^m$ the function $x \to g(x, y(x))$ is Lebesgue measurable on Ω.

Condition (ii) extends the classical Carathéodory condition. In particular, it is satisfied if $g(x, y)$ is finite, measurable in x and continuous in y. If g satisfies condition (i) and $\operatorname{int} D(g(x,\cdot)) \neq \emptyset$ a.e. $x \in \Omega$, then condition (ii) is satisfied if and only if $g(x, y)$ is measurable with respect to x for each $y \in \mathbb{R}^m$. The proof of this assertion along with other sufficient conditions for normality of convex integrands can be found in the papers [61, 63] of Rockafellar who introduced and developed the theory of convex normal integrands (see also the survey of Ioffe and Levin [32]).

Besides (i), (ii), we assume that g satisfies the following two conditions:

(iii) *g increases at least one function h on $\Omega \times \mathbb{R}^m$ of the form*

$$h(x, y) = \big(y, \alpha(x)\big) + \beta(x),$$

where $\alpha \in L_m^{p'}(\Omega)$, $((p')^{-1} + p^{-1} = 1)$ *and* $\beta \in L_m^1(\Omega)$.

(iv) *There exists at least one function* $y_0 \in L_m^p(\Omega)$ *such that* $g(x, y_0) \in L^1(\Omega)$.

It must be observed that conditions (iii) and (iv) automatically hold if g is independent of x.

For any $y \in L_m^p(\Omega)$, define the integral

$$I_g(y) = \int_\Omega g\big(x, y(x)\big)\,\mathrm{d}x. \tag{2.48}$$

More precisely, the functional I_g is defined on $L_m^p(\Omega)$ by

$$I_g(y) = \begin{cases} \int_\Omega g(x, y(x))\,\mathrm{d}x, & \text{if } g(x, y) \in L_m^1(\Omega), \\ +\infty, & \text{otherwise.} \end{cases}$$

Proposition 2.53 *Let conditions* (i), (ii), (iii) *and* (iv) *be satisfied. Then, the function* $I_g : L_m^p(\Omega) \to \overline{\mathbb{R}}^*$, $1 \le p < +\infty$, *is convex, lower-semicontinuous and* $\not\equiv +\infty$. *Moreover, for every* $y \in L_m^p(\Omega)$, *the subdifferential* $\partial I_g(y)$ *is given by*

$$\partial I_g(y) = \big\{w \in L_m^{p'}(\Omega);\ w(x) \in \partial g\big(x, y(x)\big) \text{ a.e. } x \in \Omega\big\}. \tag{2.49}$$

Proof By conditions (ii) and (iv), it follows that the integral $I_g(y)$ is well defined (either a real number or $+\infty$) for every $y \in L_m^p(\Omega)$. The convexity of I_g is a direct consequence of the convexity of $g(x,\cdot)$ for every $x \in \Omega$. To prove the lower-semicontinuity of I_g, consider a sequence $\{y_n\}$ strongly convergent to y in $L_m^p(\Omega)$. On a subsequence, again denoted $\{y_n\}$, we have

$$y_n(x) \to y(x) \quad \text{a.e. } x \in \Omega$$

and, therefore,

$$g\big(x, y_n(x)\big) - \big(y_n(x), \alpha(x)\big) - \beta(x) \to g\big(x, y(x)\big) - \big(y(x), \alpha(x)\big) - \beta(x)$$
$$\text{a.e. } x \in \Omega.$$

Then, by the Fatou Lemma

$$\liminf_{n\to\infty} I_g(y_n) \geq I_g(y)$$

because $\liminf_{n\to\infty} g(x, y_n(x)) \geq g(x, y(x))(g(x,\cdot))$ is lower-semicontinuous.

Now, let $w \in \partial I_g(y)$. By the definition of $\partial I_g(y)$, we have

$$\int_\Omega \big(g\big(x, y(x)\big) - g\big(x, u(x)\big)\big)\,\mathrm{d}x \leq \int_\Omega \big(w(x), y(x) - u(x)\big)\,\mathrm{d}x$$

for all $u \in L^p_m(\Omega)$. Let E be any measurable subset of Ω and

$$\tilde{u}(x) = \begin{cases} u, & \text{if } x \in E, \\ y(x), & \text{if } x \in \Omega \setminus E, \end{cases}$$

where u is arbitrary in $\mathbb{R}^m$. We have

$$\int_E \big(g\big(x, y(x)\big) - g(x,u) - \big(w(x), y(x) - u\big)\big)\,\mathrm{d}x \leq 0.$$

Since E is arbitrary, we may conclude that

$$g\big(x, y(x)\big) \leq g(x,u) + \big(w(x), y(x) - u\big) \quad \text{a.e. } x \in \Omega,$$

and therefore

$$w(x) \in \partial g\big(x, y(x)\big) \quad \text{a.e. } x \in \Omega,$$

as claimed. Conversely, it is easy to see that every $w \in L^{p'}_m(\Omega)$ satisfying the latter belongs to $\partial I_g(y)$. □

Remark 2.54 Under the assumptions of Proposition 2.53, the function I_g is weakly lower-semicontinuous on $L^p_m(\Omega)$ (because it is convex and lower-semicontinuous). It turns out that the convexity of $g(x,\cdot)$ is also necessary for the weak lower-semicontinuity of the function I_g (see Ioffe [29, 30]). This fact has important implications in the existence of a minimum point for I_g.

We note also that in the case $p = \infty$ the structure of $\partial I_g(y) \in (L^\infty(\Omega))^*$ is more complicated and is described in Rockafellar's work [61]. (See, also, [32].) In a few words, any element $w \in \partial I_g(y)$ is of the form $w_a + w_s$, where $w_a \in L^1(\Omega)$, $w_a(x) \in \partial g(x, y(x))$, a.e., $x \in \Omega$, and $w_s \in (L^\infty(\Omega))^*$ is a singular measure.

Now, we shall indicate an extension of Proposition 2.53 to a more general context when $\mathbb{R}^m$ is replaced by an infinite-dimensional space.

Let H be a real separable Hilbert space and $[0,T]$ a finite interval of real axis. Let $\varphi : [0,T] \to \overline{\mathbb{R}}$ be such that, for every $t \in [0,T]$, the function $x \to \varphi(t,x)$ is convex, lower-semicontinuous and $\not\equiv +\infty$. Further, we assume that φ is measurable with respect to the σ-field of subsets of $[0,T] \times H$ generated by the Lebesgue sets in $[0,T]$ and the Borel sets in H.

In accordance with the terminology used earlier, we call such a function φ a convex normal integrand on $[0,T]\times H$.

Assume, further, that there exist functions $\alpha_0 \in L^{p'}(0,T;H)$, $\beta \in L^1(0,T)$ and $x_0 \in L^p(0,T;H)$ such that $\varphi(t,x_0) \in L^1(0,T)$ and

$$\varphi(t,x) \geq \big(\alpha_0(t),x\big) + \beta(t), \tag{2.50}$$

for all $x \in H$ and $t \in [0,T]$.

Define the function $I_\varphi : L^p(0,T;H) \to \overline{\mathbb{R}}^*,\ 1 \leq p < \infty$,

$$I_\varphi(x) = \begin{cases} \int_0^T \varphi(t,x)\,\mathrm{d}t, & \text{if } \varphi(t,x) \in L^1(0,T),\\ +\infty, & \text{otherwise.}\end{cases} \tag{2.51}$$

Proposition 2.55 *The function I_φ is convex, lower-semicontinuous and $\not\equiv +\infty$ on $L^p(0,T;H)$. The subdifferential ∂I_φ is given by*

$$\partial I_\varphi(x) = \big\{w \in L^{p'}(0,T;H);\ w(t) \in \partial\varphi\big(t,x(t)\big) \text{ a.e. } t \in\,]0,T[\big\}, \tag{2.52}$$

where $\frac{1}{p} + \frac{1}{p'} = 1$.

The proof closely parallels the proof of Proposition 2.53, and so, it is left to the reader.

Example 2.56 Let Ω be a bounded and open domain of $\mathbb{R}^n$ with a smooth boundary Γ. Let $g : \mathbb{R} \to \overline{\mathbb{R}}^*$ be a lower-semicontinuous convex function and let $\beta = \partial g$ be its subdifferential. Define the function $\varphi : L^2(\Omega) \to \overline{\mathbb{R}}^* =]-\infty,+\infty]$

$$\varphi(y) = \begin{cases} \frac{1}{2}\int_\Omega |\mathrm{grad}\, y|^2\,\mathrm{d}x + \int_\Omega g(y)\,\mathrm{d}x, & \text{if } y \in H_0^1(\Omega) \text{ and } g(y) \in L^1(\Omega),\\ +\infty, & \text{otherwise.}\end{cases}$$

Proposition 2.57 *The function φ is convex, lower-semicontinuous and*

$$\begin{aligned} \partial\varphi(y) &= \big\{w \in L^2(\Omega);\, w(x) \in -\Delta y(x) + \partial g\big(y(x)\big) \text{ a.e. } x \in \Omega\big\},\\ D(\partial\varphi) &= \big\{y \in H_0^1(\Omega) \cap H^2(\Omega);\ \exists\, \tilde{w} \in L^2(\Omega), \tilde{w}(x) \in \partial g\big(y(x)\big)\\ &\qquad \text{a.e. } x \in \Omega\big\}. \end{aligned} \tag{2.53}$$

Proof We have

$$\varphi(y) = I_g(y) + I_\Delta(y), \quad \forall y \in L^2(\Omega),$$

where I_g is defined by (2.48) and $I_\Delta : L^2(\Omega) \to \overline{\mathbb{R}}^*$,

$$I_\Delta(y) = -\frac{1}{2}\int_\Omega y\Delta y\,\mathrm{d}\xi = \frac{1}{2}\int_\Omega |\nabla y|^2\,\mathrm{d}\xi, \quad \forall y \in H_0^1(\Omega).$$

This implies that φ is convex and lower-semicontinuous. If we denote by $F : L^2(\Omega) \to L^2(\Omega)$ the map defined by the right-hand side of (2.53), we see that $Fy \in \partial\varphi(y)$, $\forall y \in D(F) = \{y \in H_0^1(\Omega) \cap H^2(\Omega);\ \exists \tilde{w} \in L^2(\Omega),\ \tilde{w}(x) \in \partial g(y(x))$ a.e. $x \in \Omega\}$.

To show that $F = \partial\varphi$, it suffices to check that F is maximal monotone, that is, the range of $I + F$ is all of $L^2(\Omega)$. In other words, for each $f \in L^2(\Omega)$, the elliptic equation

$$y - \Delta y + \partial g(y) \ni f \quad \text{in } \Omega; \qquad y \in H_0^1(\Omega) \cap H^2(\Omega)$$

has solution.

One might apply for this the standard existence theory for nonlinear elliptic equations or Theorem 2.65, because, as easily seen, condition (2.89), that is,

$$\int_\Omega g\big((1 + \varepsilon A)^{-1} y\big)\,dx \le \int_\Omega g(y)\,dx, \quad \forall y \in L^2(\Omega),$$

where $A = -\Delta$, $D(A) = H_0^1(\Omega) \cap H^2(\Omega)$, is satisfied. (We assume that $g(0) = 0$.)

A similar result follows for the function $\tilde{\varphi} : L^2(\Omega) \to \overline{\mathbb{R}}$, defined by

$$\tilde{\varphi}(y) = \begin{cases} \frac{1}{2}\int_\Omega |\operatorname{grad} y|^2\,dx + \int_\Gamma g(y)\,dx, & \text{if } y \in H^1(\Omega),\ g(y) \in L^1(\Gamma), \\ +\infty, & \text{otherwise.} \end{cases}$$ □

Arguing as in the preceding example, we see that φ is convex and lower-semicontinuous. As regards its subdifferential $\partial\varphi : L^2(\Omega) \to L^2(\Omega)$, it is given by (see Brezis [11, 12])

$$\partial\varphi(y) = -\Delta y, \quad \forall y \in D(\partial g), \tag{2.54}$$

where

$$D(\partial\varphi) = \left\{ y \in H^2(\Omega);\ -\frac{\partial y}{\partial \nu} \in \beta(y) \text{ a.e. on } \Gamma \right\}.$$

In particular, if $g \equiv 0$, the domain of $\partial\varphi$ consists of all $y \in H^2(\Omega)$ with zero Neumann boundary-value conditions, that is, $\frac{\partial y}{\partial \nu} = 0$ a.e. on Γ.

2.2.3 Regularization of the Convex Functions

Let X and X^* be reflexive and strictly convex. Let $f : X \to \overline{\mathbb{R}}^*$ be a lower-semicontinuous convex function and let $A = \partial f$. Since $A : X \to X^*$ is maximal monotone, for every $\lambda > 0$ the equation

$$F(x_\lambda - x) + \lambda A x_\lambda \ni 0, \tag{2.55}$$

where $F : X \to X^*$ is the duality mapping of X, has at least one solution $x_\lambda \in D(A)$ (see Theorem 1.141). The inequality

$$\big(F(x) - F(y), x - y\big) \geq \big(\|x\| - \|y\|\big)^2 \quad \text{for all } x, y \text{ in } X$$

and the strict convexity of X and X^* then imply that the solution x_λ of (2.55) is unique. We set

$$x_\lambda = J_\lambda x, \tag{2.56}$$

$$A_\lambda x = -\lambda^{-1} F(x_\lambda - x). \tag{2.57}$$

(See Sect. 1.4.1.)

For every $\lambda > 0$, we define

$$f_\lambda(x) = \inf\left\{ \frac{\|x - y\|^2}{2\lambda} + f(y);\ y \in X \right\}, \quad x \in X. \tag{2.58}$$

Since, for every $x \in X$, the infimum defining $f_\lambda(x)$ is attained, we may infer that f_λ is convex, lower-semicontinuous and everywhere finite on X. One might reasonably expect that the function f_λ "approximates" f for $\lambda \to 0$. Theorem 2.58 given below says that this is indeed the case.

Theorem 2.58 *Let $f : X \to]-\infty, +\infty]$ be a lower-semicontinuous proper and convex function on X. Let $A = \partial f$. Then, the function f_λ is Gâteaux differentiable on X and $A_\lambda = \partial f_\lambda$ for every $\lambda > 0$. In addition,*

$$f_\lambda(x) = \left(\frac{\lambda}{2}\right) \|A_\lambda x\|^2 + f(J_\lambda x) \quad \textit{for every } x \in X, \tag{2.59}$$

$$\lim_{\lambda \to 0} f_\lambda(x) = f(x) \quad \textit{for every } x \in X, \tag{2.60}$$

$$f(J_\lambda x) \leq f_\lambda(x) \leq f(x) \quad \textit{for every } x \in X \textit{ and } \lambda > 0. \tag{2.61}$$

Proof It is readily seen that the subdifferential of the function $y \to \frac{\|x-y\|^2}{2\lambda} + f(y)$ is just the operator $y \to \lambda^{-1} F(y - x) + \partial f(y)$. This fact shows that the infimum defining $f_\lambda(x)$ is attained in a point x_λ, which satisfies the equation

$$F(x_\lambda - x) + \lambda \partial f(x_\lambda) \ni 0.$$

Thus, $x_\lambda = J_\lambda x$ and equality (2.59) is immediate. Since inequality (2.61) is obvious, we restrict ourselves to verify relation (2.60). There are two cases to be considered. If $x \in \mathrm{Dom}(f)$, then $\lim_{\lambda \to \infty} J_\lambda x = x$, by using Corollary 1.70 and Proposition 1.146. This fact, combined with the lower-semicontinuity of f and inequality (2.61), shows that $\lim_{\lambda \to 0} f_\lambda(x) = f(x)$. Now, assume that $f(x) = +\infty$. We must show that $f_\lambda(x) \to +\infty$ for $\lambda \to 0$. Suppose that this is not the case, and that, for example,

$$f_{\lambda_n}(x) \leq C \quad \text{where } \lambda_n \to 0.$$

If equality (2.59) is used again, it would follow that, under the present circumstances, $J_{\lambda_n}x \to x$ and $f(J_{\lambda_n}x) \le C$. Then the lower-semicontinuity of f would imply that $f(x) \le C$, which is a contradiction. To conclude the proof, it must be demonstrated that f is Gâteaux differentiable at every point $x \in X$ and $\partial f_\lambda(x) = A_\lambda x$. A simple calculation involving relations (2.56), (2.57), and (2.59), and the definition of ∂f gives

$$f_\lambda(y) - f_\lambda(x) \le \frac{\lambda}{2}\left(\|A_\lambda y\|^2 - \|A_\lambda x\|^2\right) + (A_\lambda y, J_\lambda y - J_\lambda x),$$

that is,

$$\begin{aligned} f_\lambda(y) - f_\lambda(x) \le (A_\lambda y, y - x) + (A_\lambda y, J_\lambda y - y) + (A_\lambda y, x - J_\lambda x) \\ + \frac{\lambda}{2}\left(\|A_\lambda y\|^2 + \|A_\lambda x\|^2\right). \end{aligned}$$

Finally,

$$0 \le f_\lambda(y) - f_\lambda(x) - (A_\lambda x, y - x) \le (A_\lambda y - A_\lambda x, y - x), \tag{2.62}$$

for all $\lambda > 0$ and x, y in X.

In inequality (2.62), we set $y = x + tu$, where $t > 0$ and divide by t. We obtain

$$\lim_{t \to 0} \frac{f_\lambda(x + tu) - f_\lambda(x)}{t} = (A_\lambda x, u) \quad \text{for every } x \in X,$$

because A_λ is demicontinuous by Proposition 1.146. Therefore, f_λ is Gâteaux differentiable at any $x \in X$ and $\partial f_\lambda(x) = A_\lambda x$. □

Corollary 2.59 *In Theorem 2.58, assume that $X = H$ is a real Hilbert space. Then, the function f_λ is Fréchet differentiable of H and its Fréchet differential $\partial f_\lambda = A_\lambda$ is Lipschitzian on H.*

Proof Denote by I the identity operator in H. Then, $F = I$ and J_λ, respectively, A_λ, can be expressed as

$$J_\lambda = (I + \lambda A)^{-1}$$

and

$$A_\lambda = \lambda^{-1}(I - J_\lambda).$$

Then, A_λ is Lipschitzian on H with the Lipschitz constant $\frac{1}{\lambda}$ (see Proposition 1.146), so that inequality (2.62) yields

$$\left| f_\lambda(y) - f_\lambda(x) - (A_\lambda x, y - x) \right| \le \frac{\|y - x\|^2}{\lambda} \quad \text{for all } \lambda > 0,$$

which, obviously, implies that f is Fréchet differentiable on H. □

Corollary 2.60 *Let X be a reflexive Banach space and let f and φ be lower-semicontinuous, convex and proper functions on X. If $\partial\varphi(x) = \partial f(x)$ for every $x \in X$, then the function $x \to \varphi(x) - f(x)$ is constant on X.*

Proof Let φ_λ and f_λ be defined by formula (2.58). Then, using Theorem 2.58, we may infer that $\partial\varphi_\lambda = \partial f_\lambda$ for every $\lambda > 0$, so that

$$\varphi_\lambda(x) - f_\lambda(x) = \text{constant}, \quad \text{for every } x \in X \text{ and } \lambda > 0,$$

because φ_λ and f_λ are Gâteaux differentiable. But this clearly implies that

$$\varphi_\lambda(x) - f_\lambda(x) = \varphi_\lambda(x_0) - f_\lambda(x_0) \quad \text{for every } x \in X \text{ and } \lambda > 0,$$

where x_0 is any element in X. Again, using Theorem 2.58, we may pass to the limit, to obtain

$$\varphi(x) - f(x) = \varphi(x_0) - f(x_0) \quad \text{for every } x \in X,$$

as claimed. □

Remark 2.61 Let $X = H$ be a Hilbert space and $g(x) = \frac{1}{2}|x|^2$. Then the function f_λ can be equivalently written as

$$f_\lambda = (f^* + \lambda g)^*.$$

2.2.4 Perturbation of Cyclically Monotone Operators and Subdifferential Calculus

It is apparent that, given two lower-semicontinuous proper convex functions f and φ from X to $]-\infty, +\infty]$, then

$$\partial f(x) + \partial\varphi(x) \subset \partial(f + \varphi)(x) \quad \text{for every } x \in D(\partial f) \cap D(\partial\varphi). \tag{2.63}$$

Thus, it may be ascertained that $\partial f + \partial\varphi = \partial(f + \varphi)$ if and only if the monotone operator $\partial f + \partial\varphi$ is again maximal. More generally speaking, the following is an interesting problem: if A and B are maximal monotone operators, is $A + B$ again a maximal monotone operator? In general, the answer has to be negative since $A + B$ can even be empty, as happens, for example, if $D(A)$ does not meet $D(B)$. The main result for the problem in this line is due to Rockafellar [60] and it states that, if at least one of the maximal monotone operators A or B has a domain with a nonempty interior and $(\operatorname{int} D(A)) \cap D(B) \neq \emptyset$ (or $(D(A) \cap \operatorname{int} D(B) \neq \emptyset)$, then $A + B$ is maximal monotone. Instead of proving this theorem in full, we generally restrict ourselves to the case when $B = \partial f$.

Theorem 2.62 *Let X be a reflexive Banach space and let A be a maximal monotone operator from X to X^*. Let $f : X \to]-\infty, +\infty]$ be a lower-semicontinuous proper and convex function on X. Assume that at least one of the following conditions is satisfied.*

$$D(A) \cap \operatorname{int} \operatorname{Dom}(f) \neq \emptyset, \tag{2.64}$$

$$\operatorname{Dom}(f) \cap \operatorname{int} D(A) \neq \emptyset. \tag{2.65}$$

Then $A + \partial f$ is a maximal monotone operator.

Proof Using the renorming theorem, we can choose in X and X^* any strictly convex equivalent norms. Without loss of generality, we may assume that $0 \in D(A)$, $0 \in A0$ and $0 \in \partial f(0)$. Moreover, according to relations (2.55) and (2.65), we may further assume that

$$0 \in D(A) \cap \operatorname{int} \operatorname{Dom}(f), \tag{2.66}$$

or

$$0 \in \operatorname{Dom}(f) \cap \operatorname{int} D(A). \tag{2.67}$$

This can be achieved by shifting the domains and ranges of A and ∂f. In view of Theorem 1.141, $A + \partial f$ is maximal monotone if and only if, for every $y^* \in Y^*$, there exists $x \in D(A) \cap D(\partial f)$ such that

$$F(x) + Ax + \partial f(x) \ni y^*. \tag{2.68}$$

To show that equation (2.68) has at least one solution, consider the approximate equation

$$Fx_\lambda + Ax_\lambda + \partial f_\lambda(x) \ni y^*, \quad \lambda > 0, \tag{2.69}$$

where f_λ is the convex function defined by (2.58). According to Theorem 2.58, the operator $\partial f_\lambda = (\partial f)_\lambda$ is monotone and demicontinuous from X to X^*. Corollary 1.140 and Theorem 1.143 are therefore applicable. These ensure us that, for every $\lambda > 0$, equation (2.69) has a solution (clearly, unique) $x_\lambda \in D(A)$. Multiplying equation (2.69) by x_λ, it yields

$$\|x_\lambda\| \le \|y^*\| \quad \text{for every } \lambda > 0, \tag{2.70}$$

because $A_\lambda, \partial f_\lambda$ are monotone and $\partial f_\lambda(0) = 0, 0 \in A)$.

First, we assume that condition (2.66) is satisfied. Since f is continuous on the interior of its effective domain $\operatorname{Dom}(f)$, there is $\rho > 0$ such that

$$f_\lambda(\rho w) \le f(\rho w) \le C \quad \text{for every } w \in X, \ \|w\| = 1,$$

where C is a positive constant independent of λ and w is in X. Then, multiplying equation (2.69) by $x_\lambda - \rho w$, it yields

$$(Fx_\lambda, x_\lambda - \rho w) + (Ax_\lambda, x_\lambda - \rho w) + f_\lambda(x_\lambda) \le (y^*, x_\lambda - \rho w) + C. \tag{2.71}$$

Let $y_\lambda^* = y^* - Fx_\lambda - \partial f_\lambda(x_\lambda) \in Ax_\lambda$. In relation (2.71), we choose

$$w = -F^{-1}\left(\frac{y_\lambda^*}{\|y_\lambda^*\|}\right)$$

to obtain

$$\rho\|y_\lambda^*\| \le C \quad \text{for all } \lambda > 0. \tag{2.72}$$

(We shall denote by C several positive constants independent of λ.) Thus, with the aid of equations (2.69) and (2.70), this yields

$$\big\|\partial f_\lambda(x_\lambda)\big\| \le C \quad \text{for all } \lambda > 0. \tag{2.73}$$

Next, we assume that condition (2.67) is satisfied. Then, according to Theorem 1.144, the operator A is locally bounded at $x = 0$, so that there is $\rho > 0$, such that

$$\sup\big\{\|z^*\|;\ z^* \in Ax;\ \|x\| \le \rho\big\} \le C. \tag{2.74}$$

Let w be any element in X such that $\|w\| = 1$.

Again, multiplying equation (2.69) by $x_\lambda - \rho w$, we obtain

$$(Fx_\lambda, x_\lambda - \rho w) + \big(\partial f_\lambda(x_\lambda), x_\lambda - \rho w\big) + (Ax_\lambda, x_\lambda - \rho w) = (y_\lambda^*, x_\lambda - \rho w).$$

Then, we put

$$w = -F^{-1}\left(\frac{\partial f_\lambda(x_\lambda)}{\|\partial f_\lambda(x_\lambda)\|}\right)$$

and use the monotonicity of A and estimate (2.74) to get

$$\big\|\partial f_\lambda(x_\lambda)\big\| \le C \quad \text{for every } \lambda > 0.$$

So far, we have shown that y_λ^*, Fx_λ and $\partial f_\lambda(x_\lambda)$ remain in a bounded subset of X^*. Since the space X is reflexive, we may assume that

$$\begin{aligned} x_\lambda &\to x \quad \text{weakly in } X,\\ Fx_\lambda + y_\lambda^* &\to z^* \quad \text{weakly in } X^*. \end{aligned} \tag{2.75}$$

To conclude the proof, it remains to be seen that $[x, z^*] \in A + F$ and $y^* - z^* \in \partial f(x)$. Let $\lambda, \mu > 0$. Subtracting the corresponding equations yields

$$(Fx_\lambda + Fx_\mu, x_\lambda - x_\mu) + (y_\lambda^* - y_\mu^*, x_\lambda - x_\mu) + \big(\partial f_\lambda(x_\lambda) - \partial f_\mu(x_\mu), x_\lambda - x_\mu\big) = 0$$

and therefore

$$\lim_{\lambda,\mu\to 0}(Fx_\lambda + y_\lambda^* - Fx_\mu - y_\mu^*, x_\lambda - x_\mu) = 0 \tag{2.76}$$

because

$$\begin{aligned}&\big(\partial f_\lambda(x_\lambda)-\partial f_\mu(x_\mu),x_\lambda-x_\mu\big)\\&\quad\ge\big(\partial f_\lambda(x_\lambda)-\partial f_\mu(x_\mu),x_\lambda-J_\lambda x\lambda-x_\mu+J_\mu x_m u\big)\\&\quad\ge-\big(\|\partial f_\lambda(x_\lambda)\|+\|\partial f_\mu(x_\mu)\|\big)\big(\lambda\|\partial f_\lambda(x_\lambda)\|+\mu\|\partial f_\mu(x_\mu)\|\big).\end{aligned}$$

Here, we have used relations (2.56), (2.57) and the monotonicity of ∂f. Extracting further subsequences, if necessary, we may assume that

$$\lim_{\lambda\to 0}\big(F(x_\lambda)+y_\lambda^*,x_\lambda\big)=\ell.$$

Then, relation (2.75) shows that $(z^*,x)=\ell$. Now, let $[u,v]$ be any element in the graph of $A+F$. We have

$$(Fx_\lambda+y_\lambda^*-v,x_\lambda-u)\ge 0,\quad \forall\lambda>0.$$

Hence,

$$(z^*-v,x-u)\ge 0, \tag{2.77}$$

because $(z^*,x)=\ell$. Since F is monotone and demicontinuous from X to X^*, it follows from Corollary 1.140 quoted above that $A+F$ is maximal monotone in $X\times X^*$. Inasmuch as $[u,v]$ was arbitrary in $A+F$, then inequality (2.77) implies that $[x,z^*]\in A+F$. In other words, $z^*\in Ax+Fx$.

Now, we fix any u in X and multiply equation (2.69) by $x_\lambda-u$. It follows from the definition of the subgradient that

$$f_\lambda(x_\lambda)\le f_\lambda(u)+(y^*,x_\lambda-u)-(x_\lambda+y_\lambda^*,x_\lambda-u) \tag{2.78}$$

and therefore

$$\limsup_{\lambda\to 0} f_\lambda(x_\lambda)\le f(u)+(y^*,x-u)-(z^*,x-u). \tag{2.79}$$

Here, we have used in particular Theorem 2.58 and relation (2.77).

Since $\{\partial f_\lambda(x_\lambda);\ \lambda>0\}$ is bounded in X^*, we have

$$\lim_{\lambda\to 0}\big(x_\lambda-J_\lambda(x_\lambda)\big)=0\quad\text{strongly in } X.$$

Hence,

$$J_\lambda(x_\lambda)\to x\quad\text{weakly in } X \text{ as } \lambda\to 0.$$

We recall that a convex function f on a topological vector space X, which is lower-semicontinuous with respect to the given topology on X, is necessarily lower-semicontinuous also with respect to the corresponding weak topology on X. Thus, the combination of relations (2.59) and (2.79) yields

$$f(x)\le f(u)+(y^*,x-u)-(z^*,x-u)$$

and therefore

$$y^* - z^* \in \partial f(x),$$

because u was arbitrary in X. Hence, x satisfies equation (2.68). The proof of Theorem 2.62 is complete. □

Corollary 2.63 *Let f and φ be two lower-semicontinuous, proper and convex functions defined on a reflexive Banach space X. Suppose that the following condition is satisfied.*

$$\mathrm{Dom}(f) \cap \mathrm{int}\,\mathrm{Dom}(\varphi) \neq \emptyset. \tag{2.80}$$

Then

$$\partial(f + \varphi) = \partial f + \partial\varphi. \tag{2.81}$$

Proof Since $D(\partial\varphi)$ is a dense subset of $\mathrm{Dom}(\varphi)$ (see Corollary 2.44), condition (2.80) implies that $\mathrm{Dom}(f) \cap \mathrm{int}\, D(\partial\varphi) \neq \emptyset$. Theorem 2.62 can therefore be applied to the present situation. Thus, the operator $\partial\varphi + \partial f$ is maximal monotone in $X \times X^*$. Since $\partial\varphi + \partial f \subset \partial(\varphi + f)$, relation (2.81) follows. □

Remark 2.64 It results that Corollary 2.63 remains valid if X is a general Banach space. An alternative proof of Corollary 2.63 in this general setting will be given in the next chapter.

We conclude this section with a maximality criterion for the case in which neither $D(A)$ nor $\mathrm{Dom}(f)$ has a nonvalid interior.

Theorem 2.65 *Let $f : H \to]-\infty, +\infty]$ be a lower-semicontinuous, proper convex function on a real Hilbert space H. Let A be a maximal monotone operator from H into itself. Suppose that, for some $h \in H$ and $C \in \mathbb{R}$,*

$$f\big((I + \lambda A)^{-1}(x + \lambda h)\big) \leq f(x) + C\lambda \quad \textit{for all } x \in H \textit{ and } \lambda > 0. \tag{2.82}$$

Then the operator $A + \partial f$ is maximal monotone and

$$\overline{D(A + \partial f)} = \overline{D(A) \cap D(\partial f)} = \overline{D(A)} \cap \overline{\mathrm{Dom}(f)}. \tag{2.83}$$

Proof To prove that $A + \partial f$ is maximal monotone, it suffices to show that for every $y \in H$ there exists $x \in D(A) \cap D(\partial f)$ such that

$$x + Ax + \partial f(x) \ni y. \tag{2.84}$$

To show that this is indeed the case, consider the equation

$$x_\lambda + A_\lambda x_\lambda + \partial f(x_\lambda) \ni y, \tag{2.85}$$

where $A_\lambda = \lambda^{-1}(I - (I - \lambda A)^{-1})$. Since A_λ is monotone and continuous on H, equation (2.85) has, for every $\lambda > 0$, a unique sol $x_\lambda \in D(\partial f)$. Let x_0 be any element in $D(A) \cap D(\partial f)$. Since $\|A_\lambda x_0\| \le \|A^0 x_0\|$ and the operators A and ∂f are monotone, we see by multiplying equation (2.85) by $x_\lambda - x_0$ that $\{\|x_\lambda\|\}$ is bounded. Next, we observe that condition (2.82) implies that

$$\begin{aligned}\big(\partial f(x), A_\lambda(x+\lambda h)\big) &= \lambda^{-1}\big(\partial f(x), x+\lambda h - (I+\lambda A)^{-1}(x+\lambda h)\big)\\ &\ge \big(\partial f(x), h\big) + \big(f(x) - f(I+\lambda A)^{-1}(x+\lambda h)\big)\lambda^{-1}\\ &\ge -C - \|h\|\,\|\partial f(x)\|. \end{aligned} \tag{2.82'}$$

Now, we write equation (2.82′) as

$$x_\lambda + A_\lambda(x_\lambda + \lambda h) + \partial f(x_\lambda) = y + A_\lambda(x_\lambda + \lambda h) - A_\lambda x_\lambda$$

and multiply it (scalarly in H) by $A_\lambda(x_\lambda + \lambda h)$. Recalling that A_λ is Lipschitzian with Lipschitz constant λ^{-1}, it follows by (2.82) that $\{\|A_\lambda x_\lambda\|\}$ is bounded for $\lambda \to 0$. We subtract the defining equations for x_λ and x_μ and then multiply by $x_\lambda - x_\mu$; we obtain

$$\|x_\lambda - x_\mu\|^2 + (A_\lambda x_\lambda - A_\mu x_\mu, x_\lambda - x_\mu) \le 0.$$

Since $A_\lambda x_\lambda \in A J_\lambda x_\lambda$ and A is monotone, we see that

$$\|x_\lambda - x_\mu\|^2 \to 0 \quad \text{as } \lambda, \mu \to 0.$$

Hence, $\lim_{\lambda\to 0} x_\lambda = 0$ exists in the strong topology of H. It remains to be shown that x satisfies equation (2.84). The techniques is similar to the one previously used, but with some simplifications. Indeed, we can extract from $\{x_\lambda\}$ a subsequence $\{x_{\lambda_n}\}$ such that

$$A_{\lambda_n} x_{\lambda_n} \rightharpoonup y_0 \quad \text{in the weak topology of } H.$$

Since A is maximal monotone, it is also demiclosed (that is, its graph is strongly–weakly closed in $H \times H$) (see Proposition 1.146). Therefore, $x \in D(A)$ and $y_0 \in Ax$. The same argument applied to ∂f shows that $y - A_\lambda x_\lambda - x_\lambda$ converges weakly to $y_1 \in \partial f(x)$. Hence, x satisfies equation (2.84). To prove (2.83), we fix any x in $\overline{D(A)} \cap \overline{\mathrm{Dom}(f)}$. Then, there exist $x_\varepsilon \in \mathrm{Dom}(f)$ such that $x_\varepsilon \to x$ strongly in H as $\varepsilon \to 0$. We set $u_\varepsilon = (I + \varepsilon A)^{-1}(x_\varepsilon + \varepsilon h)$ and observe that

$$\begin{aligned}\|u_\varepsilon - x\| &\le \big\|u_\varepsilon - (I+\varepsilon A)^{-1}x\big\| + \big\|(I+\varepsilon A)^{-1}x - x\big\|\\ &\le \|x_\varepsilon - x\| + \big\|(I+\varepsilon A)^{-1}x - x\big\| + \varepsilon\|h\|.\end{aligned}$$

Hence, $u_\varepsilon \to x$ as $\varepsilon \to 0$. Moreover, by condition (2.82), $u_\varepsilon \in D(A) \cap \mathrm{Dom}(f)$. Briefly, we have shown that $\overline{D(A)} \cap \overline{\mathrm{Dom}(f)} \subset \overline{D(A) \cap \mathrm{Dom}(f)}$. Now, we prove that $D(A) \cap \mathrm{Dom}(f) \subset \overline{D(A) \cap D(\partial f)}$. Let u be any element in $D(A) \cap \mathrm{Dom}(f)$ and let $u_\varepsilon \in D(A) \cap D(\partial f)$ be the unique solution to the equation

$$u_\varepsilon + \varepsilon A u_\varepsilon + \varepsilon \partial f(u_\varepsilon) \ni u.$$

We have

$$f(u_\varepsilon) - f(u) \le \left(\frac{u - u_\varepsilon}{\varepsilon} - Au_\varepsilon, u_\varepsilon - u\right) \le -\frac{1}{\varepsilon}\|u_\varepsilon - u\|^2 - (Au, u_\varepsilon - u),$$

which implies that $\lim_{\varepsilon\to 0} u_\varepsilon = 0$. Since u is arbitrary in $D(A) \cap \mathrm{Dom}(f)$, we may infer that $D(A) \cap \mathrm{Dom}(f) \subset \overline{D(A) \cap D(\partial f)}$, as claimed. Since $\overline{D(A)} \cap \overline{\mathrm{Dom}(f)} \subset \overline{D(A) \cap \mathrm{Dom}(f)}$, Relation (2.83) follows, and this completes the proof. □

We have shown, incidentally, in the proof of Theorems 2.62 and 2.65 that, under appropriate assumptions on A and f, the solution x of the equation

$$Ax + \partial f(x) \ni 0$$

can be obtained as a limit, as λ tends to 0 of the solutions x_λ to the approximating equations

$$Ax_\lambda + \partial f_\lambda(\lambda) \ni 0.$$

This approach to construct the solution x closely resembles the *penalty method* in constrained optimization. To be more specific, let us assume that $f = I_K$, where K is a closed convex subset of a Hilbert space H and $A = \partial\varphi$.

Thus, equation $Ax + \partial f(x) \ni 0$ assumes the form

$$\min\{\varphi(x);\ x \in K\},$$

while the corresponding approximate equation can be equivalently expressed as the following unconstrained optimization problem:

$$\min\left\{\varphi(x) + \frac{1}{2\lambda}\|x - P_K x\|^2;\ x \in H\right\},$$

because $f_\lambda(x) = \frac{1}{2\lambda}\|\partial f_\lambda(x)\|^2 + f((I + \lambda\partial f)^{-1}x)$ and $(I + \lambda\partial I_K)^{-1}x = P_K x$ ($P_K x$ is the projection of x on K).

The family of continuous functions $x \to \frac{1}{2\lambda}\|x - P_K x\|^2$, $x \in H$, for a fixed $\lambda > 0$, is a family of exterior penalty functions for the closed convex set K.

Now, we prove a mean property for convex functions.

Proposition 2.66 *Let X be a real Banach space and $f : X \to \mathbb{R}$ be a continuous convex function. If x and y are distinct points of X, then there is a point z on the open segment between x and y and $w \in \partial f(z)$ such that*

$$f(x) - f(y) = (w, x - y). \tag{2.86}$$

Proof Without loss of generality, we may assume that $y = 0$. Define the function $\varphi : R \to \mathbb{R}$

$$\varphi(\mu) = f(\mu x), \quad \mu \in \mathbb{R}.$$

Since $\partial\varphi(\mu) = (\partial f(\mu x), x)$ for all $\mu \in \mathbb{R}$, it suffices to show that there exist $\theta \in]0, 1[$ and $\zeta \in \partial\varphi(\theta)$ such that $\varphi(1) - \varphi(0) = \zeta\theta$. To this end, consider the regularization φ_λ of φ defined by formula (2.58). Since φ_λ is continuously differentiable, for every $\lambda > 0$, there exists $\theta_\lambda \in]0, 1[$, such that $\varphi_\lambda(1) - \varphi_1(0) = \partial\varphi_\lambda(\theta_\lambda)$. On a sequence $\lambda_n \to 0$ we have $\theta_{\lambda_n} \to \theta$ and $\partial\varphi_{\lambda_n}(\theta_{\lambda_n}) \to \eta \in \partial\varphi(\theta)$. Since $\varphi_\lambda \to \varphi$ for $\lambda \to 0$, we infer that $\varphi(1) - \varphi(0) = \eta \in \partial\varphi(\theta)$, as claimed (obviously, $\theta \in]0, 1[$). □

2.2.5 Variational Inequalities

Let X be a reflexive real Banach space and X^* its dual space. Let A be a linear or nonlinear monotone operator form X to X^* and let K be a closed convex set of X. We say that x satisfies a *variational inequality* if

$$x \in K, \quad (Ax - f, u - x) \geq 0 \quad \text{for all } u \in K, \tag{2.87}$$

where f is given in X^*. In terms of subdifferentials, inequality (2.87) can be written as

$$Ax + \partial I_K(x) \ni f, \tag{2.88}$$

where $I_K : X \to [0, +\infty]$ is the indicator function of K (defined by relation (2.3)).

Note that, when $K = X$ or x is an interior point of K, inequality (2.87) actually reduces to the equality

$$(Ax - f, w) = 0 \quad \text{for all } w \text{ in } X,$$

that is, $Ax - f = 0$.

It should be said that many problems in the calculus of variations naturally arise in the general form of a variational inequality such as (2.87). For instance, when A is the subdifferential of a lower-semicontinuous convex function φ on X, then any solution x of the variational inequality (2.87) is actually a solution of the optimization problem

$$\text{Minimize } \varphi(x) - (f, x) \quad \text{over all } x \in K.$$

Theorem 2.67 *Let $A : X \to X^*$ be a monotone, demicontinuous operator and let K be a closed convex subset of X. In addition, assume that either K is bounded or A is coercive on K, that is, for some $x_0 \in K$,*

$$\lim_{\{\|x\| \to +\infty,\ x \in K\}} (Ax, x - x_0)\|x\|^{-1} = +\infty. \tag{2.89}$$

Then, the variational inequality (2.87) has at least one solution. Moreover, the set of solutions is bounded, closed and convex. If A is strictly monotone, the solution to (2.87) is unique.

Proof By Corollary 1.142, the operator A is maximal monotone and by Theorem 2.62, $A + \partial I_K$ is a maximal monotone subset of $X \times X^*$. Since, by assumption, $A + \partial I_K$ is coercive, it follows by Theorem 1.143 that the range $R(A + \partial I_K)$ of $A + \partial I_K$ is all of X^*. Hence, the set C of solutions to the variational inequality (2.87) is nonempty. Since $C = (A + \partial I_K)^{-1}(0)$ and $(A + \partial I_K)^{-1}$ is maximal monotone (because so is $A + \partial I_K$), we may conclude that C is convex and closed. Using the coercivity of $A + \partial I_K$, we see that C is bounded. If A is strictly monotone, that is,

$$(Ax - Ay, x - y) = 0 \quad \text{if and only if } x = y,$$

then obviously C consists of a single point. Thus, the proof is complete. □

We pause, briefly, to point out an important generalization of Theorem 2.67 (see Brezis [10]).

The operator $A : K \to X^*$ is said to be *pseudo-monotone* if the following conditions are satisfied:

(i) *If* $\{u_n\} \subset K$ *is weakly convergent to* u *in* X *and* $\limsup_{n\to\infty}(Au_n, u_n - u) \le 0$, *then* $\liminf_{n\to\infty}(Au_n, u_n - v) \ge (Au, u - v)$ *for all* $v \in K$.
(ii) *For every* $v \in K$, *the mapping* $u \to (Au, u - v)$ *is bounded from below on every bounded subset of* K.

It is easy to show that every monotone demicontinuous operator from K to X^* is *pseudo-monotone*.

The result is that Theorem 2.67 remains valid if one merely assumes that A is pseudo-monotone and coercive from K to X^*. Other existence results for the above variational inequality could be obtained by applying the general perturbations theorems given in Sect. 2.2.4. We confine ourselves to mention the following simple consequence of Theorem 2.65.

Corollary 2.68 *Let* $X = H$ *be a real Hilbert space and* K *be a closed convex subset of* H. *Let* A *be a maximal monotone (possible) multivalued operator from* H *into itself such that*

$$(I + \lambda A)^{-1}(x + \lambda h) \in K \quad \textit{for all } x \in K \textit{ and } \lambda > 0, \tag{2.90}$$

where h *is some fixed element of* H.

If, in addition, either K *is bounded, or* A *is coercive on* K, *then the variational inequality* (2.87) *has at least one solution.*

Proof Applying Theorem 2.65, where $f = I_K$, we infer that the operator $A + \partial I_K$ is maximal monotone in $H \times H$. Since $A + \partial I_K$ is coercive, this implies that its range is all of H (see Corollary 1.140).

To be more specific, let us suppose in Theorem 2.67 that $X = V$ and $X^* = V'$ are Hilbert spaces which satisfy

$$V \subset H \subset V'$$

where H is a real Hilbert space identified with its own dual and the inclusion mapping of V into H is continuous and densely defined. We further assume that the operator $A : V \to V'$ is defined by

$$(Au, v) = a(u, v) \quad \text{for all } u, v \text{ in } V,$$

where $a(u, v)$ is a bilinear continuous form on $V \times V$, which satisfies the coercivity condition

$$a(u, u) \geq \omega \|u\|^2 \quad \text{for all } u \text{ in } V, \tag{2.91}$$

where $\omega > 0$. (As usual, $\|\cdot\|$ denotes the norm in V, and $(\cdot, \cdot)$ the pairing between V and V'.) Clearly, A is linear, continuous and positive from V to V'. Let K be a closed convex subset of V. Observe that in this case the variational inequality (2.87) becomes

$$a(u, v - u) \geq (f, v - u) \quad \text{for all } v \in K. \tag{2.92}$$

In particular, if the bilinear form a is symmetric, problem (2.92) can be equivalently expressed as

$$\min \left\{ \frac{1}{2} a(v, v) - (f, v);\ v \in K \right\}. \tag{2.93}$$

□

We deduce from Theorem 2.67 the following corollary.

Corollary 2.69 *For every $f \in V'$, the variational inequality* (2.92) *has a unique solution $u \in K$.*

It should be observed that relation (2.92) implies that the mapping $f \to u$ is Lipschitzian from V' into V with Lipschitz constant $\frac{1}{\omega}$.

The variational inequality (2.92) includes several partial differential equations with unilateral boundary conditions and free boundary-value problems of elliptic type. In applications, usually A is an elliptic differential operator on a subset of $\mathbb{R}^n$, and K incorporates various unilateral conditions on the boundary Γ or on Ω. We illustrate this by a few typical examples.

Example 2.70 (The obstacle problem) Consider in a bounded open subset Ω of $\mathbb{R}^n$, the second-order differential operator

$$Av = -\big(a_{ij}(x) v_{x_i}\big)_{x_j}, \tag{2.94}$$

where the coefficients a_{ij} are in $L^\infty(\Omega)$ and satisfy the condition ($\omega > 0$)

$$a_{ij}(x)\xi_i\xi_j \geq \omega |\xi|^2, \quad \forall \xi \in \mathbb{R}^n,\ \xi = (\xi_1, \ldots, \xi_n).$$

In equation (2.94), the derivatives are taken in the sense of distributions in Ω. More precisely, the operator A is defined from $H^1(\Omega)$ to $(H^1(\Omega))'$ by

$$(Au, v) = a(u, v) = \int_\Omega a_{ij}(x) u_{x_i} v_{x_j}\,\mathrm{d}x \quad \text{for all } u, v \in H^1(\Omega). \tag{2.94$'$}$$

Let V be a linear space such that $H_0^1(\Omega) \subset V \subset H^1(\Omega)$ and let $f \in (H^1(\Omega))'$. An element $u \in V$, which satisfies the equation

$$a(u, v) = (f, v) \quad \text{for all } v \text{ in } V,$$

is a solution to a certain boundary-value problem. For instance, the Dirichlet problem

$$-(a_{ij} u_{x_i})_{x_j} = f \quad \text{in } \Omega, \qquad u = 0 \quad \text{in } \Gamma$$

arises for $V = H_0^1(\Omega)$.

Let $V = H_0^1(\Omega)$, $f \in L^1(\Omega)$, and $K = \{v \in V; v \geq \psi \text{ a.e. in } \Omega\}$, where $\psi \in H^2(\Omega)$ is a given function such that $\psi(x) \leq 0$ a.e. $x \in \Gamma$. Then, the variational inequality (2.92) becomes

$$\int_\Omega a_{ij}(x) u_{x_i} (v - u)_{x_j}\,\mathrm{d}x \geq \int_\Omega f(v - u)\,\mathrm{d}x \quad \text{for all } v \in K. \tag{2.95}$$

According to Corollary 2.69, the latter has a unique solution $u \in K$. We shall see that u can be viewed as a solution to the following boundary-value problem (the *obstacle problem*):

$$-\big(a_{ij}(x) u_{x_i}\big)_{x_j} = f \quad \text{in } E = \big\{x \in \Omega;\ u(x) > \psi(x)\big\}, \tag{2.96}$$

$$-\big(a_{ij}(x) u_{x_i}\big)_{x_j} \geq f \quad \text{in } \Omega, \tag{2.97}$$

$$u \geq \psi \quad \text{on } \Omega, \qquad u = \psi \quad \text{in } \Omega \setminus E, \qquad u = 0 \quad \text{in } \Gamma. \tag{2.98}$$

To this end, we assume that E is an open subset. Let $\alpha \in C_0^\infty(E)$ and $\rho > 0$ be such that $u \pm \rho\alpha \geq \psi$ on Ω. Then, in (2.95), we take $v = u \pm \rho\alpha$ to get

$$\int_\Omega a_{ij} u_{x_i} \alpha_{x_j}\,\mathrm{d}x = \int_E f\alpha\,\mathrm{d}x \quad \text{for all } \alpha \in C_0^\infty(E).$$

The latter shows that u satisfies equation (2.96) (in the sense of distributions). Next, we take in (2.95) $v = \alpha + \psi$, where $\alpha \in C_0^\infty(\Omega)$ is such that $\alpha \geq 0$ on Ω, to conclude that u satisfies inequality (2.97) (again in the sense of distributions). As regards relations (2.98), they are simple consequences of the fact that $u \in K$.

Problem (2.96)–(2.98) is an elliptic boundary-value problem with the *free boundary* ∂I, where I is the incidence set $\{x \in \Omega; u(x) = \psi(x)\}$. For a detailed study of this problem, we refer the reader to the recent book [37] by Kinderlehrer and Stampacchia.

As seen earlier, in the special case $a_{ij} = a_{ji}$, the variational inequality (2.95) reduces to the minimization problem

$$\min\left\{\int_\Omega a_{ij}(x)v_{x_i}v_{x_j}\,\mathrm{d}x - \int_\Omega f\,\mathrm{d}x;\ v \in K\right\}.$$

The variational inequality (2.95) models the equilibrium configuration of an elastic membrane Ω fixed at Γ, limited from below by a rigid obstacle ψ and subject to a vertical field of forces with density f (y is the deflection of the membrane). Similar free boundary-value problems occur in hydrodynamic and plasma physics. For instance, such a free boundary problem models the water flow through an isotropic homogeneous rectangular dam (see Baiocchi [3]).

Example 2.71 Suppose now that the energy integral

$$\frac{1}{2}\int_\Omega |\operatorname{grad} v|^2\,\mathrm{d}x - \int_\Omega f v\,\mathrm{d}x$$

has to be minimized on $K = \{v \in H_0^1(\Omega); |\operatorname{grad} v| \le 1,\ \text{a.e. on } \Omega\}$. As seen earlier, this problem can be equivalently expressed as

$$\int_\Omega \operatorname{grad} u \operatorname{grad}(u - v)\,\mathrm{d}x \le \int_\Omega f(u - v)\,\mathrm{d}x \quad \text{for all } v \in K.$$

This is a variational inequality of the form (2.92) and it arises in the elasto-plastic torsion of beams of section Ω under a torque field f (see Duvaut and Lions [19]). Arguing as in Example 2.56, it follows that formally the solution u satisfies the free boundary-value problem

$$\begin{aligned} &-\Delta u = f \quad \text{on } \Omega_1, \qquad u = 0 \quad \text{on } \Gamma,\\ &|\operatorname{grad} u| = 1 \quad \text{on } \Omega_2, \end{aligned}$$

where $\Omega_1 \cap \Omega_2 = \emptyset$ and $\Omega_1 \cup \Omega_2 = \Omega$.

Example 2.72 Let $a : H^1(\Omega) \times H^1(\Omega) \to \mathbb{R}$ be the bilinear form

$$a(u, v) = \int_\Omega \operatorname{grad} u \operatorname{grad} v\,\mathrm{d}x + \int_\Omega uv\,\mathrm{d}x$$

and

$$K = \left\{u \in H^1(\Omega);\ u \ge 0 \text{ a.e on } \Gamma\right\}.$$

We recall that, by Theorem 1.133, the "trace" of $u \in H^1(\Omega)$ belongs to $H^{\frac{1}{2}}(\Gamma) \subset L^2(\Gamma)$, so that K is well defined. Invoking once again Corollary 2.69, we deduce that, for every $f \in L^2(\Omega)$, the variational inequality

$$a(u, v - u) \ge \int_\Omega f(v - u)\,\mathrm{d}x, \quad \text{for all } v \in K, \tag{2.99}$$

has a unique solution $u \in K$. Let $v = u \pm \varphi$, where $\varphi \in C_0^\infty(\Omega)$. Then, inequality (2.99) yields

$$a(u,\varphi) - \int_\Omega f\varphi \,\mathrm{d}x = 0, \quad \text{for all } \varphi \in C_0^\infty(\Omega).$$

Hence,

$$-\Delta u + u = f \quad \text{on } \Omega \tag{2.100}$$

in the sense of distributions. In particular, it follows from equation (2.100) that the outward normal derivative $\frac{\partial u}{\partial \nu}$ belongs to $H^{-\frac{1}{2}}(\Gamma)$ (see Lions and Magenes [42]). We may apply Green's formula

$$\int_\Omega (\Delta u - u)v \,\mathrm{d}x = \int_\Gamma v \frac{\partial u}{\partial \nu}\,\mathrm{d}\sigma - a(u,v) \quad \text{or all } v \in H^1(\Omega). \tag{2.101}$$

In formula (2.101), we have denoted by $\int_\Gamma v \frac{\partial u}{\partial \nu}\,\mathrm{d}\sigma$ the value of $\frac{\partial u}{\partial \nu} \in H^{-\frac{1}{2}}(\Gamma)$ at $v \in H^{\frac{1}{2}}(\Gamma)$. Thus, comparing equation (2.101) with (2.99) and (2.100), it yields

$$\int_\Gamma (v-u)\frac{\partial u}{\partial \nu}\,\mathrm{d}\sigma \geq 0 \quad \text{for all } v \in K.$$

To sum up, we have shown that the solution u of the variational problem (2.99) satisfies (in the sense of distribution) the following unilateral problem:

$$\begin{aligned} &-\Delta u + u = f \quad \text{on } \Omega, \\ &u \geq 0, \qquad \frac{\partial u}{\partial \nu} \geq 0, \qquad u\frac{\partial u}{\partial \nu} = 0 \quad \text{on } \Gamma. \end{aligned} \tag{2.102}$$

Remark 2.73 The unilateral problem (2.102) is the celebrated *Signorini's problem* from linear elasticity (see Duvaut and Lions [19]) and under our assumptions on f it follows that $u \in H^2(\Omega)$ (see Brezis [12]) and equations (2.102) hold a.e. on Ω and Γ, respectively. As a matter of fact, the variational inequality (2.99) can be equivalently written as $\partial\varphi(u) \ni f$, where $\varphi : L^2(\Omega) \to]-\infty, +\infty]$ is given by (see Example 2.56)

$$\varphi(y) = \frac{1}{2}\int_\Omega |\operatorname{grad} y|^2 \,\mathrm{d}x + \int_\Gamma g(y)\,\mathrm{d}\sigma$$

and $g(r) = 0$ for $r \geq 0$, $g(r) = +\infty$ for $r < 0$.

Similarly, if $a_{ij} \in C^1(\overline{\Omega})$ and $f \in L^2(\Omega)$, then the solution u to the variational inequality (2.95) belongs to $H_0^1(\Omega) \cap H^2(\Omega)$ and satisfies the complementarity system

$$\begin{aligned} &-\big(a_{ij}(x)u_{x_i}\big)_{x_j} - f(x)\big(u(x) - \psi(x)\big) = 0 \quad \text{a.e. } x \in \Omega, \\ &u(x) \geq \psi(x); \qquad -\big(a_{ij}(x)u_{x_i}(x)\big)_{x_j} \geq f(x) \quad \text{a.e. } x \in \Omega. \end{aligned} \tag{2.96'}$$

Indeed, by Corollary 2.68, the equation

$$A_H u + \partial I_K(u) \ni f, \tag{2.103}$$

where

$$\begin{aligned} &A_H u = Au \cap H \quad \text{for } u \in D(A_H) = H_0^1(\Omega) \cap H^2(\Omega) \quad \text{and} \\ &K = \{u \in L^2(\Omega);\ u(x) \ge \psi(x) \text{ a.e. } x \in \Omega\} \end{aligned} \tag{2.104}$$

has a unique solution $u \in K \cap D(A_H)$. (It must be noticed that condition (2.90) holds for $h(x) = (a_{ij}(x)\psi_{x_i})_{x_j}$ by the maximum principle for linear elliptic equations.) Since, by Proposition 2.53,

$$\partial I_K(u) = \{w \in L^2(\Omega);\ w(x)(u(x) - \psi(x)) = 0,\ w(x) \ge 0 \text{ a.e. } x \in \Omega\}, \tag{2.105}$$

we see that u satisfies equation (2.96′), as claimed.

Example 2.74 (Generalized complementarity problem) Several problems arising in different fields such as mathematical programming, game theory, mechanics, theory of economic equilibrium, have the same mathematical form, which may be stated as follows:

For a given map A from the Banach space X into its dual space X^*, find $x_0 \in X$ satisfying

$$x_0 \in C,\ -Ax_0 \in C^\circ, \quad (x_0, Ax_0) = 0, \tag{2.106}$$

where C is a given closed, convex cone with the vertex at 0 in X and C° is its polar, that is, $C^\circ = \{x^* \in X^*; (x, x^*) \le 0$ for all $x \in C\}$.

This problem is referred to as the *generalized complementarity problem*. In the special case, when $X = X^* = \mathbb{R}^n$, $C = \mathbb{R}^n_+$ (where $\mathbb{R}^n$ is the n-dimensional Euclidean space and $\mathbb{R}^m_+$ the set of nonnegative n-vectors), the above problem takes the familiar form

$$x_0 \ge 0,\ Ax_0 \ge 0, \quad (x_0, Ax_0) = 0. \tag{2.107}$$

The following simple lemma indicates the equivalence between problem (2.106) and a variational inequality.

Lemma 2.75 *The element $x_0 \in C$ is a solution of problem (2.106) if and only if*

$$(Ax_0, x - x_0) \ge 0 \quad \textit{for all } x \in C. \tag{2.108}$$

Proof It is obvious that every solution x_0 of the complementarity problem (2.106) satisfies the above variational inequality. Let $x_0 \in C$ be any solution of inequality (2.108). Taking $x = x_0 + y$ in (2.108), where $y \in C$, it follows that $(Ax_0, y) \ge 0$. Hence, $-Ax_0 \in C^\circ$. Also, taking $x = 2x_0$, we see that $(x_0, Ax_0) \ge 0$, while, for

$x = 0$, (2.108) implies that $(x_0, Ax_0) \le 0$. Therefore $(x_0, Ax_0) = 0$. This completes the proof. □

Now, we are ready to prove the main existence result for the complementarity problem.

Theorem 2.76 *Let X be a real reflexive Banach space, C a closed convex cone in X, and let A be a monotone, demicontinuous operator from X to X^*. If, in addition, A is coercive on C, then the generalized complementarity problem* (2.106) *has at least one solution. Moreover, the set of all solutions of this problem is bounded closed convex subset of C, which consists of a single vector if A is strictly monotone.*

Proof There is nothing left to do, except to combine Theorem 2.67 with Lemma 2.75. □

As mentioned earlier, Theorem 2.67 remains valid if the operator A is pseudo-monotone and coercive from K to X^*. In particular, this happens when the space X is finite-dimensional and A is continuous and coercive on K.

Corollary 2.77 *Let X be finite-dimensional and let A be continuous on C. If, in addition, there exists a vector $x_0 \in C$ such that*

$$\lim_{\substack{\|x\| \to +\infty \\ x \in C}} \frac{(Ax, x - x_0)}{\|x\|} = +\infty, \tag{2.109}$$

then the generalized complementarity problem (2.106) *has at least one solution.*

Before leaving the subject of complementarity problems, we should point out another existence result which can be derived on the basis of Corollary 2.68.

Corollary 2.78 *Let $X = H$ be a real Hilbert space and let A be a maximal monotone (possible) multivalued operator from H into itself, which is coercive on C. Assume further that there is $h \in H$ such that*

$$(I + \lambda A)^{-1}(x + \lambda h) \subset C \quad \text{for all } x \in C \text{ and } \lambda > 0.$$

Then, problem (2.106) *has at least one solution.*

2.2.6 ε-Subdifferentials of Convex Functions

In the following we present a generalization of subdifferential taking into account its characterization with the aid of support hyperplanes to the epigraph (see Remark 2.37). It is clear that, if $x \in D(\partial f)$, then $x \in \mathrm{Dom}(f)$ and f is lower-semicontinuous at x. Conversely, for a given proper convex lower-semicontinuous function f, the existence of support nonvertical hyperplanes passing through $(x, f(x))$ is not ensured for every $x \in \mathrm{Dom}(f)$, that is, it is possible that $x \overline{\in} D(\partial f)$.

But for any $x \in \mathrm{Dom}(f)$ there exists at least one closed hyperplane passing through $(x, f(x) - \varepsilon)$, $\varepsilon > 0$, such that epi f is contained in one of the two closed half-spaces determined by that hyperplane. These hyperplanes can be considered as the approximants of support hyperplanes passing through $(x, f(x))$. Consequently, we get a notion of approximate subdifferential.

Definition 2.79 The mapping $\partial_\varepsilon f : X \to X^*$ defined by

$$\partial_\varepsilon f(x) = \left\{x^* \in X^*;\ f(x) - f(u) \leq (x - u, x^*) + \varepsilon,\ \forall u \in X^*\right\}, \tag{2.110}$$

where f is an extended real-valued function on X, is called the ε-*subdifferential* of f at x.

It is clear that this mapping is generally multivalued and $D(\partial_\varepsilon f) = \emptyset$ if f is not proper. If f is a proper function, then we must have $\varepsilon \geq 0$ and $D(\partial_\varepsilon f) \subset \mathrm{Dom}(f)$. For $\varepsilon = 0$ we obtain the subdifferential defined by Definition 2.30. Also, we have

$$\partial f(x) = \bigcap_{\varepsilon > 0} \partial_\varepsilon f(x), \quad x \in \mathrm{Dom}(f). \tag{2.111}$$

Some properties of ε-subdifferential generalize properties of subdifferential but most of their properties are different because ∂f is a local notion while $\partial_\varepsilon f$ is a global one.

Proposition 2.80 *If f is a proper convex lower-semicontinuous function, then $\partial_\varepsilon f(x)$ is a nonvoid closed convex set for any $\varepsilon > 0$ and $x \in \mathrm{Dom}(f)$.*

Proof We have $(x, f(x) - \varepsilon) \overline{\in}$ epi f for any fixed $\varepsilon > 0$, $x \in \mathrm{Dom}(f)$. By hypothesis, epi f is a nonvoid closed convex set (see Propositions 2.36 and 2.39). Using Corollary 1.45, we get a closed hyperplane passing through $(x, f(x) - \varepsilon)$ at epi f. This hyperplane is necessarily nonvertical, that is, it can be considered of the form $(x^*, 1)$. Thus, we obtain $x^* \in \partial_\varepsilon f(x)$. □

Corollary 2.81 *For any proper convex lower-semicontinuous function f we have $D(\partial_\varepsilon f) = \mathrm{Dom}(f)$, where $\varepsilon > 0$.*

It should be observed that the reverse of Proposition 2.80 is also true. Consequently, it can be given a characterization of proper convex lower-semicontinuous functions in terms of ε-subdifferentials.

Theorem 2.82 *An extended valued function f on X is convex and lower-semicontinuous if and only if $\partial_\varepsilon f(x) \neq \emptyset$ for all $x \in \operatorname{Dom}(f)$.*

Proof According to Proposition 2.80, we must prove only the sufficiency part. First, we remark that, if there exists $\bar{u} \in X$ such that $f(\bar{u}) = -\infty$, then $\bar{u} \in \operatorname{Dom}(f)$, while $\partial_\varepsilon f(\bar{u}) = \emptyset$. Hence, f must be a proper function. Now, if $x \in \operatorname{Dom}(f)$ and $(x, \alpha) \overline{\in} \operatorname{epi} f$, then there exists $\varepsilon > 0$ such that $(x, f(x) - \varepsilon) \overline{\in} \operatorname{epi} f$. But since $\partial_\varepsilon f(x) \neq \emptyset$, we have a closed nonvertical hyperplane passing through $(x, f(x) - \varepsilon)$ such that $\operatorname{epi} f$ is contained in one of the two closed half-spaces determined by that hyperplane. Consequently, $\operatorname{epi} f$ is an intersection of closed half-spaces. Hence, $\operatorname{epi} f$ is a closed set. Therefore, f is convex and lower-semicontinuous (see Propositions 2.3, 2.5). $\square$

Proposition 2.33, concerning the relationship between the subdifferential and the conjugate, becomes the following proposition.

Proposition 2.83 *Let $f : X \to]-\infty, +\infty]$ be a proper convex function. Then the following three properties are equivalent:*

(i) $x^* \in \partial_\varepsilon f(x)$.
(ii) $f(x) + f^*(x) \leq (x, x^*) + \varepsilon$.

If, in addition, f is lower-semicontinuous, then all these properties are equivalent to the following one.

(iii) $x \in \partial_\varepsilon f^*(x^*)$.

Remark 2.84 If X is reflexive, then $\partial_\varepsilon f^* : X \to X$ is just the inverse of $\partial_\varepsilon f$, that is, (i) and (iii) are equivalent for each proper convex function f.

Remark 2.85 As follows from Definition 2.79, if $x \in \operatorname{Dom}(f)$, then $f(u) \geq f(x) - \varepsilon$ for all $u \in \operatorname{Dom}(f)$ if and only if $0 \in \partial_\varepsilon f(x)$. Therefore, for a lower-semicontinuous function f, $\partial_\varepsilon f^*(0)$ is just the set of all ε-*minimum elements* of f.

Now, to describe some properties of monotonicity of ε-subdifferential we give a weaker type of monotonicity for a multivalued mapping.

Definition 2.86 A mapping $A : X \to X^*$ is called ε-*monotone* if

$$(x - y, x^* - y^*) \geq -2\varepsilon, \quad \text{for all } x^* \in Ax,\ y^* \in Ay. \tag{2.112}$$

It is obvious that $\partial_\varepsilon f$ is ε-monotone for each $\varepsilon > 0$. But while ∂f is a maximal monotone operator, $\partial_\varepsilon f$ may be not maximal ε-monotone. In this line, we shall give the following two examples.

Example 2.87 Let f be the indicator function of the closed interval $(-\infty, 0]$. By an elementary computation for a given $\varepsilon > 0$, we find $\partial_\varepsilon f(0) = [0, \infty]$, $\partial_\varepsilon f(x) = [0, -\frac{\varepsilon}{x}]$ if $x < 0$, and $\partial_\varepsilon f(x) = \emptyset$ if $x > 0$. Thus, $-2\varepsilon \overline{\in} \partial_\varepsilon f(1)$, but for any $x \in \partial_\varepsilon f(a)$, $a \leq 0$, we obtain $(x + 2\varepsilon)(a - 1) = ax - x + 2\varepsilon a - 2\varepsilon \geq -2\varepsilon$ for all $x \leq 0$. Hence, $\partial_\varepsilon f \cup \{(-2\varepsilon, 1)\}$, $\varepsilon > 0$, is also the graph of an ε-monotone operator, that is, $\partial_\varepsilon f$ is not maximal ε-monotone.

Example 2.88 Let X be a real Hilbert space and $f : X \to \mathbb{R}$ the quadratic form defined by

$$f(x) = \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + c, \quad \text{for all } x \in X,$$

where A is one-to-one linear continuous self-adjoint operator, $b \in X$ and $c \in \mathbb{R}$. For any $\varepsilon \geq 0$, we get

$$\partial_\varepsilon f(x) = Ax + b + \left\{y \in A;\ \left\langle A^{-1}y, y\right\rangle \leq 2\varepsilon\right\}, \quad \varepsilon \geq 0,\ x \in X. \tag{2.113}$$

Indeed, if $z \in \partial_\varepsilon f(x)$, then we must have

$$\frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle - \frac{1}{2}\langle Au, u\rangle - \langle b, u\rangle \leq \langle x - u, z\rangle + \varepsilon,$$

for all $u \in X$. But, for fixed $x \in X$ and $z \in \partial_\varepsilon f(x)$, this quadratic form of u takes a maximum value on X in an element u_0 where its derivative is null, that is, $Au_0 + b - z = 0$. Thus, we have

$$\frac{1}{2}\langle Ax, x\rangle + \langle v, x\rangle - \frac{1}{2}\left\langle z - b, A^{-1}(z - b)\right\rangle + \left\langle z - b, A^{-1}(z - b)\right\rangle \leq \langle x, z\rangle + \varepsilon,$$

from which we obtain

$$\langle Ax, x\rangle + 2\langle x, b - z\rangle + \left\langle A^{-1}(z - b), z - b\right\rangle \leq 2\varepsilon,$$

and so,

$$\left\langle x - A^{-1}(z - b), b - z\right\rangle + \langle x, b - z + Ax\rangle \leq 2\varepsilon.$$

Therefore, if we denote $y = z - Ax - b$, then

$$\left\langle A^{-1}y, y\right\rangle \leq 2\varepsilon,$$

that is, equality (2.113) is completely proved.

Now, let us consider $(u, v) \in X \times X$ such that $\langle x - u, z - v\rangle \geq -2\varepsilon$, for all $z \in \partial_\varepsilon f(x)$.

According to equality (2.113), it follows that

$$\langle x - u, Ax + b + y - v\rangle \geq -2\varepsilon, \quad \text{for all } x \in X, \tag{2.114}$$

and every $y \in X$ fulfilling the inequality $\langle A^{-1}y, y\rangle \le 2\varepsilon$. But the quadratic form from (2.114) has a minimal element $x_0 \in X$ where the derivative is null, that is, $2Ax_0 + b + y - v - Au = 0$. Consequently, we have

$$\frac{1}{4}\big\langle A^{-1}(v - y - v) - u, Au + y + b - v\big\rangle \ge -2\varepsilon,$$

whenever $\langle A^{-1}y, y\rangle \le 2\varepsilon$.

Taking $z = v - Au - b$, we get

$$\big\langle A^{-1}(y - z), y - z\big\rangle \le 8\varepsilon, \quad \text{if } \big\langle A^{-1}y, y\big\rangle \le 2\varepsilon. \tag{2.115}$$

Therefore it is necessary that $\langle A^{-1}z, z\rangle \le 2\varepsilon$. Indeed, if there exists $z_0 \in X$ such that $\langle A^{-1}z_0, z_0\rangle > 2\varepsilon$, it follows that $\|A^{-\frac{1}{2}}z_0\|^2 > 2\varepsilon$. Hence, $A^{-\frac{1}{2}}z_0 = (\sqrt{2\varepsilon} + a)u_0$, where $a > 0$ and $\|u_0\| = 1$. Taking $y_0 = -\sqrt{2\varepsilon}\, A^{\frac{1}{2}}u_0$, we have $\langle A^{-1}y_0, y_0\rangle = 2\varepsilon$, but $\langle A^{-1}(y_0 - z_0), y_0 - z_0\rangle^{\frac{1}{2}} = \|A^{-\frac{1}{2}}(y_0 - z_0)\| = 2\sqrt{2\varepsilon} + a > 2\sqrt{2\varepsilon}$, which contradicts (2.115). Thus, we proved that $v = Au + b + z$, where $\langle A^{-1}z, z\rangle \le 2\varepsilon$, that is, $v \in \partial_\varepsilon f(u)$. Hence, $\partial_\varepsilon f$ is a maximal ε-monotone mapping.

Remark 2.89 Since A is a self-adjoint operator, we have

$$\big\langle A^{-1}y, y\big\rangle = \big\langle A^{-\frac{1}{2}}y, A^{-\frac{1}{2}}\big\rangle = \big\|A^{-\frac{1}{2}}y\big\|^2,$$

and so, $\langle A^{-1}y, y\rangle \le 2\varepsilon$ if and only if $y = \sqrt{2\varepsilon}\, A^{\frac{1}{2}}u$, where $\|u\| \le 1$. Consequently, (2.113) can be rewritten in the form

$$\partial_\varepsilon f(x) = Ax + b + \sqrt{2\varepsilon}\, A^{\frac{1}{2}}\big(\overline{S}(0; 1)\big), \quad \varepsilon \ge 0,\ x \in X.$$

If A is the identity operator, we obtain

$$\partial_\varepsilon\left(\frac{1}{2}\|\cdot\|^2\right)(x) = x + \sqrt{2\varepsilon}\,\overline{S}(0; 1), \quad \varepsilon \ge 0,\ x \in X. \tag{2.116}$$

It is obvious that the ε-subdifferential can be considered as an enlargement of subdifferential satisfying a weak property of monotonicity. In the sequel, we prove that the ε-subdifferential can be obtained by a special type of enlargement of subdifferential. Firstly, we define the notion of ε-enlargement which was considered by Revalski and Théra [54] in the study of some important properties of monotonicity.

Definition 2.90 Given an operator $A : X \to X^*$ and $\varepsilon \ge 0$, the ε-*enlargement* of A, denoted by A^ε, is defined by

$$A^\varepsilon x = \big\{x^* \in X^*;\ (x - y, x^* - y^*) \ge -2\varepsilon,\ \text{for all } y^* \in Ay\big\}, \quad x \in X. \tag{2.117}$$

Proposition 2.91 *Let $A : X \to X^*$ be an arbitrary operator. Then, the following properties are true*:

(i) *$A^\varepsilon x$ is convex and w^* closed for any $x \in X$.*
(ii) *$A \subset A^\varepsilon$ if and only if A is ε-monotone.*
(iii) *If A is ε-monotone, then $\overline{A}$, $\operatorname{conv}\overline{A}$, $\overline{\operatorname{conv}}A$ and A^{-1} are ε-monotone.*
(iv) *$A^{\varepsilon_1} \subset A^{\varepsilon_2}$ if $0 \le \varepsilon_1 \le \varepsilon_2$.*
(v) *If A is ε-monotone and locally bounded, then $\tilde{A}$ and $\overline{\operatorname{conv}}\tilde{A}$ are ε-monotone, where $\tilde{A} : X \to X^*$ is defined as closure of* Graph A *in $X \times X^*$ with respect to strong, weak-star topology on X and X^*, respectively.*

Proof Since properties (i)–(iv) are immediate from the definition of A^ε, we confine ourselves to prove (v). Let us consider $(x, x^*), (y, y^*) \in \tilde{A}$. Hence, there exist two nets $(x_i, x_i^*)_{i\in I} \subset A$ such that $x_i \to x$, $y_i \to y$, strongly in X and $x_i^* \to x^*$, $y_i^* \to y^*$, weak-star in X^*. Since A is an ε-monotone locally bounded operator, by passing to the limit in the equality $\langle x - y, x^* - y^*\rangle = \langle x - x_i, x_i^* - y_i^*\rangle + \langle y_j - y, x_i^* - y_j^*\rangle + \langle x_i - y_j, x_i^* - y_j^*\rangle + \langle x - y, x_i^* - x_i^*\rangle + \langle x - y, y_j^* - y^*\rangle$, we obtain $\langle x - y, x^* - y^*\rangle \ge -2\varepsilon$, that is, $\tilde{A}$ is ε-monotone. According to property (iii), $\overline{\operatorname{conv}}\tilde{A}$ is also ε-monotone. □

Concerning the maximality of an ε-monotone operator, we have the following special case.

Proposition 2.92 *If A is an ε-monotone operator, then A^ε is ε-monotone if and only if there exists a unique maximal ε-monotone operator which contains A.*

Proof If B is an ε-monotone operator which contains A, then $B \subset A^\varepsilon$, and so, if A^ε is ε-monotone, then A^ε is the unique maximal ε-monotone operator. □

Generally, A^ε is not an ε-monotone operator even if A is monotone. In the special case $A = \partial f$, where f is a subdifferentiable function, the ε-enlargement $(\partial f)^\varepsilon$ is larger than the ε-subdifferential of f, that is, $\partial_\varepsilon f \subset (\partial f)^\varepsilon$. Generally, this inclusion is strict. However, formula (2.111) remains true in the case of ε-enlargement of ∂f. Firstly, it is obvious that $x^* \in A^\varepsilon x$ for all $\varepsilon > 0$ if and only if $(x^* - y^*, x - y) \ge 0$, for every $y^* \in Ay$, and so, in the case of maximal monotone operator we have the following result.

Proposition 2.93 *If A is a maximal operator, then*

$$Ax = \bigcap_{\varepsilon>0} A^\varepsilon x, \quad \text{for all } x \in X.$$

Corollary 2.94 *If f is a proper convex lower-semicontinuous function, then*

$$\partial f(x) = \bigcap_{\varepsilon>0} (\partial f)^\varepsilon(x), \quad \text{for all } x \in X. \tag{2.118}$$

Now, we give a formula for ε-differential established by Martinez-Legaz and Théra [44]. This formula proves that the ε-subdifferential can be considered as a special type of enlargement of subdifferential.

Theorem 2.95 *Let X be a Banach space and f a lower-semicontinuous proper convex function. Then*

$$\partial_\varepsilon f(x) = \left\{ x^* \in X^*;\ (x^*, x_0 - x) + \sum_{i=0}^{m-1} (x_i^*, x_{i+1} - x_i) + (x_m^*;\ x - x_m) \le \varepsilon \right.$$

$$\left. \text{for all } x_i^* \in \partial f(x_i),\ i = 0, 1, \dots, m \right\}, \tag{2.119}$$

where $x \in \mathrm{Dom}(f)$ *and* $\varepsilon > 0$.

Proof According to the proof of Theorem 2.46, for a fixed element $x_0 \in D(\partial f)$, taking $x_0^* \in \partial f(x_0)$, we have

$$f(x) = f(x_0) + \sup\left\{ \sum_{i=0}^{n-1} (x_i^*, x_{i+1} - x_i) + (x_n^*, x - x_n);\ x_i^* \in \partial f(x_i),\ i = \overline{1, n},\ n \in N^* \right\},$$

for all $x \in \mathrm{Dom}(f)$. Therefore, for any $\eta > 0$ there exist a finite set $\{x_i;\ i = \overline{1, n}\} \subset D(\partial f)$ and $x_i^* \in \partial f(x_i)$, $i = \overline{1, n}$, such that

$$\sum_{i=0}^{n-1} (x_i^*, x_{i+1} + x_i) + (x_n^*, x - x_n) > f(x) - f(x_0) - \eta.$$

Thus, if x^* is an element belonging to the right-hand side of formula (2.119), we have

$$f(x) - f(x_0) - \eta \le (x^*, x - x_0) + \varepsilon, \quad \text{for all } \eta > 0,$$

that is,

$$f(x) - f(x_0) \le (x^*, x - x_0) + \varepsilon, \quad \text{for every } x_0 \in D(\partial f).$$

Now, since $D(\partial f)$ is a dense subset of $\mathrm{Dom}(f)$ (see Corollary 2.44), by lower-semi-continuity this inequality holds for every $x_0 \in \mathrm{Dom}(f)$, and so, $x^* \in \partial_\varepsilon f(x)$.

Conversely, if $x^* \in \partial_\varepsilon f(x)$, since ∂f is cyclically monotone (see Definition 2.45), by Definition 2.104 of the ε-subdifferential it is easy to see that x^* satisfies the inequality of the right-hand side of formula (2.119), thereby proving Theorem 2.95. □

Remark 2.96 The multivalued operator defined by the right-hand side of (2.119) can be considered the *ε-enlargement cyclically monotone* of ∂f.

2.2.7 Subdifferentiability in the Quasi-convex Case

Here, we consider the special case of *quasi-convex functions*. (See Sect. 2.1.1.) We recall that a function is quasi-convex lower-semicontinuous if and only if its level sets are closed convex sets. Thus, similarly to the convex case, if the role of epigraph is replaced by level sets, the continuous linear functionals that describe the closed semispaces whose intersection is a certain level set are candidates for the approximative quasi-subdifferentials (see Theorem 1.48). Given a function f and $\lambda \in \mathbb{R}$, we denote by $N^\lambda(f)$ the corresponding level set, that is,

$$N^\lambda(f) = \{x \in X;\ f(x) \le \lambda\}. \tag{2.120}$$

Let us consider the following sets:

$$D_\lambda f(x_0) = \{(x^*, \delta) \in X^* \times (0, \infty);\ x^*(x_0 - x) \ge \delta \text{ whenever } f(x) \le \lambda\}, \tag{2.121}$$

for every $x_0 \in X$ and $\lambda \in \mathbb{R}$.

It is obvious that, if $D_\lambda f(x_0) \ne \emptyset$, then $f(x_0) > \lambda$. Indeed, if we suppose that $f(x_0) \le \lambda$, then, for an element $(x^*, \delta) \in D_\lambda f(x_0)$, we have $0 = x^*(x_0 - x_0) \ge \delta$, which is a contradiction with the choice of δ.

Definition 2.97 The projection of $D_\lambda f(x_0)$ on X^* is called the λ-*quasi-subdifferential* of f at x_0 and is denoted by $\partial_q^\lambda f(x_0)$.

Taking into account the correspondence between the *convexity* and *quasi-convexity*, we see that this type of approximate subdifferential is proper to the quasi-convex functions.

Indeed, it is well known that a function f is convex if and only if the associated function $F_f : X \times \mathbb{R} \to \overline{\mathbb{R}}$ defined by

$$F_f(x, t) = f(x) - t, \quad (x, t) \in X \times \mathbb{R}, \tag{2.122}$$

is quasi-convex, since $N^\lambda(F_f) = -(0, \lambda) + \operatorname{epi} f$, for all $\lambda \in \mathbb{R}$. Thus, we have

$$\begin{aligned} D_\lambda F_f(x_0, t_0) = \{&(x^*, \alpha, \delta) \in X^* \times \mathbb{R} \times (0, \infty);\ x^*(x_0 - x) + \alpha(t_0 - t) \ge \delta, \\ &\text{whenever } f(x) - t \le \lambda\}. \end{aligned}$$

By a simple calculation, we find that $(x^*, \alpha, \delta) \in D_\lambda F_f(x_0, t_0)$ if $\alpha = 0$ and $\sup\{(x^*, x); x \in \operatorname{Dom}(f)\} \le x^*(x_0) - \delta$ or $\alpha < 0$ and $-\frac{x^*}{\alpha} \in \partial_{\varepsilon_0} f(x_0)$, where $\varepsilon_0 = f(x_0) - t_0 - \lambda - \frac{\delta}{\alpha}$. We recall that, necessarily, we must have $f(x_0) - t_0 = F_f(x_0, t_0) > \lambda$, $\alpha \le 0$, whenever $D_\lambda F_f(x_0, t_0) \ne \emptyset$.

Therefore, the projection on X^* contains elements of approximative subdifferential defined for convex functions. More precisely, $(x^*, -1, \delta) \in D_\lambda F_f(x_0, 0)$ if and only if $x^* \in \partial_{\varepsilon_0} f(x_0)$, $\varepsilon_0 = f(x_0) - t_0 - \lambda > 0$, $x_0 \in \operatorname{Dom}(f)$.

Now, we can establish the following characterization of quasi-convex lower-semicontinuous functions.

Theorem 2.98 *A function $f : X \to \overline{\mathbb{R}}$ is quasi-convex and lower-semicontinuous if and only if, for all $\lambda \in \mathbb{R}$ and $x_0 \in X$ such that $f(x_0) > \lambda$, the set $D_\lambda f(x_0)$ is nonempty.*

Proof According to Theorem 1.48, the function f is quasi-convex and lower-semicontinuous if and only if its level sets can be represented as an intersection of closed half-spaces.

Equivalently, for every $x_0 \overline{\in} N^\lambda(f)$ there exists a closed hyperplane strongly separating $N^\lambda(f)$ and x_0. Thus, if $f(x_0) > \lambda$, there exist $x^* \in X^* \setminus \{0\}$ and $k \in \mathbb{R}$ such that $x^*(x_0) > k$ and $x^*(x) \le k$ for all $x \in N^\lambda(f)$. Taking $\delta = x^*(x_0) - k > 0$, we obtain $x^*(x - x_0) \le -\delta$ for all $x \in N^\lambda(f)$, equivalently $(x^*, \delta) \in D_\lambda f(x_0)$. This finishes the proof of Theorem 2.98. □

Corollary 2.99 *A proper function $f : X \to \overline{\mathbb{R}}$ is quasi-convex and lower-semicontinuous if and only if $\partial_q^\lambda f(x_0) \neq \emptyset$ for all $x_0 \in X$, $\lambda \in \mathbb{R}$, with $f(x_0) > \lambda$.*

Now, it is easy to see that the λ-quasi-subdifferential of a function f can also be defined by the formula

$$\partial_q^\lambda f(x_0) = \Big\{ x^* \in X^*;\ \sup_{x \in N^\lambda f} x^*(x - x_0) < 0 \Big\}. \tag{2.123}$$

Proposition 2.100 *Let us consider $f : X \to \overline{R}$, $x_0 \in X$, $f(x_0) \neq -\infty$, $\varepsilon > 0$. Then the following properties are equivalent*:

(i) *x_0 is an ε-minimum element of f.*
(ii) *$\partial_q^\lambda f(x_0) = X^*$, whenever $\lambda < f(x_0) - \varepsilon$.*
(iii) *$0 \in \partial_q^\lambda f(x_0)$, whenever $\lambda < f(x_0) - \varepsilon$.*

Proof If there exists $x_1 \in X$ such that $f(x_1) < f(x_0) - \varepsilon$, then, taking $\lambda = f(x_1)$, we have $N^\lambda(f) \neq \emptyset$ and so, $0 \overline{\in} \partial_q^\lambda f(x_0)$. On the other hand, if $0 \in \partial_q^\lambda f(x_0)$, then, for all $\lambda < f(x_0) - \varepsilon$, we get $N^\lambda(f) = \emptyset$, that is, $f(x) \ge f(x_0) - \varepsilon$ for all $x \in X$. Also, (ii) and (iii) are obviously equivalent. □

In the following, we establish some relationships between the quasi-subdifferential defined by (2.123) and other two notions of quasi-subdifferentials introduced as extensions to the case quasi-convex of the subdifferential of a convex function. We denote

$$\partial_{\mathrm{GP}}^\lambda f(x_0) = \big\{ x^* \in X^*;\ x^*(x - x_0) < 0 \text{ if } f(x) < \lambda \big\}, \quad x_0 \in X, \tag{2.124}$$

$$\begin{aligned} \partial_{\mathrm{M\text{-}L}} f(x_0) = \big\{ x^* \in X^*;\ &\text{there exists } k \in K \text{ such that } k \circ x^* \le f \\ &\text{and } k\big(x^*(x_0)\big) = f(x_0) \big\}, \quad x_0 \in X, \end{aligned} \tag{2.125}$$

where K is a given family of functionals $k \in \mathbb{R} \to \overline{\mathbb{R}}$ closed under pointwise supremum.

If $\lambda = f(x_0) \in \mathbb{R}$, the λ-*quasi-subdifferential* (2.124) was introduced by Greenberg and Pierskalla [23] for $X = \mathbb{R}^n$, while the *quasi-subdifferential* (2.125) was introduced by Martinez-Legaz and Sach [43]. It is well known that $\partial_{\mathrm{GP}} f(x_0) = \partial_K f(x_0)$ if K is the family of all nondecreasing functions.

The λ-quasi-subdifferential associated to the quasi-subdifferential (2.125) is defined as follows:

$$\partial^{\lambda}_{\mathrm{M\text{–}L}} f(x_0) = \big\{x^* \in X^*;\ \text{there exists } k \in K \text{ such that } k \circ x^* \le f \\ \text{and } k\big((x^*)(x_0)\big) \ge \lambda\big\}. \tag{2.126}$$

Proposition 2.101 *Let K be the family of all nondecreasing functions $k : \mathbb{R} \to \overline{\mathbb{R}}$. If $f : X \to \overline{\mathbb{R}}$, $x_0 \in X$ and $\lambda \in \mathbb{R}$, then*

$$\partial^{\lambda}_{\mathrm{GP}} f(x_0) = \partial^{\lambda}_{\mathrm{M\text{–}L}} f(x_0).$$

Proof From the definition of $\partial^{\lambda}_{\mathrm{M\text{–}L}}$ given by (2.126), we obtain the inclusion $\partial^{\lambda}_{\mathrm{M\text{–}L}} f(x_0) \subset \partial^{\lambda}_{\mathrm{GP}} f(x_0)$. Conversely, if $x^* \in \partial^{\lambda}_{\mathrm{GP}} f(x_0)$, taking $k : \mathbb{R} \to \overline{\mathbb{R}}$ defined by

$$k(t) = \inf\big\{a;\ x^*(x) \ge t \text{ if } f(x) < a\big\},$$

we have $k(x^*(x)) \le a$ whenever $f(x) < a$. But k is obvious a nondecreasing function, and so $k \circ x^* \le f$. Also, $k(x^*(x_0)) \ge \lambda$. Hence, $x^* \in \partial^{\lambda}_{\mathrm{M\text{–}L}} f(x_0)$ and the proof is complete. □

Proposition 2.102 *Let K be the family of all nondecreasing lower-semicontinuous functions. If $f : X \to \overline{\mathbb{R}}$, $x_0 \in X$, $\lambda_1, \lambda_2 \in \mathbb{R}$ and $\lambda_1 > \lambda_2$, then*

(i) $\partial^{\lambda_1}_{\mathrm{M\text{–}L}} f(x_0) \subset \partial^{\lambda_2}_{q} f(x_0) \subset \partial^{\lambda_2}_{\mathrm{M\text{–}L}} f(x_0)$.
(ii) $\bigcap_{\lambda < f(x_0)} \partial^{\lambda}_{q} f(x_0) = \bigcap_{\lambda < f(x_0)} \partial^{\lambda}_{\mathrm{M\text{–}L}} f(x_0) = \partial_{\mathrm{M\text{–}L}} f(x_0)$, *if* $f(x_0) \in \mathbb{R}$.

Proof Equality (ii) follows by using (i) and the equality

$$\bigcap_{\lambda < f(x_0)} \partial^{\lambda}_{q} f(x_0) = \partial_{\mathrm{M\text{–}L}} f(x_0).$$

Now, if $x^* \in \partial^{\lambda}_{q} f(x_0)$, taking the function k defined in the proof of Proposition 2.101, we notice that k is also lower-semicontinuous. Hence, $k(x^*(x)) \le a$ if $f(x) < a$, and so, $k \circ x^* \le f$. Since $\sup_{x \in N^{\lambda}(f)} x^*(x - x_0) < 0$, it follows that $k(x^*(x_0)) \ge \lambda$. Hence, $\partial^{\lambda}_{q} f(x_0) \subset \partial^{\lambda}_{\mathrm{M\text{–}L}} f(x_0)$. On the other hand, if $\partial^{\lambda}_{\mathrm{M\text{–}L}} f(x_0) = \emptyset$ or $N^{\lambda}(f) = \emptyset$, then the inclusion of the left-hand side of (i) is obvious. Let us suppose that $N^{\lambda}(f) \ne \emptyset$. Thus, if $x^* \in \partial^{\lambda_1}_{\mathrm{M\text{–}L}} f(x_0)$, we have $k(x^*(x)) - k(x^*(x_0)) < \lambda - \lambda_1$, for all x, such that $f(x) \le \lambda$. Let us denote $\alpha = \sup_{x \in N^{\lambda}(f)} x^*(x - x_0)$ and consider a net $(x_i) \subset N^{\lambda}(f)$ such that $x^*(x_i) \to \sup_{x \in N^{\lambda}(f)} x^*(x)$. Since $k(x^*(x_i)) - k(x^*(x_0)) < \lambda - \lambda_1$, by passing to the limit we obtain

$$k\big(x^*(x_0) + \alpha\big) - k\big(x^*(x_0)\big) \le \lambda - \lambda_1 < 0.$$

Hence, $\alpha < 0$ and so, $x^* \in \partial_q^\lambda f(x_0)$. Thus, Proposition 2.102 is completely proved. □

2.2.8 Generalized Gradients

In this section, we briefly present a theory of generalized gradients for lower-semicontinuous functions of $\mathbb{R}^n$ due to Clarke [17]. This theory is still under development but some significant results have already become known.

Assume first that $f : \mathbb{R}^n \to \mathbb{R}$ is a locally Lipschitz function. According to Rademacher's theorem, f is a.e. differentiable on $\mathbb{R}^n$. By definition, the *generalized gradient* of f at x, denoted by $\partial f(x)$, is the convex hull of the set of points of the form $\{\lim_{n\to\infty} \nabla f(x + x_n)\}$, where $x_n \to 0$ and $\nabla f(x + x_n)$ (the gradient of f at $x + x_n$) exist.

In order to extend this definition to general lower-semicontinuous functions, we consider a closed subset C of $\mathbb{R}^n$ and denote by $d_C(x)$ the distance from x to C, that is,

$$d_C(x) = \inf\big\{\|x - y\|;\ y \in C\big\}.$$

Since d_C is locally Lipschitz, we may define ∂d_C. By analogy with the case when C is convex, we define the *cone of normals* to C at x, denoted $N(x; C)$, the closure of the set

$$\big\{z \in \mathbb{R}^n;\ \lambda z \in \partial d_C(x) \text{ for some } \lambda > 0\big\}. \tag{2.127}$$

We observe that, if C is convex, then, by Theorem 2.58, where $f = I_C$, it follows that d_C is differentiable outside C and

$$\nabla d_C(x) = \big(x - P_C(x)\big)\big\|x - P_C(x)\big\|^{-1}, \quad x \in C,$$

where P_C is the projection operator on C (we take the Euclidean norm on $\mathbb{R}^n$). Hence, for all $x \in \mathbb{R}^n$, we have

$$\nabla d_C(x) \in \partial I_C(P_C x)$$

and, therefore, if C is convex, then $N(x; C)$ is just the cone of normals to C at x (see Example 2.31).

It is obvious that, if f is continuously differentiable on a neighborhood of x, then $\partial f(x) = \nabla f(x)$. If f is convex, then its epigraph $E(f)$ is a convex closed subset of $\mathbb{R}^{n+1}$ and, as observed earlier, $N((x, f(x)); E(f)) = N_{E(f)}(x; f(x))$. Hence, in this case, $\partial f(x)$ is the set of all subgradients of f at x (here, $E(f) = \operatorname{epi} f$).

Given the lower-semicontinuous function $f : \mathbb{R}^n \to \mathbb{R}$, we define the *upper derivative* of f at x with respect to y, as

$$f^{\uparrow}(x, y) = \lim_{\substack{x' \to x \\ f(x') \to f(x) \\ \lambda \downarrow 0}} \ \inf_{y' \to y} \frac{f(x' + \lambda y') - f(x')}{\lambda}. \tag{2.128}$$

It should be observed that, if f is convex, then $f^{\uparrow} = f'$.

Now, let x be a point where $f(x)$ is finite.
We define

$$\partial f(x) = \{z \in \mathbb{R}^n;\ (z, -1) \in N((x, f(x)); E(f))\}$$

and call $\partial f(x)$ the *generalized gradient* of f at x.

Proposition 2.103 *The generalized gradient $\partial f(x)$ is also given by*

$$\partial f(x) = \{z \in \mathbb{R}^n;\ f^{\uparrow}(x, y) \geq (y, z),\ \forall y \in \mathbb{R}^n\}. \tag{2.129}$$

If $f^{\uparrow}(x, 0) = -\infty$, then $\partial f(x)$ is empty, but otherwise $\partial f(x) \neq \emptyset$ and one has

$$f^{\uparrow}(x, y) = \max\{(y, z);\ z \in \partial f(x),\ \forall y \in \mathbb{R}^n\}. \tag{2.130}$$

The reader will be aware of the analogy between Propositions 2.39 and 2.103. Formula (2.129) represents another way (due to Rockafellar) to define the generalized gradient. The proof of Proposition 2.103, which is quite technical, can be found in the work of Rockafellar [64] (see also [65, 66]). In this context, the works of Hirriart-Urruty [25, 26] must be also cited. The above definition of generalized gradient can be extended to infinite-dimensional Banach space. For instance, if X is a Banach space and $f : X \to \mathbb{R}$ a locally Lipschitz function, we define the generalized directional derivative of f at x in the direction z, denoted by $f^0(x, z)$ by

$$f^0(x, z) = \limsup_{\substack{x' \to x \\ \lambda \downarrow 0}} \frac{f(x' + \lambda z) - f(z)}{\lambda}.$$

If $X = \mathbb{R}^n$, then $f^0 = f^{\uparrow}$.

It is easy to see that f^0 is a positively homogeneous and subadditive function of z. Thus, by the Hahn–Banach theorem, we may infer that there exists at least one $x^* \in X^*$ satisfying

$$f^0(x, z) \geq (z, x^*) \quad \text{for all } z \in X. \tag{2.131}$$

By definition, the *generalized gradient* of f at x, denoted by $\partial f(x)$ is the set of all $x^* \in X^*$ satisfying (2.131).

It is readily seen that, for every $x \in X$, $\partial f(x)$ is a nonempty, closed, convex and bounded subset of X^*, thus $\partial f(x)$ is w^*-compact. Moreover, ∂f is w^*-upper-semicontinuous, that is, if $\eta_i \in \partial f(x)$, where $\eta_i \to \eta$ weak-star in X^* and $x_i \to x$ strongly in X, then $\eta \in \partial f(x)$ (see Clarke [18]). Note also that $f^0(x, \cdot)$ is the support functional of $\partial f(x)$, that is, for any z in X, we have (compare with (2.130))

$$f^0(x, z) = \max\{(z, x^*);\ x^* \in \partial f(x)\}.$$

For the definition and the properties of generalized gradient of vectorial functions defined on Banach spaces, we refer the reader to the work of Thibault [73].

2.3 Concave–Convex Functions

This section is concerned mainly with minimax problems for concave–convex functions. This subject is discussed in some detail in Sect. 2.3.3. Relevant to it are the closed saddle functions studied in Sect. 2.3.2.

2.3.1 Saddle Points and Mini-max Equality

Let X, Y be two nonempty sets and let F be an extended real-valued function on the product set $X \times Y$.

It is easy to prove that we always have

$$\sup_{x\in X} \inf_{y\in Y} F(x, y) \leq \inf_{y\in Y} \sup_{x\in X} F(x, y). \tag{2.132}$$

If the equality holds, the common value is called the *saddle value* of F on $X \times Y$. Furthermore, we shall require that the supremum from the left side and the infimum from the right side are actually achieved. In this case, we say that F verifies the *mini-max equality* on $X \times Y$ and we denote this by

$$\max_{x\in X} \min_{y\in Y} F(x, y) = \min_{y\in Y} \max_{x\in X} F(x, y).$$

Of course, *the mini-max equality holds if and only if the following three conditions are satisfied*:

(i) *F has saddle value, that is*, $\sup_{x\in X} \inf_{y\in Y} F(x, y) = \inf_{y\in Y} \sup_{x\in X} F(x, y)$.
(ii) *There is $\tilde{x} \in X$ such that* $\inf_{y\in Y} F(\tilde{x}, y) = \sup_{x\in X} \inf_{y\in Y} F(x, y)$.
(iii) *There is $\tilde{y} \in Y$ such that* $\sup_{x\in X} F(x, \tilde{y}) = \inf_{y\in Y} \sup_{x\in X} F(x, y)$.

Clearly, $F(\tilde{x}, \tilde{y})$ is the saddle value of F. Also, $\sup_{x\in X} F(x, \tilde{y})$ and $\inf_{y\in Y} F(\tilde{x}, y)$ are attained, respectively, at $\tilde{x}$ and $\tilde{y}$ since, from conditions (ii) and (iii), one easily obtains

$$\sup_{x\in X} \inf_{y\in Y} F(x, y) = \inf_{y\in Y} F(\tilde{x}, y) \leq F(\tilde{x}, \tilde{y}) \leq \sup_{x\in X} F(x, \tilde{y}) = \inf_{y\in Y} \sup_{x\in X} F(x, y).$$

According to condition (i), this inequality becomes an equality. Moreover, we obtain

$$\sup_{x\in X} F(x, \tilde{y}) = F(\tilde{x}, \tilde{y}) = \inf_{y\in Y} F(\tilde{x}, y)$$

from which we obtain

$$F(x, \tilde{y}) \leq F(\tilde{x}, \tilde{y}) \leq F(\tilde{x}, y), \quad \forall (x, y) \in X \times Y. \tag{2.133}$$

Definition 2.104 The pair $(\tilde{x}, \tilde{y}) \in X \times Y$ is said to be a *saddle point* for the function $F : X \times Y \to \mathbb{R}$ if relation (2.133) holds.

Thus, the mini-max equality implies the existence of a saddle point.

It is easily proven that the converse of this statement is also true. Indeed, from (2.133), we have

$$\inf_{y\in Y}\sup_{x\in X} F(x,y) \le \sup_{x\in X} F(x,\tilde{y}) \le \inf_{y\in Y} F(\tilde{x},y) \le \sup_{x\in X}\inf_{y\in Y} F(x,y),$$

which, by (2.132), implies conditions (i), (ii) and (iii). Thus, the following fundamental result holds.

Proposition 2.105 *A function satisfies the mini-max equality if and only if it has a saddle point.*

2.3.2 Saddle Functions

The purpose of this section is to present a new class of functions (that is, functions which are partly convex and partly concave), which are closely related to extremum problems.

We assume in everything that follows that X and Y are real Banach spaces with duals X^* and Y^*. For the sake of simplicity, we use the same symbol $\|\cdot\|$ to denote the norms $\|\cdot\|_X$, $\|\cdot\|_Y$, $\|\cdot\|_{X^*}$ and $\|\cdot\|_{Y^*}$ in the respective spaces X, Y, X^* and Y^*. As usual, we use the symbol $(\cdot,\cdot)$ to denote the pairing between X, X^* and Y, Y^*, respectively. If f is an arbitrary convex function on X, then we use the symbol $\mathrm{cl}\, f$ to denote its closure (see Sect. 2.1.3). For a concave function g, the closure $\mathrm{cl}\, g$ is defined by

$$\mathrm{cl}\, g = -\mathrm{cl}(-g).$$

Definition 2.106 By a *saddle function* on $X \times Y$, we mean an extended real-valued function K defined everywhere, such that $K(x,y)$ is a concave function of $x \in X$ for each $y \in Y$, and a convex function of $y \in Y$ for each $x \in X$.

Given a saddle function K on $X \times Y$, we denote by $\mathrm{cl}_1 K$ the function obtained by closing $K(x,y)$ as a concave function of x for each y. Similarly, $\mathrm{cl}_2 K$ is obtained by closing $K(x,y)$ as a convex function of y for each x.

Definition 2.107 A saddle function K is said to be *closed* if the following conditions hold:

$$\mathrm{cl}_1\,\mathrm{cl}_2\, K = \mathrm{cl}_1\, K, \qquad \mathrm{cl}_2\,\mathrm{cl}_1\, K = \mathrm{cl}_2\, K. \tag{2.134}$$

It should be observed that conditions (2.134) automatically hold if $K(x,y)$ is upper-semicontinuous in x and lower-semicontinuous in y. Two saddle functions K and K' are said to be equivalent if

$$\mathrm{cl}_1\, K = \mathrm{cl}_1\, K' \quad \text{and} \quad \mathrm{cl}_2\, K = \mathrm{cl}_2\, K'.$$

In other words, the saddle function K is closed if $\mathrm{cl}_1 K$ and $\mathrm{cl}_2 K$ are equivalent to K.

It is worth mentioning that equivalent saddle functions have the same saddle value and saddle points (if any). In fact, let K be an arbitrary saddle function on $X \times Y$. Inasmuch as the infimum of a convex function is the same as the infimum of its closure, one obtains

$$\inf\{K(x,y);\ y \in Y\} = \inf\{\mathrm{cl}_2 K(x,y);\ y \in Y\} \quad \text{for every } x \in X, \tag{2.135}$$

and, similarly,

$$\sup\{K(x,y);\ x \in X\} = \sup\{\mathrm{cl}_1 K(x,y);\ x \in X\} \quad \text{for every } y \in Y. \tag{2.136}$$

Hence, if (x_0, y_0) is a saddle point of K, that is,

$$K(x, y_0) \le K(x_0, y_0) \le K(x_0, y) \quad \text{for all } (x, y) \in X \times Y,$$

we have

$$\sup\{\mathrm{cl}_1 K(x, y_0);\ x \in X\} = K(x_0, y_0) = \inf\{\mathrm{cl}_2 K(x_0, y);\ y \in Y\}$$

and therefore for any saddle function K' equivalent with K,

$$\sup\{K'(x, y_0);\ x \in X\} = K(x_0, y_0) = \inf\{K'(x_0, y);\ y \in K\},$$

which implies that $K(x_0, y_0) = K'(x_0, y_0)$, and therefore (x_0, y_0) is a saddle point of K'.

Let K be a saddle function on $X \times Y$ and let

$$D_1(K) = \{x \in X;\ K(x,y) > -\infty \text{ for every } y \in Y\}, \tag{2.137}$$

$$D_2(K) = \{y \in Y;\ K(x,y) < +\infty \text{ for every } x \in X\}. \tag{2.138}$$

It is easy to see that $D_1(K)$ and $D_2(K)$ are convex sets. The set

$$\mathrm{dom}\, K = D_1(K) \times D_2(K) \tag{2.139}$$

is called the *effective domain* of K. Obviously, K is finite on $\mathrm{dom}\, K$ and, if K is finite everywhere, one has $\mathrm{dom}\, K = X \times Y$.

As an example, let A and B be nonempty convex sets in X and Y, respectively, and let

$$K(x,y) = \begin{cases} K_0(x,y), & \text{if } x \in A \text{ and } y \in B, \\ +\infty, & \text{if } x \in A \text{ and } y \,\overline{\in}\, B, \\ -\infty, & \text{if } x \,\overline{\in}\, A \text{ and } y \in Y, \end{cases} \tag{2.140}$$

where K_0 is any finite saddle function on $A \times B$. Then, K is a saddle function on $X \times Y$ with

$$\mathrm{dom}\, K = A \times B.$$

A saddle function $K : X \times Y \to \overline{\mathbb{R}} = [-\infty, +\infty]$ is called *proper* if $\operatorname{dom} K \neq \emptyset$.

Most of the results which are proved below closely resemble the corresponding properties of lower-semicontinuous convex functions previously established.

Theorem 2.108 *Let K be a closed proper saddle function on $X \times Y$. Then*

(i) *For every $y \in \operatorname{int} D_2(K)$, the function $K(\cdot, y)$ is concave, upper-semicontinuous and proper on X. Furthermore, its effective domain coincides with $D_1(K)$.*
(ii) *For every $y \in \operatorname{int} D_1(K)$, the function $K(x, \cdot)$ is convex, lower-semicontinuous and proper on Y, and its effective domain is $D_2(K)$.*

Proof (i) The closedness of K implies that $\operatorname{cl}_1 \operatorname{cl}_2 K = \operatorname{cl}_1 K$. Hence

$$\operatorname{cl}_1 K(x, y) = \lim_{\varepsilon \to 0} \sup_{\|x-u\| \le \varepsilon} \operatorname{cl}_2 K(u, y) \quad \text{for every } y \in D_2(K).$$

We set

$$\varphi_\varepsilon(x, y) = \sup_{\|x-u\| \le \varepsilon} \operatorname{cl}_2 K(u, y).$$

Since $\operatorname{cl}_2 K \le \operatorname{cl}_1 K$ and the function $x \to \operatorname{cl}_1 K(x, y)$, $x \in X$, is upper-semicontinuous and concave on X, we may infer that

$$\varphi_\varepsilon(x, y) < +\infty \quad \text{for every } x \in X \text{ and } y \in D_2(K). \tag{2.141}$$

Here, we have used in particular Corollary 2.6. On the other hand, $\varphi_\varepsilon(x, y)$ is lower-semicontinuous and convex as a function of y, because this is true for each of the functions $\operatorname{cl}_2 K(u, \cdot)$. Therefore, $\varphi_\varepsilon(x, y)$ is, for any $\varepsilon > 0$, a continuous function of $y \in \operatorname{int} D_2(K)$ (see Proposition 2.16). But this function majorizes the convex function $\operatorname{cl}_1 K(x, \cdot)$, and hence we may conclude that the latter is also continuous on $\operatorname{int} D_2(K)$. Of course, $\operatorname{cl}_1 K \ge K \ge \operatorname{cl}_2 K$, while the closedness of K implies that $\operatorname{cl}_2 K = \operatorname{cl}_2 \operatorname{cl}_1 K$. From the latter relation, we have

$$\operatorname{cl}_1 K(x, y) = \operatorname{cl}_2 K(x, y) \quad \text{for every } x \in X \text{ and } y \in \operatorname{int} D_2(K),$$

hence

$$K(x, y) = \operatorname{cl}_1 K(x, y) \quad \text{for every } x \in X \text{ and } y \in \operatorname{int} D_2(K).$$

Hence, $K(\cdot, y)$ is concave and upper-semicontinuous for every $y \in \operatorname{int} D_2(K)$. Obviously, the effective domain of this function includes $D_1(K)$. We shall prove that it is just $D_1(K)$. To this end, let $x_0 \in X$ be such that $K(x_0, y_0) > -\infty$, where y_0 is arbitrary but fixed in $\operatorname{int} D_2(K)$.

Therefore, the convex function $y \to \operatorname{cl}_2 K(x_0, y)$, $y \in Y$, is not identically $-\infty$ which shows that $\operatorname{cl}_2 K(x_0, y)$ is nowhere $-\infty$. This implies that $x_0 \in D_1(K)$, as claimed. The proof of part (ii) is entirely similar to that of part (i), so that it is omitted.

Given a saddle function $K : X \times Y \to \overline{\mathbb{R}}$, we denote by $\partial_y K(x, y)$ the set of all subgradients of $K(x, \cdot)$ at y and by $-\partial_x K(x, y)$ the set of all subgradients of $-K(\cdot, y)$ at x. In other words,

$$\partial_y K(x, y) = \left\{y^* \in Y^*;\ K(x, y) \leq K(x, y) + (y - v, y^*),\ \forall v \in Y\right\}, \quad (2.142)$$

$$\partial_x K(x, y) = \left\{x^* \in X^*;\ K(u, y) \leq K(x, y) + (u - x, x^*),\ \forall u \in X\right\}. \quad (2.143)$$

The multivalued operator $\partial K : X \times Y \to X^* \times Y^*$ defined by

$$\partial K(x, y) = \left\{-\partial_x K(x, y), \partial_y K(x, y)\right\}, \quad (x, y) \in X \times Y, \quad (2.144)$$

is called the *subdifferential* of the saddle function K.

It should be observed that the concave–convex function K has a saddle point (x_0, y_0) if and only if

$$(0, 0) \in \partial K(x_0, y_0). \quad (2.145)$$

$\square$

Proposition 2.109 *Let K be a proper saddle function on $X \times Y$. The multivalued mapping $\partial K : X \times Y \to X^* \times Y^*$ is a monotone operator with*

$$D(\partial K) \subset \operatorname{dom} K. \quad (2.146)$$

Proof Let $(x_1^*, y_1^*) \in \partial K(x_1, y_1)$ and $(x_2^*, y_2^*) \in \partial K(x_2, y_2)$. By definition,

$$-K(x, y_1) \geq -K(x_1, y_1) + (x - x_1, x_1^*), \quad \forall x \in X, \quad (2.147)$$

$$K(x_1, y) \geq K(x_1, y_1) + (y - y_1, y_1^*), \quad \forall y \in Y, \quad (2.148)$$

$$-K(x, y_2) \geq -K(x_2, y_2) + (x - x_2, x_2^*), \quad \forall x \in X, \quad (2.149)$$

$$K(x_2, y) \geq K(x_2, y_2) + (y - y_2, y_2^*), \quad \forall y \in Y. \quad (2.150)$$

Since (x, y) is arbitrary, we have $-K(x_1, y_1) < +\infty$ from relation (2.147) and $K(x_1, y_1) < +\infty$ from relation (2.148). Hence, $K(x_1, y_1)$ is finite, and from conditions (2.147) and (2.148), we have $(x_1, y_1) \in \operatorname{dom} K$, establishing relation (2.146). Taking $x = x_2$ in (2.147), $y = y_2$ in (2.148), $x = x_1$ in (2.149), and $y = y_1$ in (2.150), by adding the four inequalities we obtain

$$(x_1^* - x_2^*, x_1 - x_2) + (y_1^* - y_2^*, y_1 - y_2) \geq 0,$$

which means that ∂K is a monotone operator (see Sect. 1.4.1). $\square$

Corollary 2.110 *Let K be a proper closed saddle function on $X \times Y$. Then*

$$\operatorname{int} \operatorname{dom} K \subset D(\partial K) \subset \operatorname{dom} K. \quad (2.151)$$

Proof Let $(x, y) \in \operatorname{int}\operatorname{dom} K$. Thus, $x \in \operatorname{int} D_1(K)$ and $y \in \operatorname{int} D_2(K)$, so that Theorem 2.108 together with Corollary 2.38 imply that K is subdifferentiable at (x, y), establishing (2.151). □

Corollary 2.111 *Let K be a proper and closed saddle function on $X \times Y$. Then K is continuous on* $\operatorname{int}\operatorname{dom} K$.

Proof From Theorem 1.144, and Corollary 2.110, it follows that the monotone operator ∂K is locally bounded on $\operatorname{int}\operatorname{dom} K \subset \operatorname{int} D(\partial K)$. Let (x_0, y_0) be any element in $\operatorname{int}\operatorname{dom} K$. By definition, for all $(x \times Y$, one has

$$K(x_0, y_0) - K(x, y) \leq (y_0 - y, y_0^*) + (x - x_0, x^*) \tag{2.152}$$

and

$$K(x, y) - K(x_0, y_0) \leq (y - y_0, y^*) + (x_0 - x, x_0^*), \tag{2.153}$$

where $(x_0^*, y_0^*) \in \partial K(x_0, y_0)$ and $(x^*, y^*) \in \partial K(x, y)$. Since ∂K is locally bounded at (x_0, y_0), there exist $\rho > 0$ and $C > 0$ such that

$$\|x^*\| + \|y^*\| \leq C \quad \text{for } \|x - x_0\| < \rho \text{ and } \|y - y_0\| < \rho.$$

Inserting this in relations (2.152) and (2.153), it follows that

$$\big|K(x_0, y_0) - K(x, y)\big| \leq C_1\big(\|x - x_0\| + \|y - y_0\|\big),$$

for all $(x, y) \in X \times Y$ such that $\|x - x_0\| < \rho$ and $\|y - y_0\| < \rho$. Here, C_1 is a positive constant independent of x and y. Thus, we have shown that K is Lipschitzian in a neighborhood of (x_0, y_0). The proof of Corollary 2.111 is complete. □

The results presented above bring out many connections between closed saddle functions and lower-semicontinuous functions. The most important fact is stated in Theorem 2.112 below.

Theorem 2.112 *The formulas*

$$L(x, y^*) = \sup\big\{(y, y^*) - K(x, y);\ y \in Y\big\}, \tag{2.154}$$

$$K(x, y) = \sup\big\{(y, y^*) - L(x, y^*;\ y^* \in Y\big\} \tag{2.155}$$

define a one-to-one correspondence between the lower-semicontinuous proper convex functions L on the space $X \times Y^*$ and the closed saddle functions K on $X \times Y$ satisfying

$$\operatorname{cl}_2 \operatorname{cl}_1 K = K. \tag{2.156}$$

Moreover, under this correspondence, one has

$$(x^*, y^*) \in \partial K(x, y) \quad \Longleftrightarrow \quad (-x^*, y) \in \partial L(x, y^*). \tag{2.157}$$

Proof Let $L : X \times Y^* \to]-\infty, +\infty]$ be convex, lower-semicontinuous and non-identically $+\infty$ on $X \times Y^*$. Formula (2.155) says that K is he partial conjugate of L and this implies that the function $K(x, y)$ is convex and lower-semicontinuous in y on Y. Furthermore, it follows that $L(x, \cdot)$ is in turn the conjugate of $K(x, \cdot)$, establishing formula (2.154). Lastly, a simple calculation involving relation (2.155) and the convexity of L on $X \times Y^*$ implies that $K(x, y)$ is concave as a function of x on X. We leave the simple details to the reader. Now, we prove that K defined by formula (2.155) satisfies condition (2.156). To this end, we consider the conjugate $L^* : X^* \times Y \to]-\infty, +\infty]$ of L, that is,

$$L^*(x^*, y) = \sup\{(x, x^*) + (y, y^*) - L(x, y^*);\ x \in X,\ y^* \in Y^*\}.$$

According to relation (2.155), we get

$$L^*(x^*, y) = \sup\{(x, x^*) + K(x, y);\ x \in X\}. \tag{2.158}$$

Hence,

$$\mathrm{cl}_1 K(x, y) = -\sup\{(x, x^*) - L^*(x^*, y);\ x^* \in X^*\}. \tag{2.159}$$

But $L = L^{**}$, because L is lower-semicontinuous. In other words,

$$L(x, y^*) = \sup\{(x, x^*) + (y, y^*) - L^*(x^*, y);\ x^* \in X^*,\ y \in Y^*\}.$$

Hence, by equality (2.159), we must have

$$L(x, y^*) = \sup\{(y, y^*) - \mathrm{cl}_1 K(x, y);\ y \in Y\},$$

and therefore

$$\mathrm{cl}_2\, \mathrm{cl}_1 K(x, y) = \sup\{(y, y^*) - L(x, y^*);\ y^* \in Y^*\}.$$

Combining this with relation (2.155), we obtain

$$\mathrm{cl}_2\, \mathrm{cl}_1 K(x, y) = K(x, y) \quad \text{for every } (x, y) \in X \times Y,$$

as claimed.

Next, we assume that K is any closed proper saddle function on $X \times Y$ which satisfies condition (2.156). First, we note that the function L defined by formula (2.154) is convex on the product space $X \times Y^*$. Furthermore, since $\mathrm{dom}\, K \neq \emptyset$, we must have

$$L(x, y^*) > -\infty \quad \text{for every } (x, y^*) \in X \times Y^*$$

and $L \not\equiv +\infty$. It remains to be proved that L is lower-semicontinuous on $X \times Y^*$. Let L^* be the conjugate of L. One has

$$\mathrm{cl}\, L(x, y^*) = \sup\{(x, x^*) + (y, y^*) - L^*(x^*, y);\ x^* \in X^*,\ y \in Y\}.$$

Combining this with equality (2.159), we obtain

$$\begin{aligned}\operatorname{cl} L(x, y^*) &= \sup\{(y, y^*) - \operatorname{cl}_1 K(x, y);\ y \in Y\} \\ &= \sup\{(y, y^*) - \operatorname{cl}_2 \operatorname{cl}_1 K(x, y);\ y \in Y\},\end{aligned}$$

which is equivalent to

$$\operatorname{cl} L(x, y^*) = \sup\{(y, y^*) - K(x, y);\ y \in Y\} = L(x, y^*)$$

in view of relations (2.156) and (2.154). Thus, L is lower-semicontinuous on $X \times Y^*$.

In order to verify relation (2.157), we fix any (x^*, y^*) in $\partial K(x, y)$ and use the definition of $\partial_x K(x, y)$. Then

$$\begin{aligned}&-(x^*, x - x_1) + (y, y^* - y_1^*) \geq -K(x, y) + K(x_1, y) + (y, y^* - y_1^*) \\ &\quad \text{for all } x_1 \in X,\ y_1^* \in Y^*. \end{aligned} \tag{2.160}$$

From relation (2.154), we have

$$K(x_1, y) - (y, y_1^*) \geq -L(x_1, y_1^*) \tag{2.161}$$

while (2.142) implies that

$$K(x, y) + L(x, y^*) = (y, y^*) \tag{2.162}$$

because $y \to K(x, y)$ is the conjugate of the proper convex function $L(x, \cdot)$ (see Proposition 2.33). Adding relations (2.161) and (2.162) and substituting the result in (2.160), one obtains

$$-(x^*, x - x_1) + (y, y^* - y_1^*) \geq L(x, y^*) - L(x_1, y_1^*), \tag{2.163}$$

for all $x_1 \in X$ and $y_1^* \in Y^*$. In other words, we have proved that $(-x^*, y) \in \partial L(x, y^*)$. It remains to be proved that $(-x^*, y) \in \partial L(x, y^*)$ implies that $(x^*, y^*) \in \partial K(x, y)$. This follows by using a similar argument, but the details are omitted. □

Remark 2.113 The closed saddle function K associated with a convex and lower-semicontinuous function L are referred to in the following as the *Hamiltonian function corresponding to L*.

Given any closed and proper saddle function K on $X \times Y$, there always exists an equivalent closed saddle function K' which satisfies condition (2.156). An example of such a function could be $K' = \operatorname{cl}_2 K$. This fact shows that formulas (2.154) and (2.155) *define a one-to-one correspondence between the equivalence classes of closed proper saddle functions K on $X \times Y$ and lower-semicontinuous, proper convex functions L on $X \times Y^*$.*

Theorem 2.114 below may be compared most closely to Theorem 2.43.

Theorem 2.114 *Let Y be a reflexive Banach space and let $K : X \times Y \to \overline{\mathbb{R}}$ be a proper, closed saddle function on $X \times Y$. Then the operator $\partial K : X \times Y \to X^* \times Y^*$ is maximal monotone.*

Proof It should be observed that, if K' is a saddle function equivalent to K, then $\partial K' = \partial K$. Indeed, as observed earlier, $(x_0^*, y_0^*) \in \partial K(x_0, y_0)$ if and only if (x_0, y_0) is a saddle point of the function $(x, y) \to K(x, y) + (x, x_0^*) - (y, y_0^*)$ which is in turn equivalent to $(x, y) \to K'(x, y) + (x, x_0^*) - (y, y_0^*)$. Since two equivalent closed saddle functions have the same saddle points, we conclude that $(x_0^*, y_0^*) \in \partial K'(x_0, y_0)$, as claimed. Thus, replacing, if necessary, the function K by $\mathrm{cl}_2 K$, we may assume that the concave–convex function satisfies condition (2.156) in Theorem 2.112. If Y is reflexive, then $X \times Y^*$ is a Banach space, whose dual may be identified with $X^* \times Y$. Since the function L defined by formula (2.154) is convex and lower-semicontinuous on $X \times Y^*$, its subdifferential ∂L is maximal monotone (see Theorem 2.43) from $X \times Y^*$ into $X^* \times Y$. Hence, using relation (2.157), ∂K is also maximal monotone. □

Remark 2.115 Theorem 2.114 follows also in the case when X rather than Y is reflexive, by replacing K by $-K$.

Corollary 2.116 *Let X and Y be two reflexive Banach spaces, and let $K : X \times Y \to \overline{\mathbb{R}}$ be a proper, closed saddle function on $X \times Y$. Then, the domain $D(\partial K)$ of the operator ∂K is a dense subset of* $\operatorname{dom} K$.

Proof Let (x_0, y_0) be any element of $\operatorname{dom} K$, and let $(x_\lambda, y_\lambda) \in X \times Y$ be such that

$$F_1(x_\lambda - x_0) - \lambda \partial_x K(x_\lambda, y_\lambda) \ni 0, \quad \lambda > 0, \tag{2.164}$$

$$F_2(y_\lambda - y_0) - \lambda \partial_y K(x_\lambda, y_\lambda) \ni 0, \quad \lambda > 0, \tag{2.165}$$

where $F_1 : X \to X^*$ and $F_2 : Y \to Y^*$ are duality mappings of X and Y, respectively. Since ∂K is maximal monotone and the operator $(x, y) \to (F_1(x - x_0), F_2(y - y_0))$ is monotone, coercive and demicontinuous from $X \times Y$ to $X^* \times Y^*$ (without any loss of generality, we may assume that X and Y as well as their duals are strictly convex), the above equation has at least one solution $(x_\lambda, y_\lambda) \in D(\partial K)$ (see Corollary 1.140). We multiply the first equation by $x_\lambda - x_0$, the second by $y_\lambda - y_0$ and add the results; thus, we obtain

$$\begin{aligned}\big(F_1(x_\lambda - x_0), x_\lambda - x_0\big) + \big(F_2(y_\lambda - y_0), y_\lambda - y_0\big)\\ \le \lambda\big(K(x_\lambda, y_0) - K(x_0, y_\lambda)\big), \quad \text{for all } \lambda > 0.\end{aligned} \tag{2.166}$$

Inasmuch as $(x_0, y_0) \in \operatorname{dom} K$, the functions $x \to -K(x, y_0)$ and $y \to K(x_0, y)$ are convex and not identically $+\infty$ on X and Y, respectively. Thus, these functions are bounded from below by affine functions (see Proposition 2.20). This fact implies

$$\|x_\lambda - x_0\|^2 + \|y_\lambda - y_0\|^2 \le C\lambda\big(\|x_\lambda\| + \|y_\lambda\| + 1\big). \tag{2.167}$$

Therefore $x_\lambda \to x_0$ and $y_\lambda \to y_0$ as $\lambda \to 0$, thereby proving Corollary 2.116. □

Remark 2.117 It turns out that Corollary 2.116 remains true if one merely assumes that X or Y is reflexive (see Gossez [22]).

As a final (but, actually, immediate) application of Theorem 2.114, we cite a minimax result which plays a fundamental role in game theory (see, for instance, Aubin [1]).

Corollary 2.118 *Let X and Y be reflexive Banach spaces, and let A and B be two closed and convex subsets of X and Y, respectively. Let K_0 be a closed saddle function on $X \times Y$ satisfying the following condition*:

(a) *There exists some* $(x_0, y_0) \in A \times B$ *such that*

$$\lim_{\substack{\|x\|+\|y\|\to+\infty \\ x\in A,\ y\in B}} \big(K_0(x, y_0) - K_0(x_0, y)\big) = -\infty. \tag{2.168}$$

Then, the function K_0 has at least one saddle point on $A \times B$.

Proof Let $K : X \times Y \to [-\infty, +\infty]$ be the closed saddle function defined by (2.140). By Theorem 2.114, the operator $\partial K : X \times Y \to X^* \times Y^*$ is maximal monotone. Hence, for each $\lambda > 0$ $(x_\lambda, y_\lambda) \in D(\partial K) = A \times B$ such that

$$\lambda F_1(x_\lambda) - \partial_x K(x_\lambda, y_\lambda) \ni 0, \tag{2.169}$$

$$\lambda F_2(y_\lambda) + \partial_y K(x_\lambda, y_\lambda) \ni 0, \tag{2.170}$$

where $F_1 : X \to X^*$ and $F_2 : Y \to Y^*$ are dually mappings of X and Y, respectively.

Let $(x_0, y_0) \in A \times B$ be fixed as in condition (2.168). We multiply equation (2.169) by $x_\lambda - x_0$, equation (2.170) by $y_\lambda - y_0$, and use the definition of ∂K to obtain

$$\lambda\big(F_1(x_\lambda), x_\lambda - x_0\big) \le K(x_\lambda, y_\lambda) - K(x_0, y_\lambda),$$
$$\lambda\big(F_2(y_\lambda), y_\lambda - y_0\big) \le K(x_\lambda, y_\lambda) + K(x_\lambda, y_0).$$

Therefore,

$$\lambda\big(\|x_\lambda\|^2 + \|y_\lambda\|^2\big) \le \lambda\big(\|x_\lambda\|\,\|x_0\| + \|y_\lambda\|\,\|y_0\|\big) + K(x_\lambda, y_0) - K(x_0, y_\lambda).$$

According to condition (a), this inequality shows that (x_λ, y_λ) must be bounded in $X \times Y$ as λ tends to 0. Thus, without loss of generality, we may assume that

$$\begin{aligned} x_\lambda &\to \tilde{x} \quad \text{weakly in } X, \\ y_\lambda &\to \tilde{y} \quad \text{weakly in } Y, \end{aligned} \tag{2.171}$$

as $\lambda \to 0$. If we let $\lambda \to 0$ in equations (2.169) and (2.170), we may infer that

$$\lim_{\lambda\to 0} \partial K(x_\lambda, y_\lambda) = (0, 0) \quad \text{strongly in } X^* \times Y^*. \tag{2.172}$$

Since ∂K is maximal monotone, from assumptions (2.171) and (2.172) it is immediately clear that $(\tilde{x}, \tilde{y}) \in D(\partial K)$ and

$$(0, 0) \in \partial K(\tilde{x}, \tilde{y}). \tag{2.173}$$

Thus, we have shown that K has a saddle point $(\tilde{x}, \tilde{y})$ on $X \times Y$. But it is not difficult to see that $(\tilde{x}, \tilde{y})$ is a saddle point of K if and only if $(\tilde{x}, \tilde{y})$ is a saddle point of K_0 with respect to $A \times B$, that is,

$$K_0(x, \tilde{y}) \le K_0(\tilde{x}, \tilde{y}) \le K_0(\tilde{x}, y) \quad \text{for all } x \in A \text{ and } y \in B,$$

and this establishes Corollary 2.118. $\square$

Let $K^* : X^* \times Y^* \to \overline{\mathbb{R}}$ be the concave–convex conjugate of K. By analogy with the terminology used in the study of convex functions, K^* is called the *conjugate* of K. If K is closed, so is K^* and, according to Theorem 2.114, if X and Y are reflexive, then the subdifferential ∂K^* of K^* is a maximal monotone operator from $X^* \times Y^*$ into $X \times Y$. It is not difficult to see that ∂K^* is the inverse of ∂K, that is,

$$(x, y) \in \partial K^*(x^*, y^*) \quad \Longleftrightarrow \quad (x^*, y^*) \in \partial K(x, y). \tag{2.174}$$

In particular, this means that the saddle points of K are just the elements of $\partial K^*(0, 0)$. Thus, K has a saddle point, if and only if K^* has a subgradient at $(0, 0)$. In particular, this implies that the set of all saddle points of the proper closed saddle function K is a closed and convex subset of the product space $X \times Y$. Furthermore, if K^* happens to be continuous at $(0, 0)$, then this set is weakly compact in $X \times Y$. It follows that the conditions ensuring the subdifferentiability of K^* may be regarded as mini-max theorems. This subject is discussed in some detail in the sequel.

2.3.3 Mini-max Theorems

Let X, Y be two separated linear topological spaces and let $F : X \times Y \to \overline{\mathbb{R}}$. An important problem is to establish certain conditions on F, X and Y under which the mini-max equality

$$\max_{x \in X} \min_{y \in Y} F(x, y) = \min_{y \in Y} \max_{x \in X} F(x, y) \tag{2.175}$$

is true or at least a saddle value exists, that is,

$$\sup_{x \in X} \inf_{y \in Y} F(x, y) = \inf_{y \in Y} \sup_{x \in X} F(x, y). \tag{2.176}$$

All the results of this type are termed *mini-max theorems*. In view of Proposition 2.105, the mini-max equality is equivalent to the existence of a saddle point of F on $X \times Y$.

This section is concerned with the main mini-max theorems and some generalizations of the famous *mini-max theorem of von Neumann* [76].

First, we prove a general result established by Terkelsen [72].

Theorem 2.119 *Let A b e a compact set in a topological space, let B be an arbitrary set, and let F be a real-valued function defined on $A \times B$ such that $F(\cdot, y)$ is an upper-semicontinuous function on A for every $y \in B$. Then, the following statements are equivalent.*

(a) *For every $\alpha \in \mathbb{R}$ and $y_1, y_2, \ldots, y_n \in B$ such that $\alpha > \max_{x\in A} \min_{1\le i\le n} F(x, y_i)$, there is $y_0 \in B$ such that $\alpha > \max_{x\in A} F(x, y_0)$.*

(b) *F satisfies the equality*

$$\max_{x\in A} \inf_{y\in B} F(x, y) = \inf_{y\in B} \max_{x\in A} F(x, y). \tag{2.177}$$

Proof First, we notice that because A is a compact set according to the Weierstrass theorem for the upper-semicontinuous functions (see Theorem 2.8), we can take "max" instead of "sup".

We immediately obtain statement (a) from equality (2.177) by using the definition of a supremum. Let us prove that statement (a) implies (b). Let an arbitrary $\alpha \in \mathbb{R}$ be such that

$$\alpha > \max_{x\in A} \inf_{y\in B} F(x, y).$$

We write $A_y = \{x \in A;\ F(x, y) \ge \alpha\}$, for every $y \in B$, and hence $\bigcap_{y\in B} A_y = \emptyset$. By hypothesis, A_y is closed; therefore, A being a compact set, there are $y_1, \ldots, y_n \in B$ with $\bigcap_{i=1}^n A_{y_i} = \emptyset$, which implies $\min_{1\le i\le n} F(x, y_i) < \alpha$, for each $x \in X$. Thus, $\max_{x\in A} \min_{1\le i\le n} F(x, y_i) < \alpha$ and then, from statement (a) we obtain $y_0 \in B$ such that $\alpha > \max_{x\in A} F(x, y_0)$, from which it results that $\alpha > \inf_{y\in B} \max_{x\in A} F(x, y)$. Now, if α tends to $\max_{x\in A} \inf_{y\in B} F(x, y)$, we have

$$\max_{x\in A} \inf_{y\in B} F(x, y) \ge \inf_{y\in B} \max_{x\in A} F(x, y).$$

Moreover, it follows from (2.132) that equality (2.177) holds. □

Corollary 2.120 *Under the same assumptions as in the theorem, if for every $y_1, y_2 \in B$ there is $y_3 \in B$ such that $F(x, y_3) \le F(x, y_1)$ and $F(x, y_3) \le F(x, y_2)$ for every $x \in A$, then F satisfies equality (2.177).*

Corollary 2.121 *If (f_n) is a decreasing sequence of real-valued upper-semicontinuous functions on a compact set A, then*

$$\lim_{n\to\infty} \max_{x\in A} f_n(x) = \max_{x\in A} \lim_{n\to\infty} f_n(x). \tag{2.178}$$

Proof To prove this, take $B = \mathbb{N}$ and define $F(x, n) = f_n(x)$, $x \in A$, $n \in \mathbb{N}$. We have satisfied a directed condition which, obviously, implies statement (a), hence equality (2.178). □

Remark 2.122 The previous theorem is not really a mini-max theorem. If, moreover, B is a compact set and $y \to F(x, y)$ is a lower-semicontinuous function on B for every $x \in A$, then statement (a) is equivalent to the mini-max equality (2.175) because the infimum is also attained.

Property (a) is a rather natural one because, from equality (2.175), inequality (2.178) is equivalent to the following assertion:

for every $\alpha \in \mathbb{R}$ *such that* $\alpha > \max_{x\in A} \inf_{y\in B} F(x, y)$, *there is* $y_0 \in B$ *such that* $\alpha \geq \max_{x\in A} F(x, y_0)$.

Since the set A is compact and the function $F(\cdot, y)$ is upper-semicontinuous, it is "possible" to consider the infimum only on the finite subsets of B.

The natural framework for presenting mini-max theorems is that of concave–convex functions. Among the various methods used in the proof of mini-max theorems, we notice the following: the first relies on separation properties of convex sets and the second is based on the celebrated *Knaster–Kuratowski–Mazurkiewicz Theorem* [38] (Theorem 2.129 below). However, these methods can be extended to functions more general than concave–convex functions.

Definition 2.123 A function $F : X \times Y \to \overline{\mathbb{R}}$ is said to be *concave–convex-like* if the following conditions hold:

(i) For every $x_1, x_2 \in X$ and $t \in [0, 1]$ there is an $x_3 \in X$ such that

$$tF(x_1, y) + (1 - t)F(x_2, y) \leq F(x_3, y) \quad \text{for all } y \in Y, \tag{2.179}$$

whenever the left-hand side makes sense.

(ii) For every $y_1, y_2 \in Y$ and $t \in [0, 1]$, there is a $y_3 \in Y$ such that

$$F(x, y_3) \leq tF(x, y_1) + (1 - t)F(x, y_2) \quad \text{for all } x \in X, \tag{2.180}$$

whenever the right-hand side is well defined.

Definition 2.124 A function $F : X \times Y \to \overline{\mathbb{R}}$ is said to be *quasi-concave–convex* if the level sets $\{x \in X; F(x, \bar{y}) \geq \alpha\}$ and $\{y \in Y; F(\bar{x}, y) \leq \alpha\}$ are convex sets for every $\bar{y} \in Y$, $\bar{x} \in X$ and $\alpha \in \mathbb{R}$.

It is clear from condition (i) that the following property results.

(i)$'$ For every $x_1, x_2 \in X$ and $t_1, t_2, \ldots, t_n \geq 0$ with $\sum_{i=1}^{n} t_i = 1$, there is an $x_0 \in X$ such that

$$\sum_{i=1}^{n} t_i F(x_i, y) \leq F(x_0, y) \quad \text{for all } y \in Y, \tag{2.181}$$

whenever the left-hand side is well defined.

A similar statement for condition (ii) holds.

Remark 2.125 The concepts of concave–convex-like and quasi-concave–convex are independent of each other. However, a concave–convex function is at the same time concave–convex-like and quasi-concave–convex.

In the following, we assume that $A \subset X,\ B \subset Y$ are two nonempty convex sets and that F is real-valued on $A \times B$. Hence, for extended real-valued functions, the set $A \times B$ plays the role of effective domain.

Theorem 2.126 *Let X, Y be separated topological linear spaces, $A \subset X,\ B \subset Y$ compact convex sets and F a real-valued upper-semicontinuous concave–convex-like function on $A \times B$. Then F satisfies the mini-max equality on $A \times B$.*

Proof Let us prove that F has property (a) from Theorem 2.119.

Let $\alpha \in \mathbb{R}$ and $y_1, y_2, \dots, y_n \in B$ be such that

$$\alpha > \max_{x \in A} \min_{1 \le i \le n} F(x, y). \tag{2.182}$$

Now, we consider the following convex sets of $\mathbb{R}^n$:

$$C_1 = \operatorname{conv}\big\{\big(F(x, y_1), F(x, y_2), \dots, F(x, y_n)\big);\ x \in A\big\},$$
$$C_2 = \big\{(u_1, u_2, \dots, u_n);\ u_i \ge \alpha,\ i = 1, 2, \dots, n\big\}.$$

Obviously, C_2 is a cone with vertex $\bar{\alpha} = (\alpha, \alpha, \dots, \alpha) \in \mathbb{R}^n$ and $C_1 \cap C_2 = \emptyset$. Indeed, if $u = (u_1, u_2, \dots, u_n) \in C_1$, there are $x_j \in A$ and $\alpha_j \ge 0$, $j = 1, 2, \dots, m$, with $\sum_{j=1}^m a_j = 1$, such that $u_i = \sum_{j=1}^m a_j F(x_j, y_i)$ for every $i = 1, 2, \dots, n$. Now, from (i)$'$, there exists a point $x_0 \in A$ such that

$$F(x_0, y) \ge \sum_{j=1}^m a_j F(x_j, y). \tag{2.183}$$

Using (2.182), we find i_0 for which $\alpha > F(x_0, y_{i_0})$. Therefore, it follows from inequality (2.183) that $\alpha > u_{i_0}$, that is, $u = (u_1, u_2, \dots, u_n) \overline{\in} C_2$. According to Corollary 1.41, for the disjoint convex subsets C_1, C_2 we find a nonzero element $c = (c_1, c_2, \dots, c_n) \in \mathbb{R}^n$ such that

$$\sup_{u \in C_1} \sum_{i=1}^n c_i u_i \le \inf_{u \in C_2} \sum_{i=1}^n c_i u_i. \tag{2.184}$$

However, the cone C_2 contains all the points $(\alpha, \alpha, \dots, \alpha, \alpha + n, \alpha, \dots, \alpha)$, $n \in \mathbb{N}$, and therefore $c_i \ge 0$; hence, the infimum is attained at the vertex. Taking $c_i' = c_i(\sum_{j=1}^n c_j)^{-1}$ and $u_i = F(x, y_i)$, from inequality (2.184), we obtain $\sum_{i=1}^n c_i' F(x, y_i) \le \alpha$ for all $x \in A$. Combining this with property (ii) from Definition 2.123, there is a point $y_0 \in B$ such that $F(x, y_0) \le \alpha$ for every $x \in A$; hence, $\alpha \ge \max_{x \in A} F(x, y_0)$ and thus assertion (a) from Theorem 2.119 is really satisfied. Therefore relation (2.177) is true. Now, using (2.177) and the lower-semicontinuity of $F(x, \cdot)$ on the compact B for every $x \in A$, we obtain the mini-max equality (2.175). □

Corollary 2.127 *If X, Y are reflexive Banach spaces, $A \subset X$, $B \subset Y$ are bounded closed and convex sets, F is an upper-lower-semicontinuous concave–convex function on $A \times B$, then F has a saddle point on $A \times B$.*

Proof It is sufficient to recall that in a reflexive Banach space, every bounded closed convex set in weakly compact (Theorem 1.94) and the lower-(upper-)semicontinuity is equivalent to the weak lower-(upper-)semicontinuity for the class of convex (concave) functions, by virtue of Proposition 2.10. We can, therefore, apply the theorem where X, Y are endowed with their weak topologies. □

Remark 2.128 As is easily seen from the proof of Theorem 2.119, we omit the compactness condition of the set B and the lower-semicontinuity condition of the function $F(x, \cdot)$, we obtain equality (2.177).

Now, we prove similar results for quasi-concave–convex functions. As noted above, we use the following statement due to Knaster, Kuratowski and Mazurkiewicz [38].

Theorem 2.129 (Knaster–Kuratowski–Mazurkiewicz) *Let U be an arbitrary set in a finite-dimensional separated topological linear space E. To every $u \in U$, let $\mathscr{F}(u) \subset E$ be a compact set such that the convex hull of every finite subset $\{u_1, u_2, \ldots, u_n\} \subset U$ is contained in the corresponding union $\bigcup_{i=1} \mathscr{F}(u_i)$. Then, $\bigcap_{u \in U} \mathscr{F}(u) \neq \emptyset$.*

The first main result for the quasi-concave–convex functions is the following.

Theorem 2.130 *Let F be a real-valued upper-lower-semicontinuous quasi-concave–convex function on $A \times B$. If there are $y_0 \in B$ and $\alpha_0 < \inf_{y\in B} \sup_{x\in A} F(x, y)$ such that the level set $\{x \in A; F(x, y_0) \geq \alpha_0\}$ be compact, then*

$$\sup_{x\in A} \inf_{y\in B} F(x, y) = \inf_{y\in B} \sup_{x\in A} F(x, y). \tag{2.185}$$

Proof Suppose by contradiction that equality (2.185) is not true. From inequality (2.132), there is $\alpha > \alpha_0$, such that

$$\sup_{x\in A} \inf_{y\in B} F(x, y) < \alpha < \inf_{y\in B} \sup_{x\in A} F(x, y). \tag{2.186}$$

Write $A_y = \{x \in A; F(x, y) \geq \alpha\}$ and $B_x = \{y \in B; F(x, y) \leq \alpha\}$, which by hypothesis are nonempty convex and closed sets. Using (2.186), it follows that

$$\bigcap_{y\in B} A_y = \emptyset, \qquad \bigcap_{x\in A} B_x = \emptyset.$$

Since A_{y_0} is compact, there are $y_1, \ldots, y_n \in B$ such that $\bigcap_{i=1}^{n} A_{y_1} = \emptyset$. On the other hand, as the convex sets finitely generated are compact, there are $x_1, \ldots, x_m \in A$ such that

$$\bigcap_{i=1}^{m} B_{x_j} \cap \operatorname{conv}\{y_i;\ i = 1, 2, \ldots, n\} = \emptyset.$$

Let $A' = \operatorname{conv}\{x_1, x_2, \ldots, x_m\}$ and $B' = \operatorname{conv}\{y_1, y_2, \ldots, y_n\}$. Define the multi-valued mapping $\mathscr{F}$ on $A' \times B'$ by

$$\mathscr{F}(u, v) = \big\{(w, s) \in A' \times B';\ F(w, v) \geq \alpha \text{ or } F(u, s) \leq \alpha\big\}. \tag{2.187}$$

One may easily show that all the conditions of Theorem 2.129 are fulfilled. Indeed, $\mathscr{F}(u, v)$ is a compact set since F is upper-semicontinuous and $A' \times B'$, $\lambda_i \geq 0$, with $\sum_{i=1}^{p} \lambda_i = 1$ such that

$$\sum_{i=1}^{p} \lambda_i (u_i, v_i) \overline{\in} \mathscr{F}(u_j, v_j) \quad \text{for all } j = 1, 2, \ldots, p;$$

it follows that

$$F\left(\sum_{i=1}^{p} \lambda_i u_i, v_j\right) < \alpha \quad \text{and} \quad F\left(u_j, \sum_{i=1}^{p} \lambda_i v_i\right) > \alpha, \quad j = 1, 2, \ldots, p.$$

Since the sets

$$\left\{y \in B';\ F\left(\sum_{i=1}^{p} \lambda_i u_i, y\right) < \alpha\right\} \quad \text{and} \quad \left\{x \in A';\ F\left(x, \sum_{i=1}^{p} \lambda_i v_i\right) > \alpha\right\}$$

are convex, at the same time we obtain

$$F\left(\sum_{i=1}^{p} \lambda_i u_i, \sum_{i=1}^{p} \lambda_i v_i\right) < \alpha \quad \text{and} \quad F\left(\sum_{i=1}^{p} \lambda_i u_i, \sum_{i=1}^{p} \lambda_i v_i\right) > \alpha,$$

which is a contradiction. Hence,

$$\sum_{i=1}^{p} \lambda_i (u_i, v_i) \in \bigcup_{i=1}^{p} \mathscr{F}(u_i, v_i).$$

Thus, according to Theorem 2.129, there is $(x_0, y_0) \in A' \times B'$ such that $(x_0, y_0) \in \mathscr{F}(x, y)$ for all $(x, y) \in A' \times B'$, that is, $F(x_0, y_0) \geq \alpha$ or $F(x_0, y_0) \leq \alpha$ for all $x \in A'$ and $y \in B'$. On the other hand, it follows that there are i_0 and j_0 such that $x_0 \overline{\in} A_{y_{i_0}}$ and $y_0 \overline{\in} B_{x_{j_0}}$, which implies

$$\alpha < F(x_{j_0}, y_0) \leq \alpha \quad \text{or} \quad \alpha \leq F(x_0, y_{i_0}) < \alpha.$$

This is a contradiction. Therefore, equality (2.185) holds. $\square$

Remark 2.131 It is worth noting that it is sufficient to assume that $F(x, \cdot)$ is lower-semicontinuous only on the intersection of B with any finite-dimensional space. It should be emphasized that in equality (2.185) "sup" may be replaced by "max" because $F(\cdot, y)$ is upper-semicontinuous and A may be replaced by the compact set A_{y_0}.

According to Theorem 2.130, we obtain a result similar to Theorem 2.126, for the class of quasi-concave–convex functions.

Theorem 2.132 *Let A, B be two compact convex sets and let F be a real-valued upper-semicontinuous quasi-concave–convex function on $A \times B$. Then F satisfies the mini-max equality on $A \times B$.*

Remark 2.133 By Remark 2.125 and Theorem 2.126 or Theorem 2.132, we find the classical mini-max theorem for concave–convex functions. Likewise, we find again Corollary 2.118 for the semicontinuous saddle functions.

Corollary 2.134 *Let X, Y be reflexive Banach spaces, and let $A \subset X$, $B \subset X$ be closed convex sets. If F is a semicontinuous saddle function on $A \times B$ satisfying the conditions*:

(a) *A and B are bounded, or*
(b) *There is $(x_0, y_0) \in A \times B$ such that*

$$\lim_{\substack{\|x\|+\|y\|\to\infty \\ (x,y)\in A\times B}} \left\{F(x_0, y) - F(x, y_0)\right\} = \infty, \tag{2.188}$$

then F verifies the mini-max equality on $A \times B$.

Proof If F satisfies condition (a), Theorem 2.132 can be used for the work topologies on X and Y. Hence, it is sufficient to prove the corollary if F satisfies the coercivity condition (b). It is clear, from condition (b), that there exists $h > 0$ such that, for every $(x, y) \in A \times B$ with $\|x\| + \|y\| \geq h$, we have

$$F(x_0, y) - F(x, y_0) > 0. \tag{2.189}$$

We can assume that $h > \max\{\|x_0\|, \|y_0\|\}$. From the first part of the corollary applied to the function F with respect to nonempty bounded closed convex sets $A' = \{x \in A; \|x\| \leq h\}$ and $B' = \{y \in B; \|y\| \leq h\}$, it follows that there is a saddle point $(x', y') \in A' \times B'$, that is,

$$F(x, y') \leq F(x', y') \leq F(x', y), \tag{2.190}$$

for every $(x, y) \in A' \times B'$.

Particularly, since $(x_0, y_0) \in A' \times B'$, we obtain

$$F(x_0, y') \leq F(x', y') \leq F(x', y_0)$$

from which we see that (x', y') does not satisfy inequality (2.189); therefore, $\|x'\| < h$ and $\|y'\| < h$. Then, for every $y \in B$, we can choose a suitable $\lambda \in]0, 1[$ such that $\lambda y + (1 - \lambda)y' \in B'$. From the right-hand side of inequality (2.190), by virtue of the convexity of $F(x', \cdot)$, we obtain

$$F(x', y') \le F\bigl(x', \lambda y + (1 - \lambda)y'\bigr) \le \lambda F(x', y) + (1 - \lambda)F(x', y'),$$

which leads to

$$F(x', y') \le F(x', y),$$

for every $y \in B$. Similarly, from the left side of inequality (2.190) and, by virtue of the concavity of $F(\cdot, y')$, we have

$$F(x, y') \le F(x', y'),$$

for every $x \in A$. The last two inequalities imply that (x', y') is a saddle point of F on $A \times B$ and the proof is complete (Proposition 2.105). □

Remark 2.135 Condition (a) or (b) in the previous corollary may be replaced by the following conditions:

(a)′ B is bounded and there is $y_0 \in B$ such that

$$\lim_{\substack{\|x\| \to \infty \\ x \in A}} F(x, y_0) = -\infty, \tag{2.191}$$

or, by the symmetric condition

(b)′ A is bounded and there is $x_0 \in A$ such that

$$\lim_{\substack{\|y\| \to \infty \\ y \in B}} F(x_0, y) = +\infty. \tag{2.192}$$

Also, relations (2.191) and (2.192) together are sufficient.

All the results in this section can be applied to functions with values in $\overline{R}$, defined on a product of two separated topological linear spaces. It is known that, if F_0 is a real-valued function on $A \times B$, there is an extended real-valued function F defined on all space $X \times Y$ such that $F|_{\operatorname{dom} F} = F_0$ (see (2.140) from Sect. 2.3.2). Moreover, we have

$$\sup_{x \in X} \inf_{y \in Y} F(x, y) = \sup_{x \in A} \inf_{y \in B} F_0(x, y), \tag{2.193}$$

$$\inf_{y \in Y} \sup_{x \in X} F(x, y) = \inf_{y \in B} \sup_{x \in A} F_0(x, y). \tag{2.194}$$

Hence, if F_0 has a saddle value, then F has the same saddle value and reciprocally. Also, (x, y) is a saddle point of F on $X \times Y$ if and only if (x, y) is a saddle point of F_0 on $A \times B$ (provided F_0 is a proper function). On the other hand, giving an extended real-valued function $F : X \times Y \to \overline{\mathbb{R}}$, the role of A and B is played by $D_1(F)$ and $D_2(F)$. In general, relations (2.193) and (2.194) are not true. However, we can indicate a sufficiently large class of functions which satisfy these equalities.

Proposition 2.136 *If F is a proper closed saddle function on $X \times Y$, then relations* (2.193) *and* (2.194) *hold, where $A \times B = \operatorname{dom} F$.*

Proof By definition of $A = D_1(F)$, we have

$$\sup_{x\in X}\inf_{y\in Y} F(x,y) = \sup_{x\in X}\inf_{y\in Y} \operatorname{cl}_2 F(x,y) = \sup_{x\in A}\inf_{y\in Y} \operatorname{cl}_2 F(x,y).$$

On the other hand, since F is closed, by definition of $B = D_2(F)$ we have

$$\inf_{y\in Y} \operatorname{cl}_2 F(x,y) = \inf_{y\in Y} \operatorname{cl}_2 \operatorname{cl}_1 F(x,y) = \inf_{y\in Y} \operatorname{cl}_1 F(x,y) = \inf_{y\in B} \operatorname{cl}_1 F(x,y),$$

hence

$$\sup_{x\in X}\inf_{y\in Y} F(x,y) = \sup_{x\in A}\inf_{y\in B} \operatorname{cl}_1 F(x,y) \geq \sup_{x\in A}\inf_{y\in B} F(x,y).$$

Also, the converse inequality holds

$$\sup_{x\in X}\inf_{y\in Y} F(x,y) = \sup_{x\in A}\inf_{y\in Y} \operatorname{cl}_2 F(x,y) = \sup_{x\in A}\inf_{y\in Y} F(x,y) \leq \sup_{x\in A}\inf_{y\in B} F(x,y).$$

Similarly an obtains (2.194). □

2.4 Problems

2.1 Let $f : I \to \overline{\mathbb{R}}$ be a function on the real interval $I \subset \mathbb{R}$. Prove that f is quasi-convex if and only if it is either monotone or there exists $x_0 \in I$ such that f is decreasing on $(-\infty, x_0] \cap I$ and increasing on $[x_0, \infty) \cap I$.

Hint. We denote $\alpha = \inf\{f(x); x \in I\}$. Let us consider a sequence $(x_n)_{n\in\mathbb{N}^*} \subset I$ such that $f(x_n) \to \alpha$. Let $\bar{x}$ be a cluster element in $\overline{\mathbb{R}}$ of the sequence $(x_n)_{n\in\mathbb{N}^*}$ and denote by $a, b \in \overline{\mathbb{R}}$ the extremities of the interval I. The following three cases are possible: (1) $\bar{x} = a$; (2) $\bar{x} = b$; (3) $a < \bar{x} < b$. In the first case, the function f is increasing on I. Indeed, if $u, v \in I$, $u < v$ and $f(u) > f(v)$, taking $f(v) < \beta < f(u)$, we find $x_{\bar{n}}$ such that $f(x_{\bar{n}}) < \beta$, where $x_{\bar{n}} < u$, since $\alpha < \beta$. Therefore, the interval $\{x \in I; f(x) \leq \beta\}$ (see Sect. 2.1.1) contains the points $x_{\bar{n}}$ and v. Hence, it also contains the element u, that is, $f(u) \leq \beta$, which is a contradiction. Similarly, we prove that f is decreasing if $\bar{x} = b$. Now, if $a < \bar{x} < b$, then f is decreasing on $[a, \bar{x}] \cap I$ and increasing on $[\bar{x}, b] \cap I$.

2.2 Let φ be a lower-semicontinuous convex function on the Hilbert space H and let $\{x_n\}$ be defined by the following algorithm:

$$x_{n+1} + \partial\varphi(x_{n+1}) \ni x_n, \quad n \in \mathbb{N}.$$

Prove that the sequence $\{x_n\}$ is weakly convergent to a minimum point $x_e \in (\partial\varphi)^{-1}(0)$ of φ.

Hint. This is the descent step algorithm. If we set

$$K = \left\{ w - \lim_{n_k \to \infty} x_{n_k} \right\},$$

we show first that $K \subset (\partial\varphi)^{-1}(0)$ and then prove that the sequence $\{|x_n - y|^2\}_n$ is decreasing for each $y \in (\partial\varphi)^{-1}(0)$. If

$$\xi_1 = w - \lim_{n_k \to \infty} x_{n_k} \quad \text{and} \quad \xi_2 = w - \lim_{n'_k \to \infty} x_{n'_k},$$

this implies that

$$\lim_{n'_k \to \infty} |x_{n'_k} - \xi_1|^2 = \lim_{n''_k \to \infty} |x_{n''_k} - \xi_1|^2,$$
$$\lim_{n''_k \to \infty} |x_{n''_k} - \xi_2|^2 = \lim_{n'_k \to \infty} |x_{n'_k} - \xi_2|^2$$

and therefore $\xi_1 = \xi_2$, as claimed.

2.3 Let K be a closed convex subsets of $\mathbb{R}^m$ and let

$$\mathscr{K} = \left\{ y \in \left(L^p(\Omega)\right)^m;\ y(x) \in K,\ \text{a.e. } x \in \Omega \right\},$$

where $1 \le p < \infty$ and Ω is a measurable sub set of $\mathbb{R}^n$. Find the normal cone $N_{\mathscr{K}}(y) \subset (L^q(\Omega))^m$ to $\mathscr{K}$ at y, $\frac{1}{p} + \frac{1}{q} = 1$.

Hint. Apply Proposition 2.53, where $g(x, y) = 0$ if $y \in K$, $g(x, y) = +\infty$ if $y \overline{\in} K$.

2.4 Find the normal cone $N_{\mathscr{K}}$ for

$$\mathscr{K} = \left\{ y \in L^p(\Omega);\ a \le y(x) \le b,\ \text{a.e. } x \in \Omega \right\},$$
$$\mathscr{K} = \left\{ y \in \left(L^p(\Omega)\right)^m;\ \left\| y(x) \right\|_m \le \rho,\ \text{a.e. } x \in \Omega \right\},$$

where $\|\cdot\|_m$ is the Euclidean norm in $\mathbb{R}^m$.

2.5 Find the normal cone $N_{\mathscr{K}}$ to the set $\mathscr{K} = \{y \in L^2(\Omega);\ a \le y(x) \le b,$ a.e. $x \in \Omega, \int_\Omega y(x)\,dx = \ell\}$, where $am(\Omega) \le \ell \le bm(\Omega)$ (m is the Lebesgue measure).

Hint. We represent $\mathscr{K} = \mathscr{K}_1 \cap \mathscr{K}_2$ where $\mathscr{K}_1 = \{y \in L^2(\Omega);\ a \le y(x) \le b,$ a.e. $x \in \Omega\},\ \mathscr{K}_2 = \{y \in L^2(\Omega);\ \int_\Omega y(x)\,\mathrm{d}x = \ell\}$ and show that

$$N_{\mathscr{K}}(y) = N_{\mathscr{K}_1}(y) + N_{\mathscr{K}_2}(y), \quad \forall y \in \mathscr{K}.$$

Since $N_{\mathscr{K}_1}(y) + N_{\mathscr{K}_2} \subset N_{\mathscr{K}}(y)$, it suffices to show that, for every $f \in L^2(\Omega)$, the equation $y + N_{\mathscr{K}_1}(y) + N_{\mathscr{K}_2}(y) \ni f$ has a solution $y \in \mathscr{K}$. Since $N_{\mathscr{K}_2}(y) = \mathbb{R}$, the above equation reduces to $y = P_{\mathscr{K}_1}(f - \lambda)$, $\lambda \in \mathbb{R}$, where $P_{\mathscr{K}_1}$ is projection on $\mathscr{K}_1$.

2.6 Let $g : \mathbb{R} \to \mathbb{R}$ be a lower-continuous convex function such that $\lim_{|r|\to\infty} \frac{g(r)}{|r|} = +\infty$ and let $\varphi : H^{-1}(\Omega) \to \overline{\mathbb{R}}^*$ be defined by

$$\varphi(y) = \begin{cases} \int_\Omega g(y(x))\,\mathrm{d}x, & \text{if } g(y) \in L^1(\Omega), \\ +\infty, & \text{otherwise.} \end{cases}$$

Show that φ is lower-semicontinuous and that

$$\begin{aligned} \partial\varphi(y) = \big\{ & -\Delta w;\ w \in H_0^1(\Omega),\ y \in H^{-1}(\Omega) \cap L^1(\Omega), \\ & w(x) \in \partial g\big(y(x)\big) \text{ a.e. } x \in \Omega\big\}. \end{aligned} \tag{2.195}$$

Hint. Let $F(y) = \{w \in H_0^1(\Omega);\ w(x) \in \partial g(y(x)) \text{ a.e. } x \in \Omega\}$. Clearly, $F(y) \subset \partial\varphi(y)$ for each $y \in D(F)$. It suffices to show that F is maximal monotone from $(H_0^1(\Omega))' = H^{-1}(\Omega)$ to itself. Equivalently, for each $f \in H^{-1}(\Omega)$, the equation $-\Delta w + (\partial g)^{-1}(w) \ni f$ has a solution $w \in H_0^1(\Omega)$. One takes an approximating sequence $\{f_n\} \subset L^2(\Omega)$, $f_n \to f$ in $L^2(\Omega)$, and consider the corresponding solutions w_n to the equation $-\Delta w_n + (\partial g)^{-1}(w_n) \ni f_n$ in Ω, $w_n \in H_0^1(\Omega) \cap H^2(\Omega)$. Taking into account that $g^*(w_n) + g((\partial g)^{-1} w_n) = w_n (\partial g)^{-1}(w_n)$, we infer by the Dunford–Pettis theorem that $\{y_n \in (\partial g)^{-1}(w_n)\}$ is weakly compact in $L^1(\Omega)$ and therefore we may pass to the limit with w_n to prove the existence of $w \in H_0^1(\Omega)$ with $y \in (\partial g)^{-1} w \in L^1(\Omega)$.

2.7 Let $j : R \to R$ be a lower-semicontinuous convex function such that

$$\omega_2 |r|^p + c_z \le j(r) \le \omega_1 |r|^p + c_1, \quad \forall r \in \mathbb{R},$$

where $\omega_1, \omega_2 > 0$ and $p > 1$. We set $\beta = \partial j$. Consider the function $\varphi : W_0^{1,p}(\Omega) \to \overline{\mathbb{R}}^*$ defined by

$$\varphi(y) = \int_\Omega j(\nabla y)\,\mathrm{d}x.$$

Show that φ is convex, lower-semicontinuous and its subdifferential $\partial\varphi : W_0^{1,p}(\Omega) \to W^{-1,p'}(\Omega)$ is given by

$$\partial\varphi(y) = \big\{ w \in W^{-1,p'}(\Omega);\ w = -\operatorname{div} \eta,\ \eta(x) \in \partial j\big(\nabla y(x)\big),\ \text{a.e. } x \in \Omega \big\}. \tag{2.196}$$

Show that φ is lower-semicontinuous on $L^2(\Omega)$, too. Does this result remain true if $p=1$?

Hint. It suffices to show that the map defined by the second right-hand side of equation (2.196) is maximal monotone from $W_0^{1,p}(\Omega)$ to $(W_0^{1,p}(\Omega))' = W^{-1,p'}(\Omega)$, $\frac{1}{p}+\frac{1}{p'}=1$. If β is single valued, this reduces to the existence of a solution y for the nonlinear elliptic boundary-value problem $\lambda y - \operatorname{div} \partial j(\nabla y) = f$ in Ω; $y=0$ on $\partial\Omega$, where $\lambda>0$ and $f\in L^{p'}(\Omega)$. (See [4], p. 81.)

If $p=1$, then φ is no longer lower-semicontinuous on $L^2(\Omega)$ if takes $D(\varphi)=W_0^{1,1}(\Omega)$, but remains so if $D(\varphi)$ is taken to be the space of functions with bounded variation which are zero on $\partial\Omega$.

2.8 Let φ be a continuous and convex function on Hilbert space H with the norm $|\cdot|$, $\varphi(0)=0$ and let φ_t be its regularization (see (2.58)), that is,

$$\varphi_t(x) = \inf\left\{\frac{|x-y|^2}{2t} + \varphi(y);\ y\in H\right\} = S(t)\varphi, \quad t\ge 0.$$

Show that $S(t+s) = S(t)S(s)\varphi$, $\forall t,s>0$, and

$$\frac{d^+}{dt}\varphi(t,x) + \frac{1}{2}\left|\nabla_x\varphi(t,x)\right|^2 = 0, \quad \forall t>0,\ x\in H.$$

Remark 2.137 This means that $t\to S(t)\varphi$ is a continuous semigroup on the space of all continuous convex functions on H with infinitesimal generator $\varphi\to -\frac{1}{2}\,|\nabla_x\varphi(x)|^2$.

2.9 Let H be a Hilbert space and let F be a convex and continuously differentiable function on H such that

$$\lim_{|x|\to\infty}\frac{F(x)}{|x|} = +\infty, \quad \nabla F \text{ is locally Lipschitz},$$

$$\big(F'(x)-F'(y),x-y\big) \ge \omega_r|x-y|^2, \quad \forall x,y,|x|,|y|\le r.$$

We set

$$\big(S(t)\varphi\big)(x) = (\varphi^* + tF)^*(t), \quad t\ge 0,\ x\in H.$$

Show that:

(1) $\lim_{t\to 0} S(t)\varphi(x) = \varphi(x)$.
(2) $S(t+s)\varphi = S(t)S(s)\varphi$, $\forall s,t>0$.
(3) $\frac{d^+}{dt} S(t)\varphi + F(\nabla_x(S(t)\varphi)) = 0$, $\forall t>0$, $x\in H$.

Hint. Show first that $(S(t)\varphi)(x) = \varphi(y_t(x))_t + F^*(\nabla F(\partial\varphi(y_t(x)))$, where $y_t(x) = (I + t\nabla F(\partial\varphi))^{-1}(x)$ and $\nabla_x(S(t)\varphi)(x) = (\nabla F)^{-1}(t^{-1}(x-y_t(x)))$. (For details, see Barbu and Da Prato [5], p. 25.)

2.10 The unilateral (free boundary problem)

$$
\begin{aligned}
&-y''(x)+y(x)=f(x) \quad \text{in } \big[x\in[0,T];\, y(x)>\rho\big],\\
&-y''(x)+y(x)\le f(x) \quad \text{in } \big[x\in[0,1];\, y(x)=\rho\big],\\
&y(x)\ge\varphi, \quad \forall x\in[0,1], \qquad y(0)=y(1)=0,
\end{aligned}
$$

describes the equilibrium state of an elastic string fixed at $x=0,1$ and pushed against an obstacle $y=\rho<0$ by a distributed force $f(x)$. Represent it as a variational inequality and solve it for $f(x)\equiv-1$.

Hint. This is a problem of the form (2.95).

2.5 Bibliographical Notes

2.1. Most of the material on the general theory of convex functions presented in this subsection can be found in the mimeographed lecture notes of Moreau [46], the survey of Rockafellar [57] and the book [21] of Ekeland and Temam. In finite-dimensional spaces, excellent surveys on the subject are available in the Rockafellar book [56], the work of Ioffe and Tihomirov [33] and the books of Stoer and Witzgall [71] and Vainberg [74]. In infinite-dimensional spaces, the theory of conjugate functions has originally been developed by Bronsted [15] and, subsequently, studied by Bronsted and Rockafellar [16], Moreau [45, 46]. Some special types of convex function are studied by Ponstein [50] (see also the monograph of Avriel, Diewert, Schaible and Zang [2]). The first study on convex functions was published in 1945 by Popoviciu [51].

2.2. Subdifferential mappings were originally studied in Hilbert spaces by Moreau [45]. Theorem 2.43 was first proved by Moreau and later extended to a general Banach space by Rockafellar [55, 59]. Theorem 2.46 is also due to Rockafellar [55] and Theorem 2.58 is a slight extension of some results of Moreau [45] and Brezis [12]. As already noticed, Theorem 2.62 is a special case of a general perturbation theorem due to Rockafellar [60]. The idea of the proof given here comes from the work [14] by Brezis, Crandall and Pazy. Theorem 2.65 is due to Brezis [12, 13]. The theory of variational inequalities has been the subject of much development in the last fifteen years. For detailed treatments and applications, we refer the reader to the surveys of Stampacchia [70], Mosco [47], and to the books of Duvaut and Lions [19]. The nonlinear complementary problem in infinite dimension has been investigated by Karamardian [35], Habelter and Price [24], Eaves [20], Saigal [67], among others. Theorem 2.76 may be compared most closely with some results given by Karamardian [36], and Bazaraa et al. [6–9].

The concept of ε-subdifferential of convex function was introduced by Brønsted and Rockafellar [16]. The properties concerning the maximality with respect to the ε-monotonicity (Definition 2.86) considered for the first time by Vesely [75] (see also Jofré, Luc and Théra [34]) are established by Precupanu and Apetrii in [52], where some connections with the ε-enlargement of an operator defined by Revalski and Théra [54] and the special case of ε-subdifferential are investigated. A detailed treatment of calculus rules of the ε-subdifferential of a convex function is presented by Hirriart-Urruty and Phelps in [28].

The first notion of quasi-subdifferential for a quasi-convex function has been defined independently by Greenberg and Pierskalla in [23] and Zabotin, Koblev and Khabibulin in [77]. Different types of ε-quasi-subdifferential may be found in the monographs of Singer [68], Hirriart-Urruty and Lemarechal [27] and the papers of Ioffe [31], Martinez Legaz and Sach [43], Penot [49]. The concept of ε-quasi-subdifferential given by Definition 2.124 was introduced by Precupanu and Stamate in [53], where the relationship existing between this new type of quasi-subdifferential and other quasi-subdifferentials known in the literature is presented.

2.3. The results presented in Sect. 2.3.2 are essentially due to Rockafellar [58, 62] (see also [56]). The first mini-max theorem was formulated for bilinear functionals on finite-dimensional spaces by von Neumann [76]. Theorems 2.119 and 2.126 are essentially due to Terkelsen [72]. Mini-max Theorems 2.130 and 2.132 extend some classical results due to Ky Fan [40, 41], Sion [69], Kneser [39], Nikaido [48].

References

1. Aubin JP (1982) Mathematical methods of game and economic theory. North Holland, Amsterdam
2. Avriel M, Diewert W, Schaible S, Zang I (1988) Generalized concavity. Kluwer Academic, New York
3. Baiocchi C (1974) Problèmes à frontière libre en hydraulique. C R Acad Sci Paris 278:1201–1204
4. Barbu V (1996) Abstract periodic Hamiltonian systems. Adv Differ Equ 1:675–688
5. Barbu V, Da Prato G (1984) Hamilton–Jacobi equations in Hilbert spaces. Research notes in mathematics, vol 93. Pitman, Boston
6. Bazaraa MS, Goode J (1972) Necessary optimality criteria in mathematical programming in the presence of differentiability. J Math Anal Appl 40:609–621
7. Bazaraa MS, Shetty CM (1976) Foundations of optimization. Lecture notes in economics and mathematical systems, vol 122. Springer, Berlin
8. Bazaraa MS, Goode J, Nashed MZ (1972) A nonlinear complementary problem in mathematical programming in Banach spaces. Proc Am Math Soc 35:165–170
9. Bazaraa MS, Goode J, Nashed MZ (1974) On the cones of tangents with applications to mathematical programming. J Optim Theory Appl 13:389–426
10. Brezis H (1968) Equations et inéquations non linéaires dans les espaces vectoriels en dualité. Ann Inst Fourier 18:115–175
11. Brezis H (1971) Monotonicity methods in Hilbert spaces and some applications to nonlinear partial differential equations. In: Zarantonello E (ed) Contributions to nonlinear functional analysis. Academic Press, San Diego, pp 101–156
12. Brezis H (1972) Problèmes unilatéraux. J Math Pures Appl 51:1–64
13. Brezis H (1973) Opérateurs maximaux monotones et semigroupes de contractions dans les espaces de Hilbert. Math studies, vol 5. North Holland, Amsterdam

14. Brezis H, Crandall M, Pazy A (1970) Perturbations of nonlinear maximal monotone sets. Commun Pure Appl Math 23:123–144
15. Brønsted A (1964) Conjugate convex functions in topological vector spaces. Matfys Madd Dansk Vid Selsk 2 34:2–27
16. Brønsted A, Rockafellar RT (1965) On the subdifferentiability of convex functions. Proc Am Math Soc 16:605–611
17. Clarke FH (1975) Generalized gradients and applications. Trans Am Math Soc 205:247–262
18. Clarke FH (1981) Generalized gradients of Lipschitz functionals. Adv Math 40:52–67
19. Duvaut G, Lions JL (1972) Sur les inéqualitions en mécanique et en physique. Dunod, Paris
20. Eaves BC (1971) On the basic theorem of complementarity. Math Program 1:68–75
21. Ekeland I, Temam R (1974) Analyse convexe et problèmes variationnels. Dunod, Gauthier–Villars, Paris
22. Gossez JP (1972) On the subdifferential of a saddle function. J Funct Anal 11:220–230
23. Greenberg HJ, Pierskalla WP (1971) A review of quasi-convex functions. Oper Res 19:1553–1570
24. Habelter GJ, Price AL (1971) Existence theory for generalized nonlinear complementarity problem. J Optim Theory Appl 7:223–239
25. Hiriart-Urruty JB (1977) Contributions à la programmation mathématique. Thèse, Université de Clermont–Ferrand
26. Hiriart-Urruty JB (1979) Tangent cones, generalized gradients and mathematical programming in Banach spaces. Math Oper Res 4:79–97
27. Hiriart-Urruty JB, Lemarechal C (1993) Convex analysis and minimization algorithms. Springer, Berlin
28. Hiriart-Urruty JB, Phelps RR (1993) Subdifferential calculus using ε-subdifferentials. J Funct Anal 118:154–166
29. Ioffe AD (1976) An existence theorem for a general Bolza problem. SIAM J Control Optim 14:458–466
30. Ioffe AD (1977) On lower semicontinuity of integral functionals. SIAM J Control 15:521–538; 458–466
31. Ioffe AD (1990) Proximal analysis and approximate subdifferentials. J Lond Math Soc 41:1–38
32. Ioffe AD, Levin VL (1972) Subdifferential of convex functions. Trudi Mosc Mat Obsc 26:3–73 (Russian)
33. Ioffe AD, Tihomirov WM (1968) Duality of convex functions and extremal problems. Usp Mat Nauk 23:51–116 (Russian)
34. Jofré A, Luc DT, Théra M (1996) ε-subdifferential calculus for nonconvex functions and ε-monotonicity. C R Acad Sci Paris 323(I):735–740
35. Karamardian S (1971) Generalized complementarity problem. J Optim Theory Appl 8:161–168
36. Karamardian S (1972) The complementarity problem. Math Program 2:107–129
37. Kinderlehrer D, Stampacchia G (1980) An introduction to variational inequalities and their applications. Academic Press, New York
38. Knaster B, Kuratowski C, Mazurkiewicz S (1929) Eine Beweis des Fixpunktsatzes für n-dimensionale Simplexe. Fundam Math 14:132–137
39. Kneser H (1952) Sur un théorème fondamental de la théorie des jeux. C R Acad Sci Paris 234:2418–2420
40. Ky F (1953) Minimax theorems. Proc Natl Acad Sci USA 39:42–47
41. Ky F (1963) A generalization of the Alaoglu theorem. Math Z 88:48–66
42. Lions JL, Magenes E (1970) Problèmes aux limites non homogènes et applictions. Dunod, Gauthier–Villars, Paris
43. Martinez-Legaz JE, Sach PH (1999) A new subdifferential in quasi-convex analysis. J Convex Anal 6:1–11
44. Martinez-Legaz JE, Théra M (1996) ε-Subdifferentials in terms of subdifferentials. Set-Valued Anal 4:327–332

45. Moreau JJ (1965) Proximité et dualité dans un espace de Hilbert. Bull Soc Math Fr 93:273–299
46. Moreau JJ (1966–1967) Fonctionelles convexes. Séminaire sur les équations aux dérivées partielles, College de France
47. Mosco U (1970) Perturbations of variational inequality. Proc Symp Pure Math 28:182–194
48. Nikaido H (1954) On von Neumann's minimax theorem. Pac J Math 4:65–72
49. Penot JP (2000) What is quasiconvex analysis? Optimization 47:35–110
50. Ponstein J (1976) Seven kinds of convexity. SIAM Rev 9:115–119
51. Popoviciu T (1945) Les Fonctions Convexes. Hermann, Paris
52. Precupanu T, Apetrii M (2006) About ε-monotonicity of an operator. An Şt Univ AlI Cuza Iaşi, Ser I, Mat 81–94
53. Precupanu T, Stamate C (2007) Approximative quasi-subdifferentials. Optimization 56:339–354
54. Revalski JP, Théra M (2002) Enlargements of sums of monotone operators. Nonlinear Anal 48:505–519
55. Rockafellar RT (1966) Characterization of the subdifferentials of convex functions. Pac J Math 17:497–510
56. Rockafellar RT (1969) Convex analysis. Princeton Univ Press, Princeton
57. Rockafellar RT (1970) Convex functions, monotone operators and variational inequalities. In: Proc NATO Institute, Venice, Oderisi, Gubio
58. Rockafellar RT (1970) Monotone operators associated with saddle functions and minimax problems. In: Browder F (ed) Nonlinear functional analysis. Proc symp pure math, vol 18
59. Rockafellar RT (1970) On the maximal monotonicity of subdifferentials mappings. Pac J Math 33:209–216
60. Rockafellar RT (1970) On the maximality of sums of nonlinear operators. Trans Am Math Soc 149:75–88
61. Rockafellar RT (1971) Integrals which are convex functionals, II. Pac J Math 39:439–469
62. Rockafellar RT (1971) Saddle-points and convex analysis. In: Kuhn HW, Szegö GP (eds) Differential games and related topics. North-Holland, Amsterdam, pp 109–128
63. Rockafellar RT (1976) Integral functionals, normal integrands and measurable selections. In: Gossez JP et al (eds) Nonlinear operators and the calculus of variations. Lecture notes in math. Springer, Berlin
64. Rockafellar RT (1978) The theory of subgradients and its applications to problems of optimization. Lecture notes Univ Montreal
65. Rockafellar RT (1979) Directional Lipschitzian functions and subdifferential calculus. Proc Lond Math Soc 39:331–355
66. Rockafellar RT (1980) Generalized directional derivatives and subgradients of nonconvex functions. Can J Math 32:257–280
67. Saigal R (1976) Extensions of the generalized complementarity problem. CORE discussion papers 7323, Université Catholique de Louvain
68. Singer I (1997) Abstract convex analysis. Wiley, New York
69. Sion M (1958) On general minimax theorems. Pac J Math 8:171–176
70. Stampacchia G (1969) Variational inequalities. In: Ghizzetti A (ed) Theory and applications of monotone operators, Oderisi, Gubio, pp 35–65
71. Stoer J, Witzgall C (1970) Convexity and optimization in finite dimension. Springer, Berlin
72. Terkelsen F (1973) Some minimax theorems. Math Scand 31:405–413
73. Thibault L (1980) Sur les fonctions compactement Lipschitziennes et leur applications. Thèse, Université de Sciences et Techniques du Languedoc, Montpellier
74. Vainberg MM (1968) Le problème de la minimization des fonctionnelles non linéaires. Université de Moscou
75. Vesely L (1993) Local uniform boundedness principle for families of ε-monotone operators. Nonlinear Anal 24:1299–1304
76. von Neumann J (1928) Zur Theorie der Gesellschaftsspiele. Math Ann 100:295–320
77. Zabotin YaI, Korblev AI, Khabibulin RF (1973) Conditions for an extremum of a functional in the presence of constraints. Kibernetica 6:65–70 (in Russian)

Chapter 3
Convex Programming

This chapter is concerned with basic principles of convex programming in Banach spaces, that is, with the minimization of lower-semicontinuous convex functions on closed convex sets.

3.1 Optimality Conditions

As seen earlier in Sect. 2.2.1 for a proper convex function f on a Banach space X, the minimum points of f are just the solutions to the equation $0 \in \partial f(x)$. This elementary result has some specific features in the case of convex constraint minimization.

3.1.1 The Case of a Finite Number of Constraints

Let X be a real linear space and let $f : X \to \overline{\mathbb{R}}$ be a given function. Consider the minimizing problem for the function f on a subset $A_X \subset X$, that is, the problem

$$(\mathscr{P}_1) \quad \min\{f(x);\ x \in A_X\}.$$

The set A_X constitutes the constraints of Problem $\mathscr{P}_1$. We say that an element $\bar{x} \in X$ is *feasible* if $\bar{x} \in A_X \cap \mathrm{Dom}(f)$. The mathematical programming problem $\mathscr{P}_1$ is said to be *consistent* if $A_X \cap \mathrm{Dom}(f) \neq \emptyset$, that is, it has feasible elements. A feasible element x_0 is called an *optimal solution* of $\mathscr{P}_1$ if

$$f(x_0) = \inf\{f(x);\ x \in A_X\}.$$

In the theory of mathematical programming, a high degree of variability arises for the set A_X of constraints and the cost function f. Thus, if A_X and f are convex, then Problem $\mathscr{P}_1$ is a *convex programming problem*.

V. Barbu, T. Precupanu, *Convexity and Optimization in Banach Spaces*, Springer Monographs in Mathematics,
DOI 10.1007/978-94-007-2247-7_3,

The subset A_X is often defined by the solutions of a finite number of equations and inequalities as in

$$A_X = \{x \in X;\ g_i(x) \le 0,\ \forall i = 1, \dots, n;\ r_j(x) = 0,\ \forall j = 1, \dots, m\}, \tag{3.1}$$

where g_i and r_j are extended real-valued functions on X. In particular, if g_i are all identically zero, the latter reduces to a classical optimization problem with side conditions which can be solved by using the *Lagrangian function*

$$L(x, \nu) = f(x) + \sum_{j=1}^{m} \nu_j r_j(x), \quad x \in \mathrm{Dom}(f),\ \nu = (\nu_1, \nu_2, \dots, \nu_m) \in \mathbb{R}^m.$$

If certain differentiability hypotheses are present, then from classical analysis it is well known that if $x_0 \in \operatorname{int}\mathrm{Dom}(f)$ is an optimal solution for the above problem, then there exists an element $\nu^0 \in \mathbb{R}^m$ such that (x_0, ν^0) is a *critical point* for L on $\mathrm{Dom}(f) \times \mathbb{R}^m$ without side conditions. In other words, we can obtain necessary conditions for optimality by directly applying the Fermat theorem to the Lagrange function L. If $x_0 \in \operatorname{Fr}\mathrm{Dom}(f)$ or if the differentiability conditions are absent (which happens in many optimization problems), a more sophisticated treatment based on convexity theory is needed. In deriving necessary and sufficient conditions such that a given element be optimal in Problem $\mathscr{P}_1$, the convexity, which still allows a wide class of applications, avoids differentiability conditions on f and g. The first result, one of algebraic character, is concerned with the case in which the functions r_j are affine (a function is said to be *affine* if it is a sum of a linear functional and a constant function) and the functions g_i are convex. In this case, it is clear that the constraint set A_X is convex.

Theorem 3.1 *Let $f, g_1, g_2, \dots, g_n$ be proper convex functions and let $r_1, r_2, \dots, r_m$ be affine functions. If x_0 is an optimal solution of the consistent problem $\mathscr{P}_1$, where A_X is defined by (3.1), then there exist $n + m + 1$ real numbers $\lambda_0^0, \lambda_1^0, \dots, \lambda_n^0$, $\nu_1^0, \dots, \nu_m^0$, which are not all zero and have the properties*

$$\lambda_0^0 f(x_0) \le \lambda_0^0 f(x) + \sum_{i=1}^{n} \lambda_i^0 g_i(x) + \sum_{j=1}^{m} \nu_j^0 r_j(x), \quad \forall x \in X_0, \tag{3.2}$$

$$\lambda_0^0 \ge 0, \qquad \lambda_i^0 \ge 0, \qquad \lambda_i^0 g_i(x_0) = 0, \quad \forall i = 1, 2, \dots, n, \tag{3.3}$$

where

$$x_0 = \mathrm{Dom}(f) \cap \bigcap_{i=1}^{n} \mathrm{Dom}(g_i).$$

Proof Consider the subset

$$B = \{f(x) - f(x_0) + \alpha_0, g_1(x) + \alpha_1, g_2(x) + \alpha_2, \dots, g_n(x) + \alpha_n,$$
$$r_1(x), r_2(x), \dots, r_m(x);\ x \in X_0,\ \alpha_i > 0,\ i = 0, 1, \dots, n\}. \tag{3.4}$$

It is easily seen that B is a nonvoid convex set of $\mathbb{R}^{1+n+m}$ which does not contain the origin. According to Corollary 1.41, there exists a homogeneous hyperplane, that is, there exist $1+n+m$ real numbers $\lambda_0^0, \lambda_1^0, \dots, \lambda_n^0, \nu_1^0, \nu_2^0, \dots, \nu_m^0$, which are not all zero, such that

$$\lambda_0^0\big(f(x) - f(x_0) + \alpha_0\big) + \sum_{i=1}^n \lambda_i^0\big(g_i(x) + \alpha_i\big) + \sum_{j=1}^m \nu_j^0 r_j(x) \geq 0, \tag{3.5}$$

for all $x \in X_0$, $\alpha_i > 0$, $i = 0, 1, \dots, n$. Taking $x = x_0$, $\alpha_k \searrow 0$ for $k \neq i$ and $\alpha_i \nearrow \infty$ it follows that $\lambda_i^0 \geq 0$ for every $i = 0, 1, \dots, n$. Thus, relation (3.5) becomes

$$\lambda_0^0\big(f(x) - f(x_0)\big) + \sum_{i=1}^n \lambda_i^0 g_i(x) + \sum_{j=1}^m \nu_j^0 r_j(x) \geq 0, \quad \forall x \in X_0,$$

and (3.2) is proved.

Now, it is clear that $\lambda_i^0 g_i(x_0) \geq 0$, $\forall i = 1, 2, \dots, n$, since $x_0 \in A_X$. If $x = x_0$, from the above inequality we also obtain $\sum_{i=1}^n \lambda_i^0 g_i(x_0) \geq 0$, that is, $\lambda_i^0 g_i(x_0) = 0$, $\forall i = 1, 2, \dots, n$, which completes the proof. □

The numbers λ_i^0, ν_j^0 with the properties mentioned in the theorem are called the *Lagrange multipliers* of Problem $\mathscr{P}_1$. Since relations (3.2) and (3.3) are homogeneous with respect to coefficients, we can only consider $\lambda_0^0 = 0$ or 1.

Thus, it is natural to call the function

$$\begin{aligned} L(x, \lambda, \nu) &= \varepsilon f(x) + \sum_{i=1}^n \lambda_i g_i(x) + \sum_{j=1}^m \nu_j r_j(x), \quad x \in X_0, \\ \lambda &= (\lambda_i) \in \mathbb{R}_+^n, \ \nu = (\nu_j) \in \mathbb{R}^m, \end{aligned} \tag{3.6}$$

where $\varepsilon = 0$ or 1, the *Lagrange function* attached to Problem $\mathscr{P}_1$.

Remark 3.2 The necessary conditions (3.2) and (3.3) with $x_0 \in A_X$ are equivalent to the fact that the point $(x_0, \lambda_1^0, \dots, \lambda_n^0, \nu_1^0, \dots, \nu_m^0)$ is a saddle point for the Lagrange function (3.6) on $X_0 \times \mathbb{R}_+^n \times \mathbb{R}^m$, either for $\varepsilon = 0$, or for $\varepsilon = 1$, with respect to minimization on X_0 and maximization on $\mathbb{R}_+^n \times \mathbb{R}^m$, that is,

$$\varepsilon f(x_0) + \sum_{i=1}^n \lambda_i g_i(x_0) + \sum_{j=1}^m \nu_j r_j(x_0) \leq \varepsilon f(x) + \sum_{i=1}^n \lambda_i^0 g_i(x) + \sum_{j=1}^m \nu_j^0 r_j(x), \tag{3.7}$$

for every $(x, \lambda, \nu) \in X_0 \times \mathbb{R}_+^n \times \mathbb{R}^m$.

Relations (3.2) and (3.3) with $\lambda_0^0 = \varepsilon$ for $x_0 \in A_X$ clearly imply relation (3.7) because we have $g_i(x_0) \leq 0$ and $r_j(x_0) = 0$. Conversely, for $\lambda_i = 0$, $\nu_j = 0$, relation (3.7) implies inequality (3.2).

From the same relation (3.7), for $\lambda_i \nearrow \infty$ and $\nu_j \to \pm\infty$, it follows that $g_i(x_0) \le 0$, $r_j(x_0) = 0$, that is, $x_0 \in A_X$. On the other hand, for $x = x_0$, $\lambda_i = 0$, $\nu_j = 0$, relation (3.7) becomes $\sum_{i=1}^n \lambda_i^0 g_i(x_0) \ge 0$, and so, $\lambda_i^0 g_i(x_0) = 0$ for all $i = 1, 2, \ldots, n$.

Remark 3.3 The necessary optimality conditions (3.2) and (3.3) with $\lambda_0^0 \ne 0$ (that is, (3.7) with $\varepsilon = 1$) and $x_0 \in A_X$ are also sufficient for x_0 to be an optimal solution to Problem $\mathscr{P}_1$. If $\lambda_0^0 = 0$ ($\varepsilon = 0$) the optimality conditions concern only the constraints functions, without giving any piece of information for the function which is minimized.

It is natural to give certain additional conditions called *constraint qualifications* which ensure that $\lambda_0^0 \ne 0$.

The following *Slater's constraint qualification* is an instance of a constraint qualification that is easily verifiable in many particular applications.

(S) *There exists a point* $\bar{x} \in A_X$ *such that* $g_i(\bar{x}) < 0$, $\forall i = 1, 2, \ldots, n$.

For the equality constraints we consider the interiority conditions

(O) $0 \in \operatorname{int}\{(r_1(x), r_2(x), \ldots, r_m(x));\ x \in X_0\}$.

Theorem 3.4 *Let* $f, g_1, \ldots, g_n$ *be proper convex functions and let* $r_1, r_2, \ldots, r_m$ *be affine functions such that* (S) *and* (O) *are fulfilled. Then a point* $x_0 \in A_X$ *is an optimal solution for* $\mathscr{P}_1$ *if and only if there exist* $n+m$ *real numbers* $\lambda_1^0, \lambda_2^0, \ldots, \lambda_n^0$, $\nu_1^0, \nu_2^0, \ldots, \nu_m^0$, *such that*

$$f(x_0) \le f(x) + \sum_{i=1}^{r} \lambda_i^0 g_i(x) + \sum_{j=1}^{m} \nu_j^0 r_j(x), \quad \forall x \in X_0, \tag{3.8}$$

$$\lambda_i^0 \ge 0, \qquad \lambda_i^0 g_i(x_0) = 0, \quad \forall i = 1, 2, \ldots, n. \tag{3.9}$$

Proof Let x_0 be an optimal solution of $\mathscr{P}_1$. According to Theorem 3.1, there exist $\lambda_0^0, \lambda_1^0, \ldots, \lambda_n^0, \nu_1^0, \nu_2^0, \ldots, \nu_m^0$ not all zero such that (3.2) and (3.3) hold. If we suppose $\lambda_0^0 = 0$, taking $x = \bar{x} \in A_X$ from (3.2) we obtain $\sum_{i=1}^n \lambda_i^0 g_i(\bar{x}) \ge 0$. Since $\lambda_i^0 \ge 0$ and $g_i(\bar{x}) < 0$ for each i, we must have $\lambda_i^0 = 0$ for all $i = 1, 2, \ldots, n$. Hence (3.2) becomes

$$\sum_{i=1}^{m} \nu_j^0 r_j(x) \ge 0 \quad \text{for all } x \in X_0,$$

where ν_j^0 are not all zero, contradicting the interiority condition (O). Hence $\lambda_0^0 > 0$, that is, we can eventually take $\lambda_0^0 = 1$. Sufficiency follows from (3.8) since, for $x \in A_X$, we have $\lambda_i^0 g(x) \le 0$ and $r_j(x) = 0$. Moreover, $f(x_0)$ is necessarily finite. □

By virtue of Remark 3.2, we also have the following theorem.

Theorem 3.4′ *Under the hypotheses of Theorem 3.4, an element $x_0 \in X$ is an optimal solution if and only if there exist $\lambda^0 = (\lambda_1^0, \dots, \lambda_n^0) \in \mathbb{R}^n$, $\nu^0 = (\nu_1^0, \nu_2^0, \dots, \nu_m^0) \in \mathbb{R}^m$ such that (x_0, λ^0, ν^0) is a saddle point for the Lagrange function on $X_0 \times (\mathbb{R}_+^n \times \mathbb{R}^m)$, that is,*

$$f(x_0) + \sum_{i=1}^n \lambda_i g_i(x_0) + \sum_{j=1}^m \nu_j r_j(x_0) \le f(x) + \sum_{i=1}^n \lambda_i^0 g_i(x) + \sum_{j=1}^m \nu_j^0 r_j(x), \quad (3.10)$$

for all $(x, \lambda, \nu) \in X_0 \times \mathbb{R}_+^n \times \mathbb{R}^m$.

Remark 3.5 The constraint qualifications (S) and (O) ensures that the constraints are consistent on A_X and none of the equality constraints is redundant. Moreover, the trace on X_0 of the affine set from the right-hand side of (O) has dimension m.

Remark 3.6 The sufficiency is also true without the regularity. For a separated locally convex space, Theorem 3.4 can be improved to obtain the well known *Kuhn–Tucker theorem*, a classical result in programming theory.

Theorem 3.7 (Kuhn–Tucker) *Under the hypotheses of Theorem 3.4, if we further assume that the function f is lower-semicontinuous and g_i, r_j are continuous real functions, then the optimality condition (3.8) for $x_0 \in A_X$ is equivalent to the condition*

$$\begin{aligned} 0 \in \partial f(x_0) &+ \lambda_1^0 \partial g_1(x_0) + \cdots + \lambda_n^0 \partial g_n(x_0) \\ &+ \nu^0 \nabla r_1(x_0) + \nu_2^0 \nabla r_2(x_0) + \cdots + \nu_m^0 \nabla r_m(x_0). \end{aligned} \quad (3.11)$$

Proof Since conditions (3.9) are verified, condition (3.8) says that $x_0 \in A_X$ is a minimum point of the function

$$x \to f(x) + \sum_{i=1}^n \lambda_i^0 g_i(x) + \sum_{j=1}^m \nu_j^0 r_j(x) \quad \text{on } X, \quad \text{that is,}$$

$$0 \in \partial \left(f + \sum_{i=1}^n \lambda_i^0 g_i + \sum_{j=1}^m \nu_j^0 r_j \right)(x_0),$$

since inequality (3.8) is trivial on $X \setminus X_0$ and $X_0 = \mathrm{Dom}(f)$.

Using the additive of the subdifferential (see Corollary 2.63 and Remark 2.61, or Theorem 3.57), we obtain the equivalence to relation (3.11), as claimed. □

Remark 3.8 Since r_j is affine, there exist a continuous linear functional $x_j^* \in X^*$ and a real number $\alpha_j \in \mathbb{R}$ such that $r_j = x_j^* + \alpha_j$. Therefore, we have $\nabla r_j = x_j^*$ and (3.11) becomes

$$\begin{aligned} 0 \in \partial f(x_0) &+ \lambda_1^0 \partial g_1(x_0) + \lambda_2^0 \partial g_2(x_0) + \cdots + \lambda_n^0 \partial g_n(x_0) \\ &+ x_1^*(x_0) + x_2^*(x_0) + \cdots + x_m^*(x_0). \end{aligned} \quad (3.12)$$

Now, if we consider only the case of the constraint given by inequalities, that is,

$$A_x = \{x \in X;\ g_i(x) \leq 0,\ \forall i = 1, 2, \ldots, n\}, \tag{3.13}$$

the *Slater condition* is as follows.

(S) *There exists a point* $\bar{x} \in \mathrm{Dom}(f)$ *such that* $g_i(\bar{x}) < 0, \forall i = 1, 2, \ldots, n.$

From Theorem 3.7, we obtain the following result.

Theorem 3.9 *Let* f *be a proper convex lower-semicontinuous function and let* $g_1, g_2, \ldots, g_n$ *be real convex continuous functions satisfying the Slater condition* (S)*. Then a point* $x_0 \in A_X$ *(given by* (3.13)*) is an optimal solution for* $\mathscr{P}_1$ *if and only if there exists* $\lambda^0 = (\lambda_1^0, \lambda_2^0, \ldots, \lambda_n^0)$ *such that*

$$0 \in \partial f(x_0) + \lambda_1^0 \partial g_1(x_0) + \lambda_2^0 \partial g_2(x_0) + \cdots + \lambda_n^0 \partial g_n(x_0), \tag{3.14}$$

$$\lambda_i^0 \geq 0, \qquad \lambda_i^0 g_i(x_0) = 0, \quad \forall i = 1, 2, \ldots, n. \tag{3.15}$$

Corollary 3.10 *Let* $f, g_1, g_2, \ldots, g_n$ *be real convex and differentiable functions on* X *which satisfy* (S)*. Then a feasible element* x_0 *is an optimal solution of Problem* $\mathscr{P}_1$ *with* A_X *given by* (3.13) *if and only if there exist real numbers* $\lambda_1^0, \lambda_2^0, \ldots, \lambda_n^0$ *such that*

$$\nabla f(x_0) + \lambda_1^0 \nabla g_1(x_0) + \cdots + \lambda_n^0 \nabla g_n(x_0) = 0, \tag{3.16}$$

$$\lambda_1^0 \geq 0, \qquad \lambda_i^0 g_i(x_0) = 0, \quad \forall i = 1, 2, \ldots, n. \tag{3.17}$$

Remark 3.11 If X is a finite-dimensional space, $\dim X = k$, the method of Lagrange multiplier employs a simpler technique for finding the optimal solutions. First, we find the solutions $(\lambda, x) \in \mathbb{R}^n \times X$ of the system

$$f(x) + \sum_{i=1}^{n} \lambda_i \nabla g_i(x) = 0,$$

$$\lambda_i g_i(x) = 0, \quad i = 1, 2, \ldots, n.$$

This system is formed by $k + n$ equations and $k + n$ real unknowns. If $\lambda_i \geq 0$ for all $i = 1, 2, \ldots, n$, then the feasible element x corresponding to λ is an optimal solution. Hence, in the finite-dimensional case, for differentiable convex functions, the Kuhn–Tucker conditions (3.16) and (3.17) give a practical procedure to solve completely the convex programming problem.

3.1.2 Operatorial Convex Constraints

The problem studied in the preceding section, in which the constraints are given by a finite number of inequations and equations may be extended in various ways. Thus, the function g_i may be included in an operator, the natural order of real numbers being replaced by an order generated by a convex cone. Namely, we consider the programming problem

$$(\mathscr{P}_2) \quad \min\{f(x);\ x \in A,\ G(x) \in -A_Y\},$$

where X, Y are two separated locally convex spaces, A is a convex subset of X, A_Y is a closed convex cone of Y, $f : X \to \overline{\mathbb{R}}$ is a proper convex function with $\mathrm{Dom}(f) \supset A$ and $G : D(G) \to Y$ with $D(G) \subset X$ is a *convex operator*, that is, $D(G)$ is convex and

$$\begin{aligned}&G(\lambda_1 x_1 + \lambda_2 x_2) \le \lambda_1 G(x_1) + \lambda_2 G(x_2), \quad \forall x_1, x_2 \in D(G),\\ &\lambda_1 \ge 0,\ \lambda_2 \ge 0, \quad \lambda_1 + \lambda_2 = 1.\end{aligned}$$

The *ordering relation* in Y is generated by the cone A_Y, that is,

$$y_1 \ge y_2 \quad \text{if and only if } y_1 - y_2 \in A_Y.$$

Feasible elements of $\mathscr{P}_2$ are the elements of $A \cap G^{-1}(-A_Y)$. Thus, Problem $\mathscr{P}_2$ is considered *consistent* if $A \cap G^{-1}(-A_Y) \ne \emptyset$. It is clear that, in the special cases

$$\begin{aligned}&A = \mathrm{Dom}(f) \cap \{x \in X;\ r_j(x) = 0,\ \forall j = 1, 2, \dots, m\},\\ &Y = \mathbb{R}^m, \qquad A_Y = \mathbb{R}^m_+, \qquad G(x) = (g_1(x), g_2(x), \dots, g_n(x)),\\ &x \in \bigcap_{i=1}^{n} \mathrm{Dom}(g_i),\end{aligned} \tag{3.18}$$

or

$$\begin{aligned}&A = \mathrm{Dom}(f), \qquad Y = \mathbb{R}^{n+m}, \qquad A_Y = \mathbb{R}^n_+ \times \{0_{\mathbb{R}^m}\},\\ &G(x) = (g_1(x), g_2(x), \dots, g_n(x), r_1(x), r_2(x), \dots, r_m(x)),\\ &x \in \bigcap_{i=1}^{n} \mathrm{Dom}(g_i).\end{aligned} \tag{3.19}$$

Problem $\mathscr{P}_2$ reduces to the preceding problem $\mathscr{P}_1$ with A_X given in (3.1).

In general, the constraints given by equations are included in the set A, and those given by inequations are expressed in terms of ordering generated by the cone A_Y.

Now, it is natural to consider as the *Lagrange* (*Fritz John*) *multiplier* for an optimal solution x_0 of Problem $\mathscr{P}_2$ a pair of elements $(\eta_0, y_n^*) \in \mathbb{R} \times Y^*$, not both zero, which satisfies the properties

$$\eta_0 f(x_0) \le \eta_0 f(x) + (y_0^*, G(x)), \quad \forall x \in A \cap D(G), \tag{3.20}$$

$$\eta_0 \ge 0, \qquad (y_0^*, y) \ge 0, \quad \forall y \in A_Y\ (\text{i.e., } y_0^* \in -A_Y^0),\ (y_0^*, G(x_0)) = 0. \tag{3.21}$$

In the following, we are going to establish some results analogous to those given in the preceding section concerning the existence of Lagrange multipliers. With that end in view, let us first remark that the role of the set B used in the proof of Theorem 3.1 is taken, here, by the set

$$B = \{(f(x) - f(x_0) + \alpha, G(x) + y) \in \mathbb{R} \times Y;\ x \in A \cap D(G),\ y \in A_Y,\ \alpha \geq 0\}, \tag{3.22}$$

where $x_0 \in A \cap G^{-1}(-A_Y) \cap D(G) \cap \mathrm{Dom}(f)$.

It is easy to see that B is convex and contains the element $(0, G(x_0))$ if x_0 is the optimal solution for $\mathscr{P}_2$.

Lemma 3.12 *An element* $(\eta_0, y_0^*) \in \mathbb{R} \times Y^*$ *has properties* (3.20) *and* (3.21) *if and only if* $(\eta_0, y_0^*) \in -(\mathrm{cone}\, B)^0$.

Proof According to the definition of the polar of a cone, it follows that $(\eta_0, y_0^*) \in -(\mathrm{cone}\, B)^0$ if and only if

$$\eta_0(f(x) - f(x_0) + \alpha) + (y_0^*, G(x) + y) \geq 0,$$
$$\text{for all } x \in A \cap D(G),\ y \in A_Y,\ \alpha \geq 0.$$

Since $n\alpha \geq 0$ and $ny \in A_Y$ for any $n \in \mathbb{N}$ if $\alpha \geq 0$ and $y \in A_Y$, the preceding inequality is equivalent to the two properties

$$\eta_0(f(x) - f(x_0)) + (y^*, G(x)) \geq 0, \quad \text{for all } x \in A \cap D(G)$$

and

$$\eta_0 \geq 0 \text{ and } (y_0^*, y) \geq 0, \quad \text{for all } y \in A_Y,\ \text{i.e.,}\ y_0^* \in -A_Y^0.$$

We also have $(y_0^*, G(x_0)) = 0$. □

With the help of Lemma 3.12, one easily obtains the following theorem.

Theorem 3.13 *For the optimal solution* x_0 *of Problem* $\mathscr{P}_2$ *there exists a Lagrange multiplier* $(\eta_0, y_0^*) \in \mathbb{R} \times Y^*$ *if and only if*

$$\overline{\mathrm{cone}\, B} \neq \mathbb{R} \times Y. \tag{3.23}$$

Proof By virtue of Lemma 3.12 we have the result that there exist Lagrange multipliers if and only if $(\mathrm{cone}\, B)^0 \neq \{(0, 0)\}$. According to the bipolar theorem (Theorem 2.26, Chap. 2), this reduces to condition (3.23). □

In the optimization theory, an important role is played by proper Lagrange multipliers, that is those for which $\eta_0 \neq 0$. In this case, we can take $\eta_0 = 1$ and then the corresponding element y_0^* is considered as a (*proper*) *Lagrange multiplier*. The characteristic properties (3.20) and (3.21) become

$$f(x_0) \leq f(x) + (y_0^*, G(x)), \quad \forall x \in A \cap D(G), \tag{3.24}$$

$$(y_0^*, y) \geq 0, \quad \forall y \in A_Y \ \big(\text{i.e., } y_0^* \in -A_Y^0\big), \ \big(y_0^*, G(x_0)\big) = 0. \tag{3.25}$$

The *Lagrange function* is defined by

$$L(x, y^*) = f(x) + \big(y^*, G(x)\big), \quad \forall x \in A \cap D(G), \ y^* \in -A_Y^0. \tag{3.26}$$

For the existence of this type of Lagrange multiplier it is necessary to impose some regularity conditions. Among them, let us consider the simple interiority conditions

$$0 \in \operatorname{int}\big(G(A) + A_Y\big). \tag{3.27}$$

Lemma 3.14 *If* (3.27) *holds, then every Lagrange multiplier is proper.*

Proof Suppose by contradiction that $(0, y_0^*)$ is a Lagrange multiplier. Then, from the characteristic property (3.20), we obtain

$$\big(y_0^*, G(x) + y\big) \geq 0 \quad \text{for all } x \in A \cap D(G) \text{ and } y \in A_Y.$$

By virtue of hypothesis (3.27), it follows that y_0^* is nonnegative on a neighborhood of origin (an absorbent set). This implies $y_0^* = 0$, which is not possible since η_0, y_0^* are not both zero.

Next, we observe that there exists a Lagrange multiplier, that is, condition (3.23) is fulfilled, if one of these two simple conditions holds:

$$\operatorname{int} A_Y \neq \emptyset, \tag{3.28}$$

or

$$\operatorname{cone} B \text{ is closed in } \mathbb{R} \times Y. \tag{3.29}$$

Indeed, if $\operatorname{int} A_Y \neq \emptyset$, it follows that $]-\infty, 0[\times (-\operatorname{int} A_Y) \cap \overline{\operatorname{cone} B} = \emptyset$. Assume that $(\alpha_0, y_0) \in -\overline{\operatorname{cone} B}$ and $y_0 \in \operatorname{int} A_Y$. Then, according to the definition of the set B, there exist the nets $\lambda_i > 0$, $y_i \in A_Y$, $x_i \in A \cap D(G)$ and $\alpha_i \geq 0$ such that

$$\lim_i \lambda_i \big(f(x_i) - f(x_0) + \alpha_i\big) = -\alpha_0,$$

$$\lim_i \lambda_i \big(y_i + G(x_i)\big) = -y_0.$$

Since $y_0 \in \operatorname{int} A_Y$, there exists i_0 such that $\lambda_i(y_i + G(x_i)) \in -A_Y$ for all $i > i_0$ and so, $G(x_i) \in -A_Y, \forall i > i_0$. On the other hand, $x_i \in A \cap D(G_i)$ and $G(x_i) \in -A_Y$ involve $f(x_i) \geq f(x_0)$ and hence $\alpha_0 \leq 0$. Therefore, $(\alpha_0, y_0) \notin -\overline{\operatorname{cone} B}$ if $\alpha_0 > 0$ and $y_0 \in \operatorname{int} A_Y$ (that is, $\overline{\operatorname{cone} B} \neq R \times Y$). To prove that condition (3.23) is fulfilled if condition (3.29) is verified, we observe that $(]0, \infty[\times A_Y) \cap (-\overline{\operatorname{cone} B}) = \emptyset$.

For simplicity, we suppose next that

$$D(G) = X \quad \big(\text{or } D(G) \supset A\big),$$

which, eventually, leads to the restriction of the set A. $\square$

Thus, we have the following result.

Theorem 3.15 *If $\mathscr{P}_2$ satisfies conditions* (3.27) *and one of* (3.28) *or* (3.29), *then $x_0 \in A$ is an optimal solution if and only if there exists $y_0^* \in -A_Y^0$ such that*

$$f(x_0) + \bigl(G(x_0), y^*\bigr) \le f(x) + \bigl(G(x), y_0^*\bigr), \quad \forall (x, y^*) \in A \times \bigl(-A_Y^0\bigr), \tag{3.30}$$

that is, (x_0, y_0^*) *is a saddle point for the Lagrange function defined by* (3.26).

Proof It is clear that (3.30) follows from (3.24) and (3.25). Conversely, we first see that, if we take $y^* = 0$ in (3.30), we obtain $f(x_0) \le f(x)$ for every $x \in A$ such that $G(x) \in -A_Y$. Hence, it is sufficient to prove that $G(x_0) \in -A_Y$. It is easy to show that from (3.30) it follows that we have the so-called *complementary slackness condition*,

$$\bigl(G(x_0), y_0^*\bigr) = 0 \tag{3.31}$$

(taking $x = x_0$ and $y^* = \alpha y_0^*$, $\forall \alpha \ge 0$). Replacing $x = x_0$ in (3.30), it follows that $(G(x_0), y^*) \le 0$, $\forall y^* \in -A_Y^0$, that is, $G(x_0) \in -A_Y^{00} = -A_Y$ (Theorem 2.26 of bipolar), as claimed. □

Remark 3.16 The regularity condition given by (3.27) and (3.28) is equivalent to the usual *Slater condition*.

(S) *There exists $\bar{x} \in A$ such that $G(\bar{x}) \in -\operatorname{int} A_Y$.*

Other forms of regularity conditions, such as closedness conditions, are given in Sect. 3.2.3.

Now, let us observe that relation (3.30) is equivalent to relations (3.24) and (3.31). If $A = D(G) = X$, we find that x_0 is a minimum point for the function $x \to L(x, y_0)$, $x \in X$, and hence $0 \in \partial(f + y_0^* \circ G)(x_0)$. If f or G is continuous at x_0, we can apply the additivity property of the subdifferential

$$0 \in \partial f(x_0) + \partial(y_0^* \circ G)(x_0).$$

Remark 3.17 To obtain a result similar to the *Kuhn–Tucker theorem*, it is sufficient to have an equality of type

$$\partial(y_0^* \circ G)(x_0) = y_0^* \circ \partial G(x_0), \tag{3.32}$$

where the subdifferential $\partial G(x_0)$ of the convex operator G at x_0 is similarly defined with the aid of the ordering generated by the cone A_Y, that is,

$$\partial G(x_0) = \bigl\{T \in L(X, Y);\ T(x - x_0) \in G(x) - G(x_0) + A_Y,\ \forall x \in X\bigr\}. \tag{3.33}$$

But equality (3.32) does not generally hold. In this way, it is natural to say that the mapping G is *regular subdifferentiable* at x_0 if equality (3.32) is satisfied for every

$y_0^* \in A_Y^0$. A sufficient condition for the regular subdifferentiability is the following: *X is a reflexive Banach space and* A_Y *has a weakly compact base which does not contain an origin*. By virtue of the above remark, Theorem 3.15 can be restated as a theorem of Kuhn–Tucker type.

Theorem 3.18 *If* $\mathscr{P}_2$ *satisfies conditions* (3.27), (3.28) (*or* (3.29)) *and G is continuous and regularly subdifferentiable on A, then an element* $x_0 \in A$ *is an optimal solution if and only if there exists* $y_0^* \in -A_Y^0$ *such that*

$$0 \in \partial f(x_0) + y_0^* \circ \partial G(x_0), \quad \big(y_0^*, G(x_0)\big) = 0.$$

It is easily seen that, if G is continuous and differentiable (for instance, if G is Fréchet differentiable), then equality (3.32) holds (and $\partial G(x_0)$ contains a unique element). In fact, we see later that the differentiability hypothesis allows necessary optimality conditions to be obtained even for non-convex problems.

Note, finally, that the general problem $\mathscr{P}_1$ can be reformulated as a problem of minimization of a certain function over the space X, where no constraints explicitly appear. More precisely, $\mathscr{P}_1$ is equivalent to the following unconstrained problem:

$$\min\big\{f(x) + I_{A_X}(x);\ x \in X\big\}.$$

Let f be a lower-semicontinuous function and let A_X be a closed convex set of X. If $(\operatorname{int} A_X) \cap \operatorname{Dom}(f) \neq \emptyset$ or if f happens to be continuous at a point of A_X, then, by virtue of the additivity theorem for subdifferentials, we may infer that x_0 is an optimal solution if and only if

$$\partial f(x_0) \cap \big(-\partial I_{A_X}(x_0)\big) \neq \emptyset, \tag{3.34}$$

or, equivalently,

$$\partial f(x_0) \cap \big(-C^0(A_X; x_0)\big) \neq \emptyset, \tag{3.35}$$

where $C(A_X; x_0)$ represents the cone generated by $A_X - x_0$. In particular, $x_0 \in \operatorname{int} A_X$ is an optimal solution if and only if $0 \in \partial f(x_0)$ since $C^0(A_X; x_0 = \{0\})$. Therefore, the cases of special interest are those in which $x_0 \notin \operatorname{int} A_X$, that is, $x_0 \in \operatorname{Fr} A_X$.

By conjugacy, the optimality condition (3.34) is equivalent to the existence of an element $x_0^* \in X^*$, subject to

$$f(x_0) + f^*(x_0^*) + I_{A_X}^*(-x_0^*) = 0, \quad x_0 \in A_X,$$

that is,

$$f(x_0) = \inf\big\{(x_0^*, x);\ x \in A_X\big\} - f^*(x_0^*). \tag{3.36}$$

Remark 3.19 If A_X is the translate of a closed linear subspace, the optimality condition (3.35) becomes

$$\partial f(x_0) \cap (A_X - x_0)^\perp \neq \emptyset, \quad x_0 \in A_X. \tag{3.37}$$

To be more specific, we consider the *case of affine constraints*,

$$(\mathscr{P}_3) \quad \min\{f(x);\ G(x)=0\},$$

where $G : X \to Y$ is an *affine mapping*, that is,

$$G(x) = Tx + y_0, \quad \text{for all } x \in X,$$

with $T \in L(X, Y)$ and $y_0 \in Y$.

Theorem 3.20 *If f is continuous in a point of $\ker G$ and T has a closed range, then a point $x_0 \in \ker G$ is an optimal solution of $\mathscr{P}_3$ if and only if*

$$\partial f(x_0) \cap \operatorname{Range} T^* \neq \emptyset. \tag{3.38}$$

Proof We observe that $A_X - x_0 = \ker T$. On the other hand, $(\ker T)^\perp = \overline{\operatorname{Range} T^*} = \operatorname{Range} T^*$ because $\operatorname{Range} T^*$ is also closed. Hence, condition (3.38) is equivalent to the optimality condition (3.37). □

3.1.3 Nonlinear Programming in the Case of Fréchet Differentiability

In the following, we show that the results obtained in the convex case can be extended to the general case of the nonlinear programming if the Fréchet differentiability of functions involved in the problem is required.

We return to the minimization problem

$$(\mathscr{P}_2) \quad \min\{f(x);\ x \in A,\ G(x) \in -A_Y\},$$

where X, Y are two linear normed spaces, $G : X \to Y$ and $f : X \to \mathbb{R}$ are two arbitrary mappings, A is a nonvoid subset of X and A_Y is a closed convex cone of Y.

Our aim her is to obtain some necessary conditions of Kuhn–Tucker type for $\mathscr{P}_2$. With that end in view, we introduce some preliminary notions.

Definition 3.21 We call the *cone of tangents to A at $x_0 \in A$*, denoted by $TC(A; x_0)$, the set defined by

$$TC(A; x_0) = \bigcap_{V \in \mathscr{V}(x_0)} \overline{C(A \cap V; x_0)}, \tag{3.39}$$

where $\mathscr{V}(x_0)$ is a base of neighborhoods for x_0 (by $C(M; x_0)$ we denote the *cone generated* by $M - x_0$), that is, $\bigcup_{\lambda \geq 0} \lambda(M - x_0)$).

It is clearly seen that $TC(A; x_0)$ is a closed cone with vertex at zero, but need not to be convex.

Lemma 3.22 $x \in TC(A; x_0) \setminus \{0\}$ *if and only if there exist a sequence* $\{x_n\} \subset A$ *and a sequence* $\{\lambda_n\}$ *of positive numbers such that*

$$x_n \to x_0 \quad and \quad \lambda_n(x_n - x_0) \to x. \tag{3.40}$$

Proof Observe that relation (3.39) does not depend on the base of neighborhoods $\mathscr{V}(x_0)$ but, actually, it depends on the point x_0. If $\mathscr{V}(x_0) = \{S(x_0; \frac{1}{n}); n \in \mathbb{N}^*\}$, relation (3.39) is equivalent to the fact that, for every $n \in \mathbb{N}^*$, there exist $\lambda_n > 0$ and $x_n \in A \cap S(x_0; \frac{1}{n})$ subject to $\lambda_n(x_n - x_0) \in S(x; \frac{1}{n})$, thereby proving the lemma. $\square$

Definition 3.23 The *cone of pseudotangents to A at $x_0 \in A$*, denoted by $PC(A; x_0)$, is by definition the closed convex hull of the cone of tangents to A in x_0, that is,

$$PC(A; x_0) = \overline{\operatorname{cone} TC(A; x_0)}. \tag{3.41}$$

In general, $TC(A; x_0) \subset \overline{C(A; x_0)}$ and, if A is convex, then

$$TC(A; x_0) = PC(A; x_0) = \overline{C(A; x_0)}.$$

Moreover, if A is *star-shaped at* x_0, that is, $[x, x_0] \subset A$ for every $x \in A$, then

$$C(A; x_0) = C(A \cap V; x_0) \quad \text{for each } V \in \mathscr{V}(x_0).$$

Definition 3.24 We say that a set A_1 *is pseudoconvex with respect to a set A_2 at x_0* if $x - x_0 \in PC(A_2; x_0)$ for all $x \in A_1$. If $A_1 = A_2$, we say that A_1 *is pseudoconvex at* x_0.

It is clear that, if a set A is star-shaped at x_0, then it is pseudoconvex at x_0. In particular, any convex set is pseudoconvex at every one of its elements. From Definitions 3.23 and 3.24, it follows that

$$PC(A; x_0) = PC(\operatorname{conv} A; x_0) \quad \text{if } A \text{ is pseudoconvex at } x_0. \tag{3.42}$$

Indeed, if $x \in TC(\operatorname{conv} A; x_0)$, by virtue of Lemma 3.22, there exist a sequence $\{x_n\} \subset \operatorname{conv} A$ and a sequence $\{\lambda_n\}$ of positive numbers such that $x_n \to x_0$ and $\lambda_n(x_n - x_0) \to x$. But $x_n = \sum_i \alpha_n^i x_n^i$, where $\{x_n^i\} \subset A$, $\alpha_n^i > 0$, $\sum_i \alpha_n^i = 1$. Hence, $x_n - x_0 = \sum_i \alpha_n^i (x_0^i - x_0) \in PC(A; x_0)$, in view of Definition 3.24.

Since $PC(A; x_0)$ is a closed cone, it follows that $x \in PC(A; x_0)$, and thus, $TC(\operatorname{conv} A; x_0) \subset PC(A; x_0)$, hence $PC(\operatorname{conv} A; x_0) \subset PC(A; x_0)$.

In what follows, we use Fréchet differentiability conditions for the functions which will intervene. We recall that a function $\varphi : X_1 \to X_2$, where X_1 and X_2 are linear normed spaces, is *Fréchet differentiable at a point* $x_0 \in X_1$ if there exists a mapping $\varphi'_{x_0} \in L(X_1, X_2)$ such that

$$\lim_{x \to 0} \frac{\varphi(x + x_0) - \varphi(x_0) - \varphi'_{x_0}(x)}{\|x\|} = 0. \tag{3.43}$$

The mapping φ'_{x_0} is called the *Fréchet differential of the function* φ *in the point* x_0.

We easily see that the above condition can also be written as

$$\varphi(x+x_0)-\varphi(x_0)=\varphi'_{x_0}(x)+\omega(x_0;x), \quad \forall x\in X_1, \tag{3.44}$$

where

$$\lim_{x\to 0}\frac{\omega(x_0;x)}{\|x\|}=0. \tag{3.44'}$$

Lemma 3.25 *Let* $\varphi: X_1\to X_2$ *be a Fréchet differentiable function at* x_0. *Then*

$$\lim_{x\to\infty}\lambda_n\big[\varphi(x_n)-\varphi(x_0)\big]=\varphi'_{x_0}(x), \quad \forall x\in TC(A;x_0)\setminus\{0\}, \tag{3.45}$$

where $A\subset X_1$ *and* $\{x_n\},\{\lambda_n\}$ *are as in Lemma* 3.22.

Proof In relation (3.44), taking $x=x_n-x_0$, $n\in\mathbb{N}$, we obtain

$$\lambda_n\big[\varphi(x_n)-\varphi(x_0)\big]=\varphi'_{x_0}\big[\lambda_n(x_n-x_0)\big]+\big\|\lambda_n(x_n-x_0)\big\|\,\frac{\omega(x_0;x_n-x_0)}{\|x_n-x_0\|},$$

because φ'_{x_0} is linear and $\lambda_n>0$. In view of properties (3.40) and (3.44), there follows (3.45). □

Remark 3.26 According to Lemma 3.22, equality (3.45) becomes

$$\varphi'_{x_0}\big(TC(A;x_0)\big)\subset TC\big(\varphi(A);\varphi(x_0)\big). \tag{3.46}$$

Since φ'_{x_0} is linear and continuous, we also obtain

$$\varphi'_{x_0}\big(PC(A;x_0)\big)\subset PC\big(\varphi(A);\varphi(x_0)\big). \tag{3.47}$$

Definition 3.27 We say that a function $f: X\to\mathbb{R}$ is *pseudoconvex on* A *in* $x_0\in A$ if it is Fréchet differentiable in x_0 and possesses the property

$$f'_{x_0}(x-x_0)\ge 0 \quad \text{with } x\in A \text{ implies } f(x_0)\le f(x).$$

Theorem 3.28 *If* f *is a real Fréchet differentiable function in* $x_0\in A$ *and* x_0 *minimizes* f *on* A, *then* $f'_{x_0}(x)\ge 0$, $\forall x\in PC(A;x_0)$, *that is,* $f'_{x_0}\in -PC(A;x_0)^0$. *If* A *is pseudoconvex in* x_0 *and* f *is pseudoconvex on* A *in* x_0, *then the above condition is also sufficient.*

Proof For $x=0$ the condition is trivially verified. By means of Lemma 3.25, we get $f'_{x_0}(x)\ge 0$, $\forall x\in TC(A;x_0)$. This inequality may be extended to all elements x in $PC(A;x_0)$ because f'_{x_0} is a linear continuous functional. Thus, necessity is established. For the proof of sufficiency, we note that, for every $x\in A$, by virtue of Definition 3.24, we have $x-x_0\in PC(A;x_0$ and hence $f'_{x_0}(x-x_0)\ge 0$. Since f

is pseudo convex, from Definitions 3.24 and 3.27 we have the result that $f(x_0) \leq f(x)$, for all $x \in A$. This shows that x_0 is an optimal solution on A. □

Let us now establish some optimality conditions for $\mathscr{P}_2$.

If we denote by A_X the set

$$A_X = \{x \in A;\ G(x) \in -A_Y\} = A \cap G^{-1}(-A_Y), \tag{3.48}$$

then $\mathscr{P}_2$ reduces to the minimization of f on A_X.

A first auxiliary result is concerned with the cone of tangents to A at x_0.

Lemma 3.29 *If* $\operatorname{int} A_Y \neq \emptyset$ *and* G *is Fréchet differentiable in* $x_0 \in A_X$, *then we have*

$$TC(A; x_0) \cap G_{x_0}^{\prime -1}\big(\operatorname{int}(-A_Y) - G(x_0)\big) \subset TC(A_X; x_0).$$

Proof Since $TC(A_X; x_0)$ includes the origin, we can reduce our consideration to the case in which $x \in TC(A; x_0) \cap G_{x_0}^{\prime -1}(\operatorname{int}(-A_Y) - G(x_0))$, with $x \neq 0$. According to Lemma 3.25, we have

$$\lim_{n\to\infty} \lambda_n\big[G(x_n) - G(x_0)\big] = G'_{x_0}(x),$$

where $\{x_n\} \subset A$ and $\{\lambda_n\} \subset \mathbb{R}_+$ have properties (3.40).

But $-G'_{x_0}(x) \in \operatorname{int} A_Y + G(x_0)$ and $\operatorname{int} A_Y + G(x_0)$ is an open set. Thus, there exists $n_0 \in \mathbb{N}$ such that

$$\lambda_n\big[G(x_n) - G(x_0)\big] \in -\operatorname{int} A_Y - G(x_0), \quad \forall n > n_0.$$

Therefore,

$$\lambda_n G(x_n) \in (\lambda_n - 1)G(x_0) - A_Y, \quad \forall n > n_0.$$

However, we can assume that $\lambda_n > 1$ because, necessarily, $\lambda_n \to \infty$ for $n \to \infty$ (otherwise, from property (3.40), $x = 0$ results). Since A_Y is cone and $G(x_0) \in -A_Y$, we obtain $G(x_n) \in -A_Y$, that is, $x_n \in A_X$, $\forall n > n_0$. Hence, $x \in TC(A_X; x_0)$ by virtue of Lemma 3.22 and Definition (3.48) of A_X. □

Theorem 3.30 *If* $\operatorname{int} A_Y \neq \emptyset$, x_0 *is a solution of Problem* $\mathscr{P}_2$ *and* f *and* G *are Fréchet differentiable in* x_0, *then there exist a real number* η_0 *and an element* $y_0^* \in Y^*$, *not both equal to zero, such that*

$$\eta_0 f'_{x_0}(x) + \big(y_0^*, G'_{x_0}(x)\big) \geq 0, \quad \forall x \in K, \tag{3.49}$$

$$\eta_0 \geq 0, \qquad (y_0^*, y) \geq 0, \quad \forall y \in A_Y,\ \big(y_0^*, G(x_0)\big) = 0, \tag{3.50}$$

where $K \subset TC(A; x_0)$ *is an arbitrary convex cone with vertex at origin.*

Proof Let x_0 be an optimal solution of $\mathscr{P}_2$. By virtue of Theorem 3.28, it follows that f'_{x_0} takes its minimum value on $TC(A_X; x_0)$ at the origin. According to Lemma 3.29, we obtain so much the more that the origin is an optimal solution of the problem

$$\min\{f'_{x_0}(x);\ x \in TC(A; x_0), G'_{x_0}(x) + G(x_0) \in -\overline{\operatorname{int} A_Y}\}.$$

Since f'_{x_0} is linear and continuous, we obtain the result that the origin is also an optimal solution for the following problem with operatorial convex constraints:

$$\min\{f'_{x_0}(x);\ x \in K,\ G'_{x_0}(x) + G(x_0) \in -A_Y\}. \tag{3.51}$$

Taking into account the results of the preceding section, we see that if $\operatorname{int} A_Y \neq \emptyset$, there exists a Lagrange multiplier (η_0, y_0^*), that is, (3.20) and (3.21) hold

$$\eta_0 f'_{x_0}(x) + \big(y_0^*, G'_{x_0}(x) + G(x_0)\big) \geq 0, \quad \forall x \in K,$$
$$\eta_0 \geq 0, \qquad (y_0^*, y) \geq 0, \quad \forall y \in A_Y.$$

(Here, the role of A and $G(x)$ is played by K and $G'_{x_0}(x) + G(x_0)$, respectively.) Now, the proof is finished if we observe that the complementarity slackness condition, $(y_0^*, G(x_0)) = 0$, is also verified. Indeed, for $x = 0$ we obtain $(y_0^*, G(x_0)) \geq 0$. But $G(x_0) \in -A_Y$ and $y_0^* \in -A_Y^0$ yield $(y_0^*, G(x_0)) \leq 0$, that is, $(y_0^*, G(x_0)) = 0$. □

Remark 3.31 In view of the definition of the polar of a cone, we observe that relations (3.49) and (3.50) may be, equivalently, written as

$$\eta_0 f'_{x_0} + y_0^* \circ G'_{x_0} \in -K^0, \tag{3.52}$$

$$\eta_0 \geq 0, \qquad y_0 \in -A_Y^0, \quad \big(y_0^*, G(x_0)\big) = 0, \tag{3.53}$$

respectively.

Let us now consider the problem

$$(\mathscr{P}_2) \quad \min\{f(x);\ G(x) \in A_Y\},$$

which may be obtained from $(\mathscr{P}_2)$ by taking $A = X$.

Since $TC(X; x_0) = X$ and $X^0 = \{0\}$, we can take $K = X$ and so, in this case, Condition (3.49) becomes

$$\eta_0 f'_{x_0}(x) + y_0^*\big(G'_{x_0}(x)\big) = 0, \quad \forall x \in X,$$

that is,

$$\eta_0 f'_{x_0} + y_0^* \circ G'_{x_0} = 0.$$

Remark 3.32 If $A_Y = \{0\}$, relation (3.49) is again satisfied if the range of G'_{x_0} is closed. Moreover, if G'_{x_0} is surjective or if $f'_{x_0} \in \operatorname{Range} G'_{x_0}$, then both η_0 and y_0^* may be chosen to be nonzero (see Norris [81]). The affine case is considered in the next Theorem 3.35.

As we easily see from the proof of the above theorem, the interiority condition, $\operatorname{int} A_Y \neq \emptyset$, is essential only to use Lemma 3.25, which ensures us that the origin is an optimal solution for problem (3.51). However, it can also be shown that, in the case of closed affine sets, the results continue to remain valid even if the interiority condition is violated.

Certainly, a remarkable case is that for which $\eta_0 \neq 0$, for example, if there exists $\bar{x} \in -K$ such that $G'_{x_0}(\bar{x}) \in G(x_0) + \operatorname{int} A_Y$ (according to Theorem 3.15 and Remark 3.16, via (3.51)). In such a case, we can suppose, without loss of generality, that $\eta_0 = 1$. In what follows, we point out a situation in which this fact is possible.

We set

$$K = \bigl\{x \in X;\ G'_{x_0}(x) \in PC\bigl(-A_Y; G(x_0)\bigr)\bigr\}, \tag{3.54}$$

$$\begin{aligned} H &= \bigl\{x^* \in X^*;\ x^* = y^* \circ G'_{x_0},\ y^* \in PC\bigl(-A_Y; G(x_0)\bigr)^0\bigr\} \\ &= G'^*_{x_0}\bigl(PC\bigl(-A_Y; G(x_0)\bigr)^0\bigr), \end{aligned} \tag{3.55}$$

and observe that both sets are convex cones with the vertex at the origin. Moreover, K is closed.

Theorem 3.33 *Let f and G be two Fréchet differentiable functions in $x_0 \in A_X$ and let H be w^*-closed. If there exists a closed convex cone $K_1 \subset X$ subject to $K \cap K_1 \subset PC(A_X; x_0)$ and $K^0 + K_1^0$ is w^*-closed in X^*, a necessary condition for x_0 to be an optimal solution to Problem $\mathscr{P}_2$ is the existence of an element $y_0^* \in Y^*$ which satisfies the properties*

$$(y_0^*, y) \le 0, \quad \forall y \in PC\bigl(-A_Y; G(x_0)\bigr), \tag{3.56}$$

$$f'_{x_0}(x) + \bigl(y_0^*, G'_{x_0}(x)\bigr) \ge 0, \quad \forall x \in K_1. \tag{3.57}$$

Furthermore, if A_X is pseudoconvex in x_0, f is pseudoconvex on A_X in x_0 and $A_X \subset x_0 + K_1$, then the above conditions are also sufficient for optimality in Problem $\mathscr{P}_2$.

Proof Let x_0 be an optimal solution of $\mathscr{P}_2$. From Theorem 3.28, it follows that $f'_{x_0} \in -PC(A_X; x_0)^0$.

We easily observe that, because K and K_1 are cones, from the definition of the polar, we have $(K \cap K_1)^0 = K^0 + K_1^0 \supset PC(A_X; x_0)^0$ (because $K^0 + K_1^0$ is w^*-closed). But, by hypothesis, according to Theorem 3.28, we have $-f'_{x_0} \in K^0 + K_1^0$, which says that there exists $x_0^* \in K^0$ such that $f'_{x_0} + x_0^* \in -K_1^0$. Now, let us show that $K^0 \subset H$, or equivalently (by virtue of the bipolar theorem, see Theorem 2.26) that $H^0 \subset K$. Indeed, if $x \in H^0$, from (3.55) and the definition of the polar, it follows that

$$(y^* \circ G'_{x_0}, x) \leq 0, \quad \forall y^* \in PC\big(-A_Y; G(x_0)\big)^0,$$

that is,

$$G'_{x_0}(x) \in PC\big(-A_Y; G(x_0)\big)^{00}.$$

Since $PC(-AQ_Y; G(x_0))$ is a closed convex set and inasmuch as it contains the origin, from the same bipolar theorem we have $PC(-A_Y; G(x_0))^{00} = PC(-A_Y; G(x_0))$. Hence $G'_{x_0}(x) \in PC(-A_Y; G(x_0))$ which implies that $x \in K$ by virtue of relation (3.54). Therefore, $K^0 \subset H$, that is, $x_0^* \in H$. On the other hand, from relation (3.55) we have the result that there exists $y_0^* \in PC(-A_Y; G(x_0))^0$ such that $x_0^* = y_0^* \circ G'_{x_0}$. Consequently, $f'_{x_0} + y^*_{x_0} \circ G'_{x_0} \in -K_1^0$ with $y_0^* \in PC(-A_Y; G(x_0))^0$. In this way, necessity of conditions (3.56) and (3.57) is completely proved.

Let us now prove sufficiency. Since $x - x_0 \in K_1$, for all $x \in A_X$, from inequality (3.56) we obtain

$$f'_{x_0}(x - x_0) + \big(y_0^*, G'_{x_0}(x - x_0)\big) \geq 0, \quad \text{for all } x \in A_X. \tag{3.58}$$

Because A_X is pseudoconvex in x_0, we have $x - x_0 \in PC(A_X; x_0)$, for all $x \in A_X$. By virtue of Remark 3.26 and noting that $G(A_X) \subset -A_Y$, we obtain

$$G'_{x_0}\big(PC(A_X; x_0)\big) \subset PC\big(-A_Y; G(x_0)\big).$$

Thus, making use of the pseudoconvexity of the set A_X, we obtain $G'_{x_0}(x - x_0) \in PC(-A_Y; G(x_0))$, $\forall x \in A_X$. From inequality (3.56) we have the result $(y^*, G'_{x_0}(x - x_0)) \leq 0$, $\forall x \in A_X$ which, by virtue of relation (3.58), implies $f'_{x_0}(x - x_0) \geq 0$, $\forall x \in A_X$. Since f is pseudoconvex on A_X in x_0, the latter yields $f(x_0) \leq f(x)$, for all $x \in A_X$, that is, $x_0 \in A_X$ is an optimal solution of Problem $\mathscr{P}_2$.

Now, we divert our attention to the important case $K = PC(A_X; x_0)$, that is, $G'^{-1}_{x_0}(PC(-A_Y; G(x_0))) = PC(A_X; x_0)$. Observe that $K_1 = PC(A_X; x_0)$ satisfies the required conditions by hypothesis. This implies that, if H is w^*-closed, inequality (3.57) holds on $PC(A_X; x_0)$.

On the other hand, it is natural to ask what the connections are between minimizing on A_X and minimizing on $x_0 + PC(A_X; x_0)$. A partial answer is given by Theorem 3.34 below. □

Theorem 3.34 *If x_0 minimizes a convex continuous function f on a subset A, then x_0 also minimizes f on $x_0 + PC(A; x_0)$.*

Proof Because of the convexity and the continuity of f it is sufficient to show that x_0 minimizes f on $x_0 + TC(A; x_0)$. If $TC(A; x_0) = \{0\}$, the assertion is trivial. Thus, let $TC(A; x_0) \neq \{0\}$. Assume that an element $x \in x_0 + TC(A; x_0)$ exists such that $f(x) < f(x_0)$. From Lemma 3.22 there exists $\{x_n\} \subset A$ with $x_n \to x_0$ and $\{\lambda_n\} \subset$

$\mathbb{R}_+$ with $\lambda_n(x_n - x_0) \to x - x_0$. From continuity we have the result that there exists $n_0 \in \mathbb{N}$, such that $f(x_0 + \lambda_n(x_n - x_0)) < f(x_0)$, $\forall n > n_0$. We can assume $\lambda_n > 1$, $\forall n > n_0$, since $\lambda_n \to \infty$. From the convexity, we obtain

$$\begin{aligned} f(x_n) &= f\left[\frac{1}{\lambda_n}\big(\lambda_n x_n + (1-\lambda_n)x_0\big) + \left(1 - \frac{1}{\lambda_n}\right)x_0\right] \\ &\le \frac{1}{\lambda_n} f\big(\lambda_n x_n + (1-\lambda_n)x_0\big) + \left(1 - \frac{1}{\lambda_n}\right) f(x_0) < f(x_0), \end{aligned}$$

which is impossible because $x_n \in A$. □

We note that Theorem 3.33 has several virtues in comparison to Theorem 3.30. The first and most important consists of the fact that we can disregard the regularity condition $\operatorname{int} A_Y \neq \emptyset$. In particular, this allows us to use the *Kuhn–Tucker optimality conditions* for constraints given by equalities.

For example, if $T \in L(X, Y)$, $k \in Y$, then the problem with affine constraints

$$(\mathscr{P}_4) \quad \min\big\{f(x);\ T(x) = k\big\}$$

may be obtained from $\mathscr{P}_2$ for $A = X$, $A_Y = \{0\}$ and $G(x) = Tx - k$, $\forall x \in X$.

The Fréchet differentiability condition is satisfied and $G'_{x_0} = T$. We also observe that $PC(-A_Y; G(x_0)) = \{0\}$; hence, $K = \ker T$, $H = \operatorname{Range}(T^*)$ and $PC(A_X; x_0) = \ker T$. Therefore, the hypotheses of Theorem 3.33 are satisfied for $K_1 = X$ noting that $\operatorname{Range} T^*$ is w^*-closed if $\operatorname{Range} T$ is closed.

Theorem 3.35 *Let f be a Fréchet differentiable function in x_0 and T with the closed range. If x_0 is an optimal solution of Problem $\mathscr{P}_4$, then there exists $y_0^* \in Y^*$ such that*

$$f'_{x_0}(x) + (y_0^*, Tx) = 0, \quad \forall x \in X. \tag{3.59}$$

If f is pseudoconvex on $T^{-1}(k)$ in $x_0 \in T^{-1}(k)$ and there exists $y_0^* \in Y^*$ subject to condition (3.59), then x_0 is an optimal solution for $\mathscr{P}_4$.

Now, let us consider the *Lagrange function* associated to Problem $\mathscr{P}_2$, $L : A \times (-A_Y^0) \to \overline{\mathbb{R}}$, defined by

$$L(x, y^*) = f(x) + \big(y^*, G(x)\big), \quad \forall [x, y^*] \in A \times \big(-A_Y^0\big). \tag{3.60}$$

We establish the relationship between the solutions of Problem $\mathscr{P}_2$ and the existence of the saddle points of L with respect to the minimization on A and the maximization on $-A_Y^0$, that is, the problem of the existence of a pair $(x_0, y_0^*) \in A \times (-A_Y^0)$ such that

$$L(x_0, y^*) \le L(x_0, y^*) \le L(x, y_0^*), \quad \forall (x, y^*) \in A \times \big(-A_Y^0\big). \tag{3.61}$$

Theorem 3.36 *Let* (x_0, y^*) *be a saddle point of* L *on* $A \times (-A_Y^0)$. *Then,* x_0 *is an optimal solution of* $\mathscr{P}_2$. *Moreover, if* f *and* G *are Fréchet differentiable in* x_0, *then conditions* (3.49) *and* (3.50) *are satisfied with* $\eta_0 = 1$.

Proof If (x_0, y_0^*) is a saddle point of L on $A \times (-A_Y^0)$, we have

$$f(x_0) + \big(y^*, G(x_0)\big) \le f(x_0) + \big(y_0^*, G(x_0)\big) \le f(x) + \big(y_0^*, G(x)\big),$$

for every $x \in A$ and $y^* \in -A_Y^0$. But $y^* + y_0^* - A_Y^0$ for all $y^* \in -A_Y^0$ because A_Y^0 is a cone. Replacing y^* by $y^* + y_0^*$ in the left-hand side of this last inequality, we obtain $(y^*, G(x_0)) \le 0, \forall y^* \in -A_Y^0$, and hence $G(x_0) \in (-A_Y^0)^0 = -A_Y$, that is, $x_0 \in A_X$. In particular, $(y_0^*, G(x_0)) \le 0$. Also, the converse inequality is valid (taking $y^* = 0$) and so, we obtain $(y_0^*, G(x_0)) = 0$. Since $x_0 \in A_X$ and since, from the relation on the right-hand side of the inequality mentioned above, we have $f(x_0) \le f(x)$, we find that x_0 is an optimal solution of $\mathscr{P}_2$. Moreover, relations (3.50) hold. Relation (3.49) can be obtained from the right-hand side of the same inequality using the Fréchet differentiability definition. □

Remark 3.37 When A is convex and f is convex (consequently, L becomes convex-concave), conditions (3.49) and (3.50), satisfied with $\eta_0 = 1$, are sufficient in order that (x_0, y_0^*) be a saddle point of L; in particular, x_0 is an optimal solution to Problem $\mathscr{P}_2$. Finally, we note that condition (3.50) may be written as (3.56) because $(y_0^*, y) \le 0$, for all $y \in -A_Y$, implies $(y_0^*, y) \le 0, \forall u \in -A_Y - G(x_0)$ (we recall that $(y_0^*, G(x_0)) = 0$). Since y_0^* is a linear continuous functional, we obtain $(y_0^*, u) \le 0$, for all $u \in PC(-A_Y; G(x_0))$, which implies the desired relation (3.56).

A refinement of the results of the Kuhn–Tucker type for the non-convex case can be obtained using the concept of *tangent cone*, *in Clarke's sense*, given by

$$\widetilde{TC}(A; x_0) = \{y;\ \forall \lambda_n \nearrow \infty \text{ and } \{x_n\} \subset A,\ x_n \to x_0 \text{ there exists } \{y_n\} \subset A \text{ such that } \lambda_n(y_n - x_n) \to y\}.$$

This tangent cone is always closed, convex and it contains the origin. We also have $\widetilde{TC}(A; x_0) \subset TC(A; x_0)$. A special role is played by the so-called *tangentially regular points* x_0 for which we have $\widetilde{TC}(A; x_0) = TC(A; x_0)$.

3.2 Duality in Convex Programming

Roughly speaking, the duality method reduces the infimum problem $\inf\{f + g\}$ to a similar problem formulated in terms of conjugate functions f^* and g^*. In this section, we present the basic results of this theory.

3.2.1 Dual Convex Minimization Problems

Consider the equation

$$\partial\varphi(x^*) + \partial\psi(x^*) \ni 0,$$

where φ and ψ are lower semicontinuous convex functions. Clearly, the above equation can be rewritten as

$$\partial\varphi^*(y^*) - \partial\psi^*(-y^*) \ni 0.$$

As seen earlier, if $\operatorname{int} D(\varphi) \cap D(\psi)$, x is a solution to the minimization problem

$$(\mathrm{P}_0) \quad \operatorname{Min}\{\varphi(x) + \psi(x)\},$$

while $y^* \in \partial\psi(x^*)$ is a solution to

$$(\mathrm{P}_0^*) \quad \operatorname{Min}\{\varphi^*(y) + \psi^*(-y)\},$$

where φ^* and ψ^* are conjugate of φ and ψ, respectively. We have obtained, therefore, a close relationship between P_0 and P_0^*, which is called dual of P_0.

In the sequel, starting from this simple example, we present a general way to define the dual of a given problem, which relies on the conjugate duality of functions.

Let X, Y be real Banach spaces and X^*, Y^*, respectively, their duals or, more generally, two dual systems. In both cases we denote the duality functional by $(\cdot, \cdot)$, understanding in each case that we consider the duality (X, X^*) or (Y, Y^*). Suppose that the spaces are equipped with compatible topologies with respect to the dual systems.

Naturally, we obtain a duality between $X \times Y$ and $X^* \times Y^*$ given by

$$\big((x, y), (x^*, y^*)\big) = (x, x^*) + (y, y^*), \quad \forall (x, y) \in X \times Y,\ (x^*, y^*) \times X^* \times Y^*. \tag{3.62}$$

Let $F : X \times Y \to \overline{\mathbb{R}}$ be a function subject to

$$F(x, 0) = f(x), \quad \forall x \in X, \tag{3.63}$$

where $f : X \to \overline{\mathbb{R}}$ is a given function.

We consider the minimization problem

$$(\mathscr{P}) \quad \min\{f(x);\ x \in X\}.$$

Definition 3.38 The maximalization problem

$$(\mathscr{P}^*) \quad \max\{-F^*(0, y^*);\ y^* \in Y^*\},$$

where F^* is the conjugate function of F with respect to the duality given by (3.62), is called the *dual problem* of $\mathscr{P}$ with respect to the family of perturbations generated by F.

Thus, we recall that, by virtue of the definition of the conjugate function (see Chap. 2, Sect. 2.1.4), we have

$$F^*(x^*, y^*) = \sup\big\{(x, x^*) + (y, y^*) - F(x, y);\ (x, y) \in X \times Y\big\}, \tag{3.64}$$

and hence

$$F^*(0, y^*) = \sup\big\{(y, y^*) - F(x, y);\ (x, y) \in X \times Y\big\}. \tag{3.65}$$

Similarly, we can define the *bidual problem* of $\mathscr{P}$ as

$$(\mathscr{P}^{**}) \quad \min\big\{F^{**}(x, 0);\ x \in X\big\}.$$

Since $F^{***} = F^*$, the duals of higher order of $\mathscr{P}$ identify either with $\mathscr{P}^*$ or with $\mathscr{P}^{**}$. If $\mathscr{P}^{**}$ identifies with $\mathscr{P}$, that is $F^{**}(x, 0) = F(x, 0)$, $\forall x \in X$ (for instance, if F is a proper, lower-semicontinuous convex function on $X \times Y$), then we have a complete duality between $\mathscr{P}$ and $\mathscr{P}^*$, because they are dual to each other.

A first remarkable result concerns the relationship existing between the values of the two problems.

Proposition 3.39

$$-\infty \leq \sup \mathscr{P}^* \leq \inf \mathscr{P} \leq +\infty. \tag{3.66}$$

Proof From relation (3.65), by virtue of (3.63), we obtain

$$F^*(0, y^*) \geq (0, y^*) - F(x, 0) = -F(x, 0) = -f(x)$$

for all $x \in X$ and $y^* \in Y^*$, which, obviously, implies relation (3.66). □

Definition 3.40 Problem $\mathscr{P}$ is called *normal* if

$$-\infty < \inf \mathscr{P} = \sup \mathscr{P}^* < +\infty. \tag{3.67}$$

Consider, for every $y \in Y$, the minimization problem

$$(\mathscr{P}_y) \quad \min\big\{F(x, y);\ x \in X\big\},$$

called the *perturbed problem* of $\mathscr{P}$.

It is clear that $\mathscr{P}_0 = \mathscr{P}$. Hence, according to condition (3.63), the function F can be considered as a source of perturbations for Problem $\mathscr{P}$.

The function $h : Y \to \overline{\mathbb{R}}$, defined by

$$h(y) = \inf \mathscr{P}_y = \inf\big\{F(x, y);\ x \in X\big\}, \tag{3.68}$$

is called the *value function* of the family $\{\mathscr{P}_y;\ y \in Y\}$.

According to condition (3.63), we have

$$h(0) = \inf \mathscr{P}. \tag{3.69}$$

Lemma 3.41 *If F is convex on $X \times Y$, then its value function h is convex on Y.*

Proof If $y_1, y_2 \in \mathrm{Dom}(h)$, from relation (3.68) we have the result that for any $\varepsilon > 0$ there exist $x_1, x_2 \in X$ such that

$$h(y_i) \leq F(x_i, y_i) \leq h(y_i) + \varepsilon, \quad \forall i = 1, 2.$$

Thus, we have

$$\begin{aligned} h(\lambda_1 y_1 + \lambda_2 y_2) &= \inf\big\{F(x, \lambda_1 y_1 + \lambda_2 y_2);\ x \in X\big\} \\ &\leq F(\lambda_1 x_1 + \lambda_2 x_2, \lambda_1 y_1 + \lambda_2 x_2) \leq \lambda_1 F(x_1, y_1) + \lambda_2 F(x_2, y_2) \\ &\leq \lambda_1 h(y_1) + \lambda_2 h(y_2) + \varepsilon, \end{aligned}$$

for all $\lambda_1, \lambda_2 \geq 0$ with $\lambda_1 + \lambda_2 = 1$ and $\varepsilon > 0$. Since $\varepsilon > 0$ is arbitrary, this implies the convexity of the function h.

We easily see that

$$h^*(y^*) = F^*(0, y^*), \quad \forall y^* \in Y^*. \tag{3.70}$$

Indeed, we have

$$\begin{aligned} h^*(y^*) &= \sup_{y \in Y}\big\{(y, y^*) - h(y)\big\} = \sup_{y \in Y}\big\{(y, y^*) - \inf_{x \in X} F(x, y)\big\} \\ &= \sup_{(x,y) \in X \times Y}\big\{(y, y^*) - F(x, y)\big\} = F^*(0, y^*). \end{aligned}$$

From relation (3.70) it follows, in particular, that $\mathscr{P}^*$ is straightforwardly related to h; more precisely, the following relation holds:

$$\sup \mathscr{P}^* = h^{**}(0), \tag{3.71}$$

because from the definition of the conjugate of a function and from relation (3.70) we have

$$\sup \mathscr{P}^* = \sup\big\{-F^*(0, y^*);\ y^* \in Y^*\big\} = \sup\big\{(0, y^*) - h^*(y^*);\ y^* \in Y^*\big\} = h^{**}(0).$$

$\square$

Remark 3.42 Since $\inf \mathscr{P} = h(0)$, we observe that inequality (3.66) actually reduces to the obvious inequality $h^{**}(0) \leq h(0)$. This fact allows one to find several examples to demonstrate that in inequality (3.66) all the cases could occur.

Theorem 3.43 *Problem $\mathscr{P}$ is normal if and only if $h(0)$ is finite and h is lower-semicontinuous at the origin.*

Proof Since h is a proper convex function we may infer that $\mathrm{cl}\, h = \liminf h$. On the other hand, from Corollary 2.23, in Chap. 2, we have $h^{**} = \mathrm{cl}\, h$. Therefore, $\mathscr{P}$ is normal, that is, we have $h^{**}(0) = h(0) \in \mathbb{R}$, if and only if $h(0)$ is finite and $h(0) = \liminf_{y\to 0} h(y)$ (here, we have used relations (3.63) and (3.71)). □

Now, let us study the relationship between the normality of the primal problem $\mathscr{P}$ and the normality of the dual problem $\mathscr{P}^*$. Using Definition 3.40, we see that $\mathscr{P}^*$ is normal if and only if

$$\sup \mathscr{P}^* = \inf \mathscr{P}^{**}, \tag{3.72}$$

where the common value is finite.

We have already seen that if F is a lower-semicontinuous proper convex function on $X \times Y$, then $\mathscr{P}^{**}$ coincides with $\mathscr{P}$ because under these conditions $F^{**} = F$.

We summarize this in the next proposition.

Proposition 3.44 *If F is a proper, lower-semicontinuous convex function on $X \times Y$, then $\mathscr{P}$ is normal if and only if $\mathscr{P}^*$ is normal.*

Since $\mathscr{P}^*$ represents the maximization of an upper-semicontinuous, concave function, it is natural to expect that the properties of the dual problem are more intricate than those of this primal problem. In fact, one has the following proposition.

Proposition 3.45 *The set of solutions to the dual problem $\mathscr{P}^*$ coincides with $\partial h^{**}(0)$.*

Proof The element $y_0^* \in Y^*$ is a solution of $\mathscr{P}^*$ if and only if

$$-F^*(0, y_0^*) \geq -F(0, y^*), \quad \forall y^* \in Y^*.$$

According to relation (3.70), we obtain

$$h^*(y_0^*) \leq h^*(y^*), \quad \forall y^* \in Y^*,$$

which shows that y_0^* is a minimum point of h^* on Y^* or, equivalently, $0 \in \partial h^*(y_0^*)$. Because h^* is convex and lower-semicontinuous, using Proposition 2.2 in Chap. 2, we may express this condition as $y_0^* \in \partial h^{**}(0)$. □

Definition 3.46 Problem $\mathscr{P}$ is said to be *stable* if it is normal and $\mathscr{P}^*$ has at least one solution.

Remark 3.47 We easily see that if F is a convex function, so an element $(x_0, y_0^*) \in X \times Y^*$ constitutes a pair of solutions to $\mathscr{P}$ and $\mathscr{P}^*$ which satisfy the normality condition (3.67), if and only if the following relation holds:

$$F(x_0, 0) + F^*(0, y_0^*) = 0. \tag{3.73}$$

Indeed, $F(x_0, 0) = \inf \mathscr{P}$ and $-F^*(0, y_0^*) = \sup \mathscr{P}^*$ if and only if (x_0, y_0^*) is a pair of solutions. Since $((x_0, 0), (0, y_0^*)) = 0$, it follows from a characteristic property of the subdifferential (see Proposition 2.2, Chap. 2) that relation (3.73) is equivalent to

$$(0, y_0^*) \in \partial F(x_0, 0). \tag{3.74}$$

Moreover, if F is a lower-semicontinuous function, this relation is also equivalent to

$$(x_0, 0) \in \partial F^*(0, y_0^*). \tag{3.75}$$

In the following text, the following condition is required.

(A) *F is a proper, convex lower-semicontinuous function on $X \times Y$.*

As we have already seen, this hypothesis ensures the coincidence of Problems $\mathscr{P}$ and $\mathscr{P}^{**}$; hence, Problems $\mathscr{P}$ and $\mathscr{P}^*$ are dual to each other.

Theorem 3.48 *If the function F satisfies Hypothesis* (A), *then the stability of Problem $\mathscr{P}$ is equivalent to the subdifferentiability at the origin of the function h, that is, $\partial h(0) \neq \emptyset$.*

Proof In view of the above proposition, Problem $\mathscr{P}^*$ has solutions if and only if $\partial h^{**}(0) \neq \emptyset$. By virtue of Theorem 3.43 and Definition 3.46 it remains to be proven that, if h is lower-semicontinuous at the origin, that is, $h(0) = h^{**}(0)$, then $\partial h(0) = \partial h^{**}(0)$.

Indeed, it is well known that $y^* \in \partial h(0)$ if and only if $h(0) + h^*(y^*) = (0, y^*) = 0$ or, equivalently, $h^{**}(0) + h^{***}(y) = (0, y^*)$, that is, $y^* \in \partial h^{**}(0)$. $\square$

Now, let us attach to Problem $\mathscr{P}$ (with respect to the perturbation function F) the *Hamiltonian* $H : X \times Y^* \to \overline{\mathbb{R}}$ defined by

$$H(x, y^*) = \sup\big\{(y, y^*) - F(x, y);\ \ y \in Y\big\}. \tag{3.76}$$

We observe that, for each $x \in X$, the function $H(x, \cdot)$ is the convex conjugate of $F(x, \cdot)$. Thus, the function $y^* \to H(x, y^*)$, $y^* \in Y^*$, is convex and lower-semicontinuous on Y^*. On the other hand, the function $x \to H(x, y^*)$, $x \in X$, is concave and closed on X for every $y^* \in Y^*$.

In the following, we show that, under Hypothesis (A), Problems $\mathscr{P}$ and $\mathscr{P}^*$ arise as dual problems in minimax form generated by the Hamiltonian H.

We recall that the pair $(x_0, y_0^*) \in X \times Y^*$ is a saddle point for the concave-convex Hamiltonian function H if and only if

$$H(x, y_0^*) \le H(x_0, y_0^*) \le H(x_0, y^*), \quad \forall (x, y^*) \in X \times Y^*. \tag{3.77}$$

(See Sect. 2.3.1, Chap. 2.)

Theorem 3.49 *If F satisfies Condition* (A), *then the following statements are equivalent*:

(i) (x_0, y_0^*) *is a saddle point of H on $X \times Y^*$.*
(ii) x_0 *is an optimal solution of* $\mathscr{P}$, y_0^* *is an optimal solution of* $\mathscr{P}^*$ *and their values are equal.*

Proof We have

$$\begin{aligned}\sup \mathscr{P}^* &= \sup_{y^* \in Y^*} \left\{-F^*(0, y^*)\right\} \\ &= -\inf_{y^* \in Y^*} \sup_{x \in X} \sup_{y \in Y} \left\{(y, y^*) - F(x, y)\right\} \\ &= -\inf_{y^* \in Y^*} \sup_{x \in X} H(x, y^*).\end{aligned} \tag{3.78}$$

Since F is convex and lower-semicontinuous on $X \times Y$, it follows that the function $F_x(\cdot) = F(x, \cdot)$ is also convex and lower-semicontinuous on Y for each $x \in X$. According to the bipolar theorem (see Theorem 2.26, Chap. 2), we obtain

$$\begin{aligned}F(x, 0) = F_x^{**}(0) &= \sup_{y^* \in Y^*} \left\{(0, y^*) - F_x^*(y^*)\right\} \\ &= -\inf_{y^* \in Y^*} \sup_{y \in Y} \left\{(y, y^*) - F(x, y)\right\} \\ &= -\inf_{y^* \in Y^*} H(x, y^*).\end{aligned} \tag{3.79}$$

Hence

$$\inf \mathscr{P} = \inf_{x \in X} F(x, 0) = -\sup_{x \in X} \inf_{y^* \in Y^*} H(x, y^*). \tag{3.80}$$

□

Now, we conclude the proof as a direct consequence of relations (3.78) and (3.70).

Corollary 3.50 *If* $\mathscr{P}$ *is stable, then* $x_0 \in X$ *is a solution of* $\mathscr{P}$ *if and only if there exists* $y_0^* \in Y^*$ *such that* (x_0, y_0^*) *is a saddle point of the Hamiltonian.*

Corollary 3.51 *The Hamiltonian H has at least one saddle point if and only if both* $\mathscr{P}$ *and* $\mathscr{P}^*$ *are stable.*

3.2.2 Fenchel Duality Theorem

Consider now the special case when the perturbations are generated by translations Let the *primal problem* be defined by

$$(\mathscr{P}_1) \quad \min\big\{f(x) - g(Ax);\ x \in X\big\},$$

where X, Y are real Banach spaces, $f : X \to]-\infty, +\infty]$ is proper, convex and lower-semicontinuous function, $g : Y \to [-\infty, +\infty[$ is a proper, concave and upper-semicontinuous function and $A : X \to Y$ is a linear continuous operator.

As a perturbation function $F : X \times Y \to \overline{\mathbb{R}}$, we take

$$F(x, y) = f(x) - g(Ax - y). \tag{3.81}$$

In this way, it is clear that we can apply the duality results just presented in the preceding section.

First, we determine the conjugate of F. We have

$$\begin{aligned} F^*(x^*, y^*) &= \sup_{(x,y)\in X\times Y} \big\{(x, x^*) + (y, y^*) - f(x) + g(Ax - y)\big\} \\ &= \sup_{x\in X} \sup_{z\in Y} \big\{(x, x^*) + (Ax, y^*) - f(x) + g(z) - (z, y^*)\big\} \\ &= \sup_{x\in X} \big\{(x, x^*) + (x, A^*y^*) - f(x)\big\} - \inf_{z\in Y} \big\{(z, y^*) - g(z)\big\} \\ &= f^*(A^*y^* + x^*) - g^*(y^*), \end{aligned}$$

where f^* is the convex conjugate of f, while g^* is the concave conjugate of g, and A^* is the adjoint of A.

Therefore,

$$F^*(0, y^*) = f^*(A^*y^*) - g^*(y^*), \quad \forall y^* \in Y^*. \tag{3.82}$$

Hence, the *dual problem* is given by

$$(\mathscr{P}_1^*) \quad \max\big\{g^*(y^*) - f^*(A^*y^*);\ y^* \in Y^*\big\}.$$

We note that $\mathscr{P}_1$ is *consistent* if at least one element $x \in X$ exists, such that $f(x) < \infty$ and $g(Ax) > -\infty$, that is,

$$A\big(\mathrm{Dom}(f)\big) \cap \mathrm{Dom}(g) \neq \emptyset. \tag{3.83}$$

Similarly, $\mathscr{P}_1^*$ is consistent if

$$A^*\big(\mathrm{Dom}(g^*)\big) \cap \mathrm{Dom}(f^*) \neq \emptyset. \tag{3.84}$$

From Proposition 3.39 it follows that, if $\mathscr{P}_1$ and $\mathscr{P}_1^*$ are consistent, then their values are both finite.

Using the convexity and semicontinuity of f and g, we obtain a complete duality between $\mathscr{P}_1$ and $\mathscr{P}_1^*$ since f, f^* and g, g^*, respectively, are mutually conjugate. Moreover, the dual problem $\mathscr{P}_1^*$ is equivalent to a minimization problem of type $\mathscr{P}_1$. Indeed, $\mathscr{P}_1^*$ can be rewritten as

$$\min\{f_1(y^*) - g_1(A_1 y^*);\ y^* \in Y^*\},$$

where $f_1(y^*) = -g^*(y^*)$, $g_1(x^*) = -f^*(x^*)$ and $A_1 = A^*$ (changing X by Y^* and Y by X^*). Therefore, the results established for $\mathscr{P}_1$ can be transposed by the above change to the dual problem $\mathscr{P}_1^*$. In our case, Condition (A) for F is superfluous because $\mathscr{P}_1$ and $\mathscr{P}_1^*$ are mutual duals, that is, $\mathscr{P}_1^{**} = \mathscr{P}_1$.

As we have seen earlier (Theorems 3.43 and 3.48), the properties of Problems $\mathscr{P}_1$ and $\mathscr{P}_1^*$ depend on the properties of the convex function $h : Y \to \overline{\mathbb{R}}$ given by

$$h(y) = \inf\{f(x) - g(Ax - y);\ x \in X\}, \quad \forall y \in Y. \tag{3.85}$$

For instance, if $\mathscr{P}_1$ is consistent and h is lower-semicontinuous at the origin, then $\mathscr{P}_1$ is normal or, equivalently, $\mathscr{P}_1^*$ is normal (see Proposition 3.44).

Lemma 3.52 *If there exists $x_0 \in X$, such that $f(x_0) < +\infty$ and g is continuous at Ax_0, then h is continuous in a neighborhood of the origin.*

Proof Since h is convex (Lemma 3.41), it suffices to prove that h is upper-bounded on a certain neighborhood of the origin (Theorem 2.14, Chap. 2). Applying the continuity, we have the result that the concave function g is bounded from below on a neighborhood of Ax_0. Hence, an open neighborhood V_0 of the origin exists in Y such that $g(Ax_0 - y) \geq M$, $\forall y \in V_0$. However, $h(y) \leq f(x_0) - g(Ax_0 - y) \leq f(x_0) - M$, $\forall y \in V_0$, which implies the continuity of h on V_0, as claimed. □

Theorem 3.53 *Under the hypothesis of Lemma 3.52, Problem $\mathscr{P}_1$ is stable, in other words the equality*

$$\inf\{f(x) - g(Ax);\ x \in X\} = \max\{g^*(y^*) - f^*(A^* y^*);\ y^* \in Y^*\} \tag{3.86}$$

holds.

Also, the following two properties are equivalent:

(i) $(x_0, y_0^*) \in X \times Y^*$ *is a couple of solutions for $\mathscr{P}_1$ and $\mathscr{P}_1^*$.*
(ii) $x_0 \in X$ *and* $y_0^* \in Y^*$ *verify the system*

$$0 \in \partial f(x_0) - A^* y_0^*, \qquad 0 \in y_0^* - \partial g(Ax_0). \tag{3.87}$$

Proof Since h is continuous at the origin, it is also subdifferentiable at this point (see Proposition 2.36, Chap. 2). By virtue of Theorem 3.48, Problem $\mathscr{P}_1$ is stable. On the other hand, in view of Theorem 3.49, every couple of solutions (x_0, y_0^*) for $\mathscr{P}_1$ and $\mathscr{P}_1^*$ is a saddle point of the Hamiltonian

$$H(x, y^*) = \sup\{(y, y^*) + g(Ax - y) - f(x);\ y \in Y\} = (Ax, y^*) - g^*(y^*) - f(x).$$

But we know (see Sect. 2.3.2, Chap. 2) that the saddle points of H coincide with the solutions of equation $(0,0) \in \partial H(x, y^*)$. Making an elementary calculation, we obtain the equivalence of properties (i) and (ii). We note that, in (ii), by ∂g we mean the subdifferential in the sense of concave functions, that is, $\partial g = -\partial(-g)$. $\square$

As a consequence of Theorem 3.53, taking $X = Y$ and A as identity operator, we obtain a remarkable result in duality theory known in the literature as the *Fenchel duality theorem*.

Theorem 3.54 (Fenchel) *Let f and $-g$ be proper convex lower-semicontinuous functions on X. If there exists an element $\bar{x} \in \mathrm{Dom}(f) \cap \mathrm{Dom}(g)$ such that either f or g is continuous at $\bar{x}$, then the following equality holds*:

$$\inf\{f(x) - g(x);\ x \in X\} = \max\{g^*(x^*) - f^*(x^*);\ x^* \in X^*\}. \tag{3.88}$$

Remark 3.55 From relations (3.86) and (3.88) we see that the dual problem always has solutions; but this is not always the case with the primal problem. Furthermore, if in a point x_0 we have

$$\partial f(x_0) \cap \partial g(x_0) \neq \emptyset,$$

then in relation (3.88) the infimum is attained. The points of $\partial f(x_0) \cap \partial g(x_0)$ are optimal solutions to the dual problem $\mathscr{P}_1^*$.

Remark 3.56 A characterization of the elements $x \in X$ such that (3.88) holds will be established in the next section.

Under more general conditions, namely without the reflexivity properties of the space, we can prove, as a consequence of the Fenchel theorem, the additivity theorem of the subdifferential (see Corollary 2.63 and Remark 2.64).

Theorem 3.57 *If the functions f_1 and f_2 are finite at a point in which at least one is continuous, then*

$$\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x), \quad \forall x \in X. \tag{3.89}$$

Proof First, we prove that for every $x^* \in X^*$ there exists $u_0^* \in X^*$ such that

$$(f_1 + f_2)^*(x^*) = f_1^*(x^* - u_0^*) + f_2^*(u_0^*). \tag{3.90}$$

Indeed, if we take in the Fenchel theorem $f = f_2$ and $g = x^* - f_1$, we have $f^* = f_2^*$ and

$$\begin{aligned} g^*(u^*) &= \inf_{x \in X}\{(u^*, x) - g(x)\} \\ &= \inf_{x \in X}\{(u^* - x^*, x) + f_1(x)\} \end{aligned}$$

$$= -\sup_{x\in X}\big\{(x^* - u^*, x) - f_1(x)\big\}$$
$$= -f_1^*(x^* - u^*).$$

From equality (3.88), we obtain

$$\inf_{x\in X}\big\{f_1(x) + f_2(x) - (x^*, x)\big\} = \max_{u^*\in X^*}\big\{-f_1^*(x^* - u^*) - f_2(u^*)\big\}$$

which yields relation (3.90).

Now, consider x^* in $\partial(f_1 + f_2)(x)$. By virtue of Proposition 2.33, we have

$$(f_1 + f_2)(x) + (f_1 + f_2)^*(x^*) = (x^*, x).$$

Using relation (3.90), we obtain

$$f_1(x) + f_1^*(x^* - u_0^*) + f_2(x) + f_2^*(u_0^*) = (x^*, x) = (x^* - u_0^*, x) + (u_0^*, x). \quad (3.91)$$

On the other hand, it is clear that

$$f_1(x) + f_1^*(x^* - u_0^*) \ge (x^* - u_0^*, x),$$
$$f_2(x) + f_2^*(u_0^*) \ge (u_0^*, x).$$

However, according to relation (3.91), the equality sign must hold in both inequalities. By virtue of the same Proposition 2.33, it follows that $x^* - u_0^* \in \partial f_1(x)$ and $u_0^* \in \partial f_2(x)$, that is, $x^* \in \partial f_1(x) + \partial f_2(x)$. Hence $\partial(f_1 + f_1)(x) \subset \partial f_1(x) + \partial f_2(x)$. The converse inclusion is always true without supplementary hypotheses concerning the functions f_1 and f_2. Thus, relation (3.89) is completely proved. □

We note that the hypotheses of Lemma 3.52, which was essential for establishing the preceding results, imply a consistency condition stronger than condition (3.83), namely

$$A\big(\mathrm{Dom}(f)\big) \cap \operatorname{int}\mathrm{Dom}(g) \ne \emptyset. \quad (3.92)$$

In this case, we say that $\mathscr{P}_1$ is *strongly consistent*.

Since X and Y are Banach spaces, if $\mathscr{P}_1$ is strongly consistent and normal, we can conclude that $\mathscr{P}_1$ is also stable because the normality implies the lower-semicontinuity of the function h at $0 \in \operatorname{int}\mathrm{Dom}(h)$. From Proposition 2.16, we obtain the continuity of the function h at the origin, which clearly implies that $\partial h(0) \ne \emptyset$. Now, from Theorem 3.48, we immediately obtain the desired conclusion.

In the finite-dimensional case, because the restriction of a convex function to its effective domain is continuous in the relative interior points (see Proposition 2.17), the stability is provided only by the interiority condition

$$A\big(\operatorname{ri}\mathrm{Dom}(f)\big) \cap \operatorname{ri}\mathrm{Dom}(g) \ne \emptyset. \quad (3.93)$$

Finally, let us determine the Hamiltonian of Problem $\mathscr{P}_1$. By virtue of relations (3.76) and (3.81), we obtain

$$\begin{aligned} H(x, y^*) &= \sup\{(y, y^*) - f(x) - g(Ax - y);\ y \in Y\} \\ &= -f(x) - \inf\{(-Ax, y^*) + (u, y^*) - g(u);\ u \in Y\} \\ &= -f(x) + (Ax, y^*) - g^*(y^*) \\ &= -K(x, y^*), \end{aligned} \tag{3.94}$$

where

$$K(x, y^*) = f(x) + g^*(y^*) - (Ax, y^*), \quad (x, y^*) \in \mathrm{Dom}(f) \cap \mathrm{Dom}(g^*) \tag{3.95}$$

is just the *Kuhn–Tucker function* associated with the problems $\mathscr{P}_1$ and $\mathscr{P}_1^*$. Hence, K is convex-concave. Theorem 3.49, together with Corollary 3.50, yields the next theorem.

Theorem 3.58 *The Kuhn–Tucker function attached to Problems $\mathscr{P}_1$ and $\mathscr{P}_1^*$ admits a saddle point if and only if $\mathscr{P}_1$ and $\mathscr{P}_1^*$ are stable. A point $(x_0, y_0^*) \in X \times Y^*$ is a pair of solutions for $\mathscr{P}_1$ and $\mathscr{P}_1^*$ with the same extremal values if and only if it is a saddle point for the Kuhn–Tucker function. Furthermore, we have*

$$\min \mathscr{P}_1 = \max \mathscr{P}_1^* = \min_{x \in X} \max_{y^* \in Y^*} K(x, y^*) = \max_{y^* \in Y^*} \min_{x \in X} K(x, y^*).$$

We observe that the condition of the saddle point can be explicitly rewritten as

$$\begin{aligned} &\min_{x \in X} \sup_{y^* \in Y^*} \inf_{y \in Y} \{f(x) - g(Ax - y) - (y, y^*)\} \\ &\quad = \max_{y^* \in Y^*} \inf_{x \in X} \inf_{y \in Y} \{f(x) - g(Ax - y) - (y, y^*)\} \\ &\quad = \max_{y^* \in Y^*} \inf_{(x,y) \in X \times Y} \{f(x) - g(y) - (Ax - y, y^*)\}. \end{aligned}$$

Thus, it is natural to consider the problem of saddle points on $(X \times Y) \times Y^*$ of the Lagrangian

$$\mathscr{L}(x, y; y^*) = f(x) - g(y) - (Ax - y, y^*), \quad \forall (x, y) \in X \times Y,\ y^* \in Y^*. \tag{3.96}$$

It is easy to prove that a point (x_0, y_0^*) is a saddle point of H on $X \times Y^*$ if and only if (x_0, Ax_0, y_0^*) is a saddle point of $\mathscr{L}$ on $(X \times Y) \times Y^*$. Now, if we take as the perturbation function $F_r : X \times Y \to \overline{\mathbb{R}}^*$, $r \in \mathbb{R}$, defined by

$$F_r(x, y) = f(x) - g(Ax - y) + \frac{1}{2} r \|y\|^2 = F(x, y) + \frac{1}{2} r \|y\|^2,$$

we obtain the Hamiltonian

$$H_r(x, y^*) = \sup\{(y, y^*) - F_r(x, y);\ y \in Y\}.$$

The corresponding Lagrangian

$$\mathscr{L}_r(x, y; y^*) = \mathscr{L}(x, y; y^*) + \frac{1}{2} r \|Ax - y\|^2,$$

called the *augmented Lagrangian*, has the same saddle points as $\mathscr{L}$. The Hamiltonian H_r and the corresponding Lagrangian $\mathscr{L}_r$ are differentiable with respect to y^* for every $r > 0$. Thus, convenient algorithms in the finding of saddle points can be given. A detailed treatment of the methods generated by augmented Lagrangian has been given by Rockafellar [102].

3.2.3 Optimality Through Closedness

In this section, we characterize the global optimality of a family of optimalization problems in terms of closedness. In this way, we can obtain various optimality conditions using some criteria for closedness of the image of a closed set by a multivalued function (see Sect. 1.1.4).

Let us consider the following general family of minimization problems:

$$(\mathscr{P}_y) \quad \min\big\{F(x, y);\ x \in X\big\}, \quad y \in Y,$$

where Y is a topological space and $F : X \times Y \to \overline{\mathbb{R}}$.

Let us denote by

$$h(y) = \inf\big\{F(x, y);\ x \in X\big\}, \quad y \in Y, \tag{3.97}$$

and

$$H = \big\{(y, a) \in Y \times \mathbb{R};\ \text{there exists } \bar{x} \in X \text{ such that } F(\bar{x}, y) \le a\big\}. \tag{3.98}$$

Lemma 3.59 *Problems $(\mathscr{P}_y)_{y \in Y}$ have optimal solutions, whenever $h(y)$ is finite, and the function h is lower-semicontinuous on Y if and only if the set H is closed in $Y \times \mathbb{R}$.*

Proof Let a be a real number such that $a > h(y)$. Then, there exists an element $x_a \in X$ such that $F(x_a, y) \le a$, that is $(y, a) \in H$. If H is closed, we also have

$$\lim_{a \searrow h(y)} (y, a) = \big(y, h(y)\big) \in H.$$

Therefore, from definition (3.98) of the set H, there exists $\bar{x} \in X$ such that $F(\bar{x}, y) \le h(y)$, which say that $\bar{x}$ is an optimal solution for $\mathscr{P}_y$. On the other hand, from definition (3.97) of the function h, it is easy to observe that we have the inclusion relation $H \subset \operatorname{epi} h \subset \overline{H}$. Hence, if H is closed, it follows that h is lower-semicontinuous on Y.

Conversely, let $(y, a) \in Y \times \mathbb{R}$ be a cluster element of H. Since h is lower-semicontinuous, we have $(y, a) \in \overline{H} = \overline{\operatorname{epi} h} = \operatorname{epi} h$, and so $h(y) \leq a$. Therefore, $h(y) < \infty$. Now, if $h(y)$ is finite, by hypothesis there exists an optimal solution $x_1 \in X$, that is, $F(x_1, y) = h(y)$. Hence, $(y, a) \in H$. If $h(y) = -\infty$, by definition of h there exist elements $x \in X$ such that $F(x, y) \leq a$, which says that $(y, a) \in H$. Therefore, the set H is closed. □

Remark 3.60 This optimality result can be extended to the case of optimality only for the elements of a subset A of Y. Indeed, from the proof it follows that the value function h is lower-semicontinuous on $A \subset Y$ and each problem $\mathscr{P}_y$ has optimal solutions, whenever $y \in A$ and $h(y) > -\infty$, if and only if

$$H \cap (A \times \mathbb{R}) = \overline{H} \cap (A \times \mathbb{R}).$$

It is obvious that the set H given by (3.98) is a set of epigraph type, that is, $(x, \lambda') \in H$, whenever $(x, \lambda) \in H$ and $\lambda' \geq \lambda$. Consequently, the closedness property in H can by fulfilled if its section in Y and $\mathbb{R}$, respectively, are closed sets. Let us denote

$$H_y = \{a \in \mathbb{R};\ (y, a) \in H\}, \tag{3.99}$$

$$H_a = \{y \in Y;\ (y, a) \in H\}. \tag{3.100}$$

Thus, Lemma 3.59 can be refined as follows.

Lemma 3.61

(i) *Problems $(\mathscr{P}_y)_{y \in Y}$ have optimal solutions whenever $h(y) \in \mathbb{R}$ if and only if the sets H_y, $y \in Y$, are closed in $\mathbb{R}$.*

(ii) *If the sets H_a, $a \in \mathbb{R}$, are closed in Y, then the value function h is lower-semicontinuous on the parameter space Y.*

(iii) *If the set H_y, $y \in Y$, are closed in $\mathbb{R}$ for every $y \in Y$ and the value function h is lower-semicontinuous on Y, then the set H_a, $a \in \mathbb{R}$, are closed in Y.*

(iv) *The set H is closed in $Y \times \mathbb{R}$ if and only if the sets H_y, $y \in Y$, and H_a, $a \in \mathbb{R}$, are closed in $\mathbb{R}$ and Y, respectively.*

(v) *If the function F has the property*

$$y \in \overline{H}_{h(y)} \quad \text{if } h(y) \text{ is finite}, \tag{3.101}$$

then Problems $(\mathscr{P}_y)_{y \in Y}$ have optimal solutions whenever $h(y)$ is finite and the value function is lower-semicontinuous if and only if the sets H_a, $a \in \mathbb{R}$, are closed in Y.

Proof (i) We take $Y = \{y\}$ if $h(y)$ is finite and applies Lemma 3.59.

(ii) We have $\{y;\ h(y) \leq a\} = \bigcap_{\varepsilon > 0} H_{a+\varepsilon}$.

(iii) and (iv) follow from Lemma 3.59 by taking into account (i).

(v) By Lemma 3.59, optimality and lower-semicontinuity ensure that H is closed, and so H_a is closed for every $a \in \mathbb{R}$. Conversely, if H_a, $a \in \mathbb{R}$, are closed, by (ii) h

is lower-semicontinuous. Moreover, property (3.101) proves that $y \in H_{h(y)}$, that is, $\mathscr{P}_y$ has optimal solutions if $h(y) \in \mathbb{R}$. □

If the set H is also convex (this is possible even if F is not convex), the above result can be restated as a duality result.

With that end in view, we consider the duals by conjugacy of Problems $(\mathscr{P}_y)_{y \in Y}$, that is, the following maximization problems:

$$(\mathscr{D}_y) \quad \max\left\{(y, y^*) - F^*(0, y^*);\ y^* \in Y^*\right\}, \quad y \in Y,$$

where (X, X^*), (Y, Y^*) are two dual systems endowed with compatible topologies.

We observe that the family $\{\mathscr{D}_y;\ y \in Y\}$ coincides with the family of all the linear perturbations of Fenchel–Rockafellar duals of Problem $\mathscr{P}_0$.

By an elementary calculation involving the definition of the conjugate of h, we obtain

$$\operatorname{val} \mathscr{D}_y = h^{**}(y), \quad \text{for all } y \in Y. \tag{3.102}$$

Lemma 3.62 *Suppose that at least one of the problems* $(\mathscr{P}_y)$ *has a finite value. Then*, $\operatorname{val} \mathscr{P}_y = \operatorname{val} \mathscr{D}_y \neq -\infty$, *and* $\mathscr{P}_y$ *has an optimal solution, for any* $y \in Y$, *if and only if the set* H *is closed and convex.*

Proof If H is closed and convex, by Lemma 3.59 it follows that h is convex and lower-semicontinuous. On the other hand, according to Proposition 2.9, we have $h(y) \neq -\infty$, $\forall y \in Y$. Thus, by the theorem of bipolar, Lemma 3.59 and the equality (3.102), it follows that $\operatorname{val} \mathscr{P}_y = h(y) = h^{**}(y) = \operatorname{val} \mathscr{D}_y \neq -\infty$ and $\mathscr{P}_y$ has an optimal solution. Conversely, if $\operatorname{val} \mathscr{P}_y = \operatorname{val} \mathscr{D}_y$, for all $y \in Y$, we obtain $h = h^{**}$, that is, h is convex and lower-semicontinuous. By Lemma 3.59, it follows that the set H is closed. Moreover, the equality $H = \operatorname{epi} h$ implies the convexity of H. □

Remark 3.63 The above result gives a characterization of the global stability of the family $(\mathscr{D}_y)_{y \in Y}$. Now, if we apply Lemma 3.62 for Problems $\mathscr{P}$ and $\mathscr{P}^*$ considered in the preceding section (the hypothesis (A) being satisfied), we easily obtain the sufficient stability conditions: $\mathscr{P}^*$ *has a finite value and the set*

$$H^* = \left\{(x^*, a) \in X^* \times \mathbb{R};\ \textit{there exists } \bar{y}^* \in Y^* \textit{ such that } F^*(x^*, \bar{y}^*) \leq a\right\}$$

is closed in $X^* \in \mathbb{R}$.

At the same time, this constitutes a sufficient subdifferentiability condition for the function h at the origin (see Theorem 3.48).

Theorem 3.64 *Let* $F : X \times Y \to \overline{\mathbb{R}}$ *be a positively homogeneous and lower-semicontinuous function satisfying the following coercivity condition*:

$$F(x, 0) > 0 \quad \textit{for any } x \in X \setminus \{0\}. \tag{3.103}$$

Then, if epi F *is locally compact, every Problem* $\mathscr{P}_y$ *has an optimal solution whenever its value is finite.*

Proof It is easy to observe that $H = \mathrm{Proj}_{Y\times\mathbb{R}}(\mathrm{epi}\, F)$. By hypothesis, epi h is a closed cone and so $(\mathrm{epi}\, F)_\infty = \mathrm{epi}\, F$. Therefore, it suffices to use Corollary 1.60 for $T = \mathrm{Proj}_{Y\times\mathbb{R}}$ and $A = \mathrm{epi}\, F$, taking into account that the separation condition (1.41) of Corollary 1.60 may be written as condition (3.103). □

Remark 3.65 We can omit the condition that F is positively homogeneous, by using the *recession function* associated to F, which is defined by

$$F_\infty(x, y) = \sup_{\varepsilon>0} \liminf_{[u,v]\to[x,y]} \inf_{0\le\lambda\le\varepsilon} \lambda F\left(\frac{u}{\lambda}, \frac{v}{\lambda}\right), \quad \forall [x, y] \in X \times Y.$$

It is clear that F_∞ is *positively homogeneous and lower-semicontinuous*. We also have

$$F_\infty(x, 0) > 0 \quad \text{for any } x \in X \setminus \{0\} \tag{3.104}$$

and epi F must be asymptotically compact. For example, the last property holds if *there exists* $s > 0$ *such that the origin of* $X \times Y$ *has a relatively compact neighborhood in the induced topology of* $X \times Y$ *on the set*

$$\left\{(x, y) \in X \times Y;\ \inf_{0\le\lambda\le s} \lambda F\left(\frac{x}{\lambda}, \frac{y}{\lambda}\right) \le s\right\}.$$

(For details, see Precupanu [86].)

Next, consider the family of *Fenchel–Rockafellar problems*

$$(\mathscr{P}_y) \quad \min\bigl\{f(x) - g(Ax + y);\ x \in X\bigr\}, \quad y \in Y,$$

where f *and* $-g$ *are two proper lower-semicontinuous functions on* X *and* Y, *respectively, and* $A : X \to Y$ *is a linear continuous operator.*

In this case, the set H and the dual problems can be written as

$$H = \bigl\{\bigl(Ax - y, f(x) - g(y) + r\bigr) \in Y \times \mathbb{R};\ x \in \mathrm{Dom}(f),\ y \in \mathrm{Dom}(g),\ r \ge 0\bigr\}, \tag{3.105}$$

$$(\mathscr{D}_y) \quad \max\bigl\{g^*(y^*) - f^*(A^*y^*) + (y, y^*);\ y^* \in Y^*\bigr\}, \quad y \in Y.$$

If we consider the associated operator $\widetilde{A} : X \times \mathbb{R} \to Y \times \mathbb{R}$ defined by

$$\widetilde{A}(x, t) = (Ax, t) \quad \text{for all } (x, t) \in X \times \mathbb{R}, \tag{3.106}$$

then the set H can be rewritten in the simple form

$$H = \widetilde{A}(\mathrm{epi}\, f) - \mathrm{hypo}\, g. \tag{3.107}$$

Theorem 3.66 *If $\mathscr{P}_y$ has a finite value at least an element of Y, then*

$$\min\left\{f(x) - g(Ax+y);\ x \in X\right\} = \sup\left\{g^*(y^*) - f^*(A^*y^*) + (y, y^*);\ y^* \in Y^*\right\},$$

whenever val $\mathscr{P}_y$ *is finite, if and only if H is closed and convex.*

Proof Apply Lemma 3.62. □

Corollary 3.67 *If f, $-g$ are proper convex and lower-semicontinuous and $\mathscr{P}_0$ is consistent, then*

$$\begin{aligned}&\inf\left\{f(x) - g(Ax) - (x, x^*);\ x \in X\right\}\\&\quad = \max\left\{g^*(y^*) - f^*(A^*y^* + x^*);\ y^* \in Y^*\right\}, \quad x^* \in X^*,\end{aligned}$$

whenever the left-hand side is finite, if and only if the set

$$\begin{aligned}H^* = \big\{&\big(A^*y^* - x^*, f^*(x^*) - g^*(y^*) + r\big) \in X^* \times \mathbb{R};\\&x^* \in \operatorname{Dom}(f^*),\ y^* \in \operatorname{Dom}(g^*),\ r \geq 0\big\}\end{aligned} \tag{3.108}$$

is closed in $X^* \times \mathbb{R}$.

Proof Take $-g^*, -f^*, A^*$ instead of f, g, A, respectively. Thus, the corresponding set of H is even H^*. It is clear that H^* is always convex. As initial family we consider

$$(\widetilde{\mathscr{P}}_{x^*}) \quad \min\left\{-g^*(y^*) + f^*(A^*y^* + x^*);\ y^* \in Y\right\}$$

and so, as dual family (according to the theorem of bipolar), we have

$$(\widetilde{\mathscr{D}}_{x^*}) \quad \max\left\{-f(x) + g(Ax) + (x, x^*);\ x \in X\right\}.$$

Also, it is clear that val $\widetilde{\mathscr{P}}_0 \neq -\infty$ if $\mathscr{P}_0$ is consistent since $-\widetilde{\mathscr{P}}_0$ is the dual of $\mathscr{P}_0$; hence, $-$ val $\widetilde{\mathscr{P}}_0 \leq$ val $\mathscr{P}_0 \neq \infty$. The proof is complete. □

Here, the similar form of formula (3.107) is

$$H^* = \widetilde{A}^*\big(\operatorname{epi}(-g^*)\big) - \operatorname{hypo}(-f^*).$$

Obviously, this set can be, equivalently, replaced by the set

$$\widetilde{H}^* = \operatorname{epi} f^* - \widetilde{A}(\operatorname{hypo} g^*). \tag{3.109}$$

Remark 3.68 If the two functions f, g are arbitrary, then by Lemma 3.59 we obtain the following result. *The associated Fenchel–Rockafellar problems $\mathscr{P}_y$, $y \in Y$, have optimal solutions whenever $h(y)$ is finite and h is lower-semicontinuous on Y if and only if the set* (3.105) *is closed.* Here, we admit that $\infty + a = \infty$ for all $a \in \overline{\mathbb{R}}$. If, in addition, the set (3.105) is convex, then the duality properties from Lemma 3.61 hold.

If A is the identity operator, it is clear that

$$H = \operatorname{epi} f - \operatorname{hypo} g, \tag{3.110}$$

$$\widetilde{H}^* = \operatorname{epi} f^* - \operatorname{hypo} g^*. \tag{3.111}$$

Thus, we obtain the following result of Fenchel–Rockafellar type.

Corollary 3.69 *If $\mathscr{P}_0$ or $\mathscr{D}_0$ has a finite value and* epi $f^* -$ hypo g^* *is closed in $X^* \times \mathbb{R}$, then $\mathscr{P}_0$ is stable, that is,*

$$\inf\big\{f(x) - g(x);\ x \in X\big\} = \max\big\{g^*(x^*) - f^*(x^*);\ x^* \in X^*\big\}.$$

It is easy to observe that the set H defined by (3.105) can be decomposed into a difference of two sets in various forms. From this, we consider the following four cases:

$$H = M_i - N_i, \quad i = 1, 2, 3, 4, \tag{3.112}$$

where

$$\begin{aligned} M_1 &= \big\{\big(Ax, f(x) + r\big);\ x \in \operatorname{Dom}(f),\ r \ge 0\big\}, \\ N_1 &= \big\{\big(y, g(y)\big);\ y \in \operatorname{Dom}(g)\big\}, \end{aligned} \tag{3.113}$$

$$\begin{aligned} M_2 &= \big\{\big(Ax, f(x) + r\big);\ x \in \operatorname{Dom}(f)\big\}, \\ N_2 &= \big\{\big(y, g(y) - r\big);\ y \in \operatorname{Dom}(g),\ r \ge 0\big\}, \end{aligned} \tag{3.114}$$

$$M_3 = M_1, \qquad N_3 = N_2, \tag{3.115}$$

$$\begin{aligned} M_4 &= \big\{\big(Ax - y, f(x) - g(x)\big);\ x \in \operatorname{Dom}(f),\ y \in \operatorname{Dom}(g)\big\}, \\ N_4 &= \{0\} \times \mathbb{R}_-. \end{aligned} \tag{3.116}$$

If f an d g are arbitrary functions, the above sets can be non-closed and non-convex (except the set N_4 which is a closed convex cone). We observe that

$$N_1 = \operatorname{graph} g, \qquad N_2 = \operatorname{hypo} g, \qquad M_4 = M_2 - N_1,$$

and

$$M_1 = \widetilde{A}(\operatorname{epi} f), \qquad M_2 = \widetilde{A}(\operatorname{graph} f). \tag{3.117}$$

Thus, the four decompositions can be rewritten as follows:

$$H = \widetilde{A}(\operatorname{epi} f) - \operatorname{graph} g, \tag{3.118}$$

$$H = \widetilde{A}(\operatorname{graph} f) - \operatorname{hypo} g, \tag{3.119}$$

$$H = \widetilde{A}(\operatorname{epi} f) - \operatorname{hypo} g, \tag{3.120}$$

$$H = \widetilde{A}(\operatorname{graph} f) - \operatorname{graph} g - \{0\} \times \mathbb{R}_-. \tag{3.121}$$

Now, we can obtain closedness conditions for the set H if $(\operatorname{epi} f)_\infty$ (or $(\operatorname{graph} f)_\infty$, $(\operatorname{epi} g)_\infty$, $(\operatorname{graph} g)_\infty$) is asymptotically compact using Theorem 1.59 or Corollary 1.60. Thus, in the special case of positively homogeneous functions, the separation condition (1.41) from Theorem 1.59 for the decompositions (3.118)–(3.120) is the same, namely,

(c) $f(x) \le g(Ax)$ implies $x = 0$,

and the separation condition (1.42) from Corollary 1.60 becomes:

(c_1) $f(x) \le g(Ax)$ implies $Ax = 0$ (for (3.118))
(c_2) $f(x) \le g(Ax)$ implies $Ax = 0$ and $f(x) = 0$ (for (3.119) and (3.120))
(c_3) $f(x) \le g(Ax)$ implies $f(x) = g(Ax)$ (for (3.121)).

Theorem 3.70 *Let f, $-g$ be proper positively homogeneous functions. Each of the following properties is sufficient for the closedness of the set H:*

(i) $\widetilde{A}(\operatorname{epi} f)$ *is a locally compact and either g has closed graph and* (c_1) *holds or g is upper-semicontinuous and* (c_2) *holds.*
(ii) $\operatorname{epi} f$ *is locally compact, condition* (c) *is satisfied and g is either upper-semicontinuous or has closed graph.*
(iii) $\widetilde{A}(\operatorname{graph} f) - \operatorname{graph} g$ *is closed and* (c_3) *holds.*
(iv) $\operatorname{graph} g$ *is locally compact,* (c_1) *is satisfied and* $\widetilde{A}(\operatorname{epi} f)$ *is closed.*
(v) $\operatorname{hypo} g$ *is locally compact,* (c_2) *is satisfied and* $\widetilde{A}(\operatorname{epi} f)$ *or* $\widetilde{A}(\operatorname{graph} f)$ *is closed.*

Proof The sufficiency of properties (i), (iii), (iv) and (v) follows using Corollary 1.60. To obtain (ii), we apply Theorem 1.59. In fact, according to Corollary 1.60, condition (ii) is stronger than condition (i). □

Remark 3.71 If $\operatorname{graph} f$ is locally compact, then $\operatorname{epi} f$ is also locally compact since $\operatorname{epi} f = \operatorname{graph} f + \{0\} \times \mathbb{R}_+$. If $\operatorname{graph} f$ is closed, the converse is also true.

The local compactness conditions can be ensured by dual interiority conditions taking into account that a *closed cone is locally compact if the interior of polar cone is nonvoid, with respect to Mackey topology.*

In the sequel, we consider the homogeneous program

$$\min\bigl\{a(x);\ x \in P,\ Ax + y \in Q\bigr\}, \quad y \in Y, \tag{3.122}$$

where $a : X \to]-\infty, +\infty]$ is a positively homogeneous lower-semicontinuous function, $A : X \to Y$ is a linear continuous operator, y is a fixed element of Y and $P \subset X$, $Q \subset Y$ are closed cones.

This minimization problem is of $\mathscr{P}_y$ type for

$$f = a + I_P, \qquad g = -I_Q,$$

and so, as dual problem we have

$$(\mathscr{D}'_y) \quad \max\{\langle y, y^*\rangle;\ y^* \in Q^0,\ A^*y^* \in \partial(a + I_P)(0)\}, \quad y \in Y. \tag{3.123}$$

Here, the set H defined by (3.123) is

$$H = \{(Ax - y, a(x) + r);\ x \in P \cap \mathrm{Dom}(a),\ y \in Q,\ r \geq 0\}.$$

The coercivity conditions (c_1), $i = 1, 2, 3$, and (c) become:

(c_1^0) *If* $x \in P \cap A^{-1}(Q)$ *and* $a(x) \leq 0$, *then* $Ax = 0$
(c_2^0) *If* $x \in P \cap A^{-1}(Q)$ *and* $a(x) \leq 0$, *then* $Ax = 0$ *and* $a(x) = 0$
(c_3^0) *If* $x \in P \cap A^{-1}(Q)$ *and* $a(x) \leq 0$, *then* $a(x) = 0$
(c^0) *If* $x \in P \cap A^{-1}(Q)$ *and* $a(x) \leq 0$, *then* $x = 0$.

We also have

$$\mathrm{epi}\, f = \mathrm{epi}\, a|_P; \qquad \mathrm{graph}\, f = \mathrm{graph}\, a|_P, \tag{3.124}$$

$$\mathrm{hypo}\, g = Q \times \mathbb{R}_-; \qquad \mathrm{graph}\, g = Q \times \{0\}, \tag{3.125}$$

where we denote by $a|_P$ the restriction of a to $P \cap \mathrm{Dom}(a)$.

By Theorem 3.70, we obtain the following result.

Theorem 3.72 *The homogeneous program* (3.122) *has an optimal solution for every* $y \in Y$, *whenever its value is finite, if one of the following conditions holds*:

(i) $\widetilde{A}(\mathrm{epi}\, a|_P)$ *is locally compact and* (c_1^0) *is satisfied*
(ii) $\widetilde{A}(\mathrm{graph}\, a|_P) - Q \times \{0\}$ *is closed and* (c_3^0) *is satisfied*
(iii) Q *is locally compact*, $\widetilde{A}(\mathrm{epi}\, a|_P)$ *is closed and* (c_1^0) *is satisfied*
(iv) Q *is locally compact*, $\widetilde{A}(\mathrm{graph}\, a|_P)$ *is closed and* (c_2^0) *is satisfied.*

The local compactness and closedness of $\widetilde{A}(\mathrm{epi}\, a|_P)$ and $\widetilde{A}(\mathrm{graph}\, a|_P)$ which appear in the above conditions can be derived by Corollaries 1.60 and 1.61, by using the coercivity conditions (c^0). Thus, for example, if $\mathrm{epi}\, a|_P$ or $\mathrm{graph}\, a|_P$ is locally compact and (c^0) is fulfilled, then (i) holds. Also, it is sufficient that $P \cap \mathrm{Dom}(a)$ to be locally compact and $\mathrm{graph}\, a|_P$ to be closed since $\widetilde{A}(\mathrm{graph}\, a|_P) = (A \times a)(P \cap \mathrm{Dom}(a))$.

In the linear case, $a \in X^*$, all these conditions have a more simple form since $\partial(a + I_P)(0) = a + P^0$. Let us remark that the linear case is not different from the positively homogeneous case because every positively homogeneous program can be reduced to a linear program. Indeed, the positively homogeneous program (3.122) is equivalent to

$$\min\{t;\ (x, t) \in \mathrm{epi}\, a,\ x \in P,\ Ax + y \in Q\}. \tag{3.126}$$

This program is linear and of the same type as (3.122), where X, Y, A, P, Q are replaced by $X \times \mathbb{R}$, Y, A_1, $(\mathrm{epi}\, a) \cap (P \times \mathbb{R})$, Q, respectively, with the operator $A_1 : X \times \mathbb{R} \to Y$ defined by $A_1(x, t) = Ax$, for all $(x, t) \in X \times \mathbb{R}$. Since

$A_1^* y^* = (A^* y^*, 0) \in X^* \times \mathbb{R}$, for all $y^* \in Y^*$, $((\operatorname{epi} a) \cap (P \times \mathbb{R}))^0 = P^0 \times \{0\} \cup \operatorname{cone}(\partial a|_P(0) \times \{-1\}$ and the cost functional of problem (3.126) can be identified with $(0, 1) \in X^* \times \mathbb{R}$, the duals of problems (3.122) and (3.126) are the same.

Finally, we remark that Theorem 3.72 can be completed as a dual result, that is, programs (3.122) and (3.123) have equal values if the set H is also convex.

3.2.4 Non-convex Optimization and the Ekeland Variational Principle

As seen earlier, the central problem of the topics discussed so far is the minimization problem

$$\operatorname{Min}\big\{f(x);\ x \in M\big\}, \tag{3.127}$$

where $f : X \to \overline{\mathbb{R}}$ is a given lower-semicontinuous function on a Banach space X and M is a closed subset of X. By the Weierstrass theorem, a sufficient condition for existence of a minimum in (3.127) is that M be compact and f lower-semicontinuous with respect to a certain topology on X, for instance, the weak topology of X, and the latter holds if f is convex, lower-semicontinuous and M is bounded.

If these conditions are absent, one cannot prove the existence in problem (3.128). A second important question related to problem (3.128) is to find the minimum points (if any) by a first order condition

$$\partial f(x) \ni 0, \tag{3.128}$$

where ∂f is the gradient of f in some generalized sense.

The Ekeland variational principle [37, 38] to be briefly presented below is a sharp instrument to give a partial answer to these questions.

Theorem 3.73 *Let X be a complete metric space and let $f : X \to (-\infty, +\infty]$ be a lower-semicontinuous function, nonidentically $+\infty$ and bounded from below. Let $\varepsilon > 0$ be arbitrary and let $x \in X$ be such that*

$$f(x) \geq \inf\big\{f(u);\ u \in X\big\} + \varepsilon.$$

Then, there exists $x_\varepsilon \in X$ such that

$$\begin{aligned} f(x_\varepsilon) \leq f(x), \qquad & d(x_\varepsilon, x) \leq 1, \\ f(u) > f(x_\varepsilon) - \varepsilon d(x_\varepsilon, u), \qquad & \forall u \neq x_\varepsilon. \end{aligned} \tag{3.129}$$

Here, $d : X \times X \to \mathbb{R}$ is the distance on X.

Proof We take $x_0 = x$ and define inductively the sequence $\{x_n\}$ as follows.

If x_{n-1} is known, then either

$$f(u) > f(x_{n-1}) - \varepsilon d(x_{n-1}, u), \quad \forall u \neq x_{n-1}, \tag{3.130}$$

and, in this case, take $x_n = x_{n-1}$, or there exists $u \neq x_{n-1}$ such that

$$f(u) \le f(x_{n-1}) - \varepsilon d(x_{n-1}, u). \tag{3.131}$$

In the latter case, denote by S_n the set of all $u \in X$ satisfying (3.131) and choose $x_n \in S_n$ such that

$$f(x_n) - \inf\bigl\{f(u);\ u \in S_n\bigr\} \le \frac{1}{2}\bigl(f(x_{n-1}) - \inf\bigl\{f(u);\ u \in S_n\bigr\}\bigr).$$

We prove that the sequence $\{x_n\}$ so defined is convergent. If (3.130) happens for all n, then $\{x_n\}$ is stationary; otherwise, it follows by (3.131) that

$$d(x_{n-1}, x_n) \le f(x_{n-1}) - f(x_n)$$

and, therefore,

$$\varepsilon d(x_{n-1}, x_m) \le f(x_{n-1}) - f(x_m), \quad \forall m \ge n - 1. \tag{3.132}$$

The sequence $\{f(x_n)\}$ is bounded from below and monotonically decreasing.

Hence, $\{f(x_n)\}$ is convergent and, by (3.132), it follows that so is $\{x_n\}$. Hence, $\lim_{n\to\infty} x_n = x_\varepsilon$ exists. We have

$$f(x) \ge f(x_1) \ge \cdots \ge f(x_{n-1}) \ge f(x_n) \ge \cdots$$

and we may conclude that

$$f(x) \ge \lim_{n\to\infty} f(x_n) = \ell \ge f(x_\varepsilon),$$

because f is lower-semicontinuous. We get

$$d(x, x_m) \le f(x) - f(x_m) \le f(x) - \inf\bigl\{f(u);\ u \in X\bigr\} \le \varepsilon.$$

Then, letting n tend to $+\infty$, we get $d(x_\varepsilon, x) \le 1$. To prove the last relation, we assume that there exists $u \neq x_\varepsilon$ such that

$$f(u) \le f(x_\varepsilon) - \varepsilon d(x_\varepsilon, u)$$

and we argue from this to a contradiction. Since $f(x_\varepsilon) \le f(x_{n-1})$ for all n, the latter yields

$$f(u) \le f(x_{n-1}) - \varepsilon d(x_\varepsilon, u) + \varepsilon d(x_\varepsilon, x_{n-1}) \le f(x_{n-1}) - \varepsilon d(x_{n-1}, u).$$

Hence, $u \in S_n$ for all n. On the other hand, we have

$$2f(x_n) - f(x_{n-1}) \le \inf\bigl\{f(v);\ v \in S_n\bigr\} \le f(u).$$

Hence, $f(u) \geq \ell \geq f(x_\varepsilon)$. The contradiction we arrived at proves the desired relation. □

Corollary 3.74 *Let X be a complete metric space and let $f : X \to (-\infty, +\infty]$ be a lower-semicontinuous which is bounded from below and $\not\equiv +\infty$. Let $\varepsilon > 0$ and $x \in X$ be such that*

$$f(x) \leq \inf\big\{f(u);\ u \in X\big\} + \varepsilon.$$

Then, there exists $x_\varepsilon \in X$ such that

$$f(x_\varepsilon) \leq f(x), \qquad d(x_\varepsilon, x) \leq \varepsilon^{\frac{1}{2}}, \qquad f(x_\varepsilon) < f(u) + \varepsilon^{\frac{1}{2}} d(x_\varepsilon, u), \quad \forall u \neq x_\varepsilon.$$

Proof One applies Theorem 3.73 on the space X endowed with the metric $\varepsilon^{\frac{1}{2}} d$. □

In the special case, where X is a Banach space and f is Gâteaux differentiable, we have

Corollary 3.75 *Let X be a Banach space and let $f : X \to \mathbb{R}$ be Gâteaux differentiable and bounded from below. Then, for each $\varepsilon > 0$, there exists x_ε such that*

$$f(x_\varepsilon) \leq \inf\big\{f(u);\ u \in X\big\} + \varepsilon,$$
$$\big\|\nabla f(x_\varepsilon)\big\| \leq \sqrt{\varepsilon}.$$

Proof It suffices to take in (3.129) $u = x_\varepsilon \pm \lambda h$, divide by λ and let λ go to zero. □

Corollary 3.76 *Let $f : X \to \mathbb{R}$ be Gâteaux differentiable and bounded from below on the Banach space X. If f satisfies the Palais–Smale condition, then it attains its infimum on X.*

Proof We recall that a Gâteaux differentiable function f on X is said to satisfy the Palais–Smale condition if every sequence $\{x_n\} \subset X$ such that $\lim_{n\to\infty} f(x_n)$ exists, and $\lim_{n\to\infty} \nabla f(x_n) = 0$ contains a convergent subsequence.

Now, let $\{x_n\} \subset X$ be such that

$$f(x_n) \leq \inf\big\{f(u);\ u \in X\big\} + n^{-1},$$
$$\big\|\nabla f(x_n)\big\| \leq n^{-\frac{1}{2}}.$$

Then, there exists $\{x_{n_k}\} \subset \{x_n\}$ such that $\lim_{n_k\to\infty} x_{n_k} = x$. Clearly, $f(x) = \inf\{f(u); u \in X\}$, as claimed. □

The Ekeland variational principle may also be viewed as an existence result for an approximating minimum point of a lower-semicontinuous bounded from below function. Indeed, by Corollary 3.74, we have

$$x = \arg\inf\big\{f(u) + \varepsilon^{\frac{1}{2}} d(x_\varepsilon, u);\ u \in X\big\}.$$

On these lines, a sharper result was established by Stegall [108].

Theorem 3.77 *Let M be convex, weakly compact subset of a Banach space X and let $f : M \to \mathbb{R}$ be a lower-semicontinuous function which is bounded from below. Then, for every $\varepsilon > 0$, there exists $w_\varepsilon \in X^*$, $\|w_\varepsilon\| \le \varepsilon$, such that the function $x \to f(x) + (x, w_\varepsilon)$ attains its infimum on K.*

In particular, if the space X is reflexive, then we may take in Theorem 3.77 any convex, closed and bounded subset M of X.

We omit the proof.

3.2.5 Examples

In this section, we illustrate the general results by discussing some specific examples of optimization problems associated with partial differential equations, stress being laid on the formulation as well as on the explicit determination of the dual problem and of the optimality conditions.

Example 3.78 Here and throughout in the sequel, we denote by Ω a bounded open domain of $\mathbb{R}^n$ with the boundary Γ, an $(n-1)$-dimensional variety of class C^∞.

Consider the problem

$$\min\left\{\frac{1}{2}\int_\Omega |\operatorname{grad} u|^2\,dx - \int_\Omega hu\,dx;\ u \in K\right\}, \tag{3.133}$$

when $h \in L^2(\Omega)$ and $K = \{u \in H_0^1(\Omega); u \ge 0 \text{ a.e. on } \Omega\}$.

Let us take in the Frenchel duality theorem (see Theorem 3.54) $X = H_0^1(\Omega)$, $X^* = H^{-1}(\Omega)$, $g = -I_K$ (the indicator function of the set K) and

$$f(u) = \frac{1}{2}\int_\Omega |\operatorname{grad} u|^2\,dx - \int_\Omega hu\,dx, \quad u \in H_0^1(\Omega).$$

In other words, $f(u) = \frac{1}{2}\|u\|^2_{H_0^1(\Omega)} - (h, u)$, where $(\cdot,\cdot)$ denotes the duality bilinear functional between $H_0^1(\Omega)$ and its dual $H^{-1}(\Omega)$ (respectively, the inner product in $L^2(\Omega)$).

Thus, we have

$$\begin{aligned} f^*(p^*) &= \sup\left\{(p^*, u) - \frac{1}{2}\|u\|^2 + (h, u);\ u \in H_0^1(\Omega)\right\} \\ &= \sup\left\{(p^* + h, u) - \frac{1}{2}\|u\|^2;\ u \in H_0^1(\Omega)\right\} = \frac{1}{2}\|p^* + h\|^2_{H^{-1}(\Omega)}. \end{aligned}$$

On the other hand,

$$g^*(p^*) = \inf\{(u, p^*);\ u \in K\} = \begin{cases} 0, & \text{if } p^* \in K^*, \\ -\infty, & \text{if } p^* \overline{\in} K^*, \end{cases}$$

where $K^* = \{p^* \in H^{-1}(\Omega); (p^*, u) \geq 0,\ \forall u \in K\} = \{p^* \in H^{-1}(\Omega); p^* \geq 0\}$. We note that the relation $p^* \geq 0$ must be understood in the same sense of distributions.

We also remark that $-K^*$ is just the polar cone associated to the cone K. Therefore, the dual problem associated to problem (3.133) can be written as

$$\max\left\{-\frac{1}{2}\|p^* + h\|^2_{H^{-1}(\Omega)};\ p^* \in H^{-1}(\Omega),\ p^* \geq 0\right\}; \tag{3.134}$$

hence, we have the equality

$$\begin{aligned} &\inf \frac{1}{2}\left\{\int_\Omega |\operatorname{grad} u|^2 \,\mathrm{d}x - \int_\Omega hu \,\mathrm{d}x;\ u \in K\right\} \\ &\quad = -\inf\left\{\frac{1}{2}\|p^* + h\|^2_{H^{-1}(\Omega)};\ p^* \in K^*\right\}. \end{aligned} \tag{3.135}$$

Since the function f is coercive and strictly convex, problem (3.133) admits a unique solution $\bar{u} \in K$. Similarly, problem (3.134) admits a unique solution $\bar{p}^* \in K^*$. Since the Hamiltonian function $H(u, p^*)$ associated with our problem is given by

$$H(u, p^*) = \sup_{v \in K}\left\{-(p^*, v) - f(u) + (p^*, u)\right\} = I_{K^*}(p^*) + (p^*, u) - f(u),$$

we may infer (see Theorem 3.53) that the pair $(\bar{u}, \bar{p}^*)$ verifies the optimality system

$$\bar{p}^* \in \partial f(\bar{u}), \qquad \bar{p}^* + \partial I_K(\bar{u}) \ni 0.$$

Hence,

$$\begin{aligned} &-\Delta \bar{u} = h + \bar{p}^* \quad \text{on } \Omega, \\ &\bar{p}^* + \partial I_K(\bar{u}) \ni 0. \end{aligned} \tag{3.136}$$

From the definition of ∂I_K, it follows that $\bar{p}^* = 0$ on the set $\{x \in \Omega; \bar{u}(x) > 0\}$, and $\bar{p}^* \geq 0$ on the complementary set.

Problem 3.136 can be restated as

$$\begin{aligned} &-\Delta \bar{u} \geq h \quad \text{on } \Omega, \\ &-\Delta \bar{u} = h \quad \text{on } \left\{x \in \Omega;\ \bar{u}(x) > 0\right\}, \\ &\bar{u} \geq 0 \quad \text{on } \Omega, \qquad \bar{u} = 0 \quad \text{on } \Gamma. \end{aligned} \tag{3.137}$$

We have, therefore, obtained a free boundary-value problem studied in Chap. 2 (the obstacle problem).

Example 3.79 This example deals with an abstract control problem.

Let V, H be a pair of real Hilbert spaces such that V is dense in H and the injection of V in H is continuous. In other words, if we denote by V' the dual of V, we have

$$V \subset H \subset V'$$

which is meant both in the algebraic and the topological sense.

We denote by $\|\cdot\|$ and $|\cdot|$ the norm in V, and in H, respectively, and by $\|\cdot\|_*$ the norm in V'. We denote also by $(\cdot,\cdot)$ the duality bilinear functional between V and V' (the inner product in H, respectively).

Let U be another real Hilbert space and let B be a continuous linear operator from U to V'. Finally, let $A \in L(V, V')$ be a continuous linear operator subject to

$$(Au, u) \geq \omega \|u\|^2, \quad \forall u \in V, \tag{3.138}$$

where $\omega > 0$.

Consider the optimization problem

($\mathscr{P}$) *Minimize the function*

$$\frac{1}{2}\|y - y_0\|^2 + \varphi(u) \tag{3.139}$$

on the set of all points $y \in V$, $u \in U$, *which satisfy the equation*

$$Ay - Bu = 0. \tag{3.140}$$

Here, φ is a lower-semicontinuous convex function from U to $]-\infty, +\infty]$ and y_0 is a fixed element of V.

The above problem, which was not formulated in the most general framework, represents a typical example of control problem. The parameter u is called *control*, while the solution y is called *state*.

Since relation (3.138) implies $A^{-1} \in L(V', V)$, Problem $\mathscr{P}$ can be expressed as

$$\min\left\{\varphi(u) + \frac{1}{2}\left\|A^{-1}Bu - y_0\right\|^2;\ u \in U\right\}. \tag{3.141}$$

Denote $g(y) = -\frac{1}{2}\|y - y_0\|^2$. Then, problem (3.141) becomes

$$\inf\left\{\varphi(u) - g\left(A^{-1}Bu\right);\ u \in U\right\}$$

and the associated dual problem can be written as

$$\max\left\{(y_0, y^*) - \frac{1}{2}\|y^*\|_*^2 - \varphi^*\left(B^*(A^*)^{-1}y^*\right);\ y^* \in V'\right\}. \tag{3.142}$$

If we write $p = (A^*)^{-1}y^*$, then it is obvious that the dual problem (3.142) may be also regarded as an optimal control problem. Thus, we are confronted with the following problem.

($\mathscr{P}^*$) *Maximize the function*

$$-\varphi^*(B^*p) - \frac{1}{2}\|y^*\|_*^2 + (y_0, y^*) \tag{3.143}$$

on the set of all pairs of points $p \in V$ and $y^* \in V'$ which satisfy the equation

$$A^*p = y^*. \tag{3.144}$$

By virtue of the same theorem, the pair $(\bar{u}, \bar{y}^*)$ is optimal if and only if the system

$$\left(A^{-1}B\right)^*\bar{y}^* \in \partial\varphi(\bar{u}), \quad \bar{y}^* \in \partial g\left(A^{-1}B\bar{u}\right)$$

is verified, that is,

$$B^*\bar{p} \in \partial\varphi(\bar{u}), \qquad \bar{y}^* + JA^{-1}B\bar{u} = -Jy_0,$$

where $J : V \to V'$ is the canonical isomorphism from V into V'.

Consequently, $(\bar{y}, \bar{u})$ is an optimal solution of the problem and $(\bar{p}, \bar{y}^*)$ is an optimal solution of the dual problem if and only if the system

$$\begin{aligned} &A\bar{y} = B\bar{u}, \qquad A^*\bar{p} = \bar{y}^*, \\ &B^*\bar{p} \in \partial\varphi(\bar{u}), \qquad \bar{y}^* + J\bar{y} + Jy_0 = 0, \end{aligned} \tag{3.145}$$

is verified.

At the same time, we note that system (3.145) allows an explicit calculation of the optimal controls $\bar{u}$ and $\bar{y}^*$ with respect to the primal optimal state $\bar{y}$ and the adjoint state $\bar{p}$.

The previous statements will be illustrated by the following example:

Minimize the functional

$$\frac{1}{2}\left(\int_\Omega \left|\operatorname{grad}(y - y_0)\right|^2 \mathrm{d}x + \int_\Omega |u|^2 \,\mathrm{d}x\right) \tag{3.146}$$

on the class of all functions $y \in H_0^1(\Omega)$ and $u \in K = \{u \in L^2(\Omega), u \geq 0$, a.e., on $\Omega\}$ which verify the equation

$$\begin{aligned} &-\Delta y \quad \text{on } \Omega, \\ &y = 0 \quad \text{on } \Gamma. \end{aligned} \tag{3.147}$$

Here y_0 is a given function in $H_0^1(\Omega)$.

This problem may be written as a problem of type $\mathscr{P}$ by taking

$$V = H_0^1(\Omega), \qquad H = L^2(\Omega), \qquad U = L^2(\Omega), \qquad B = I, \qquad A = -\Delta,$$

and

$$\varphi(u) = \frac{1}{2}\int_\Omega |u|^2 \,\mathrm{d}x + I_K(u), \quad u \in L^2(\Omega),$$

where I_K is the indicator function of the convex cone K. It is clear that the conjugate function φ^* is defined by

$$\varphi^*(u) = (u, P_K u) - \frac{1}{2}|P_K u|^2,$$

where $P_K u = (I + I_K)^{-1}u$, $u \in L^2(\Omega)$ is the projection operator in $L^2(\Omega)$ on the set K. In other words,

$$P_K u(x) = \max\{0, u(x)\} = u^+(x), \quad \text{a.e. on } \Omega.$$

Then, the dual problem can be written as

$$\min \frac{1}{2}\|v^*\|^2_{H^{-1}(\Omega)} - (v^*, y_0) + \frac{1}{2}\int_\Omega |p^+|^2 \,\mathrm{d}x,$$

where $v^* \in H^{-1}(\Omega)$ and $p \in H_0^1(\Omega)$ satisfy the equation

$$\begin{aligned} -\Delta p &= v^* \quad \text{on } \Omega, \\ p &= 0 \quad \text{on } \Gamma. \end{aligned}$$

In our case, the optimality system associated with Problem $\mathscr{P}$ becomes

$$\begin{aligned} &\Delta \bar{y} + \bar{u} = 0, \qquad \Delta \bar{p} + \bar{v}^* = 0 \quad \text{on } \Omega, \\ &\bar{y} + \bar{p} = 0, \qquad \bar{u} = \bar{p}^+ \quad \text{on } \Omega, \\ &\bar{v}^* - \Delta \bar{y} - \Delta y_0 = 0 \quad \text{on } \Omega, \end{aligned} \tag{3.148}$$

or, equivalently,

$$\begin{aligned} &\Delta \bar{y} + \bar{p}^+ = 0, \quad \text{on } \Omega, \\ &\Delta \bar{p} - \bar{p}^+ + \Delta y_0 = 0 \quad \text{on } \Omega, \\ &\bar{p} = \bar{y} = 0 \quad \text{on } \Gamma. \end{aligned}$$

The optimal controls $\bar{u}$ and $\bar{v}^*$ are defined by

$$\bar{u} = \bar{p}^+ = \max(0, \bar{p}), \qquad \bar{v}^* = \Delta y_0 - \bar{p}^+.$$

Like problem (3.142), the latter is a unilateral elliptic problem of the same type as that previously studied in Sect. 2.2.5.

Example 3.80 Consider now the problem

$$\min_{u \in H_0^1(\Omega)} \left\{ \frac{1}{2}\int_\Omega |\operatorname{grad} u|^2 \,\mathrm{d}x + \int_\Omega |\operatorname{grad} u| \,\mathrm{d}x - \int_\Omega hu \,\mathrm{d}x \right\}, \tag{3.149}$$

where the function $h \in L^2(\Omega)$ is given. Problem (3.149) arises in the study of the nonnewtonian fluids and in other problems of physical interest (see Duvaut and

Lions [35]). To construct the dual problem, we apply Theorem 3.53 in which $X = H_0^1(\Omega)$, $X^* = H^{-1}(\Omega)$, $Y = (L^2(\Omega))^n$

$$Au = \left(\frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_n}\right) = \operatorname{grad} u,$$

$$f(u) = \frac{1}{2}\int_\Omega |\operatorname{grad} u|^2 \, dx - \int_\Omega hu \, dx$$

and $g : (L^2(\Omega))^n \to]-\infty, +\infty[$ is defined by

$$g(y) = -\int_\Omega |y| \, dx, \quad y = (y_1, \dots, y_n).$$

Then, we obtain (see Example 3.78)

$$f^*(p^*) = \frac{1}{2}\|p^* + h\|^2_{H^{-1}(\Omega)}, \quad p^* \in H^{-1}(\Omega),$$

$$g^*(y^*) = \inf\left\{\int_\Omega \big((y^*, y) + |y|\big) \, dx;\ y \in \big(L^2(\Omega)\big)^n\right\},$$

where $y^* \in (L^2(\Omega))^n$. Thereby, we obtain

$$g^*(y^*) = \begin{cases} 0, & \text{if } |y^*(x)| \le 1, \text{ a.e. on } \Omega, \\ -\infty, & \text{otherwise.} \end{cases}$$

Hence, the dual problem associated with problem (3.149) is

$$\sup\left\{-\frac{1}{2}\|h - \operatorname{div} p^*\|^2_{H^{-1}(\Omega)};\ p^* \in \big(L^2(\Omega)\big)^n,\ |p^*| \le 1, \text{ a.e. on } \Omega\right\}. \tag{3.150}$$

Now, let us find the optimality system. Using again Theorem 3.53, it follows that $(\bar{u}, \bar{p}^*) \in H_0^1(\Omega) \times (L^2(\Omega))^n$ is an optimal pair for problem (3.149), respectively, for its dual (3.150) if and only if

$$A^*\bar{p}^* \in \partial f(\bar{u}), \quad \bar{p}^* \in \partial g(A\bar{u}). \tag{3.151}$$

On the other hand, we have

$$\partial f(u) = -\Delta\bar{u} - h,$$

$$A^*\bar{p}^* = -\sum_{i=1}^n \frac{\partial}{\partial x_i}\,\bar{p}_i^*, \quad \bar{p}^* = (\bar{p}_1^*, \dots, \bar{p}^*).$$

Then, the first equation of relation (3.151) can be written as

$$-\Delta\bar{u} + \sum_{i=1}^n \frac{\partial}{\partial x_i}\,\bar{p}_i^* = h \quad \text{on } \Omega.$$

From the second equation of relation (3.151), it follows that

$$\int_{\Omega} \bar{p}^*(\operatorname{grad}\bar{u} - v)\,\mathrm{d}x \le \int_{\Omega} \big(|v| - |\operatorname{grad}\bar{u}|\big)\,\mathrm{d}x, \quad \forall v \in \big(L^2(\Omega)\big)^n.$$

Hence,

$$\int_{\Omega} \big(\bar{p}^*\operatorname{grad}\bar{u} + |\operatorname{grad}\bar{u}|\big)\,\mathrm{d}x \le \int_{\Omega} \big(|v| + \bar{p}^* v\big)\,\mathrm{d}x.$$

Since v is arbitrarily chosen in $(L^2(\Omega))^n$ and

$$|\bar{p}^*| = \left(\sum_{i=1}^{n} |\bar{p}_i^*|^2\right)^{\frac{1}{2}} \le 1 \quad \big(\text{because } \bar{p}^* \in \operatorname{Dom}(g^*)\big)$$

the above inequality implies

$$\bar{p}^*\operatorname{grad}\bar{u} + |\operatorname{grad}\bar{u}| = 0, \quad \text{a.e. in } \Omega.$$

Consequently, $\bar{u}$ is a solution of problem (3.149) and $\bar{p}^*$ is a solution of the dual problem (3.150) if and only if they verify the system

$$\begin{aligned}
&-\Delta\bar{u} + \operatorname{div}\bar{p}^* = h \quad \text{on } \Omega,\\
&\sum_{i=1}^{n} \bar{p}_i^* \frac{\partial\bar{u}}{\partial x_i} + |\operatorname{grad}\bar{u}| = 0 \quad \text{a.e. on } \Omega,\\
&\bar{u} = 0 \quad \text{on } \Gamma,\\
&\bar{p}^* \in \big(L^2(\Omega)\big)^n \quad \text{and} \quad \sum_{i=1}^{n} |\bar{p}_i^*|^2 \le 1 \quad \text{a.e. on } \Omega.
\end{aligned}$$

It is interesting to note that the system $\bar{p}^* = (\bar{p}_1^*, \ldots, \bar{p}_n^*) \in (L^2(\Omega))^n$ can be regarded as a system of Lagrange multipliers for problem (3.149).

Example 3.81 Detection filter problem (Fortmann, Athans [42]). This problem can be expressed as the following maximization problem:

$$\begin{aligned}
&\max\big\{\langle u, x\rangle : u \in L^2[0,T]\big\} \quad \text{subject to}\\
&\langle u, s_t\rangle \le \varepsilon\langle u, s\rangle, \qquad \delta \le |t| \le T,\\
&-\langle u, s_t\rangle \le \varepsilon\langle u, s\rangle, \qquad \delta \le |t| \le T,\\
&\|u\| \le 1,
\end{aligned} \tag{3.152}$$

where s is the signal function.

Suppose that $s : R \to \mathbb{R}$ is continuous with $\operatorname{supp} s \subset [0,T]$ and

$$\|s\|^2 = \int_{-\infty}^{\infty} s^2(t)\,\mathrm{d}t = \int_0^T s^2(t)\,\mathrm{d}t = 1, \tag{3.153}$$

that is, the energy of s equals 1.

Problem (3.152) can be considered as a problem of type $\mathscr{P}_2$ (see Sect. 3.1.2) by taking $X = A = L^2[0, T]$, $f(u) = -\langle u, s\rangle$, $Y = C[-T, T] \times C[-T, T] \times \mathbb{R}$ and $G = (G_1, G_2, G_3) : X \to Y$ defined by $(G_i u)(t) = -(-1)^i \langle u, s_t\rangle - \varepsilon\langle u, s\rangle$, $t \in [-T, T]$, $i = 1, 2$, and $G_3 u = \|u\|^2 - 1$.

Now, it suffices to consider the cone $A_Y = C^-[-T, T] \times C^-[-T, T] \times \mathbb{R}$, where $C^-[-T, T] = \{x \in C[-T, T];\ x(t) \le 0, \delta \le |t| \le T\}$.

It is clear that problem (3.152) has a unique optimal solution since the unit ball of a Hilbert space is weakly compact and its indicator function is strictly convex. If we suppose that the Slater condition holds, that is, there exists $u \in L^2[-T, T]$ such that $G(u) \in \operatorname{int} A_Y$, then there exists the Lagrange multiplier $y_0^* = (p_1^0, p_2^0, p_3^0)$. Here, p_1^0, p_2^0 are positive regular Borel measures on $[-T, T]$ which are equal to zero on $(-\delta, \delta)$ and $p_3^0 \in \mathbb{R}_+$. Therefore, if $u^0 \in L^2[-T, T]$ is the optimal solution of Problem (3.152), then the Kuhn–Tucker conditions become: u^0 *minimizes*

$$\begin{aligned} L\big(u; p_1^0, p_2^0, p_3^0\big) = -\langle u, s\rangle &+ \int_{-T}^{T} \big[\langle u, s_t\rangle - \varepsilon\langle u, s\rangle\big]\, \mathrm{d}p_1^0 \\ &+ \int_{-T}^{T} \big[-\langle u, s_t\rangle - \varepsilon\langle u, s\rangle\big]\, \mathrm{d}_2^0 + p_3^0\big[\|u\|^2 - 1\big] \\ &\text{on } L^2[-T, T], \end{aligned} \tag{3.154}$$

$$\begin{aligned} p_1^0 G_1\big(u^0\big) &= \int_{-T}^{T} \big[\big\langle u^0, s_t\big\rangle - \varepsilon\big\langle u^0, s\big\rangle\big]\, \mathrm{d}p_1^0 = 0, \\ p_2^0 G_2\big(u^0\big) &= \int_{-T}^{T} \big[-\big\langle u^0, s_t\big\rangle - \varepsilon\big\langle u^0, s\big\rangle\big]\, \mathrm{d}p_2^0 = 0, \\ p_3^0 G_3\big(u^0\big) &= p_3^0\big[\big\|u^0\big\|^2 - 1\big] = 0. \end{aligned} \tag{3.155}$$

It is easily seen that $p_3^0 > 0$, hence $\|u^0\| = 1$.

Also, (u^0, p^0) is a saddle point of the Lagrangian. We note that the Lagrangian can be written as

$$L(u; p_1, p_2, p_3) = -\big\langle u, \hat{u}(p_1, p_2)\big\rangle + p_3\big(\|u\|^2 - 1\big),$$

where $\hat{u}(p_1, p_2) \in L^2[0, T]$ is defined by

$$\begin{aligned} \hat{u}(p_1, p_2) &= s + \int_{-T}^{T} (-s_t + \varepsilon s)\, \mathrm{d}p_1 + \int_{-T}^{T} (s_t + \varepsilon s)\, \mathrm{d}p_2 \\ &= s\big[1 + \varepsilon\big(\|p_1\| + \|p_2\|\big)\big] - \int_{-T}^{T} s_t\, \mathrm{d}p_1 + \int_{-T}^{T} s_t\, \mathrm{d}p_2. \end{aligned}$$

Thus, we have

$$L(u;p_1,p_2,p_3)=p_3\left\|u-\frac{1}{2p_3}\,\hat{u}(p_1,p_2)\right\|^2-\frac{1}{4p_3}\left\|\hat{u}(p_1,p_2)\right\|^2-p_3.$$

Using the minimax duality generated by the Lagrangian, let us find now the dual problem of (3.152).

Since $u=\frac{1}{2p_3}\,\hat{u}(p_1,p_2)$ minimizes L, we obtain (see Sect. 3.2.3) the following dual problem:

$$\max\left\{-p_3-\frac{1}{4p_3}\left\|\hat{u}(p_1,p_2)\right\|^2;\ p_1\ge 0,\ p_2\ge 0,\ p_3>0\right\}. \tag{3.156}$$

Hence, the derivative with respect to p_3 must be equal to zero

$$-1+\frac{\|\hat{u}(p_1,p_2)\|^2}{4p_3^2}=0,$$

that is,

$$p_3^0(p_1,p_2)=\frac{1}{2}\|\hat{u}(p_1,p_2)\|.$$

Therefore, the dual problem (3.156) becomes

$$\min\left\|\hat{u}(p_1,p_2)\right\|^2;\quad p_1,p_2\in M_0[-T,T],$$

where $M_0[-T,T]$ is the space of all regular Borel measures which are zero on $(-\delta,\delta)$.

Since the constraints are not simultaneously active, the measures p_1,p_2 are mutually singular. Taking $p=p_1-p_2$, we have

$$\hat{u}(p_1,p_2)=\hat{u}(p)=s+\varepsilon s\|p\|-\int_{-T}^{T}s_t\,\mathrm{d}p.$$

Thus, we obtain the following final form for the dual problem:

$$\min\left\{\left\|\hat{u}(p)\right\|^2;\ p\in M_0[-T,T]\right\}.$$

3.3 Applications of the Duality Theory

We discuss below a few applications of the duality theory on some specific convex optimization problems.

3.3.1 Linear Programming

Let $b \in \mathbb{R}^n$ and $c \in \mathbb{R}^m$ be fixed and let $A : \mathbb{R}^n \to \mathbb{R}^m$ be a linear operator. Denote by $\langle \cdot, \cdot \rangle_n$ and $\langle \cdot, \cdot \rangle_m$ the usual inner products in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively.

The basic problem of the finite-dimensional linear programming can be expressed as

$$(\mathscr{P}_c) \quad \min\left\{\langle x, b\rangle_n;\ x \in \mathbb{R}^n,\ x \geq 0,\ Ax \geq c\right\}$$

and is termed the *canonical form*, or can be written in *standard form*

$$(\mathscr{P}_s) \quad \min\left\{\langle x, b\rangle_n;\ x \in \mathbb{R}^n,\ x \geq 0,\ Ax = c\right\}.$$

The equivalence of these forms follows from the fact that, on the components, every equality can be replaced by two inequalities and conversely, each inequality becomes an equality by introducing a new nonnegative variable.

In the following text, we only use the canonical form of linear programming problems because in this case we may impose some interiority conditions. It is easy to see that the minimizing problem $\mathscr{P}_c$ with operational constraints is a problem of the type $\mathscr{P}_1$ described in the preceding section.

From this, it is enough to take $X = X^* = \mathbb{R}^n$ and $Y = Y^* = \mathbb{R}^m$ and to choose the functions f and g as

$$f(x) = \begin{cases} \langle x, b\rangle_n, & \text{if } x \geq 0, \\ +\infty, & \text{otherwise,} \end{cases}$$

$$g(y) = \begin{cases} 0, & \text{if } y \geq c, \\ -\infty, & \text{otherwise.} \end{cases}$$

Thus, to determine the dual of problem $\mathscr{P}_c$ as discussed in Sect. 3.2.2, the conjugates of the functions f and g are needed. We have

$$\begin{aligned} f^*(u) &= \sup\left\{\langle u, x\rangle_n - f(x);\ x \in \mathbb{R}^n\right\} = \sup\left\{\langle u - b, x\rangle_n;\ x \geq 0\right\} \\ &= \begin{cases} 0, & \text{if } x \geq 0, \\ +\infty, & \text{otherwise,} \end{cases} \end{aligned}$$

for all $u \in \mathbb{R}^n$,

$$\begin{aligned} g^*(v) &= \inf\left\{\langle v, y\rangle_m - g(y);\ y \in \mathbb{R}^m\right\} = \inf\left\{\langle v, y\rangle_n;\ y \geq c\right\} \\ &= \begin{cases} \langle v, c\rangle_m, & \text{if } v \geq 0, \\ -\infty, & \text{otherwise,} \end{cases} \end{aligned}$$

for all $v \in \mathbb{R}^m$. Therefore,

$$g^*(v) - f^*(A^*v) = \begin{cases} \langle v, c\rangle_m, & \text{if } v \geq 0 \text{ and } A^*v \leq b, \\ -\infty, & \text{otherwise.} \end{cases}$$

Thus, the dual problem is

$$(\mathscr{P}_c^*) \quad \max\big\{\langle v, c\rangle_m;\ v \in \mathbb{R}^m,\ v \geq 0,\ A^* v \leq b\big\},$$

which is a maximization problem, of the same type as $\mathscr{P}_c$, on $\mathbb{R}^m$.

In the standard form, the *dual problem* is expressed as

$$(\mathscr{P}_s^*) \quad \max\big\{\langle v, c\rangle_m;\ v \in \mathbb{R}^m,\ A^* v \leq b\big\}.$$

It is obvious that the programs $\mathscr{P}_c$, $\mathscr{P}_c^*$ and $\mathscr{P}_s$, $\mathscr{P}_s^*$, respectively, are dual to each other.

If the consistency condition (3.93) is satisfied for $\mathscr{P}_c$ and $\mathscr{P}_c^*$, then, from Theorem 3.53, it follows that both problems $\mathscr{P}_c$ and $\mathscr{P}_c^*$ have optimal solutions. In this way, we have obtained the following general result.

Theorem 3.82 *A feasible element x_0 of $\mathscr{P}_c$ is an optimal solution if and only if there exists a feasible element v_0 of $\mathscr{P}_c^*$ such that*

$$\langle x_0, b\rangle_s = \langle v_0, c\rangle_m.$$

On the other hand, the *extremality conditions* (see Theorem 3.54) for a point $(x_0, v_0) \in \mathbb{R}^n \times \mathbb{R}^m$ are the following:

$$\begin{aligned} &x_0 \geq 0,\ b - A^* v_0 \geq 0, \quad \langle x_0, b\rangle_n = \langle x_0, A^* v_0\rangle_n, \\ &A x_0 - c \geq 0,\ v_0 \geq 0, \quad \langle A x_0, v_0\rangle_m = \langle c, v_0\rangle_m. \end{aligned}$$

The Kuhn–Tucker function is

$$K(x, v) = \langle x, b\rangle_n + \langle v, c\rangle_m, \quad \forall x \geq 0,\ v \geq 0.$$

It is easy to see that an element is a saddle point if and only if it satisfies the above extremality conditions.

In general, if (x, v) is a pair of feasible elements of $\mathscr{P}_c$ and $\mathscr{P}_c^*$, then the relation

$$\langle v, c\rangle_m \leq \langle x, b\rangle_n$$

holds.

Finally, we remark that, if one of Problems $\mathscr{P}_c$ or $\mathscr{P}_c^*$ is inconsistent, then the other is inconsistent or unbounded. If both $\mathscr{P}_c$ and $\mathscr{P}_c^*$ are consistent, then they have an optimal solution and their extreme values are equal.

As another simple utilization of the Fenchel duality, we can consider the *infinite-dimensional linear programming*. Thus, we have the basic problem

$$(\mathscr{P}) \quad \min\big\{(x_0^*, x);\ x \in P,\ y_0 - Ax \in Q\big\},$$

where X, Y are two Banach spaces, $P \subset X$, $Q \subset Y$ are closed convex cones, $A : X \to Y$ is a linear continuous operator and $x_0^* \in X^*$, $y_0 \in Y$.

We easily see that $\mathscr{P}$ can be obtained from $\mathscr{P}_1$, by taking $f = x_0^* + I_P$ and $g = -I_{y_0 - Q}$. Hence, we obtain

$$\begin{aligned} f^*(x^*) &= \sup\{(x^* - x_0^*, x);\ x \in P\} = I_{P^0}(x^* - x_0^*), \\ g^*(y^*) &= \{(y^*, y);\ y \in y_0 - Q\} = (y^*, y_0) - \sup\{(y^*, y);\ y \in Q\} \\ &= (y^*, y_0) - I_{Q^0}(y^*). \end{aligned}$$

Therefore, a dual problem associated with $\mathscr{P}$ is

$$(\mathscr{P}^*) \quad \max\{(y^*, y_0);\ y^* \in Q^0,\ A^* y^* - x_0^* \in P^0\}.$$

By virtue of Proposition 3.39, it follows that, if x is a feasible element of $\mathscr{P}$ and y is the feasible element of $\mathscr{P}^*$, then

$$(y^*, y_0) \le (x_0^*, x).$$

This relation becomes an equality only for pairs of optimal solutions.

The two problems $\mathscr{P}$ and $\mathscr{P}^*$ have finite and equal extreme values if and only if the function

$$h(y) = \inf\{(x_0^*, x);\ x \in P,\ y_0 + y - Ax \in Q\}$$

is finite and lower-semicontinuous at the origin of Y. Moreover, according to Theorem 2.22 it follows that $\mathscr{P}$ is stable, that is, $\mathscr{P}^*$ has optimal solutions and $\inf \mathscr{P} = \sup \mathscr{P}^* \in \mathbb{R}$, if the consistency condition

$$(y_0 - \operatorname{int} Q) \cap A(P) \neq \emptyset$$

holds. Also, the existence of optimal solutions for $\mathscr{P}$ is guaranteed by the dual consistency condition

$$\left(x_0^* + \operatorname{int} P^0\right) \cap A^*\left(Q^0\right) \neq \emptyset.$$

For the Kuhn–Tucker function, we obtain

$$K(x, y^*) = (x_0^*, x) + (y^*, y_0) - (Ax, y^*), \quad \forall (x, y^*) \in P \times Q^0.$$

Hence, $(x_0, y_0^*) \in P \times Q^0$ is a pair of solutions of $\mathscr{P}$ and $\mathscr{P}^*$, if and only if (x_0, y_0^*) is a saddle point of K on $P \times Q^0$. On the other hand, the existence of optimal solutions of $\mathscr{P}$, for every $y_0 \in Y$, can be characterized by using Theorem 3.72.

Theorem 3.83 *Suppose that $P \cap A^{-1}(\bar{y} - Q) \neq \emptyset$ for at least element $\bar{y} \in Y$. Then, $\mathscr{P}$ has optimal solutions for every $y_0 \in Y$ and its value is equal to dual value* (*that is*, *$\mathscr{P}^*$ is stable*) *if and only if the set*

$$H = (A \times x_0^*)(P) + Q \times \mathbb{R}_+, \tag{3.157}$$

is closed in $Y \times \mathbb{R}$, where $A \times x_0^* : X \to Y \times \mathbb{R}$ is defined by $(A \times x_0^*)(x) = (Ax, (x_0^*, x))$.

Using some closedness criteria of the sum of two closed convex sets, we obtain various optimality conditions. We obtain a special case if P or Q is also locally compact.

Let us consider the following properties:

(i) $x \in P \cap A^{-1}(Q)$ *and* $(x_0^*, x) \le 0$ *implies* $x = 0$
(ii) $x \in P \cap A^{-1}(Q)$ *and* $(x_0^*, x) \le 0$ *implies* $Ax = 0$ *and* $(x_0^*, x) = 0$
(iii) $x \in P \cap A^{-1}(Q)$ *and* $(x_0^*, x) \le 0$ *implies* $Ax = 0$.

Theorem 3.84 *The set* H *given by* (3.157) *is closed if one of the following three conditions is fulfilled*:

(c$_1$) P *is a locally compact cone and* (i) *is satisfied*
(c$_2$) Q *is a locally compact cone*, $(A \times x_0^*)(P)$ *is closed and* (ii) *is satisfied*
(c$_3$) Q *is a locally compact cone*, (iii) *is satisfied and the set*

$$\big\{\big(Ax, (x_0^*, x) + r\big) \in Y \times \mathbb{R};\ x \in P,\ r \ge 0\big\}$$

is closed.

Proof This result is a consequence of closedness Dieudonné's criterion (see Theorem 1.59 and its Corollaries 1.60, 1.61).

Now, we consider the problem

$$(\overline{\mathscr{P}}) \qquad \min\{t;\ tx_1 + x_2 \in K\},$$

where $K \subset X$ is a closed convex set which contains the origin, and $x_1, x_2 \in X$ are two fixed elements. Taking $f(t) = t$, $t \in \mathbb{R}$, $g(x) = -I_{K-x}$, $x \in X$, and $A(t) = tx_1$, $t \in \mathbb{R}$, we find the dual problem

$$(\overline{\mathscr{P}}^*) \qquad \max\big\{(x_2, x^*) - p_{K^0}(-x^*);\ (x_1, x^*) = -1\big\}. \qquad \square$$

It is clear that, if $\overline{\mathscr{P}}$ is consistent and $\inf \overline{\mathscr{P}} > -\infty$, then it has optimal solutions; the extreme value of $\overline{\mathscr{P}}$ is the lower bound of a real segment. As a consequence of Theorem 3.53, we obtain Theorem 3.85.

Theorem 3.85 *Assume that* $\operatorname{int} K \ne \emptyset$ *and* $\bar{t} \in \mathbb{R}$ *exists such that* $\bar{t}x_1 + x_2 \in \operatorname{int} K$. *If* $\overline{\mathscr{P}}$ *has finite value, then* $\overline{\mathscr{P}}^*$ *has optimal solutions and the two extreme values are equal.*

In fact, one and only one of the following three assertions is true:

(i) $\overline{\mathscr{P}}$ *and* $\overline{\mathscr{P}}^*$ *are consistent.*
(ii) *One, and only one, of* $\overline{\mathscr{P}}$ *and* $\overline{\mathscr{P}}^*$ *is consistent and has an infinite value.*
(iii) *Both problems are inconsistent.*

Remark 3.86 If K is a cone, the dual problem $\overline{\mathscr{P}}^*$ becomes

$$\max\big\{(x_2, x^*);\ x^* \in -K^0,\ (x_1, x^*) = -1\big\}.$$

3.3.2 The Best Approximation Problem

Let C be a convex subset of a real linear normed space X. An element $z \in C$ is called a *best approximation* to $x \in X$ from C if

$$\|x - z\| \le \|x - u\|, \quad \text{for all } u \in C. \tag{3.158}$$

Denote

$$P_C(x) = \{z \in C;\ \|x - z\| \le \|x - u\|,\ \text{for all } u \in C\}.$$

The multivalued mapping $x \to P_C(x)$, $x \in X$, is called the *projection mapping* of the space X into C. Obviously, if C is convex, then $P_C(x)$ is a convex subset of C (possible empty) for every $x \in X$ and $\|x - z\| = d(x; C)$ for all $z \in P_C(x)$.

Now, we establish a simple property concerning the best approximation elements of an element in semistraight line determined by x and $z \in P_c(x)$.

Theorem 3.87 *If C is a nonvoid set in X, $x \overline{\in} C$, and $z \in P_C(x)$, then*

(i) $z \in P_C(\lambda x + (1-\lambda)z)$ *for all* $\lambda \in (0, 1)$
(ii) $z \in P_C(\lambda x + (1-\lambda)z)$ *for all* $\lambda > 1$, *whenever the set C is convex.*

Proof (i) If $\lambda \in (0, 1)$, we have

$$\begin{aligned}\|\lambda x + (1-\lambda)z - z\| &= \|x - z\| - (1-\lambda)\|x - z\| \\ &\le \|x - y\| - \|x - (\lambda x + (1-\lambda)z\| \\ &\le \|x - y - x + (\lambda x + (1-\lambda)z)\| = \|\lambda x + (1-\lambda)z - y\|,\end{aligned}$$

for all $y \in C$, that is, $z \in P_C(\lambda x + (1-\lambda)z)$ if $\lambda \in (0, 1)$.

(ii) If $z \in P_C(x)$ and $\lambda > 1$, we have

$$\begin{aligned}\|\lambda x + (1-\lambda)z - z\| &= \lambda\|x - z\| \le \lambda\left\|x - \left(\frac{1}{\lambda} y + \left(1 - \frac{1}{\lambda}\right)z\right)\right\| \\ &= \|\lambda x + (1-\lambda)z - y\|, \quad \text{for all } y \in C,\end{aligned}$$

since, by convexity of C, it follows that

$$\frac{1}{\lambda} y + \left(1 - \frac{1}{\lambda}\right)z \in C.$$

This proves that $z \in P_C(\lambda x + (1-\lambda)z)$. □

It is obvious that an element of the best approximation is an optimal solution of the minimization problem

$$\min\left\{\frac{1}{2}\|u - x\|^2 + I_C(u);\ u \in X\right\},$$

where x is the given element of X.

According to Remark 3.2, we may infer that z is a best approximation to x from C, if and only if there exists $x_0^* \in X^*$ subject to

$$f(z) + f^*(x_0^*) \le (x_0^*, u), \quad \text{for all } u \in C, \tag{3.159}$$

where $f(u) = \frac{1}{2}\|u - x\|^2$, $u \in X$. As a consequence of this fact, we obtain the following theorem.

Theorem 3.88 *An element $z \in C$ is a best approximation to $x \in X$ from elements of the convex set C if and only if there exists $x_0^* \in X^*$ such that*

(i) $\|x_0^*\| = \|z - x\|$
(ii) $(x_0^*, u - x) \ge \|z - x\|^2, \ \forall u \in C$.

Proof Indeed, we have

$$\begin{aligned} f^*(x_0^*) &= \sup\left\{(x_0^*, u) - \frac{1}{2}\|u - x\|^2;\ u \in X\right\} \\ &= (x_0^*, x) + \sup\left\{(x_0^*, u) - \frac{1}{2}\|u\|^2;\ u \in X\right\} = (x_0^*, x) + \frac{1}{2}\|x_0^*\|^2 \end{aligned}$$

and the optimality condition (3.159) becomes

$$\frac{1}{2}\|z - x\|^2 + \frac{1}{2}\|x_0^*\|^2 \le \left(x_0^*, u - x\right), \quad \forall u \in C. \tag{3.160}$$

In particular, for $u = z$, we obtain $(\|z - x\| - \|x_0^*\|) \le 0$, which implies condition (i). Consequently, from inequality (3.160), condition (ii) follows, as claimed. Conversely, it is clear that conditions (i) and (ii) imply that z is a best approximation, because we have

$$\|z - x\|^2 \le (x_0^*, u - x) \le \|x_0^*\|\,\|u - x\| = \|z - x\|\,\|u - x\|, \quad \forall u \in C,$$

and therefore we must have (3.158). □

Corollary 3.89 *If $z \in C$ is a best approximation of $x \in X$ by elements of the convex set C, then the minimax relation*

$$\|x - z\| = \min_{u \in C}\ \max_{\|x^*\|=1}\ (x^*, u - x) = \max_{\|x^*\|=1}\ \min_{u \in C}\ (x^*, u - x) \tag{3.161}$$

holds.

Proof This follows with clarity if we use the relationship between the solutions to Problem $\mathscr{P}$ and property (2.133) of the saddle points and relation (1.36). To this end, it suffices to remark that the point x_0, the existence of which is ensured by the above theorem, is just the solution of the dual problem. □

Remark 3.90 Generally, if x is a point situated at the distance $d > 0$ from the convex set C, we obtain a weak minimax relation by replacing "min" by "inf" because in such a case only the dual problem has solutions (see Theorem 3.70).

Next, we note several special cases in which conditions (i) and (ii) have a simplified form. Namely, if C is a convex cone with vertex in the origin, then condition (ii) is equivalent to the following pair of conditions:

(ii′) $(x_0^*, u) \le 0$, $\forall u \in C$, that is, $x_0^* \in C^0$
(ii″) $(x_0^*, x) = \|x - z\|^2$.

Here is the argument. From condition (ii), replacing x_0^* by $-x_0^*$, we obtain $(x_0^*, x - nu) \ge \|x - z\|^2$, $\forall u \in C$, $\forall n \in \mathbb{N}$, because C is a cone. Therefore, we cannot have $(x_0^*, u) > 0$ for some element $u \in C$, that is, (ii′) holds.

Moreover, from Properties (ii) and (ii′) it follows that

$$\|x - z\|^2 \le (x_0^*, x - z) \le \|x_0^*\| \, \|x - z\| = \|x - z\|^2;$$

hence $(x_0^*, x - z) = \|z - x\|^2$. Thus, we have

$$0 \ge (x_0^*, z) = (x_0^*, x) - (x_0^*, x - z) = (x_0^*, x) - \|x - z\|^2.$$

On the other hand, from (ii), for $-x_0^*$ and $u = 0$, we obtain the inequality

$$(x_0^*, x) \ge \|x - z\|^2,$$

which implies Property (ii″). The reciprocal is obvious.

When C is a linear space, Condition (ii′) is equivalent to

$$(x_0^*, u) = 0, \quad \forall u \in C,$$

because, in this case, $C = -C$.

It should be mentioned that the best approximation belongs to $C \cap \overline{S}(x; d)$ and it exists if and only if there exist separating hyperplanes which meet C. Moreover, the set of all the best approximations is convex and coincides with the intersection of the set with any separating hyperplanes. When this intersection is nonempty, the separating hyperplanes is a support hyperplane and is given by the equation

$$(x_0^*, u - x) = \|x - z\|^2, \quad u \in X.$$

Now, let us study the existence of the best approximations. Let

$$d = \inf\big\{\|u - x\|;\ u \in C\big\}. \tag{3.162}$$

Firstly, it is obvious that if a minimizing sequence, that is $(u_n)_{n\in\mathbb{N}} \subset C$ and $\|u_n - x\| \to d$, has a convergent subsequence in C, then its limit is a best approximation element. The set of this type is called an *approximatively compact set*. Any approximatively compact set is necessarily closed. For instance, any closed convex set in a Banach uniform convex space is approximatively compact. Indeed, using Proposition 1.59, it follows that any minimizing sequence is a Cauchy sequence, and therefore it is convergent.

We easily see that

$$\inf\{\|u-x\|;\ u\in C\}=\inf\{\|u-x\|;\ u\in C\cap\overline{S}(x;d+\varepsilon)\}, \tag{3.163}$$

where $\overline{S}(x;d+\varepsilon)=\{y\in X;\ \|y-x\|\le d+\varepsilon\},\ \varepsilon>0$.

Theorem 3.91 *If the convex set C is such that there exists an $\varepsilon>0$, for which the set $C\cap\overline{S}(x;d+\varepsilon)$ is weakly compact, then x has a best approximation in C.*

Proof According to relation (3.163), it suffices to recall that a lower-semicontinuous function on a compact set attains its infimum. In our case, the function is obviously weakly lower-semicontinuous (see Proposition 1.73) on the weakly compact set $C\cap\overline{S}(u;d+\varepsilon)$. □

Corollary 3.92 *In a reflexive Banach space, every element possesses at least one best approximation with respect to every closed convex set.*

Proof The set $C\cap\overline{S}(u;d+1)$ is convex closed and bounded and, hence, it is weakly compact by virtue of the reflexivity (see Theorem 1.94). □

Corollary 3.93 *In a Banach space, every element possesses at least one best approximation with respect to every closed, convex and finite-dimensional set.*

Proof In a finite-dimensional space, the bounded closed convex sets are compact and, hence, weakly compact. □

Remark 3.94 The existence of the best approximations for closed convex sets is a characteristic property of reflexive spaces: *a Banach space has the property that every element possesses best approximations with respect to every closed convex set if and only if it is reflexive.* It is clear that this characterization is equivalent to the property that every continuous linear functional attains its supremum on the unit ball (see James [56, 57]).

In the uniqueness of the best approximations, a crucial role is played by strictly convex spaces, while in the existence problem the same role is played by reflexive spaces.

Theorem 3.95 *If X is strictly convex, then each element $x\in X$ possesses at most one best approximation with respect to a convex set $C\subset X$.*

Proof Assume by contradiction that there exist two distinct best approximations, z_1, z_2 in C. Since the set of best approximations is convex, it follows that $\frac{1}{2}(z_1+z_2)$ is also a best approximation.

Hence, if $d = d(x; C)$, we have

$$0 < d = \|x - z_1\| = \|x - z_2\| = \left\| x - \frac{1}{2}(z_1 + z_2) \right\|$$

and, thereby

$$\left\| \frac{1}{d}(x - z_1) \right\| = \left\| \frac{1}{d}(x - z_2) \right\| = 1.$$

In view of the strict convexity (see Proposition 1.103(ii)), we have

$$1 > \left\| \frac{1}{2d}(x - z_1) + \frac{1}{2d}(x - z_2) \right\| = \frac{1}{d} \left\| x - \frac{1}{2}(z_1 + z_2) \right\| = 1,$$

which is a contradiction. □

Remark 3.96 This property is characteristic of the strictly convex spaces: *if, in a Banach space X, every element possesses at most a best approximation with respect to every convex set* (*it is enough for the segments*), *then X is strictly convex.*

Indeed, if we assume that X is not strictly convex, then there exist $x, y \in X$, $x \neq y$, with $\|x\| = \|y\| = \|\frac{1}{2}(x + y)\| = 1$. Furthermore, $\|\alpha x + (1 - \alpha)y\| = 1$, $\forall \alpha \in [0, 1]$. Hence, the origin has as the best approximation with respect to the closed convex set $[x, y]$ every element of this set, and this, clearly, contradicts the uniqueness.

From Corollary 3.92 and Theorem 3.91, it follows that in a reflexive strictly convex Banach space, for every closed convex set C, the domain of projection mapping P_C is whole X, that is, $P_C(x) \neq \emptyset$ for any $x \in X$.

Theorem 3.97 *If C is a closed locally compact convex set of a strictly convex space X, then the projection function is continuous on X.*

Proof If $x_n \to x$, for every $\varepsilon > 0$, then there exists $n_0(\varepsilon) \in \mathbb{N}$ such that $\|x_n - x\| < \varepsilon$ for all $n > n_0(\varepsilon)$. Denote

$$d_n = d(x_n; C) = \inf_{u \in C} \|x_n - u\|, \quad d = d(x; C) = \inf_{u \in C} \|x - u\|.$$

We have

$$d_n \leq \inf_{u \in C} \big\{ \|x - u\| + \|x_n - x\| \big\} < d + \varepsilon, \quad \forall n > n_0(\varepsilon),$$

hence,

$$\|x - P_C x_n\| \leq \|x_n - P_C x_n\| + \|x_n - x\| < d_n + \varepsilon < d + 2\varepsilon.$$

Since the locally compact convex set $C \cap \overline{S}(x; d + \varepsilon)$ does not contain any half-line, it follows that it is compact (see, for instance, Köthe [61], p. 340). Thus,

$$\bigcap_{\varepsilon > 0} \overline{S}(x; d + \varepsilon) \cap C \neq \emptyset$$

and any subsequence of $P_C x_n$ has a cluster point z which satisfies $\|x - z\| = d$. Because X is strictly convex, this point is unique and so, $P_C x_n \to z = P_C x$, as claimed. □

A set C is called *proximinal* if every element of X has a best approximation in C. That is, the set C is proximinal if the problem

$$\min\{\|x - u\|;\ u \in C\}$$

has solutions for every $x \in X$. Thus, by Theorem 3.64, we obtain the following characterization of proximinal sets.

Theorem 3.98 *A nonempty set C of a linear normed space X is proximinal if and only if* $\operatorname{epi}\|\cdot\| + C \times \{0\}$ *is closed in $X \times \mathbb{R}$. Moreover, if C is a convex set which contains the origin, we have*

$$\min\{\|x - u\|;\ u \in C\} = \max\{(x^*, x) - P_{C^0}(x^*);\ x^* \in X^* \cap C^0\},$$

for every $x \in X$.

Proof Taking, in Theorem 3.72, $f = I_C$, $g = -\|\cdot\|$, $A = I$, we observe that

$$H = \{(u + x, \|x\| + r) \in X \times \mathbb{R};\ u \in C,\ x \in X,\ r \geq 0\} = \operatorname{epi}\|\cdot\| + C \times \{0\},$$

as claimed. □

Finally, we establish a simple characterization of proximinal sets.

Theorem 3.99 *A nonempty set C in a linear normed space is proximinal if and only if $C + \overline{S}(0; r)$ is closed for any $r \geq 0$.*

Proof Let C be a proximinal set. If $x \in \overline{C + \overline{S}(0; r)}$, there exists a sequence $(a_n + b_n)_{n \in \mathbb{N}}$ convergent to x such that $(a_n)_{n \in \mathbb{N}} \subset C$, $\|b_n\| \leq r$, for all $n \in \mathbb{N}$. Hence, $d(a_n + b_n; C) \leq r$, $n \in \mathbb{N}$, and so, $d(x; C) \leq r$. Now, if $\bar{x} \in P_C(x)$, then $\|x - \bar{x}\| = d(x; C) \leq r$. Therefore, taking $\bar{y} = x - \bar{x}$, we have $x = \bar{x} + \bar{y} \in C + \overline{S}(0; r)$, that is, $C + \overline{S}(0; r)$ is closed.

Conversely, we consider an arbitrary element x and we denote $d(x; C) = r$. If $r = 0$, then $x \in C$ since C is closed. Hence, $x \in P_C(x)$. Therefore, we can suppose $r > 0$. Let $(x_n)_{n \in \mathbb{N}} \subset C$ be an approximant sequence $\|x - x_n\| \leq r + \frac{1}{n}$, $n \in \mathbb{N}^*$. If we denote

$$\frac{r}{r + \frac{1}{n}}\,(x - x_n) = y_n,$$

then $y_n \in \overline{S}(0; r)$ and

$$x = x_n + y_n + \frac{1}{nr} y_n, \quad n \in \mathbb{N}^*.$$

Since $x_n + y_n \in C + \overline{S}(0; r)$ and $\frac{1}{nr} y_n \to 0$, it follows that $x \in \overline{C + \overline{S}(0; r)}$. Thus, if $C + \overline{S}(0, r)$ is closed, we get $x \in C + \overline{S}(0, r)$, that is, there exists $\bar{x} \in C$ such that $\|x - \bar{x}\| \le r$. Therefore, $\bar{x} \in P_C(x)$, which proves that the set C is proximinal. □

Remark 3.100 In the case of linear closed subspaces, this result is due to Godini [45] in the following equivalent form: the image of unit closed ball by quotient operator is closed in quotient space with respect to that linear closed subspace.

3.3.3 Additivity Criteria for Subdifferentials of Convex Functions

The Fenchel duality theory can be used to get sharp additivity criteria for subdifferentials besides that established in Sect. 2.2.4 or in Sect. 3.2.2.

To begin with, we mention the following theorem concerning the pointwise additivity of subdifferential.

Theorem 3.101 *Let f_1, f_2 be two proper convex lower-semicontinuous functions defined on a locally convex space X. If $x \in \mathrm{Dom}(f_1) \cap \mathrm{Dom}(f_2)$, then the following statements are equivalent*:

(i) $\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x)$
(ii) *For every $x^* \in \partial(f_1 + f_2)(x)$ there exists $x_1^* \in X^*$ such that*

$$(f_1 + f_2)^*(x^*) = f_1^*(x_1^*) + f_2^*(x^* - x_1^*), \tag{3.164}$$

(iii) *The minimization problem*

$$\min\{f_1^*(u^*) + f_2^*(x^* - u^*);\ u^* \in X^*\} \tag{$\mathscr{P}_{x^*}$}$$

has optimal solutions for any $x^* \in \partial(f_1 + f_2)(x)$ and its optimal value is $(f_1 + f_2)^*(x^*)$.

Proof (i) → (ii) Let us consider an arbitrary element $x^* \in \partial(f_1 + f_2)(x)$. By (i), there exists $x_1^* \in \partial f_1(x)$ such that $x - x_1^* \in \partial f_2(x)$. Therefore, we have

$$x_1^*(x) = f_1(x) + f_1^*(x_1^*), \tag{3.165}$$

$$(x^* - x_1^*)(x) = f_2(x) + f_2^*(x^* - x_1^*), \tag{3.166}$$

$$x^*(x) = (f_1 + f_2)(x) + (f_1 + f_2)^*(x^*). \tag{3.167}$$

Adding (3.165) and (3.166), it follows that we have equality (3.164).

(ii) → (iii) By the Young inequality, for any $u^* \in X^*$, we have

$$f_1^*(u^*) \geq u^*(v) - f_1(v) \quad \text{for all } v \in X,$$
$$f_2^*(x^* - u^*) \geq (x^* - u^*)(v) - f_2(v), \quad \text{for all } v \in X,$$

and so,

$$f_1^*(u^*) + f_2^*(x^* - u^*) \geq x^*(v) - (f_1 + f_2)(v), \quad \text{for all } v \in X.$$

Using the definition of the conjugate, we get

$$f_1^*(u^*) + f_2^*(x^* - u^*) \geq (f_1 + f_2)^*(x^*).$$

Taking $x_1^* \in X^*$ such that equality (3.164) holds, it follows that x_1^* is an optimal solution of $(\mathscr{P}_{x^*})$. Moreover, its optimal value is equal to $(f_1 + f_2)^*(x^*)$, that is, (iii) is completely proved. In fact, it is obvious that (iii) → (ii), and so, (ii) and (iii) are equivalent.

(ii) → (i) Let us have x^* an arbitrary element of $\partial(f_1 + f_2)(x)$ and $x_1^* \in X^*$ such that equality (3.164) holds. Using again the Young inequality, we have

$$f_1^*(x_1^*) \geq x_1^*(x) - f_1(x),$$
$$f_2^*(x^* - x_1^*) \geq (x^* - x_1^*)(x) - f_2(x). \qquad \square$$

Since (3.167) also holds, it follows that both inequalities are just equalities, that is, $x_1^* \in \partial f_1(x)$ and $x^* - x_1^* \in \partial f_2(x)$. Consequently, the pointwise additivity equality (i) is true.

Now, we can establish a characterization of the *pointwise additivity of a subdifferential* by a closedness property.

Theorem 3.102 *Given an element $x \in \partial(f_1) \cap D(f_2)$, where f_1, f_2 are two proper convex lower-semicontinuous functions on a locally convex space X, then*

$$\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x), \tag{3.168}$$

if and only if the following closedness condition is fulfilled:

$$\begin{aligned}(\operatorname{epi} f_1^* + \operatorname{epi} f_2^*) &\cap \left[\partial(f_1 + f_2)(x) \times \mathbb{R}\right] \\ &= \overline{(\operatorname{epi} f_1^* + \operatorname{epi} f_2^*)} \cap \left[\partial(f_1 + f_2)(x) \times \mathbb{R}\right].\end{aligned} \tag{3.169}$$

Proof According to the previous characterization, Theorem 3.101(iii), we can apply Lemma 3.59 and Remark 3.60, where the spaces X, Y are equal to the dual space X^*, $A = \partial(f_1 + f_2)(x)$ and $F(u^*, x^*) = f_1^*(u^*) + f_2^*(x^* - u^*)$, $u^*, x^* \in X^*$. By a standard computation, it follows that $h^{**}(x^*) = (f_1 + f_2)^*(x^*)$, and so, statement (iii) in Theorem 3.101 holds if and only if Problems $\mathscr{P}_{x^*}$ have optimal solutions and h is lower-semicontinuous on $\partial(f_1 + f_2)(x)$ since, by hypothesis,

$h(x^*) > -\infty$ for all $x^* \in \partial(f_1 + f_2)(x)$. On the other hand, the set H in Lemma 3.59 becomes

$$\begin{aligned} H &= \{(x^*, t) \in X^* \times \mathbb{R};\ \text{there exists } u^* \in X^* \text{ such that } F(x^*, u^*) \le t\} \\ &= \{(x^*, t) \in X^* \times \mathbb{R};\ \text{there exists } u^* \in X^* \\ &\qquad \text{such that } f_1^*(u^*) + f_2^*(x^* - u^*) \le t\} \\ &= (\operatorname{epi} f_1^* + \operatorname{epi} f_2^*) \cap (X^* \times \mathbb{R}). \end{aligned}$$

Thus, the equality $H \cap (A \times \mathbb{R}) = \overline{H} \cap (A \times \mathbb{R})$ in Remark 3.60 is just equality (3.171), and the theorem is proved. □

Corollary 3.103 *If* epi f_1^* + epi f_2^* *is a closed set, then the subdifferential is additive, that is* (3.168) *holds for all* $x \in X$.

Remark 3.104 In the previous proof, we supposed that $\partial(f_1 + f_2)(x) \neq \emptyset$. In fact, if $\partial(f_1 + f_2)(x) = \emptyset$, then $\partial f_1(x) = \emptyset$, or $\partial f_2(x) = \emptyset$, and so, the additivity property (3.168) is fulfilled.

Remark 3.105 We recall that the *convolution* of two extended real-valued functions f, g on a liner space U is defined by

$$(f \nabla g)(u) = \inf\{f(v) + g(u - v);\ v \in U\}, \quad u \in U.$$

If "inf" is attended on U for every element of a set $A \subset U$, we have an *exact convolution* on A (see, for instance, [70]). Property (iii) in Theorem 3.101 proves that the (infimal) *convolution* $f_1^* \nabla f_2^*$ is exact on $\partial(f_1 + f_2)(x)$. Also, the equality (3.164) can be rewritten as $(f_1^* \nabla f_2^*)(x^*) = (f_1 + f_2)^*(x^*)$ for all $x^* \in \partial(f_1 + f_2)(x)$. By Corollary 3.103, it follows that, if epi f_1^* + epi f_2^* is a closed set, then the convolution $(f_1 \nabla f_2)(x)$ is exact for any $x \in \operatorname{Range} \partial(f_1^* + f_2^*)$. We recall that $(f_1 \nabla f_2)^* = f_1^* + f_2^*$, but $f_1 \nabla f_2$ cannot be proper and lower-semicontinuous (see, for instance, Laurent [70]). Also, there exists a strong connection between property (iii) and the Fenchel duality theorem. Consequently, property (iii) in Theorem 3.101 can be reformulated in the following two equivalent forms:

(iv) *The convolution* $f_1^* \nabla f_2^*$ *is exact and lower-semicontinuous on* $\partial(f_1 + f_2)(x)$
(v) *The duality Fenchel theorem is true for the functions* f_1, *and* $x^* - f_2$ *for every* $x^* \in \partial(f_1 + f_2)(x)$.

3.3.4 Toland Duality Theorem

Surprisingly enough, the Fenchel duality theorem extends to a non-convex minimization problem of the form

$$\inf_{u \in X} \{g(u) - f(u)\}, \tag{3.170}$$

where f and g are convex, proper and lower-semicontinuous functions on a linear topological space X. More precisely, one has the following theorem known in the literature as the *Toland duality theorem*. (See Toland [111].)

Theorem 3.106 (Toland) *Let X and X^* be linear topological spaces in duality through the bilinear functional $\langle \cdot, \cdot \rangle$. Let $f : X \to \overline{\mathbb{R}}$ and $g : X \to \overline{\mathbb{R}}$ be two lower-semicontinuous, convex and proper functions and let $f^* : X^* \to \overline{\mathbb{R}}$, $g^* : X^* \to \overline{\mathbb{R}}$ be their conjugates. Then*

$$\inf_{u\in X}\big\{g(u) - f(u)\big\} = \inf_{v\in X^*}\big\{f^*(v) - g^*(v)\big\}. \tag{3.171}$$

Proof We note first that by the definition of f^* and g^* that, for each $\lambda \in \mathbb{R}$, if $g(u) - f(u) \geq \lambda$ for all $u \in X$, then $f^*(v) - g^*(v) \geq \lambda$, $\forall v \in X^*$. Hence,

$$\inf_{u\in X}\big\{g(u) - f(u)\big\} \leq \inf_{v\in X^*}\big\{f^*(v) - g^*(v)\big\}.$$

Conversely, if $f^*(v) - g^*(v) \geq \lambda$ for all $v \in X^*$, we have $f^*(v) \geq g^*(v) + \lambda$, $\forall v \in X^*$, and therefore (see Proposition 2.19)

$$f(u) = f^{**}(u) \leq g^{**}(u) - \lambda \leq g(u) - \lambda, \quad \forall u \in X.$$

We have, therefore,

$$g(u) - f(u) \leq \lambda, \quad \forall u \in X,$$

which yields

$$\inf_{u\in X}\big\{g(u) - f(u)\big\} \geq \inf_{v\in X^*}\big\{f^*(v) - g^*(v)\big\},$$

thereby completing the proof of (3.171). □

Theorem 3.107 *Under the conditions of Theorem 3.106, assume that u_0 is a solution to Problem* (P), *that is,*

$$u_0 = \arg\inf_{u\in V}\big\{g(u) - f(u)\big\}.$$

Then, any $u_0 \in \partial f(u_0)$ is a solution to the dual problem

$$\inf_{v\in X^*}\big\{f^*(v) - g^*(v)\big\}, \tag{3.172}$$

that is,

$$u_0^* = \arg\inf_{v\in X^*}\big\{f^*(v) - g^*(v)\big\}. \tag{3.173}$$

Moreover, one has, in this case,

$$\begin{aligned} &0 \in \partial f(u_0) - \partial g(u_0), \\ &0 \in \partial f^*(u_0^*) - \partial g^*(u_0^*). \end{aligned}$$

Proof We have

$$g(u_0) - f(u_0) \leq g(u) - f(u), \quad \forall u \in X.$$

This yields

$$f(u) - f(u_0) \leq g(u) - g(u_0), \quad \forall u \in X,$$

and, since $u_0^* \in \partial f(u_0)$, we have

$$\langle u - u_0, u_0^* \rangle + g(u_0) \leq g(u), \quad \forall u \in X.$$

Hence, $u_0^* \in \partial g(u_0)$. We have, therefore,

$$g(u_0) + g^*(u_0^*) = \langle u_0, u_0^* \rangle,$$
$$f(u_0) + f^*(u_0^*) = \langle u_0, u_0^* \rangle.$$

One might suspect, by Theorem 3.106, that if u_0 is a solution to problem (3.170), then

$$0 \in \partial g(u_0) - \partial f(u_0)$$

and that

$$0 \in \partial g^*(u_0^*) - \partial f^*(u_0^*),$$

where $u_0^* \in \partial f(u_0)$.

It turns out that this is, indeed, the case, even in the case where f and g are not convex. □

Theorem 3.108 *Assume that u_0 is a solution of* (P) *and that $\partial f(u_0) \neq \emptyset$. Then, any $u_0^* \in \partial f(u_0)$ is a solution to the dual problem*

$$\inf_{v \in X^*} \left\{ f^*(v) - g^*(v) \right\}.$$

Moreover, one has

$$f(u_0) + f^*(u_0^*) = \langle u_0, u_0^* \rangle, \tag{3.174}$$
$$g(u_0) + g^*(u_0^*) = \langle u_0, u_0^* \rangle. \tag{3.175}$$

Proof We have $g(u_0) - f(u_0) \leq g(u) - f(u)$, $\forall u \in X$, and, since $u_0^* \in \partial f(u_0)$, we have

$$\langle u - u_0, u_0^* \rangle + g(u_0) \leq g(u), \quad \forall u \in X.$$

Hence, $v_0^* \in \partial g(u_0)$ and (3.174) and (3.175) follow. □

3.3.5 The Farthest Point Problem

The aim of this section is to establish characterizations of remotal sets by a property of closedness of some associated sets. We also examine the connection between the farthest point problem and the best approximation problem (see Sect. 3.3.2).

Let X be a linear normed space and let A be a given nonvoid set A in X. Let us consider the optimization problem

$$(A_1) \qquad \max_{y\in A} \|x-y\|, \qquad x\in X,$$

called the *farthest point problem*. The corresponding *best approximation problem* is

$$(A_2) \qquad \min_{y\in A} \|x-y\|, \qquad x\in X,$$

but most of the properties of this two problems are different.

First, we remark that Problem (A_1) can be considered only for convex sets A, because it has solutions if and only if there exist solutions in its convex hull, conv A (see, for instance, Hiriart–Urruty [52]). Even in the case of convex sets A when Problem (A_2) is a *convex optimization problem*, Problem (A_1) is not convex being a typical *d.c. optimization problem*, that is, the minimization of a difference of two convex functions. It is easy to remark that the farthest point Problem (A_1) can be also given in one of the following types: P_1 (*the maximization of a convex function on a convex set*) and P_2 (*the minimization of a convex function on the complement of a convex set*). Indeed, Problem (A_1) can be, equivalently, written as

$$(A_1') \qquad \min_{y\in A}\big\{I_A(x)-\|x-y\|\big\}, \qquad x\in X,$$

or

$$(A_1'') \qquad \min_{\|x-y\|\ge t}\big\{I_A(x)-t\big\}, \qquad x\in X.$$

Consequently, we obtain some optimality conditions using the normal cone of A and the ε-subdifferential of the norm (see, for example, Hiriart-Urruty [51]).

In the sequel, we recall some concepts associated to the farthest point problem which are similar to some known concepts of the best approximation theory.

We denote

$$\Delta_A(x)=\sup_{y\in A}\|x-y\|, \qquad x\in X, \tag{3.176}$$

called the *farthest distance function* of the set A,

$$Q_A(x)=\big\{\bar{x}\in A;\ \|x-\bar{x}\|=\Delta_A(x)\big\}, \qquad x\in X, \tag{3.177}$$

called the *farthest point mapping* (or *antiprojection*) with respect to A. The elements of $Q_A(x)$ are called *farthest points* of x through elements of the set A. The set A is called a *remotal set* if $Q_A(x)\ne\emptyset$ for $x\in X$.

It is obvious that the mapping $\varDelta_A$ is a continuous convex function. Moreover,

$$\big|\varDelta_A(x_1)-\varDelta_A(x_2)\big|\le\|x_1-x_2\|,\quad x_1,x_2\in X. \tag{3.178}$$

Consequently, $\varDelta_A$ is subdifferentiable and $\partial\varDelta_A(x)\subset\overline{S}_{X^*}(0;1)$ for every $x\in X$.

On the other hand, we notice that by Toland duality (see Sect. 3.3.4), we have the following equality:

$$\varDelta_A(x)=\sup\big\{x^*(x)-s_A(x^*);\ \|x^*\|\le 1\big\},$$

where s_A is the *support functional* of A, that is,

$$s_A(x^*)=\sup\big\{x^*(u);\ u\in A\big\}.$$

We also recall some simple convexity properties.

$$\varDelta_A(x)=\varDelta_{\overline{\operatorname{conv}A}}(x), \tag{3.179}$$

$$\varDelta_A\big(\lambda x+(1-\lambda)y\big)=\lambda\varDelta_A(x)\quad\text{for all } y\in Q_A(x), \tag{3.180}$$

$$y\in Q_A\big(\lambda x+(1-\lambda)y\big)\quad\text{if } y\in Q_A(x)\text{ and }\lambda>1. \tag{3.181}$$

Theorem 3.109 *A nonvoid bounded set A in a linear normed space X is remotal if and only if the associated set*

$$K_d=A+cS(0;d), \tag{3.182}$$

is closed for every $d>0$.

Proof Let x be an adherent element of $A+cS(0;d)$, that is, there exist a sequence $(x_n)_{n\in\mathbb{N}}$ convergent to x and a sequence $(u_n)_{n\in\mathbb{N}}\subset A$ such that $\|x_n-u_n\|\ge d$ for all $n\in\mathbb{N}$. Thus, for every $\varepsilon>0$ there exists $n_\varepsilon\in\mathbb{N}$ such that $\|x-u_n\|>d-\varepsilon$ for all $n\ge n_\varepsilon$. Now, if A is remotal, taking an element $\bar{x}\in Q_A(x)$, we find that $\|x-\bar{x}\|\ge\|x-u_n\|$, $n\in\mathbb{N}$, and so $\|x-\bar{x}\|>d-\varepsilon$, for every $\varepsilon>0$. Consequently, $\|x-\bar{x}\|\ge d$, that is, $x\in A+cS(0;d)$.

Conversely, for an arbitrary element $x\in X$, we take $d=\varDelta_A(x)$. We can suppose $d>0$ since $\varDelta_A(x)=0$ if and only if $A=\{x\}$ when A is obviously remotal. For every $n\in\mathbb{N}^*$ there exists $u_n\in A$ such that $\|x-u_n\|\ge d-\frac{1}{n}$. But we have

$$\frac{1}{n}\left(d-\frac{1}{n}\right)^{-1}(x-u_n)+x\in u_n+cS(0;d)\subset A+cS(0;d),$$

for all $n\in\mathbb{N}^*$, such that $n>\frac{1}{d}$. Since $(u_n)_{n\in\mathbb{N}^*}$ is bounded, passing to the limit we get $x\in\overline{A+cS(0;d)}$. Therefore, if $A+cS(0;d)$ is closed there exists $\bar{x}\in A$ such that $\|x-\bar{x}\|\ge d$, that is, $\bar{x}\in Q_A(x)$. Hence, the set A is remotal. $\square$

Remark 3.110 It is easy to see that

$$cK_d = \bigcap_{a \in A} S(a; d) \tag{3.183}$$

and so, the set K_d is always the complement of a convex bounded set. Consequently, *a nonvoid bounded set A is remotal if and only if the convex set $\bigcap_{a \in A} S(a; d)$ is open for any $d > 0$.*

Corollary 3.111 *Any closed ball in a linear normed space is remotal.*

Remark 3.112 We denote $r(A) = \inf\{\Delta_A(x);\ \ x \in X\}$, usually called the *radius* of A. If $d < r(A)$, then $K_d = X$, and so, in Theorem 3.109, it suffices to consider only the case $d \geq r(A)$. Obviously, the set K_d is nonvoid complement of a convex set for any $d \geq 0$ and $K_d \neq X$ if $d > r(A)$.

Remark 3.113 We say that a set A is *d-remotal* (*d-proximinal*) if $Q_A(x) \neq \emptyset$ ($P_A(x) \neq \emptyset$) whenever $\Delta_A(x) = d$ ($d(A; x) = d$). From the proof of Theorem 3.109 it follows that a set A is d-remotal if K_d is closed. Generally, the converse statement is not true. But if a set is d-remotal for any $d \geq d_0$, then the sets K_d are closed for all $d \geq d_0$. Therefore, the property of d-remotability is different of the property of the set K_d to be closed. Similar statements for d-proximinality hold. A relationship between d-proximinality and d-remotability will be given in Remark 3.119.

The above characterization of remotal sets is similar to the one of proximinal sets (Theorem 3.97) characterized by the closedness of the associated set

$$H_d = A + \overline{S}(0; d). \tag{3.184}$$

Taking into account that the associated sets H_d and K_d have the properties of symmetry with respect to the sets A and $\overline{S}(0; 1)$, respectively, $cS(0; 1)$, we obtain some properties of duality between proximinality and remotability.

Theorem 3.114 *Let A be a closed bounded convex set such that $0 \in \operatorname{int} A$. Then*

(i) *A is proximinal if and only if $\overline{S}(0; 1)$ is proximinal with respect to p_A*
(ii) *A is remotal if and only if $cS(0; d)$ is proximinal, for any $d > 0$, with respect to p_A*

where p_A is the Minkowski functional associated to the set A.

Proof (i) By hypothesis, the Minkowski functional p_A is an equivalent norm in X, generally asymmetric and

$$\overline{S}_{p_A}(0; d) = dA, \qquad S_{p_A}(0; 1) = \operatorname{int} A.$$

Therefore, in the linear (asymmetric) normed space (X, p_A) the closed set $\overline{S}(0;1)$ is proximinal with respect to p_A if and only if $\overline{S}(0;1) + \overline{S}_{p_A}(0;d) = d(A + \overline{S}(0;\frac{1}{d}))$ is closed for all $d > 0$, that is, A is proximinal in X. The other assertions can be proved using the corresponding above theorems. □

Corollary 3.115 *If in a linear normed space X there exists a remotal set, then X can be endowed with an equivalent norm, generally asymmetric, such that there exists a bounded, symmetric, convex body whose complement is proximinal.*

Proof By Theorem 3.114, the set $S(0;1)$ has the required properties. □

Remark 3.116 Theorem 3.114 is also true in any asymmetric normed space.

Therefore, the proximinality, respectively, remotability, are dependent of the topological properties of a pair of sets (A, B) such that $A + B$, $A + cB$ is closed. If A, B are convex sets, then $A + B$ is also convex, while $A + cB$ is not convex being the complement of a convex set. Thus, the case $A + cB$ is little difficult, even in the case of weakly compact sets.

Property (ii) in Theorem 3.114 can be also presented in a pointwise form which, at the same time, establishes a relationship between two d.c. optimization problems. Let us consider the special farthest point problem

$$(A_9) \quad \max\big\{\|x - y\|_1;\ y \in \overline{S}_{\|\cdot\|_2}(0;1)\big\}, \quad x \in X,$$

and the associated best approximation problem

$$(A_{10}) \quad \min\big\{\|x - \alpha y\|_2;\ y \in cS_{\|\cdot\|_1}(0;1)\big\}, \quad x \in X,$$

where $\|\cdot\|_1$, $\|\cdot\|_2$ are two equivalent norms in X and

$$(A_{11}) \quad \alpha = \sup\big\{\|x - y\|_1;\ y \in \overline{S}_{\|\cdot\|_2}(0;1)\big\}.$$

Theorem 3.117 *Problem (A_9) has an optimal solution if and only if Problem (A_{10}) has an optimal solution.*

Proof Let $\bar{y}$ be an optimal solution of (A_9), that is, $\alpha = \|x - \bar{y}\|_1$ and $\|\bar{y}\|_2 \le 1$.

Obviously, $\alpha > 0$ and $\|\bar{y}\|_2 = 1$. Taking $x - \bar{y} = \alpha\bar{z}$, we have $\|\bar{z}\|_1 = 1$ and $\|x - \alpha y\|_2 \ge 1$ for any $y \in cS_{\|\cdot\|_1}(0;1)$. Indeed, in the contrary case, it follows that there exists $y_1 \in X$ such that $\|y_1\|_1 \ge 1$ and $\|x - \alpha y_1\|_2 < 1$. Since $\|x - y\|_1 < \alpha$ for any $y \in S_{\|\cdot\|_2}(0;1)$ (the solutions of (A_9) are boundary elements of $S_{\|\cdot\|_2}(0;1)$, it follows that $\alpha\|y_1\|_1 = \|x - (x - \alpha y_1)\|_1 < \alpha$, that is, $\|y_1\| < 1$, which is a contradiction.

Conversely, if $\bar{z}$ is an optimal solution of (A_{10}), we denote $x - \alpha\bar{z} = \bar{y}$. But necessarily it follows that $\|\bar{z}\|_1 = 1$ and therefore $\|x - \bar{y}\|_1 = \alpha$. On the other hand, for every $\varepsilon > 0$ there exists $y_\varepsilon \in \overline{S}_{\|\cdot\|_2}(0;1)$ such that $\|x - y_\varepsilon\|_1 > \alpha - \varepsilon$ and so

$$\left\|x - \frac{\alpha}{\alpha - \varepsilon}(x - y_\varepsilon)\right\|_2 \ge \|x - \alpha\bar{z}\|_2,$$

which implies

$$\left\| y_\varepsilon - \frac{\varepsilon}{\alpha-\varepsilon}(x-y_\varepsilon)\right\|_2 \geq \|x-\alpha\bar{z}\|_2 = \|\bar{y}\|_2.$$

But

$$\left\| y_\varepsilon - \frac{\varepsilon}{\alpha-\varepsilon}(x-y_\varepsilon)\right\|_2 \leq \|y_\varepsilon\|_2 + \frac{\varepsilon}{\alpha-\varepsilon}\|x-y_\varepsilon\|_2 \leq 1 + \frac{\varepsilon}{\alpha-\varepsilon}\big(\|x\|_2+1\big).$$

Therefore,

$$\|\bar{y}\|_2 \leq 1 + \frac{\varepsilon}{\alpha-\varepsilon}\big(1+\|x\|_2\big),$$

for any $\varepsilon > 0$. Consequently, for $\varepsilon \searrow 0$ we obtain $\|\bar{y}\|_2 \leq 1$. Since $\|x-\bar{y}\|_1 = \alpha$, it follows that $\bar{y}$ is an optimal solution of (A_9). □

Remark 3.118 In fact, we have

$$\operatorname{val}(A_9) = \operatorname{val}(A_{10}) = \alpha$$

and both problems are d.c. optimization problems of type P_1 and P_2, respectively.

Remark 3.119 According to Remark 3.113, we find that Theorem 3.117 can be reformulated as follows: *if* $\overline{S}_{\|\cdot\|_2}(0;1)$ *is* d*-remotal with respect to* $\|\cdot\|_1$*, then* $cS_{\|\cdot\|_1}(0;d)$ *is* d*-proximinal for* $\|\cdot\|_2$ (see, also, (3.182) in this special case).

3.4 Problems

3.1 Find the dual of the function $\varphi : L^2(\Omega) \to \overline{\mathbb{R}}^*$ defined in Example 2.56.

Hint. We have

$$\varphi^*(p) = \sup_{y\in H_0^1(\Omega)}\left\{\int_\Omega py\,\mathrm{d}\xi - \frac{1}{2}\int_\Omega |\nabla y|^2\,\mathrm{d}\xi - \int_\Omega g(y)\,\mathrm{d}\xi\right\}$$

and by Theorem 3.54 we have (see, also, Example 3.78)

$$\varphi^*(p) = \sup\left\{\int_\Omega g^*(p)\,\mathrm{d}\xi - \frac{1}{2}\|p+y\|^2_{H^{-1}(\Omega)}\right\}.$$

3.2 Let $\{f_n\}$ be a sequence of lower-semicontinuous convex functions on a reflexive Banach space X. The sequence is said to be M-convergent to f (convergent in the sense of Mosco) if the following conditions hold:

(a) For each sequence $\{u_n\}$ weakly convergent to u, we have

$$\liminf_{n\to\infty} f_n(u_n) \ge f(u).$$

(b) For each $u \in X$, $\exists\{u_n\}$ strongly convergent to u such that

$$\lim_{n\to\infty} f_n(u_n) = f(u).$$

Show that, if $\{f_n\}$ is M-convergent to f, then for each $\lambda > 0$ and $u \in X$

$$u_n = \arg\inf_X \left\{ f_n(x) + \frac{1}{2\lambda}\|x-u\|^2 \right\} \to u^* \quad \text{as } n \to \infty,$$

where

$$u^* = \arg\inf_X \left\{ \frac{1}{2\lambda}\|x-u\|^2 + f(x) \right\}.$$

Hint. We have $f_n(u_n) + \frac{1}{2\lambda}\|u_n - u\|^2 \le f_n(u)$, $\forall n$, and this yields the desired result.

3.3 Show that $\{f_n\}$ is M-convergent to f if and only if $\partial f_n \xrightarrow{G} \partial f$, that is, for each $(u, v) \in \partial f$, there are $(u_n, v_n) \in \partial f_n$, $n \in \mathbb{N}^*$, such that $u_n \to u$, $v_n \to v$ strongly in X and X^*, respectively, as $n \to \infty$.

Hint. If $f_n \to f$ (M-convergent to f) and $(u, v) \in \partial f$, we have by (3.1) that

$$u_n = \arg\inf \left\{ f_n(x) - (v, x) + \frac{1}{2}\|x-u\|^2 \right\} \to \arg\inf_X \left\{ \frac{1}{2}\|x-u\|^2 + f(x) - (v, x) \right\}$$

and ($F : X \to X^*$ is the duality mapping)

$$\partial f_n(u_n) - v + F(u_n - u) = 0.$$

Hence, $v_n = v - F(u_n - u) \in \partial f_n(u_n)$ is strongly convergent to v.

3.4 Show that if $F : X \to X$ is a mapping on the complete metric space X with the distance d such that $d(x, F(x)) \le \varphi(x) - \varphi(F(x))$, $\forall x \in X$, where $\varphi : X \to \mathbb{R}$ is an lower-semicontinuous function bounded from below, then F has a fixed point (the *Caristi fixed point theorem*).

Hint. One applies Corollary 3.74, where $f = \varphi$ and $\varepsilon = \frac{1}{2}$.

3.5 Prove that a nonvoid w^*-closed set A in the dual X^* of a linear normed space X is proximinal.

Hint. By Theorem 1.81, the closed balls in the dual are w^*-compact. On the other hand, the norm of dual is w^*-lower-semicontinuous (Proposition 2.5) and so the problem of best approximation has at least a solution by the Weierstrass theorem (Theorem 2.8).

3.6 Prove that an element $\bar{x} \in A$ is a farthest element of an element $x \in X$ with respect to the convex set A if and only if there exists an element $x_0^* \in X^*$ such that $\|x_0^*\| = 1$ and $x_0^*(\bar{x} - x) = \sup\{\|u - x\|;\ u \in A\}$.

Hint. If $\|u - x\| \le x_0^*(\bar{x} - x)$ for all $u \in A$ and $\|x_0^*\| = 1$, then it is obvious that $\|u - x\| \le \|\bar{x} - x\|$ for all $u \in A$ since $x_0^*(\bar{x} - x) \le \|\bar{x} - x\|$. Hence $\bar{x}$ is a farthest point in A for $x \in X$. Conversely, if $\bar{x}$ is a farthest point, we take $x_0^* \in X^*$ such that $\|x_0^*\| = 1$ and $x_0^*(\bar{x} - x) = \|\bar{x} - x\|$ (see (1.36)). (For other dual characterizations, see [107].)

3.7 Let $f : X \to \overline{\mathbb{R}}$ be a proper lower-semicontinuous convex function on the locally convex space X. Prove that ∂f is surjective on the domain of its conjugate and f^* is continuous on $\operatorname{dom} f^*$ if and only if the set

$$H_f = \big\{\big(x^*, \lambda - x^*(x)\big);\ (x, \lambda) \in \operatorname{epi} f,\ x^* \in X^*\big\}$$

is closed in $X^* \times \mathbb{R}$.

Hint. Take $F(x, x^*) = f(x) - x^*(x)$, $(x^*, x) \in X^* \times \mathbb{R}$ and its corresponding family of minimization $\min\{f(x) - x^*(x); x \in X\}$, $x^* \in X^*$. It is obvious that we have the value function $h(x^*) = -f^*(x^*)$ and so, by Lemma 3.59, we obtain $x^* \in \partial f(\bar{x})$ for an element $\bar{x} \in \operatorname{dom} f$, that is, the corresponding problem has the optimal solution $\bar{x}$ whenever its value is not $-\infty$ (see the Young inequality (2.16) and Proposition 2.25), and h is lower-semicontinuous (equivalently, f^* is upper-semicontinuous) if and only if the set H_f is closed. On the other hand, by Proposition 2.19(i), it follows that f^* is just continuous on $\operatorname{Dom} h = \operatorname{Dom} f^*$. Moreover, necessarily, $\operatorname{dom} f^*$ is an open set in X^* (see [92]).

3.5 Bibliographical Notes

3.1. The results presented in the first part of Sect. 3.1 have an algebraic character and are direct consequences of separation properties of convex sets in finite-dimensional spaces. The regularity condition $(S) + (O)$ is known in the literature as the Uzawa constraint qualification condition. In finite-dimensional spaces many other qualification conditions are known (see the monographs of Bazaraa and Shetty [14], El-Hodiri [40], Stoer and Witzgall [110], Hestenes [49]). The extension of the Kuhn–Tucker theorem on separated locally convex spaces has been given by Rockafellar [98, 99].

For operatorial constraints, results of this type have been obtained by many authors under different assumptions on the set B defined by (3.22) (see Theorem 3.13). The existence of multipliers can be regarded as a consequence of Farkas' lemma extended in several directions by Altman [2], Arrow, Hurwicz and Uzawa [3], and Zălinescu [120], among others.

As regards the regularity of subdifferential mappings, we refer to the works of Ioffe and Levin [55], Kutateladze [66], Valadier [117], Zowe [123]. The general optimality condition (3.37) is due to Pshenichny [93]. The concept of tangent cone (Definition 3.21) given by Abadie [1] arises as a natural extension of the Kuhn–Tucker notion of feasible direction. The notion of pseudotangent cone (Definition 3.23) has been used first by Guinard [46]. In Theorem 3.30, which is due to Nagahisa and Sakawa [79], the interiority condition $\operatorname{int} A_Y \neq \emptyset$ is not essential. As a matter of fact, it remains true under any conditions, which ensures the conclusion of Lemma 3.29. Theorem 3.33, due to Guinard [46], extends some earlier results of Varaiya [119]. Other results in this direction have been obtained by Borwein [18, 19], Bazaraa and Shetty [14], Hiriart-Urruty [50], Ursescu [114, 115]. In the differentiable case, the following two conditions are usually imposed: (1) $x_0 \in \operatorname{int} A$ and G'_{x_0} is surjective; (2) $\exists x \in C(A, x_0)$ such that $G'_{x_0} \in \operatorname{int} C(-A_Y, G(x_0))$. Kurcyusz (see Zowe and Kurcyusz [124]) uses the condition $G'_{x_0}(C(A;x_0)) - C(-A_Y;G(x_0)) = Y$. This condition is quite general if one takes into account that, if the problem admits Lagrange multipliers, then $\overline{G'_{x_0}(C(A,x_0)) - C(-A_Y;G(x_0))} = Y$. Moreover, in several special cases, the Kurcyusz regularity condition is necessary as happens for instance if $\dim Y < \infty$ or $A_Y = A_1 \times A_2$ with $\dim A_1 < \infty$ and $\operatorname{int} A_2 \neq \emptyset$. Other regularity conditions have been given by Bazaraa and Goode [13], Bazaraa, Goode and Nashed [15], Halkin and Neustadt [48], Kurcyusz [65], Mangasarian and Fromovitz [75], Robinson [97], Tuy [112, 113]. Asymptotic Kuhn–Tucker conditions have been studied by Zlobek [122] and the characterization of critical points on affine sets was given by Norris [81]. In the latter paper, it was proved that, if $f'_{x_0} \neq 0$, then Theorem 3.35 remains valid without assuming that T has a closed range.

3.2. For the most part, the results presented in this section are due to Rockafellar [99–101]. For a detailed discussion of Rockafellar's duality theory, we refer the reader to the recent book of Ekeland and Temam [39]. Regarding the duality by Lagrangian in infinite-dimensional spaces, the papers of Varaiya [119], Ritter [96], Arrow, Hurwicz and Uzawa [3], and Rockafellar [102] may be cited.

Duality results given in minimax form may be found in the works of Karamardian [59], Stoer [109], and Mangasarian and Ponstein[76] for finite-dimensional problems and in the papers of Arrow, Hurwicz and Uzawa [3], Claesen [24], Brans and Claesen [22], Precupanu [84], in an infinite-dimensional setting. As has been suggested by Moreau [78], the bilinear form which occurs in conjugate duality theory can be replaced by a nonbilinear form without invalidating many of the essential properties. Results of this type have been given by Balder [11], Deumlich and Elster [30], Dolecki and Kurcyusz [33]. Other duality schemes have been studied by Ekeland [38], Linberg [74], Rosinger [103], Toland [111], Rockafellar [102], Singer [106].

In the study of the optimization problems, the condition that certain sets are closed arises frequently. Closedness conditions (generally sufficient) are determined by the following simple result: an extended real-valued function f on a set X has a point of minimum in X and the minimum value is finite if and only if the set $\{\lambda \in \mathbb{R}; \exists x \in X \text{ such that } f(x) \leq \lambda\}$ is closed and proper in $\mathbb{R}$. A detailed discussion of the optimality with the aid of sets of this kind may be found in the papers of Slyke and Wets [118], Levine and Pomerol [71–73]. For the linear case, some sufficient optimality conditions viewed as closedness conditions were established by Kretschmer [64], Krabs [62], Duffin [34], Dieter [31], Nakamura and Yamasaki [80], and for the convex case were established by Dieter [31], Levine and Pomerol [72, 73], Arrow, Hurwicz and Uzawa [3], Zălinescu [120, 121]. Furthermore, Levine and Pomerol proved that, if the stability of the family of all problems obtained by linear perturbations is demanded, then the closedness condition is also necessary. It must be emphasized that this result can be generalized to a family of non-convex problems (Lemmata 3.59, 3.61 and 3.62), the convexity property being necessary only for the equality of the values of primal and dual problems (see Precupanu [85–88] and Precupanu and Precupanu [92]). Other sufficient optimality conditions could be obtained by applying new criteria for the closedness of the image of a closed subset by a multivalued mapping. In this context, the works of Dieudonné [32], Ky Fan [67], Dedieu [28, 29], Gwinner [47], Asimow and Simoson [5], Mennicken and Sergaloff [77], Ursescu [116], Beattie [16], Asimow [4] and Floret [41] can be cited.

Theorems 3.70, 3.72 extend some results for linear cases due to Nakamura and Yamasaki [80] (see Precupanu [87]). In recent years, notable results have been obtained in the non-convex and non-smooth optimization theory. The presentation of these results is beyond the scope of this book. However, we mention in this directions the works of Ekeland [37], Clarke [25, 26], Hiriart-Urruty [50], Aubin and Clarke [7].

3.3. This section is mainly concerned with some implications of Fenchel's duality theorem in linear programming. In this respect, several results have been obtained in the absence of the interiority condition by Ky Fan [68], Levine and Pomerol [71], Zălinescu [120], Kallina and Williams [58], Kortanek and Soyster [60], Nakamura and Yamasaki [80] and Raffin [94]. The problem $\overline{\mathscr{P}}$ has been previously considered by Schechter [104].

A large number of papers have been written on the best approximation, regarded as an optimum problem (Holmes [53], Krabs [63], Laurent [70], Balakrishnan [10], Singer [105]). Dual properties of best approximation elements was established by Garkavi [44]. The characterization of proximinality via closedness condition (Theorem 3.98) is due to Precupanu [83] (see also Precupanu and Precupanu [92]).

Many examples of convex programming problems arising in mechanics and other fields may be found in the books of Duvaut and Lions [35], Ekeland and Temam [39], Balakrishnan [10] and Holmes [53, 54].

The usual additivity criterion for the subdifferential is contained in Theorem 3.4 (Rockafellar [98]). Other general criteria were established by Boţ and

Wanka [20], Boţ, Grad and Wanka [21], Revalski and Théra [95], Burachik and Jeyakumar [23]. But it is possible that the additivity of subdifferential to be true only in certain points. This pointwise property is investigated in Sect. 3.3.3. The characterizations contained in Theorems 3.101 and 3.102 was established by Precupanu and Precupanu [92] (see also [91]).

The farthest point problem has been the subject of much development in the last years. For different special results concerning the existence of farthest points or remotal sets, we refer the reader to the papers of Asplund [6], Balaganskii [8, 9], Franchetti and Papini [43], Cobzaş [27], Baronti and Papini [12], Blatter [17], Edelstein [36], Lau [69], Panda and Kapoor [82], and the book of Singer [107].

The characterization of remotal sets (Theorem 3.109) was established by Precupanu [89, 90]. Also, the connection between the farthest point problem and the best approximation problem was investigated in [90].

References

1. Abadie M (1965) Problèmes d'Optimisation. Institut Blaise Pascal, Paris
2. Altman M (1970) A general separation theorem for mappings, saddle-points, duality and conjugate functions. Stud Math 36:131–167
3. Arrow KJ, Hurvicz L, Uzawa H (1958) Studies in linear and non-linear programming. Stanford University Press, Stanford
4. Asimow L (1978) Best approximation by gauges on a Banach space. J Math Anal Appl 62:571–580
5. Asimow L, Simoson A (1979) Decomposability and dual optimization in Banach spaces. Preprint, Univ of Wyoming
6. Asplund E (1966) Farthest points in reflexive locally uniformly rotund Banach spaces. Isr J Math 4:213–216
7. Aubin JP, Clarke FH (1977) Multiplicateurs de Lagrange en optimisation non convexe et applications. C R Acad Sci Paris 285:451–453
8. Balaganskii VS (1995) On approximation properties of sets with convex complements. Math Notes 57:26–29
9. Balaganskii VS (1998) On nearest and farthest points. Math Notes 63:250–252
10. Balakrishnan AV (1971) Introduction to optimization theory in a Hilbert space. Springer, Berlin
11. Balder EJ (1977) An extension of duality–stability relations. SIAM J Control Optim 15:329–343
12. Baronti M, Papini PL (2001) Remotal sets revisited. Taiwan J Math 5:367–373
13. Bazaraa MS, Goode J (1972) Necessary optimality criteria in mathematical programming in the presence of differentiability. J Math Anal Appl 40:609–621
14. Bazaraa MS, Shetty CM (1976) Foundations of optimization. Lecture notes in economics and mathematical systems, vol 122. Springer, Berlin
15. Bazaraa MS, Goode J, Nashed MZ (1972) A nonlinear complementary problem in mathematical programming in Banach spaces. Proc Am Math Soc 35:165–170
16. Beattie R (1980) Continuous convergence and the closed–graph theorem. Math Nachr 99:87–94
17. Blater J (1969) Werteste punkte and Nachste punkte. Rev Roum Math Pures Appl 4:615–621
18. Borwein J (1977) Proper efficient points for maximizations with respect to cones. SIAM J Control Optim 15:57–63

19. Borwein J (1978) Weak tangent cones and optimization in a Banach space. SIAM J Control Optim 16:512–522
20. Boţ I, Wanka G (2006) A weaker regularity condition for subdifferential calculus and Fenchel duality in infinite dimensional spaces. Nonlinear Anal, Theory Methods Appl 64(12):2787–2804
21. Boţ I, Grad S, Wanka G (2009) Generalized Moreau–Rockafellar results for composed convex functions. Optimization 58:917–933
22. Brans JP, Claesen G (1970) Minimax and duality for convex–concave functions. Cah Cent étud Rech Opér 12:149–163
23. Burachik R, Jeyakumar V (2005) A dual condition for the convex subdifferential sum formula with applications. J Convex Anal 12:279–290
24. Claesen G (1974) A characterization of the saddle points of convex–concave functions. Cah Cent étud Rech Opér 14:127–152
25. Clarke FH (1973) Necessary conditions for nonsmooth problems in optimal control and the calculus of variations. Thesis, Univ Washington
26. Clarke FH (1975) Generalized gradients and applications. Trans Am Math Soc 205:247–262
27. Cobzaş St (2005) Geometric properties of Banach spaces and the existence of nearest and farthest points. Abstr Appl Anal 3:259–285
28. Dedieu J-P (1977) Cône asymptote d'un ensemble non convexe. Applications à l'optimisation. C R Acad Sci Paris 185:501–503
29. Dedieu J-P (1978) Critères de femeture pour l'image d'un fermé non convexe par une multiplication. C R Acad Sci Paris 287:941–943
30. Deumlich R, Elster KH (1980) Duality theorems and optimality conditions for nonconvex problems. Math Operforsch Stat, Ser Optim 11:181–219
31. Dieter U (1966) Optimierungsaufgaben in topologische Vectorräumen. I. Dualitatstheorie. Z Wahrscheinlichkeitstheor Verw Geb 5:89–117
32. Dieudonné J (1966) Sur la séparation des ensembles convexes. Math Ann 163:1–3
33. Dolecki S, Kurcyusz S (1978) On Φ-convexity in extremal problems. SIAM J Control Optim 16:277–300
34. Duffin J (1973) Convex analysis treated by linear programming. Math Program 4:125–143
35. Duvaut G, Lions JL (1972) Sur les inéqualitions en mécanique et en physique. Dunod, Paris
36. Edelstein M (1966) Farthest points of sets in uniformly convex Banach space. Isr J Math 4:171–176
37. Ekeland I (1974) On the variational principle. J Math Anal Appl 47:324–353
38. Ekeland I (1979) Nonconvex minimization problems. Bull Am Math Soc 1:443–474
39. Ekeland I, Temam R (1974) Analyse convexe et problèmes variationnels. Dunod, Gauthier-Villars, Paris
40. El-Hodiri MA (1971) Constrained extrema. Introduction to the differentiable case with economic applications. Lecture notes in oper res and math systems. Springer, Berlin
41. Floret K (1978) On the sum of two closed convex sets. Math Methods Oper Res 39:73–85
42. Fortmann TE, Athans M (1974) Filter design subject to output sidelobe constraints: theoretical considerations. J Optim Theory Appl 14:179–198
43. Franchetti C, Papini PL (1981) Approximation properties of sets with bounded complements. Proc R Soc Edinb A 89:75–86
44. Garkavi L (1961) Duality theorems for approximation by elements of convex sets. Usp Mat Nauk 16:141–145 (Russian)
45. Godini G (1973) Characterizations of proximinal subspaces in normed linear spaces. Rev Roum Math Pures Appl 18:900–906
46. Guinard M (1969) Generalized Kuhn–Tucker conditions for mathematical programming problems in a Banach space. SIAM J Control 7:232–241
47. Gwinner J (1977) Closed images of convex multivalued mappings in linear topological spaces with applications. J Math Anal Appl 60:75–86
48. Halkin H, Neustadt LW (1966) General necessary conditions for optimizations problems. Proc Natl Acad Sci USA 56:1066–1071

49. Hestenes MR (1975) Optimization theory: the finite dimensional case. Wiley, New York
50. Hiriart-Urruty JB (1977) Contributions à la programmation mathématique. Thèse, Université de Clermont–Ferrand
51. Hiriart-Urruty JB (1989) From convex optimization to nonconvex optimization. Necessary and sufficient conditions for global optimality. In: Clarke FH, Demyanov VF, Giannesi F (eds) Nonsmooth optimization and related topics. Plenum, New York, pp 219–239
52. Hiriart-Urruty JB (2005) La conjecture des points les plus éloingnés revisitée. Ann Sci Math Qué 29:197–214
53. Holmes RB (1972) A course on optimization and best approximation. Lecture notes in oper res and math systems. Springer, Berlin
54. Holmes RB (1975) Geometric functional analysis and its applications. Springer, Berlin
55. Ioffe AD, Levin VL (1972) Subdifferential of convex functions. Trudi Mosc Mat Obsc 26:3–73 (Russian)
56. James RC (1957) Reflexivity and the supremum of linear functionals. Ann Math 66:159–169
57. James RC (1964) Characterization of reflexivity. Stud Math 23:205–216
58. Kallina C, Williams AC (1971) Linear programming in reflexive spaces. SIAM Rev 13:350–376
59. Karamardian S (1967) Strictly quasi-convex (concave) functions and duality in mathematical programming. J Math Anal Appl 20:344–358
60. Kortanek KO, Soyster AL (1972) On refinements of some duality theorems in linear programming over cones. Oper Res 20:137–142
61. Köthe G (1969) Topological vector spaces. I. Springer, Berlin
62. Krabs W (1969) Duality in nonlinear approximation. J Approx Theory 2:136–151
63. Krabs W (1979) Optimization and approximation. Wiley, Chichester
64. Kretschmer KS (1961) Programmes in paired spaces. Can J Math 13:221–238
65. Kurcyush S (1976) On the existence and nonexistence of Lagrange multiplier in Banach space. J Optim Theory Appl 20:81–110
66. Kutateladze SS (1977) Formulas for computing subdifferentials. Dokl Akad Nauk SSSR 232:770–772 (Russian)
67. Ky F (1953) Minimax theorems. Proc Natl Acad Sci USA 39:42–47
68. Ky F (1969) Asymptotic cones and duality. J Approx Theory 2:152–169
69. Lau K-S (1975) Farthest points in weakly compact sets. Isr J Math 2:165–174
70. Laurent PJ (1972) Approximation and optimization. Herman, Paris
71. Levine L, Pomerol JCh (1974) Infinite programming and duality in topological vector spaces. J Math Anal Appl 46:75–81
72. Levine L, Pomerol JCh (1976) C-closed mappings and Kuhn–Tucker vectors in convex programming. CORE Disc Papers 7620, Univ Louvain
73. Levine L, Pomerol JCh (1979) Sufficient conditions for Kuhn–Tucker vectors in convex programming. SIAM J Control Optim 17:689–699
74. Linberg PO (1980) Duality from LP duality. Math Operforsch Stat, Ser Optim 11:171–180
75. Mangasarian OL, Fromovitz S (1967) The Fritz John necessary optimality conditions in the presence of equality and inequality constraints. J Math Anal Appl 17:37–47
76. Mangasarian OL, Ponstein J (1965) Minimax and duality in nonlinear programming. J Math Anal Appl 11:504–518
77. Mennicken R, Sergaloff B (1979/1980) On Banach's closed range theorem. Arch Math 33:461–465
78. Moreau JJ (1966–1967) Fonctionelles convexes. Séminaire sur les équations aux dérivées partielles. College de France
79. Nagahisa Y, Sakawa Y (1969) Nonlinear programming in Banach spaces. J Optim Theory Appl 4:182–190
80. Nakamura T, Yamasaki M (1979) Sufficient conditions for duality theorem in infinite linear programming problems. Hiroshima Math J 9:323–334
81. Norris DO (1971) A generalized Lagrange multiplier rule for equality constraints in normed linear spaces. SIAM J Control 9:561–567

82. Panda BB, Kapoor OP (1978) On farthest points of sets. J Math Anal Appl 62:345–353
83. Precupanu T (1980) Duality in best approximation problem. An St Univ Iaşi 26:23–30
84. Precupanu T (1981) Some duality results in convex optimization. Rev Roum Math Pures Appl 26:769–780
85. Precupanu T (1982) On the stability in Fenchel–Rockafellar duality. An St Univ Iaşi 28:19–24
86. Precupanu T (1984) Closedness conditions for the optimality of a family of nonconvex optimization problems. Math Operforsch Stat, Ser Optim 15:339–346
87. Precupanu T (1984) Global sufficient optimality conditions for a family of non-convex optimization problems. An St Univ Iaşi 30:51–58
88. Precupanu T (1994) Sur l'existence des solutions optimales pour une famille des problèmes d'optimisation. An St Univ Iaşi 40:359–366
89. Precupanu T (2007) Some mappings associated to the farthest point problem and optimality properties. An St Univ Timişoara 45:125–133
90. Precupanu T (2011) Relationships between the farthest point problem and the best approximation problem. An St Univ Iaşi 57:1–12
91. Precupanu T (2012) Characterizations of pointwise additivity of subdifferential (to appear)
92. Precupanu A, Precupanu T (2000) Proximinality and antiproximinality for a family of optimization problems. In: Proc Natl Conf Math Anal Appl, pp. 295–308, Timisoara
93. Pshenichny BN (1965) Convex programming in linear normed spaces. Kibernetika 1:46–54 (Russian)
94. Raffin VL (1969) Sur les programmes convexes définis dans des espaces vectoriels topologiques. C R Acad Sci Paris 268:738–741; Ann Inst Fourier 20:457–491 (1970)
95. Revalski J, Théra M (1999) Generalized sums of monotone operators. C R Acad Sci, Ser 1 Math 329:979–984
96. Ritter K (1967) Duality for nonlinear programming in a Banach space. SIAM J Appl Math 15:294–302
97. Robinson SM (1976) Regularity and stability of convex multivalued functions. Math Oper Res 1:130–143
98. Rockafellar RT (1966) Extension of Fenchel's duality theorems for convex functions. Duke Math J 33:81–90
99. Rockafellar RT (1967) Duality and stability in extremum problems involving convex functions. Pac J Math 21:167–187
100. Rockafellar RT (1969) Convex analysis. Princeton University Press, Princeton
101. Rockafellar RT (1971) Saddle-points and convex analysis. In: Kuhn HW, Szegö GP (eds) Differential games and related topics. North-Holland, Amsterdam, pp 109–128
102. Rockafellar RT (1974) Augmented Lagrange multiplier functions and duality in nonconvex programming. SIAM J Control 12:268–285
103. Rosinger R (1978) Multiobjective duality without convexity. J Math Anal Appl 66:442–450
104. Schechter M (1972) Linear programs in topological linear spaces. J Math Anal Appl 37:492–500
105. Singer I (1971) Best approximation in normed linear spaces by elements of linear subspaces. Springer, Berlin
106. Singer I (1980) Maximization of lower semi-continuous convex functionals on bounded subsets on locally convex spaces. II. Quasi-Lagrangian duality theorems. Results Math 3:235–248
107. Singer I (2006) Duality for nonconvex approximation and optimization. Springer, Berlin
108. Stegall C (1978) Optimization of functions on certain subsets of Banach spaces. Math Annalen 236:171–176
109. Stoer J (1963) Duality in nonlinear programming and the minimax theorems. Numer Math 5:371–379
110. Stoer J, Witzgall C (1970) Convexity and optimization in finite dimension. Springer, Berlin
111. Toland JF (1978) Duality in nonconvex optimization. J Math Anal Appl 66:339–354
112. Tuy H (1964) Sur une classe des programmes nonlinéaires. Bull Acad Pol 12:213–215

113. Tuy H (1977) Stability property of a system of inequalities. Math Operforsch Stat, Ser Optim 8:27–39
114. Ursescu C (1973) Sur une généralisation de la notion de différentiabilité. Atti Accad Naz Lincei, Rend Cl Sci Fis Mat Nat 54:199–204
115. Ursescu C (1975) A differentiable dependence on the right-hand side of solutions of ordinary differential equations. Ann Pol Math 31:191–195
116. Ursescu C (1975) Multifunctions with closed convex graph. Czechoslov Math J 25:438–441
117. Valadier M (1972) Sous-différentiabilité de fonctions convexes à valeurs dans un espace vectoriel ordonné. Math Scand 30:65–74
118. van Slyke R, Wets R (1968) A duality theory for abstract mathematical programs with applications to optimal control theory. J Math Anal Appl 22:679–706
119. Varaiya (1967) Nonlinear programming in Banach space. SIAM J Appl Math 15:284–293
120. Zălinescu C (1978) A generalization of Farkas lemma and applications to convex programming. J Math Anal Appl 66:651–678
121. Zălinescu C (1983) Duality for vectorial nonconvex optimization by convexification and applications. An ştiinţ Univ Al I Cuz Iaşi 29:15–34
122. Zlobek (1970) Asymptotic Kuhn–Tucker conditions for mathematical programming in a Banach space. SIAM J Control 8:505–512
123. Zowe J (1974) Subdifferentiability of convex functions with values in an ordered vector space. Math Scand 34:63–83
124. Zowe J, Kurcyusz S (1979) Regularity and stability for the mathematical programming problems in Banach spaces. Appl Math Optim 5:49–62

Chapter 4
Convex Control Problems in Banach Spaces

This chapter is concerned with the optimal convex control problem of Bolza in a Banach space. The main emphasis is put on the characterization of optimal arcs (the maximum principle) as well as on the synthesis of optimal controllers. Necessary and sufficient conditions of optimality, generalizing the classical Euler–Lagrange equations, are obtained in Sect. 4.1 in terms of the subdifferential of the convex cost integrand. The abstract cases of distributed and boundary controls are treated separately. The material presented in this chapter closely parallels that exposed in Chap. 3 and, as a matter of fact, some results given here can be obtained formally from those of Chap. 3. However, there are significant differences and we have a greater number of specific things that can be said or done in the case of the optimal control problem than in the case of the constrained optimization problems considered earlier.

4.1 Distributed Optimal Control Problems

This section is devoted to the presentation of optimal control problems with convex cost criterion and governed by linear infinite-dimensional differential systems in Banach spaces.

4.1.1 Formulation of the Problem and Basic Assumptions

To begin with, we present the abstract settings of distributed control systems we shall work with in the sequel. To this purpose, we frequently refer to the notation and concepts exposed in Sect. 1.4.

From now on, E and U are two real Banach spaces with norms denoted by $|\cdot|$, $\|\cdot\|$ and with duals E^* and U^*, respectively. The norms of E^* and U^*, which are always dual norms, are again denoted by $|\cdot|$ and $\|\cdot\|$, respectively. We denote by $(\cdot,\cdot)$ and $\langle\cdot,\cdot\rangle$ the duality between E, E^* and U, U^*, respectively.

V. Barbu, T. Precupanu, *Convexity and Optimization in Banach Spaces*, Springer Monographs in Mathematics,
DOI 10.1007/978-94-007-2247-7_4,

Consider in E the linear evolution process described by the differential equation

$$x'(t) = A(t)x(t) + (Bu)(t) + f(t), \quad 0 \le t \le T, \tag{4.1}$$

where $x : [0, T] \to E$ is the unknown function and $u : [0, T] \to U$, $f : [0, T] \to E$ are given.

The function u is the *input* or the *controller* of the state system (4.1) and x is the *output* or the *state*. System (4.1) is called a *controlled system* (or a *state system*). Roughly speaking, the object of control theory is to modify a given dynamical system (of the form (4.1)) by adjustment of a certain control parameter u in order to achieve a desired behavior of the motion x. In this context, we may speak about a control approach to the dynamical system:

$$x'(t) = A(t)x(t) + f(t). \tag{4.2}$$

Usually, the parameter control u is selected from a certain admissible class according to some optimum principle.

In the following, we use also the term of *control* instead of *controller* and take $L^p(0, T; U)$, $1 \le p \le \infty$, as our space of controllers. The function f is fixed in $L^1(0, T; E)$.

As regard the operators $A(t)$ and B, the following assumptions are in effect throughout this section.

(A) $\{A(t); 0 \le t \le T\}$ *generates an evolution operator* $U(t, s)$; $0 \le t \le T$, *on* E *and the adjoint* $U^*(t, s)$ *of* $U(t, s)$ *is strongly continuous on* $\Delta = \{0 \le s \le t \le T\}$.

(B) B *is a linear continuous operator from* $L^p(0, T; U)$ to $L^p(0, T; E)$. *Furthermore,* B *is "causal", that is,* $\chi_t Bu = \chi_t B \chi_t u$ *for all* $u \in L^p(0, T; U)$ *and a.e.* $t \in]0, T[$.

Here, $\chi_t = \mathbb{1}_{[0,t]}$ is the characteristic function of the interval $[0, t]$.

By solution to (4.1) we mean, of course, a continuous function $x : [0, T] \to E$ which satisfies (4.1) in the *mild* sense (1.116), that is,

$$x(t) = U(t, 0)x(0) + \int_0^t U(t, s)\big((Bu)(s) + f(s)\big)\, \mathrm{d}s, \quad \forall t \in [0, T]. \tag{4.1$'$}$$

Perhaps the simplest and the most frequent example of such an operator B is the (memoryless) operator

$$(Bu)(t) = B(t)u(t) \quad \text{a.e. } t \in]0, T[, \tag{4.3}$$

where $B(t) \in L(U, E)$ for all $t \in [0, T]$ and the function $B(t) : [0, T] \to L(U, E)$ is strongly measurable (that is, $B(t)u$ is measurable for every $u \in U$). Assumption (B) is satisfied in this case if one assume that

$$\|B(t)\|_{L(U,E)} \le \eta(t) \quad \text{a.e. } t \in]0, T[,$$

where $\eta \in L^\infty(0, T)$. In other situations, B arises as a Volterra integral operator or is defined by certain linear operator equations.

Certainly, (4.1) is not the most general class of distributed control systems. It is true, however, that many important classes of system have such a representation. In applications to partial differential equations, E is usually the space $L^p(\Omega)$ on a bounded and open subset Ω of $\mathbb{R}^n$, $A(t)$ is an elliptic differential operator on Ω and the forcing term $B(t)u$ acts on all of Ω (this means that u is a distributed control). However, other dynamical systems such as hyperbolic equations and functional differential equations can be written in this form.

In the sequel, we assume that the spaces E and U are reflexive and strictly convex along with their duals E^* and U^*, and we denote by $\Phi : E \to E^*$ and $\Psi : U \to U^*$ the duality mappings of E and U, respectively. As noticed earlier, our assumptions on E and U imply that Φ and Ψ are single-valued and demicontinuous.

Very often, the control approach to a dynamical system of the form (4.1) can be expressed as the problem of minimization of a certain functional (*cost functional*) defined on the set of admissible controllers and the states of system (4.1). Now, we formulate a general class of such problems.

Problem (P) Find a pair $(x^*, u^*) \in C([0, T]; E) \times L^p(0, T; U)$ which minimizes the functional

$$\int_0^T L\big(t, x(t), u(t)\big)\,\mathrm{d}t + \ell\big(x(0), x(T)\big)$$

in the class of all the functions $(x, u) \in C([0, T]; E) \times L^p(0, T; U)$ subject to (4.1) and to the state constraint

$$x(t) \in K \quad \text{for all } t \in [0, T]. \tag{4.4}$$

Here, $L : (0, T) \times E \times U \to \overline{\mathbb{R}}^*$ and $\ell : E \times E \to \overline{\mathbb{R}}^*$ are given functions and K is a subset of E.

A pair (x^*, u^*), for which the infimum in Problem (P) is attained, is called an *optimal pair* of Problem (P). The state function x^* is then called *optimal arc* and the corresponding control u^* is called *optimal controller* or *optimal control*.

According to terminology coming from classical mechanics, the function L is called a *Lagrangian*. In fact, in the special case where $U = E$, $A(t) \equiv I$, $B(t) \equiv I$, $f \equiv 0$, Problem (P) reduces to the classical problem of the calculus of variations, that is,

$$\operatorname{Min}\left\{\int_0^T L\big(t, x(t), x'(t)\big)\,\mathrm{d}t + \ell\big(x(0), x(T)\big)\right\}.$$

Problem (P) is studied here under the following assumptions.

(C) *The functions ℓ and $L(t)$, $0 \le t \le T$, are lower-semicontinuous and convex on $E \times E$ (resp. $E \times U$) with values in $\overline{\mathbb{R}}^*$. Furthermore, the following conditions hold.*

(i) *For all* $(y, v) \in E \times U$, *the functions* $L(t, y, v) : [0, T] \to \overline{\mathbb{R}}^* =]-\infty, +\infty]$ *and* $J_\lambda^L(t, y, v) : [0, T] \to E \times U$ *are measurable. There exists* $v_0 \in L^p(0, T; U)$ *such that* $L(t, 0, v_0) \in L^p(0, T)$.

(ii) *There exist* $r_0 \in L^2(0, T; E^*)$, $s_0 \in L^\infty(0, T; U^*)$ *and* $g_0 \in L^1(0, T)$ *such that, for all* $(y, v) \in E \times U$, *one has*

$$L(t, y, v) \geq \big(y, r_0(t)\big) + \big\langle v, s_0(t)\big\rangle + g_0(t) \quad \text{a.e. } t \in]0, T[. \tag{4.5}$$

(iii) *For each* $x_0 \in E$ *there are a neighborhood* $\mathcal{O}$ of x_0, *functions* $\alpha, \beta \in L^p(0, T)$ *and a map* $\Sigma : [0, T] \times \mathcal{O} \to U$ *such that* $t \to \Sigma(t, y(t))$ *is measurable on* $[0, T]$ *for every measurable function* $y \in [0, T] \to \mathcal{O}$ *and*

$$L\big(t, y, \Sigma(t, y)\big) \leq \alpha(t) \quad \text{a.e. } t \in]0, T[, \ \forall y \in \mathcal{O}, \tag{4.6}$$

$$\big\|\Sigma(t, y)\big\| \leq \beta(t) \quad \text{a.e. } t \in]0, T[, \ \forall y \in \mathcal{O}. \tag{4.7}$$

Here, $J_\lambda^L(t, y, v) = (y_\lambda, v_\lambda) \in E \times U$ denotes the solution to the equation (see (1.109))

$$\big\{\Phi(y_\lambda - y), \Psi(v_\lambda - v)\big\} + \lambda \partial L(t, y_\lambda, v_\lambda) \ni 0, \tag{4.8}$$

where $\partial L(t) : E \times U \to E^* \times U^*$ is the subdifferential of $L(t)$.

The function $x \in C([0, T]; E)$ is said to be *feasible* for Problem (P) if there exists $u \in L^p(0, T; U)$ such that x is a solution to (4.1) and

$$L(t, x, u) \in L^1(0, T); \qquad x(t) \in K \quad \text{for } t \in [0, T].$$

We say that an end-point pair $(x_0, x_T) \in E \times E$ is *attainable* for L if there is a feasible function x such that $x(0) = x_0$ and $x(T) = x_T$. The set of all attainable point pairs is denoted by K_L. The last two assumptions concern K and K_L only.

(D) *K is a closed and convex subset of E. There is at least one feasible arc x such that*

$$\big(x(0), x(T)\big) \in \mathrm{Dom}(\ell), \qquad x(t) \in \operatorname{int} K \quad \text{for } t \in [0, T].$$

Here, $\mathrm{Dom}(\ell)$ denotes, as usual, the effective domain of ℓ.

(E) *There is at least one attainable pair* $(x_0, x_T) \in K_L \cap \mathrm{Dom}(\ell)$ *such that one of the following two conditions holds*:

$$x_T \in \operatorname{int}\big\{h \in E;\ (x_0, h) \in K_L\big\}, \tag{4.9}$$

$$x_T \in \operatorname{int}\big\{h \in E;\ (x_0, h) \in \mathrm{Dom}(\ell)\big\}. \tag{4.10}$$

While the role played by the above hypotheses will become clear later on, some comments here might be in order.

First, we note that condition (i) in Hypothesis (C) implies that $L(t, y(t), v(t))$ is a Lebesgue measurable function of t whenever $y(t)$ and $v(t)$ are E-valued (resp. U-valued) Lebesgue measurable functions. It turns out that, if E and U are separable

Hilbert spaces, then condition (i) is satisfied if and only if L is a *convex normal integrand* in the sense of the definition given in Sect. 2.2 (see Example 2.52).

We notice also that conditions (i) and (ii) of Assumption (C) imply, in particular, that for every $(x,u) \in L^2(0,T;E) \times L^p(0,T;U)$ the integral $\int_0^T L(t,x(t),u(t))\,\mathrm{d}t$ is well defined (unambiguously $+\infty$ or a real number).

If L is independent of t, then both conditions (i) and (ii) automatically hold. In this case, r_0, s_0 and g_0 may be taken to be constant functions.

Condition (iii), while seemingly complicated, is implied by others that are more easily verified.

1. If L is independent of t, then (iii) is implied by the following condition.

The spaces E^* and U are uniformly convex and the Hamiltonian function H associated with L is finite on $E \times U^*$. (If E^* and U are separable, the condition that E^* and U are uniformly convex is no longer necessary.)

We recall that (see (2.155))

$$H(y,p) = \sup\big\{(p,y^*) - L(y,y^*);\ y^* \in U^*\big\}.$$

Let $\partial H(y,p) = \{-\partial_y H(y,p), \partial_p H(y,p)\}$ be the subdifferential of H at $(y,p) \in E \times U^*$. Since the above condition on H implies that the map $\partial H : E \times U^* \to E^* \times U$ is monotone and, therefore, locally bounded (see Theorem 1.144), we may infer that for every $y_0 \in E$ there exists a neighborhood of y_0 and a real constant C such that $-H(y,0) \le C$ for every $y \in \mathscr{O}$ and

$$\sup\big\{\|v\|;\ v \in \partial_p H(y,0)\big\} \le C, \quad \forall y \in \mathscr{O}.$$

Then, by virtue of the conjugacy relation between $L(y,\cdot)$ and $H(y,\cdot)$, we have (see Theorem 2.112)

$$L(y,v) = -H(y,0) \quad \text{for all } y \in E \text{ and } v \in \partial_p H(y,0).$$

Let $\Gamma y = (\Gamma_1 y, \Gamma_2 y) \in E^* \times U$ be the element of minimum norm in $\partial H(y,0)$, that is, $\Gamma y = (\partial H(y,0))^\circ$.

By Proposition 1.146, $(\partial H)_\lambda(y,0) \to \Gamma y$ strongly in $E^* \times U$ for $\lambda \to 0$ and $(\partial H)_\lambda$ is continuous from $E \times U^*$ to $E^* \times U$. Therefore, we may conclude that, for any measurable function $y : [0,T] \to E$, the function $t \to \Gamma_2 y(t)$ is measurable on $[0,T]$. Hence, the mapping $\Sigma(t,y) = \Gamma_2 y$ satisfies all the conditions required in Hypothesis (C)(iii), because, as noticed above, $L(y,\Gamma_2 y) = -H(y,0)$ and the function H is locally bounded. It should be emphasized that the condition $H(y,p) < +\infty$ for all $y \in E$ and $p \in U^*$ is implied by the following growth condition:

$$\lim_{\|u\| \to \infty} \frac{L(y,u)}{\|u\|} = +\infty, \quad \forall y \in E. \tag{4.11}$$

The rest of the condition, that is, $H(y,p) > -\infty$ for all $(y,p) \in E \times U^*$, amounts to saying that there is no $y \in E$ such that $L(y,\cdot) = +\infty$. As a matter of fact, condition (iii) also implies this stringent requirement.

2. $L(t,y,v)=\varphi(t,y)+\psi(t,v)$; $y\in E$, $v\in U$, $t\in[0,T]$, where $\varphi(t):E\to\mathbb{R}$, $\psi(t):U\to\overline{\mathbb{R}}^*$ are lower-semicontinuous and convex.

In this case, condition (iii) is obviously implied by the following one.

There exists $v_0\in L^p(0,T;U)$ such that $\psi(t,v_0)\in L^p(0,T)$ and the mapping $y\to\varphi(t,y)$ is locally bounded from E to $L^p(0,T)$.

3. If the spaces E and U are both finite-dimensional, then Assumption (C) is implied by the following one.

The Hamiltonian function $H(t,x,q)$ is (finite and) L^p-summable on $[0,T]$ as a function of t for each $x\in E$ and $q\in U^*$.

For the proof, the reader is referred to Rockafellar's paper [39].

As regards Assumption (E), part (4.9), it has a severe implication on the state system (4.1). In fact, (4.9) requires that at least for one $x_0\in E$ the attainable set

$$\Omega_T=\left\{\int_0^T U(T,x)\big((Bu)(s)+f(s)\big)\,\mathrm{d}s+U(T,0)x_0;\ u\in L^p(0,T;U)\right\}$$

has a nonempty interior in E. However, it is known (see, e.g., Balakrishnan [1], Fattorini [26]) that $\operatorname{int}\Omega_T=\emptyset$ unless $A(t)\equiv A$ generates a group on the space E and B is onto. Thus, from the point of view of applications in infinite dimensions, assumption (4.10) is more convenient.

We pause briefly to observe that our hypotheses on $L(t)$ do not prevent us from treating some apparently unmanageable cases as that of the end-point constraint

$$\big(x(0),x(T)\big)\in D$$

or control constraint

$$u(t)\in U_0(t)\quad \text{a.e. } t\in\,]0,T[.$$

(Here, D is a closed, convex subset of $E\times E$ and $U_0(t)$ a family of closed and convex subsets of U.) In fact, these situations can, implicitly, be incorporated into Problem (P) by defining (or redefining, as the case may be)

$$\ell(x_1,x_2)=+\infty\quad \text{if } (x_1,x_2)\,\overline{\in}\, D$$

and

$$L(t,x,u)=+\infty\quad \text{if } u\,\overline{\in}\, U_0(t).$$

Formally, also the state constraint (4.4) can be incorporated into Problem (P) by redefining $L(t)$ as

$$L(t,x,u)=+\infty\quad \text{if } x\,\overline{\in}\, K.$$

However, as remarked earlier, condition (iii) in Assumption (C) precludes this situation, so that it is better to keep the state constraint explicit and separate.

4.1.2 Existence of Optimal Arcs

The existence of an optimal arc in Problem (P) is a delicate problem under the general assumptions presented above. The main reason is that the convex function $\phi(u) = \int_0^T L(x^u(t), u(t))\,\mathrm{d}t + \ell(x^u(0), x^u(T))$, where x^u is given by (4.1), generally is not coercive on $L^p(0,T;U)$ and so, the standard existence result (Theorem 2.11, for instance) is not applicable, that is why some additional hypotheses must be imposed.

We study here the existence of an optimal pair in Problem (P) under the following assumptions on $L(t)$ and ℓ.

(a) *The functions $L(t)$ and ℓ satisfy condition* (i) *of Assumption* (C).
(b) *There exists a continuous convex, nondecreasing function $\omega : \mathbb{R}^+ \to \mathbb{R}^+$ such that*

$$\omega(0) = 0, \qquad \lim_{r\to\infty} \frac{\omega(r^p)}{r} = +\infty \quad \textit{and} \quad L(t,x,u) \geq \omega\big(\|u\|^p\big) - \beta_0|x| + \gamma(t),$$

where $\gamma \in L(0,T)$ and β_0 is a real constant.
(c) *There exists a nondecreasing function $j : \mathbb{R}^+ \to \mathbb{R}^+$ such that*

$$\lim_{r\to\infty} \frac{j(r)}{r} = +\infty \quad \textit{and} \quad \ell(x_1, x_2) \geq j\big(|x_1|\big) - \eta|x_2| \quad \textit{for all } (x_1, x_2) \in E \times E,$$

where η is a real constant.
(d) *K is a closed convex subset of E and $K_L \cap \mathrm{Dom}(\ell) \neq \emptyset$.*

Proposition 4.1 *Let E and U be reflexive Banach spaces and let Assumptions* (A), (B) *and* (a)–(d) *hold. Then, for $1 \leq p < \infty$, Problem* (P) *has at least one solution, $(x,u) \in C([0,T];E) \times L^p(0,T;U)$.*

Proof We set

$$I(x,u) = \int_0^T L\big(t, x(t), u(t)\big)\,\mathrm{d}t + \ell\big(x(0), x(T)\big),$$

where $(x,u) \in C([0,T];E) \times L^p(0,T;U)$ satisfy (4.1). Since B is continuous from $L^p(0,T;U)$ to $L^1(0,T;E)$, we have

$$\int_0^T |B\chi_t u|\,\mathrm{d}s \leq \|B\| \bigg(\int_0^T \|\chi_t u\|^p\,\mathrm{d}s\bigg)^{\frac{1}{p}}, \qquad \forall u \in L^p(0,T;U),$$

where $\|B\|$ is the operator norm of B. On the other hand, since B is "causal", $\chi_t B(\chi_t u) = \chi_t Bu$ and, therefore,

$$\int_0^t \big|(Bu)(s)\big|\,\mathrm{d}s \leq \|B\| \bigg(\int_0^t \big\|u(s)\big\|^p\,\mathrm{d}s\bigg)^{\frac{1}{p}}, \quad \forall t \in [0,T].$$

Then, by (4.3), we see that

$$|x(t)| \le C\left(1 + |x(0)| + \left(\int_0^t \|u(s)\|^p \, ds\right)^{\frac{1}{p}}\right), \quad 0 \le t \le T. \tag{4.12}$$

We have, therefore,

$$\begin{aligned}
I(x,u) &= \int_0^T L(t,x,u)\,dt + \ell\big(x(0),x(T)\big) \\
&\ge \int_0^T \omega\big(\|u(t)\|^p\big)\,dt - \beta_0 \int_0^T |x(t)|\,dt + \int_0^T \gamma(t)\,dt \\
&\quad + j\big(|x(0)|\big) - \eta|x(T)| \\
&\ge T\omega\left(T^{-1}\int_0^T \|u(t)\|^p\,dt\right) - \beta_0 \int_0^T |x(t)|\,dt \\
&\quad + \int_0^T \gamma(t)\,dt + j\big(|x(0)|\big) - \eta|x(T)| \\
&\ge T\omega\big(T^{-1}\|u\|_p^p\big) - C_1\|u\|_p + j\big(|x(0)|\big) - C_2\big(|x(0)|\big) + C_3.
\end{aligned} \tag{4.13}$$

This implies that $\inf I(x,u) > -\infty$ and, by (d), $I \not\equiv +\infty$. Thus, $d = \inf I(x,u) < +\infty$. Let $\{(x_n,u_n)\} \subset C([0,T];E) \times L^p(0,T;U)$ be such that

$$d \le I(x_n,u_n) \le d + n^{-1}. \tag{4.14}$$

By (4.13) and (c), we see that $\{u_n\}$ is bounded in $L^p(0,T;U)$ and $\{x_n(0)\}$ is bounded in E. Hence, $\{x_n\}$ is bounded in $C([0,T];E)$ and $\{u_n\}$ is weakly compact in $L^p(0,T;U)$ if $p > 1$. If $p = 1$, for every measurable subset Ω of $[0,T]$, we have

$$\begin{aligned}
d + 1 \ge d + n^{-1} &\ge I(x_n,u_n) \\
&\ge \int_0^T \omega\big(\|u_n(s)\|\big)\,ds - \beta_0 \int_0^T |x_n(s)|\,ds \\
&\quad + \int_0^T \gamma(t)\,dt + j\big(|x_n(0)|\big) - \eta|x_n(T)| \\
&\ge \int_\Omega \omega\big(\|u_n(s)\|\big)\,ds + C.
\end{aligned}$$

Hence, by the Jensen inequality,

$$|d + 1 - C| \ge \int_\Omega \omega\big(\|u_n(s)\|\big)\,ds \ge m(\Omega)\omega\big(m(\Omega)\big)^{-1} \int_\Omega \|u_n(s)\|\,ds$$

and, therefore,

$$\int_{\Omega} \|u_n(s)\| \, \mathrm{d}s \le \sup\left\{\lambda \ge 0;\ m(\Omega)\omega\left(\frac{\lambda}{m(\Omega)}\right) \le |d+1-C|\right\}. \tag{4.15}$$

We may conclude that the family $\{\int_{\Omega} \|u_n(s)\| \, \mathrm{d}s;\ \Omega \subset [0,T]\}$ is equibounded and equicontinuous. Thus, by the Dunford–Pettis criterion in $L^1(0,T;U)$ (see Theorem 1.121)), $\{u_n\}$ is weakly compact in $L^1(0,T;U)$. Hence, without loss of generality, we may assume that there exists some $u \in L^p(0,T;U)$ such that

$$u_n \to u \quad \text{weakly in } L^p(0,T;U).$$

Since $\{x_n\}$ are uniformly bounded on $[0,T]$ and E is reflexive, we may assume that $x_n(0) \to x_1$ weakly in E and, by (4.1′), we see that

$$x_n(t) \to x(t) = U(t,0)x_1 + \int_0^t U(t,s)\big((Bu)(s) + f(s)\big) \, \mathrm{d}s$$

weakly in E for every $t \in [0,T]$. Since ℓ is weakly lower-semicontinuous on $E \times E$ (because it is convex and lower-semicontinuous), we have

$$\liminf_{n\to\infty} \ell\big(x_n(0), x_n(T)\big) \ge \ell\big(x(0), x(T)\big). \tag{4.16}$$

Next, our assumption on $L(t)$ implies (see Proposition 2.19) that the function $(y,v) \to \int_0^T L(t, y(t), v(t)) \, \mathrm{d}t$ is convex and lower-semicontinuous on $L^1(0,T;E) \times L^1(0,T;U)$. Hence, this function is weakly lower-semicontinuous, so that we have

$$\liminf_{n\to\infty} \int_0^T L(t, x_n, u_n) \, \mathrm{d}t \ge \int_0^T L(t, x, u) \, \mathrm{d}t.$$

Along with (4.13) and (4.16), the latter implies that $I(x,u) = d$, thereby completing the proof. □

Remark 4.2 From the preceding proof, it is apparent that, for $p > 1$, the condition on ω in assumption (b) can be weakened to

$$\liminf_{r\to\infty} \frac{\omega(r^p)}{r} > 0.$$

Notice also that the weak lower-semicontinuity of I on $L^1(0,T;E) \times L^1(0,T;U)$ was essential for the proof of the existence. Hence, without the convexity of $L(t,\cdot,\cdot)$ (or of $L(t,y,\cdot)$ if the map $u \to y$ is compact), there is little motivation to study Problem (P) (since it might have no solution). However, using the Ekeland variational principle (see Theorem 2.43 and Corollaries 3.74, 3.75), one might show even in this case that, for each $\varepsilon > 0$, there is an approximation minimum,

$$u_\varepsilon = \arg\inf\left\{I\big(x^u, u\big) + \varepsilon\left(\int_0^T |u - u_\varepsilon|^p \, \mathrm{d}t\right)^{\frac{1}{p}}\right\}$$

(see Remark 2.54).

Remark 4.3 Proposition 1.12 is a particular case of a general result established in [38] by Popescu. For other sharp existence results in Problem (P), we refer the reader to the works of Rockafellar [43], Ioffe [29, 30], Olech [36, 37].

4.1.3 The Maximum Principle

We present here some optimality theorems of the maximum principle type for Problem (P). The main theorem of this section, Theorem 4.5 below, characterizes the optimal arcs of Problem (P) as generalized solutions to a certain Euler–Lagrange system associated to Problem (P).

We denote by $\mathscr{K}$ the closed convex subset of $C([0,T];E)$ defined by

$$\mathscr{K}=\big\{x\in C\big([0,T];E\big);\ x(t)\in K,\ \forall t\in[0,T]\big\}.$$

Given a function $w:[0,T]\to E^*$ of bounded variation on $[0,T]$, we denote, as usual (see Sect. 1.3.3), by dw the E^*-valued Stieltjes–Lebesgue measure on $[0,T]$ corresponding to w.

Definition 4.4 We say that a pair $(x^*,u^*)\in C([0,T];E)\times L^p(0,T;U)$ is *extremal* for Problem (P) if there exist the functions $q\in L^1(0,T;E^*)$, $w\in BV([0,T];E^*)$ and $p^*:[0,T]\to E^*$ satisfying with x^*,u^* the equations

$$x^*(t)=U(t,0)x^*(0)+\int_0^t U(t,s)\big((Bu^*)(s)+f(s)\big)\,\mathrm{d}s,\quad 0\le t\le T, \tag{4.17}$$

$$p^*(t)=U^*(T,t)p_T^*-\int_t^T U^*(s,t)q(s)\,\mathrm{d}s-\int_t^T U^*(s,t)\,\mathrm{d}w(s), \tag{4.18}$$

$$\int_0^T\big(dw(t),x^*(t)-y(t)\big)\ge 0\quad\text{for all } y\in\mathscr{K}, \tag{4.19}$$

$$\big(q(t),(B^*p^*)(t)\big)\in\partial L\big(t,x^*(t),u^*(t)\big)\quad\text{a.e. } t\in\,]0,T[, \tag{4.20}$$

$$\big(p^*(0),-p_T^*\big)\in\partial\ell\big(x^*(0),x^*(T)\big). \tag{4.21}$$

Here, $\int_t^T U^*(s,t)\,\mathrm{d}w(s)$ stands for the Riemann–Stieltjes integral of $U^*(\cdot,t):[t,T]\to L(E^*,E^*)$ with respect to the function of bounded variation $w:[t,T]\to E^*$ (see Sect. 1.3.3), U^* is the adjoint of U and B^* is the adjoint of B.

Such a function p^* is called a *dual extremal arc* of Problem (P).

Let $M(0,T;E^*)$ be the dual space of $C([0,T];E)$ and let $\mu_w\in M(0,T;E^*)$ be defined by

$$\mu_w(x)=\int_0^T(\mathrm{d}w,x),\quad x\in C\big([0,T];E\big).$$

Then, (4.19) may be rewritten as

$$\mu_w \in \mathscr{N}(x^*, \mathscr{K}),$$

where $\mathscr{N}(x^*, \mathscr{K})$ is the cone of normals to $\mathscr{K}$ at x^*, that is,

$$\mathscr{N}(x^*, \mathscr{K}) = \{\mu \in M(0, T; E^*);\ \mu(x^* - y) \geq 0,\ \forall y \in \mathscr{K}\}.$$

By analogy with (4.1), we may say that the dual extremal arc p^* is a solution to the differential equation

$$(p^*)' + A^*(t)p^* = q + \mathrm{d}w \quad \text{on } [0, T], \qquad p^*(T) = p_T,$$

but the exact sense of this equation is given by (4.18).

Note that, in contrast to the solution to (4.1), $p^*(t)$ does not need to be continuous, unless $K = E$ (in this case, $\mathscr{N}(x^*, K) = \{0\}$ and w is constant on $[0, T]$). However, since the function $t \to \int_t^T U^*(s, t)\,\mathrm{d}w(s)$ is of bounded variation on $[0, T]$, the function p^* arises under the form $p_1 + p_2$, where $p_1(t) = U^*(T, t)p_T^* - \int_t^T U^*(s, t)q(s)\,\mathrm{d}s$ is continuous and p_2 is of bounded variation on $[0, T]$. Hence, $p^*(t + 0)$ and $p^*(t - 0)$ exist at every point $t \in [0, T]$ (we make the convention $p^*(0 - 0) = p^*(0)$ and $p^*(T + 0) = p^*(T)$).

We see, by (4.18), that the points of discontinuity for p^* are just the points t, where w is discontinuous and, as we shall see later, these points belong to the set of all t for which $x^*(t)$ lies on the boundary of K. As a matter of fact, we may take the function continuous from the left on $]0, T]$ and regard p_T^* as $p^*(T)$. Parenthetically, we note that in terms of the Hamiltonian function $H(t)$ associated to $L(t)$, (4.20) can be written in the classical form (see formula (2.157))

$$\begin{aligned} q(t) &\in -\partial_x H\big(t, x^*(t), (B^*p^*)(t)\big), \\ u^*(t) &\in \partial_p H\big(t, x^*(t), (B^*p^*)(t)\big). \end{aligned} \tag{4.22}$$

Equations (4.17), (4.18), (4.19), and (4.22) represent the Hamiltonian form of the generalized Euler–Lagrange equations. If $K = E$, then (4.17)–(4.20) can be written as

$$\begin{aligned} x^{*\prime}(t) - A(t)x^*(t) &\in B\partial_p H\big(t, x^*(t), (B^*p^*)(t)\big) + f(t), \\ p^{*\prime}(t) + A^*(t)p^*(t) &\in -\partial_x H\big(t, x^*(t), (B^*p^*)(t)\big), \end{aligned} \tag{4.23}$$

which resemble the classical Hamiltonian equations.

Observe that the set $\partial_p H(t, x^*, (B^*p^*))$ consists of the control vectors $u \in U$ for which the supremum of $\{(u, (B^*p^*)(t)) - L(t, x^*, u)\}$ is attained. This clarifies the equivalence between the above optimality conditions and the well-known maximum principle.

The main result is the following theorem.

Theorem 4.5 *Let Assumptions* (A), (B), (C), (D) *and* (E) *be satisfied, where* $2 \leq p < \infty$, *the spaces* E *and* U *are reflexive and strictly convex together with their* *duals and E is separable. Then, the pair $(x^*, u^*) \in C([0,T]; E) \times L^p(0,T; U)$ is optimal in Problem* (P) *if and only if it is extremal. If, in addition, B is given by* (4.3), *then the function q in* (4.18) *belongs to $L^p(0,T; E^*)$.*

Let $\dot{w} \in L^1(0,T; E^*)$ be the weak derivative of w, and $w_s \in BV([0,T]; E^*)$ be the singular part of the function w, that is,

$$w(t) = \int_0^t \dot{w}\,\mathrm{d}s + w_s(t), \quad 0 \le t \le T. \tag{4.24}$$

As noticed before, in Sect. 1.3.3 (see (1.88)), the measure dw_s is the *singular part* of $\mathrm{d}w$.

Let us denote by $N_K(x) \subset E^*$ the cone of normals to K at x, that is, $N_K(x) = \partial I_K(x)$. In terms of w and w_s, Theorem 4.5 can be made more precise as follows.

Theorem 4.6 *Under the assumptions of Theorem* 4.5, *the pair (x^*, u^*) is optimal in Problem* (P) *if and only if there exist functions $q \in L^1(0,T; E^*)$, $w \in BV([0,T]; E^*)$, and p^* satisfying along with x^*, u^*,* (4.17), (4.18), (4.20), (4.21), *and*

$$\dot{w}(t) \in N_K\big(x^*(t)\big) \quad a.e.\ t \in]0,T[, \tag{4.25}$$

$$\mathrm{d}w_s \in \mathscr{N}(x^*, \mathscr{K}). \tag{4.26}$$

If B is defined by (4.3), *then $q \in L^p(0,T; E^*)$.*

Remark 4.7 The condition that E is separable is not absolutely necessary, but it has been imposed in order to simplify the proof. Anyway, as we see in the proof of Theorem 4.5 in the unconstrained case $K = E$ it is obviously superfluous.

It is interesting that, by Theorem 4.5, it follows under certain circumstances that the optimal controllers u^* are continuous functions, though control functions u which are merely p-summable have been admitted. For instance, if $K = E$, B is given by (4.3) with $B(t)$ continuous in t and $\partial_p H$ is single-valued and continuous, then we see by (4.22) that u^* is continuous on $[0,T]$. This is a "smoothing effect" of optimality on the control input. Other information on the optimal controller u^* is contained in (4.22).

Now, let us consider the particular case in which $E = H$ is a Hilbert space and $\{A(t); 0 \le t \le T\}$ is a family of linear continuous operators form V to V' satisfying conditions (j), (jj) and (jjj) of Proposition 1.149.

Here, V is a real Hilbert space continuously and densely imbedded in H and V' is its dual ($V \subset H \subset V'$).

We further assume that B is defined by (4.3), $f \in L^2(0,T; H)$ and $p = 2$.

Then, by Proposition 1.149, the solution x to (4.1) belongs to the space

$$W(0,T) = \big\{x \in L^2(0,T; V);\ x' \in L^2(0,T; V')\big\}. \tag{4.27}$$

Next, if $K = H$, then also the dual arc p^* belongs to $W(0,T)$ and the extremality system (4.17), (4.18) can be written in the following more precise form:

$$\begin{aligned} &x^{*\prime}(t) = A(t)x^*(t) + B(t)u^*(t) + f(t) \quad \text{a.e. } \in]0,T[, \\ &p^{*\prime}(t) = -A^*(t)p^*(t) + q(t). \end{aligned} \tag{4.28}$$

This functional setting is appropriate to describe the distributed control systems of parabolic type, and more will be said about it in Sect. 4.1.9.

Another situation in which x^* and p^* are strong solutions to (4.17) and (4.18) is that when $A(t) \equiv A$ is the infinitesimal generator of an analytic semigroup and $x^*(0) \in D(A)$, $p^*(T) \in D(A^*)$ (see Proposition 1.148).

4.1.4 Proof of Theorem 4.5

It is convenient to reformulate Problem (P) as that of minimizing a certain functional F over the space $C([0,T];E) \times L^p(0,T;U)$ where no constraints appear explicitly. Let $\mathscr{H}$ be the subset of $C([0,T];E) \times L^p(0,T;U)$ defined by

$$\mathscr{H} = \big\{(y,v) \in C\big([0,T];E\big) \times L^p(0,T;U);\ y' = A(t)y + Bv + f \text{ on } [0,T]\big\}. \tag{4.29}$$

It is elementary that $\mathscr{H}$ is a closed convex subset of $C([0,T];E) \times L^p(0,T;U)$. Now, let $F : C([0,T];E) \times L^p(0,T;U) \to \overline{\mathbb{R}}^*$ be the convex function defined by

$$F(y,v) = \begin{cases} \int_0^T L(t,y(t),v(t))\,\mathrm{d}t + \ell(y(0),y(T)), & \text{if } (y,v) \in \mathscr{H},\ y \in \mathscr{K}, \\ +\infty, & \text{otherwise.} \end{cases} \tag{4.30}$$

We note that Assumption (C) (part (i), (ii)) guarantees that the integral $\int_0^T L(t,y(t),v(t))\,\mathrm{d}t$ is well defined (that is, nowhere $-\infty$) for all $(y,v) \in \mathscr{H}$. Moreover, $F \not\equiv +\infty$ by Assumption (E) and F is convex and lower-semicontinuous on $c([0,T];E) \times L^p(0,T;U)$. The latter is easily deduced from the Fatou Lemma and condition (4.4) in Assumption (C) (see Propositions 2.53, 2.55).

In terms of the function F defined above, we can express the control problem (P) as

$$\text{Min } F(y,v) \text{ over all } (y,v) \in C\big([0,T];E\big) \times L^p(0,T;U). \tag{4.31}$$

If we denote by $J(v,y_0)$ the solution to $y' = A(t)y + Bv + f$, $y(0) = y_0$, and set $J(v,y_0) = F(y(v,y_0),v)$, we can rewrite (4.31) as

$$\text{Min}\big\{J(v,y_0);\ v \in L^p(0,T;U),\ y_0 \in F\big\} \tag{4.32}$$

and so (x^*,u^*) is optimal in Problem (P) if and only if

$$\partial J\big(u^*,x^*(0)\big) \ni 0.$$

Thus, the maximum principle formally reduces to the exact description of the subdifferential

$$\partial J : L^p(0,T;U) \times E \to L^q(0,T;U^*) \times E^*.$$

However, since the general rules presented in Chap. 3 to calculate subdifferentials are not applicable to the present situation, we proceed in a direct way with the analysis of the minimization problem (4.31).

Now, we prove *sufficiency* of the extremality conditions (4.17)–(4.21) for optimality. Let x^*, u^*, p^* and $q \in L^1(0,T;E^*)$, $w \in BV([0,T];E^*)$ satisfy (4.17)–(4.21).

By (4.20) and the definition of the "subgradient", we have for all $(x,u) \in C([0,T];E) \times L^p(0,T;U)$

$$\begin{aligned} L\big(t, x^*(t), u^*(t)\big) \le L\big(t, x(t), u(t)\big) + \big(q(t), x^*(t) - x(t)\big) \\ + \big\langle (B^*p^*)(t), u^*(t) - u(t)\big\rangle \quad \text{a.e. } t \in]0,T[\end{aligned}$$

and, by (4.21),

$$\ell\big(x^*(0), x^*(T)\big) \le \ell\big(x(0), x(T)\big) + \big(p^*(0), x^*(0) - x(0)\big) - \big(p^*(T), x^*(T) - x(T)\big).$$

Hence,

$$\begin{aligned} F(x^*,u^*) \le F(x,u) + \int_0^T \Big(\big(q(t), x^*(t) - x(t)\big) + \big(p^*(t), (Bu^*)(t) - (Bu)(t)\big)\Big)\,\mathrm{d}t \\ + \big(p^*(0), x^*(0) - x(0)\big) - \big(p^*(T), x^*(T) - x(T)\big), \end{aligned} \tag{4.33}$$

for all $(x,u) \in \mathscr{H}$ and $x \in \mathscr{K}$.

Now, using (4.1′) and (4.17), we have

$$\begin{aligned} \int_0^T \big(q(t), x^*(t) - x(t)\big)\,\mathrm{d}t = \left(\int_0^T U^*(t,0)q(t)\,\mathrm{d}t, x^*(0) - x(0)\right) \\ + \int_0^T \left(q(t), \int_0^t U(t,s)\big((Bu^*)(s) - (Bu)(s)\big)\,\mathrm{d}s\right)\mathrm{d}t. \end{aligned}$$

Interchanging the order of integration, which is easily justified by the hypotheses and Fubini's theorem, yields

$$\begin{aligned} &\int_0^T \big(q(t), x^*(t) - x(t)\big)\,\mathrm{d}t \\ &\quad = \left(x^*(0) - x(0), \int_0^T U^*(t,0)q(t)\,\mathrm{d}t\right) \\ &\qquad + \int_0^T \left((Bu^* - Bu)(s), \int_0^T U^*(t,s)q(t)\,\mathrm{d}t\right)\mathrm{d}s \end{aligned}$$

$$
\begin{aligned}
&= \big(x^*(0) - x(0), U^*(T,0)p^*(T)\big) - \big(x^*(0) - x(0), p^*(0)\big) \\
&\quad - \left(x^*(0) - x(0), \int_0^T U^*(s,0)\,\mathrm{d}w(s)\right) \\
&\quad - \int_0^T \big((Bu^* - Bu)(s), p^*(s)\big)\,\mathrm{d}s \\
&\quad - \int_0^T \left((Bu^* - Bu)(s), \int_0^T U^*(t,s)\,\mathrm{d}w(t)\right)\mathrm{d}s \\
&\quad + \int_0^T \big((Bu^* - Bu)(s), U^*(T,s)p^*(T)\big)\,\mathrm{d}s.
\end{aligned}
\tag{4.34}
$$

Here, we have also used (4.18). Then, by Proposition 1.125, we have

$$
\begin{aligned}
&\int_0^T \left(B(u^* - u)(s), \int_s^T U^*(t,s)\,\mathrm{d}w(t)\right) \\
&\quad = \int_0^T \left(\mathrm{d}w(t), \int_0^t U(t,s)B(u^* - u)(s)\,\mathrm{d}s\right) \\
&\quad = \int_0^T \big(\mathrm{d}w(t), x^*(t) - x(t)\big) - \int_0^T \big(\mathrm{d}w(t), U(t,0)\big(x^*(0) - x(0)\big)\big),
\end{aligned}
$$

while, by (4.19), we have

$$\int_0^T (\mathrm{d}w, x^* - x) \ge 0 \quad \text{for all } x \in \mathscr{K}.$$

Along with (4.33) and (4.34), the latter yields

$$F(x^*, u^*) \le F(x,u) \quad \text{for all } (x,u) \in \mathscr{H},\ x \in \mathscr{K},$$

as claimed.

Necessity The proof of necessity is more complicated and it is divided into several steps. The underlying idea behind the method to be used below is to approximate a solution (x^*, u^*) to Problem (P) by a sequence $\{(x_\lambda, u_\lambda)\}$ consisting of optimal pairs in a family of smooth optimal control problems. Roughly speaking, we use a variant of the penalty functions method mentioned earlier.

Before proceeding further, we must introduce some notation and give a brief account of the background required.

We denote by L_λ, ℓ_λ and φ_λ the regularizations of L, ℓ and $\varphi = I_K$ (see (2.58)). In other words,

$$
\begin{aligned}
&L_\lambda(t,x,u) = \inf\big\{(2\lambda)^{-1}\big(|y - x|^2 + \|v - u\|^2\big) + L(t,y,v);\ (y,v) \in E \times U\big\}, \\
&\quad \lambda > 0,
\end{aligned}
$$

$$\ell_\lambda(x_1,x_2)=\inf\bigl\{(2\lambda)^{-1}\bigl(|x_1-y_1|^2+|x_2-y_2|^2\bigr)+\ell(y_1,y_2);\ y_1,y_2\in E\bigr\}$$

and

$$\varphi_\lambda(x)=\inf\bigl\{(2\lambda)^{-1}|x-y|^2;\ y\in K\bigr\}.$$

We notice that Hypothesis (C), part (ii), when used in equalities (see Theorem 2.58)

$$L_\lambda(t,x,u)=L\bigl(t,J_\lambda^L(t,x,u)\bigr)+\frac{\lambda}{2}\bigl\|\partial L_\lambda(t,x,u)\bigr\|^2_{E^*\times U^*} \tag{4.35}$$

yields

$$L_\lambda(t,x,u)\ge\bigl(x,r_0(t)\bigr)+\bigl\langle u,s_0(t)\bigr\rangle+\xi(t)+M\lambda\bigl(1+\bigl|r_0(t)\bigr|^2\bigr),$$
$$\forall(x,u)\in E\times U,\ \lambda>0, \tag{4.36}$$

where $r_0\in L^2(0,T;E^*)$, $s_0\in L^\infty(0,T;U^*)$, $\xi\in L^1(0,T)$ and M is a positive constant independent of λ. Likewise, we have

$$\ell_\lambda(x_1,x_2)\ge(x_1,x_1^*)+(x_2,x_2^*)+M_1,\quad \text{for all } x_1,x_2\in E, \tag{4.37}$$

where x_1^*,x_2^* are some elements of E^*. In the sequel, we use the same symbol I to designate the identity operator in E, U and $E\times U$.

We note that Assumption (C), part (i), implies that, for every $\{y,v\}\in E\times U$, the function $\partial L_\lambda(t,y,v)$ is measurable on $[0,T]$. We see, by (4.35), that, for all $(y,v)\in E\times U$, $L_\lambda(t,y,v)$ is measurable on $[0,T]$. Finally, we may conclude that, for any pair of measurable functions $y:[0,T]\to E$ and $v:[0,T]\to U$, $L_\lambda(t,y(t),v(t))$ is measurable on $[0,T]$.

Moreover, $(y,v)\to L(t,y,v)$ is Gâteaux differentiable and its gradient ∇L_λ is just the subdifferential ∂L_λ; similarly for ℓ_λ and φ_λ.

Having summarized these elementary properties of L_λ and ℓ_λ, we establish now the first auxiliary result of the proof.

Let $(x^*,u^*)\in C([0,T];E)\times L^p(0,T;U)$ be a fixed optimal pair of Problem (P).

Lemma 4.8 *For every $\lambda>0$, there exist $(x_\lambda,u_\lambda)\in\mathscr{H}$, $q_\lambda\in L^p(0,T;E^*)$ and $p_\lambda\in C([0,T];E^*)$ satisfying the equations*

$$x_\lambda(t)=U(t,0)x_\lambda(0)+\int_0^t U(t,s)\bigl((Bu_\lambda)(s)+f(s)\bigr)\,\mathrm{d}s,\quad 0\le t\le T, \tag{4.38}$$

$$p_\lambda(t)=U^*(T,t)p_\lambda(T)-\int_t^T U^*(s,t)\bigl(q_\lambda+\partial\varphi_\lambda(x_\lambda)\bigr)(s)\,\mathrm{d}s, \tag{4.39}$$

$$(B^*p_\lambda)(t)+\Psi\bigl(u^*(t)-u_\lambda(t)\bigr)\bigl\|u^*(t)-u_\lambda(t)\bigr\|^{p-2}$$
$$=\partial_u L_\lambda\bigl(t,x_\lambda(t),u_\lambda(t)\bigr)\quad \textit{a.e. } t\in\,]0,T[, \tag{4.40}$$

$$q_\lambda(t)=\partial_x L_\lambda\bigl(t,x_\lambda(t),u_\lambda(t)\bigr)\quad \textit{a.e. } t\in\,]0,T[, \tag{4.41}$$

$$\bigl\{p_\lambda(0)+\Phi\bigl(x^*(0)-x_\lambda(0)\bigr),-p_\lambda(T)\bigr\}=\partial\ell_\lambda\bigl(x_\lambda(0),x_\lambda(T)\bigr). \tag{4.42}$$

Furthermore, for $\lambda \to 0$,

$$u_\lambda \to u^* \quad \textit{strongly in } L^p(0,T;U), \tag{4.43}$$

$$x_\lambda \to x^* \quad \textit{in } C\big([0,T];E\big). \tag{4.44}$$

Proof Let $F_\lambda : L^p(0,T;U) \times E \to]-\infty,+\infty]$ be the convex function defined by

$$F_\lambda(u,h) = \int_0^T \left(L_\lambda\big(t,x(t),u(t)\big) + \varphi_\lambda\big(x(t)\big) + \frac{1}{p}\big\|u(t)-u^*(t)\big\|^p \right) \mathrm{d}t$$
$$+ \ell_\lambda\big(x(0),x(T)\big) + \frac{1}{2}\big|x(0)-x^*(0)\big|^2, \quad u \in L^p(0,T;U),\ h \in E,$$

wherein

$$x(t) = U(t,0)h + \int_0^t U(t,s)\big((Bu)(s)+f(s)\big)\,\mathrm{d}s, \quad t \in [0,T]. \tag{4.45}$$

In particular, it follows by Assumption (C), part (ii), that there exists $v_0 \in L^p(0,T;U)$ such that $L(t,0,v_0) \in L^p(0,T)$. Since, by the definition of L_λ,

$$L_\lambda(t,x,u) \le (2\lambda)^{-1}\big(|x|^2 + \|u-v_0\|^2\big) + L(t,0,v_0),$$

we may infer that $-\infty < F_\lambda < +\infty$.

Moreover, we may infer by Proposition 1.8 that F_λ attains its infimum on $L^p(0,T;U) \times E$ in a unique point (u_λ,h_λ) (unique because F_λ is strictly convex). We set

$$x_\lambda(t) = U(t,0)h_\lambda + \int_0^t U(t,s)\big((Bu_\lambda)(s)+f(s)\big)\,\mathrm{d}s$$

and define

$$p_\lambda(t) = U^*(T,t)p_\lambda^T - \int_t^T U^*(s,t)\big(\partial_x L_\lambda\big(s,x_\lambda(s),u_\lambda(s)\big) + \partial\varphi_\lambda\big(x_\lambda(s)\big)\big)\,\mathrm{d}s, \tag{4.46}$$

wherein

$$p_\lambda^T = p_\lambda(T) = -y_\lambda^2; \qquad \big(y_\lambda^1,y_\lambda^2\big) = \partial\ell_\lambda\big(x_\lambda(0),x_\lambda(T)\big).$$

Since (u_λ,h_λ) is a minimum point of F_λ and the functions $\|\cdot\|$, L_λ, ℓ_λ and φ_λ are Gâteaux differentiable, we have

$$\begin{aligned}
&\int_0^T \Big(\big(\partial_x L_\lambda(t,x_\lambda,u_\lambda),z\big) + \big\langle \partial_u L_\lambda(t,x_\lambda,u_\lambda),v\big\rangle \\
&\quad + \big(\partial\varphi_\lambda(x_\lambda),z\big) + \big\langle \Psi(u_\lambda-u^*)\|u_\lambda-u^*\|^{p-2},v\big\rangle\Big)\,\mathrm{d}t \\
&\quad + \Big(\big(\partial\ell_\lambda\big(x_\lambda(0),x_\lambda(T)\big),\big(z(0),z(T)\big)\big)\Big) \\
&\quad + \big(\Phi\big(x_\lambda(0)-x^*(0)\big),z(0)\big) = 0,
\end{aligned} \tag{4.47}$$

for all $v \in L^p(0,T;U)$, where $((\cdot,\cdot))$ is the duality between $E \times E$ and $E^* \times E^*$, while $z \in C([0,T];E)$ is a solution to

$$z' = A(t)z + (Bv)(t), \quad 0 \le t \le T,$$

that is,

$$z(t) = z(0) + \int_0^t U(t,s)(Bv)(s)\,\mathrm{d}s, \quad 0 \le t \le T.$$

Using once again (4.35), we obtain the estimate (without loss of generality, we may assume that $L \ge 0$)

$$\begin{aligned} &\lambda \big\| \partial L_\lambda(t,x,u) \big\|^2_{E^* \times U^*} \le \lambda^{-1}\big(|x|^2 + \|u - v_0\|^2\big) + 2L(t,0,v_0), \\ &\quad \text{for all } (x,u) \in E \times U, \end{aligned} \tag{4.48}$$

where $v_0 \in L^p(0,T;U)$ has been chosen as above.

This implies that $\partial L_\lambda(x_\lambda,u_\lambda) \in L^p(0,T;E^*) \times L^p(0,T;U^*)$, and this justifies (4.47).

Now, in (4.47) we substitute for $\partial_u L_\lambda(t,x_\lambda,u_\lambda)$, $\partial_x L(t,x_\lambda,u_\lambda)$ and $\partial\varphi_\lambda(x_\lambda)$ by their expressions (4.40), (4.41), and (4.42). By straightforward calculation, we find that

$$\begin{aligned} &\int_0^T \big\langle \partial_u L_\lambda(t,x_\lambda,u_\lambda) - B^* p_\lambda + \Psi(u_\lambda - u^*)\|u_\lambda - u^*\|^{p-2}, v\big\rangle\,\mathrm{d}t \\ &\quad + \big(y^1_\lambda - p_\lambda(0) + \Phi\big(x_\lambda(0) - x^*(0)\big), z(0)\big) = 0. \end{aligned}$$

Since $v \in L^p(0,T;U)$ and $z(0) \in E$ were arbitrary, we find (4.40) and (4.42), as claimed.

In particular, it follows, by Assumption (E), that there exists at least one feasible pair $(x_0,u_0) \in C([0,T];E) \times L^p(0,T;U)$. Then, by the inequality

$$F_\lambda(x_\lambda,u_\lambda) \le F_\lambda(x_0,u_0) \le C, \quad \lambda > 0, \tag{4.49}$$

we may infer that $\{u_\lambda\}$ is bounded in $L^p(0,T;U)$ and $\{x_\lambda(0)\}$ is bounded in E for $\lambda \to 0$. (Here, we have also used inequalities (4.36) and (4.37).) Thus, taking weakly convergent subsequences, we may assume, for $\lambda \to 0$, that

$$\begin{aligned} u_\lambda &\to u^1 \quad \text{weakly in } L^p(0,T;U), \\ x_\lambda(0) &\to x^0 \quad \text{weakly in } E. \end{aligned} \tag{4.50}$$

Keeping in mind (4.38), we see that

$$x_\lambda(t) \to x^1(t) = U(t,0)x^0 + \int_0^t U(t,s)\big(\big(Bu^1\big)(s) + f(s)\big)\,\mathrm{d}s \tag{4.51}$$

weakly in E for $t \in [0,T]$.

The well-known equality (see Theorem 2.58)

$$\varphi_\lambda(x_\lambda) = \frac{\lambda}{2}\left|\partial\varphi_\lambda(x_\lambda)\right|^2 + \varphi\left(J_\lambda^\varphi x_\lambda\right) \geq \frac{\lambda}{2}\left|\partial\varphi_\lambda(x_\lambda)\right|^2$$

implies that $\{\lambda|\partial\varphi_\lambda(x_\lambda)|^2\}$ is bounded in $L^1(0,T)$.

Since $\partial\varphi_\lambda(x_\lambda) = \lambda^{-1}\Phi(x_\lambda - J_\lambda^\varphi x_\lambda)$, this implies that $x_\lambda - J_\lambda^\varphi x_\lambda \to 0$ in $L^1(0,T;E)$. Thus, on some subsequence, $x_\lambda(t) - J_\lambda^L(t) \to 0$ a.e. $t \in]0,T[$. Since $J_\lambda^\varphi x_\lambda(t) \in K$, for every $t \in [0,T]$ we may infer by (4.51) that $x_1(t) \in K$, $\forall t \in [0,T]$. On the other hand, by (4.35), we see that $\{\lambda\|\partial L_\lambda(t,x_\lambda,u_\lambda)\|^2_{E^*\times U^*}\}$ is bounded in $L^1(0,T)$ and, therefore,

$$\lim_{\lambda\to 0}\left((x_\lambda,u_\lambda) - J_\lambda^L(t,x_\lambda,u_\lambda)\right) = 0 \quad \text{strongly in } L^2(0,T;E\times U)$$

and

$$\liminf_{\lambda\to 0}\int_0^T L_\lambda\left(t,x_\lambda(t),u_\lambda(t)\right)\mathrm{d}t \geq \liminf_{\lambda\to 0}\int_0^T L\left(t,J_\lambda^L\left(t,x_\lambda(t),u_\lambda(t)\right)\right)\mathrm{d}t.$$

On the other hand, it follows by (4.50) and (4.51) that

$$\lim_{\lambda\to 0} J_\lambda^L(t,x_\lambda,u_\lambda) = \left(u^1,x^1\right) \quad \text{weakly in } L^2(0,T;E\times U).$$

Since the convex function $(y,v) \to \int_0^T L(t,y,v)\,\mathrm{d}t$ is weakly lower-semicontinuous on $L^2(0,T;E\times U)$ (because it is convex and lower-semicontinuous), we have

$$\liminf_{\lambda\to 0}\int_0^T L_\lambda\left(t,x_\lambda(t),u_\lambda(t)\right)\mathrm{d}t \geq \int_0^T L\left(t,x^1(t),u^1(t)\right)\mathrm{d}t. \tag{4.52}$$

Similarly, from the equality

$$\ell_\lambda(x_1,x_2) = \frac{\lambda}{2}\left\|\partial\ell_\lambda(x_1,x_2)\right\|^2_{E^*\times E^*} + \ell\left(J_\lambda^\ell(x_1,x_2)\right)$$

we find by the same reasoning that $\{(x_\lambda(0),x_\lambda(T)) - J_\lambda^\ell(x_\lambda(0),x_\lambda(T))\} \to 0$ in $E\times E$ and, therefore,

$$\liminf_{\lambda\to 0}\ell_\lambda\left(x_\lambda(0),x_\lambda(T)\right) \geq \ell\left(x^1(0),x^1(T)\right). \tag{4.53}$$

By $J_\lambda^\ell(x_1,x_2)$ we have denoted, as usual, the solution (y_1,y_2) to the equation $(\Phi(y_1-x_1),\Phi(y_2-x_2)) + \lambda\partial\ell(y_1,y_2) \ni 0$ (see Sect. 2.2.3).

By (4.52) and (4.53), we have

$$\begin{aligned}\liminf_{\lambda\to 0}&\int_0^T L_\lambda(t,x_\lambda,u_\lambda)\,\mathrm{d}t + \ell_\lambda\left(x_\lambda(0),x_\lambda(T)\right)\\ &\geq \int_0^T L\left(t,x^1,u^1\right)\mathrm{d}t + \ell\left(x^1(0),x^2(T)\right).\end{aligned}$$

On the other hand, we have

$$F_\lambda(x_\lambda, u_\lambda) \le F_\lambda(x^*, u^*) \le \int_0^T L(t, x^*, u^*)\,\mathrm{d}t + \ell\big(x^*\big(0 < x^*(T)\big)\big)$$

because $L_\lambda \le L$ and $\ell_\lambda \le \ell$ for all $\lambda > 0$.

Since $F(x^*, u^*) \le F(x^1, u^1)$, we may infer that

$$\lim_{\lambda \to 0} \int_0^T \|u_\lambda - u^*\|^p\,\mathrm{d}t = 0.$$

Hence, $u^1 = u^*$, $x^1 = x^*$, and by (4.51), (4.44) follows, thereby completing the proof of Lemma 4.8. □

For the sake of simplicity, in the subsequent proof we take $f \equiv 0$, throughout.

Lemma 4.9 *There exists $C > 0$ independent of λ such that*

$$\big|p_\lambda(T)\big| \le C. \tag{4.54}$$

Proof We define on $E \times E$ the function

$$\begin{aligned}\varLambda(h_1, h_2) = \inf\big\{&G(x, u);\ (x, u) \in \mathscr{H},\ x(0) = h_1,\\ &x(T) = h_2,\ x(t) \in K \text{ for } t \in [0, T]\big\},\end{aligned}$$

where

$$G(x, u) = \int_0^T \big(L(t, x, u) + p^{-1}\|u - u^*\|^p\big)\,\mathrm{d}t + \frac{1}{2}\big|x(0) - x^*(0)\big|^2.$$

We have already seen that the function $(x, u) \to \int_0^T L(t, x, u)\,\mathrm{d}t$ is convex and lower-semicontinuous on $L^1(0, T; E) \times L^p(0, T; U)$. Since G is also coercive on $\mathscr{H}$, we may infer that, for every choice of h_1, h_2, the infimum defining $\varLambda(h_1, h_2)$ is attained. This fact implies by a standard argument involving the convexity and the weak lower-semicontinuity of the convex integrand L that $\varLambda$ is convex and lower-semicontinuous on $E \times E$. Furthermore, the effective domain $D(\varLambda)$ of $\varLambda$ is the very set K_L defined in Sect. 4.1.1. To prove estimate (4.54), we use Assumption (E). First, let us suppose that condition (4.9) is satisfied. Then there exists $y \in C([0, T]; E)$ such that $\ell(y(0), y(T)) < +\infty$ and $y(T) \in \operatorname{int} D(\varLambda(y(0), \cdot))$. This implies that the function $h \to \varLambda(y(0), h)$ is locally bounded at $h = y(T)$ and, therefore, there exist some positive constants ρ and C such that

$$\varLambda\big(y(0), y(T) + \rho h\big) \le C \quad \text{for all } h \in E,\ |h| = 1. \tag{4.55}$$

Now, let $(z, v) \in \mathscr{H}$ be such that $z \in \mathscr{K}$, $z(0) = y(0)$ and $z(T) = y(T) + \rho h$, where h is fixed and $|h| = 1$. Again, using (4.38)–(4.41), we find, after some calculations,

$$\big(p_\lambda(T), x_\lambda(T) - y(T) - \rho h\big) - \big(p_\lambda(0), x_\lambda(0) - y(0)\big) \ge G_\lambda(x_\lambda, u_\lambda) - G_\lambda(z, v),$$

where

$$G_\lambda(x,u)=\int_0^T\big(L_\lambda(t,x,u)+\varphi_\lambda(x)+p^{-1}\|u-u^*\|^p\big)\,dt+\frac{1}{2}\big|x(0)-x^*(0)\big|^2.$$

Inasmuch as $(y(0), y(T)+\rho h)\in K_L$, we may choose the pair (z,v) in such a way that $G(z,v)=\Lambda(y(0),y(T)+\rho h)$. Since $G_\lambda(z,v)\le G(z,v)$, by (4.55) we may infer that

$$G_\lambda(z,v)\le C$$

and, therefore,

$$\big(p_\lambda(T),x_\lambda(T)-y(T)-\rho h\big)-\big(p_\lambda(0),x_\lambda(0)-y(0)\big)\ge C, \tag{4.56}$$

for all $h\in E$, $|h|=1$. (We denote by C several positive constants independent of λ.) To obtain the latter, we have also used the fact, already noticed in the proof of Lemma 4.8, that $F_\lambda(x_\lambda,u_\lambda)$ and, consequently, $G_\lambda(x_\lambda,u_\lambda)$ are bounded from below with respect to λ.

Since, by (4.42) and the definition of $\partial\ell_\lambda$,

$$\begin{aligned}
&\big(p_\lambda(0),x_\lambda(0)-y(0)\big)-\big(p_\lambda(T),x_\lambda(T)-y(T)\big)\\
&\quad\ge \ell_\lambda\big(x_\lambda(0),x_\lambda(T)\big)-\ell_\lambda\big(y(0),y(T)\big)\\
&\qquad+\frac{1}{2}\big(\big|x_\lambda(0)-x^*(0)\big|^2-\big|y(0)-x^*(0)\big|^2\big),
\end{aligned}$$

while, by (4.37), $\ell_\lambda(x_\lambda(0),x_\lambda(T))$ is bounded from below and $\ell_\lambda(y(0),y(T))\le \ell(y(0),y(T))<+\infty$, we see by (4.56) that $p_\lambda(T)$ is bounded in E^*.

Now, assume that condition (4.10) is satisfied. In other words, there is $(y,v)\in K$, $y\in\mathscr{K}$, such that $(y(0),y(T))\in K_L\cap \mathrm{Dom}(\ell)$ and $y(T)\in \mathrm{int}\{h\in E;\ (y(0),h)\in \mathrm{Dom}(\ell)\}$. Hence, there exist some positive constants ρ and C such that (see Proposition 2.16)

$$\ell\big(y(0),y(T)+\rho h\big)\le C\quad \text{for all } h\in E,\ |h|=1.$$

Next, by (4.42) we have

$$\begin{aligned}
&\big(p_\lambda(0),x_\lambda(0)-y(0)\big)-\big(p_\lambda(T),x_\lambda(T)-y(T)-\rho h\big)\\
&\quad\ge \ell_\lambda\big(x_\lambda(0),x_\lambda(T)\big)-\ell_\lambda\big(y(0),y(T)+\rho h\big)\\
&\qquad+\frac{1}{2}\big(\big|x_\lambda(0)-x^*(0)\big|^2-\big|y(0)-x^*(0)\big|^2\big).
\end{aligned}$$

Now, using once again (4.38)–(4.41), we find that

$$\big(p_\lambda(T),x_\lambda(T)-y(T)\big)-\big(p_\lambda(0),x_\lambda(0)-y(0)\big)\ge G_\lambda(x_\lambda,u_\lambda)-G_\lambda(y,v)\ge C$$

for all $\lambda>0$.

Hence,

$$\rho\big(p_\lambda(T), h\big) \le C \quad \text{for all } \lambda > 0,\ |h| = 1,$$

wherein C is independent of λ and h. We have, therefore, proved the boundedness of $\{|p_\lambda(T)|\}$ in both situations, thereby completing the proof of Lemma 4.9. □

Now, we continue the proof of Theorem 4.5 with further a priori estimates on p_λ. By Assumption (D), there exists at least one pair $(x^0, u^0) \in K$ such that $x^0(t) \in \operatorname{int} K$ for $t \in [0,T]$, $L(t, x^0, u^0) \in L^1(0,T)$ and $\ell(x^0(0), x^0(T)) < +\infty$. Since x^0 is continuous on $[0,T]$, there exists $\rho > 0$ such that

$$x^0(t) + \rho h \in K \quad \text{for all } h,\ |h| = 1;\ t \in [0,T].$$

By the definition of $\partial\varphi_\lambda$, we have

$$\big(\partial\varphi_\lambda(x_\lambda), x_\lambda - x^0 - \rho h\big) \ge \varphi_\lambda(x_\lambda) - \varphi_\lambda\big(x^0 + \rho h\big) = \varphi_\lambda(x_\lambda),$$

whereupon we get

$$\rho \int_0^T \big|\partial\varphi_\lambda(x_\lambda)\big|\,\mathrm{d}t \le \int_0^T \big(\partial\varphi_\lambda(x_\lambda), x_\lambda - x^0\big)\,\mathrm{d}t. \tag{4.57}$$

On the other hand, once again using (4.38)–(4.41), one obtains

$$\begin{aligned}
&\int_0^T \big(\partial\varphi_\lambda(x_\lambda), x_\lambda - x^0\big)\,\mathrm{d}t \\
&\quad \le \int_0^T \Big(L_\lambda\big(t, x^0, u^0\big) - L_\lambda(t, x_\lambda, u_\lambda) + p^{-1}\big(\big\|u^0 - u^*\big\|^p - \|u_\lambda - u^*\|^p\big)\Big)\,\mathrm{d}t \\
&\qquad + \big(p_\lambda(T), x_\lambda(T) - x^0(T)\big) - \big(p_\lambda(0), x_\lambda(0) - x^0(0)\big) \le C,
\end{aligned}$$

because $L_\lambda(t, x^0, u^0) \le L(t, x^0, u^0)$, $\{\int_0^T L_\lambda(t, x_\lambda, u_\lambda)\,\mathrm{d}t\}$ is bounded from below (by (4.36)) and

$$\begin{aligned}
&\big(p_\lambda(T), x_\lambda(T) - x^0(T)\big) - \big(p_\lambda(0), x_\lambda(0) - x^0(0)\big) \\
&\quad \ge \big(\Phi\big(x^*(0) - x_\lambda(0)\big), x_\lambda(0) - x^0(0)\big) \\
&\qquad + \ell_\lambda\big(x^0(0), x^0(T)\big) - \ell_\lambda\big(x_\lambda(0), x_\lambda(T)\big) \le C.
\end{aligned}$$

Hence,

$$\int_0^T \big|\partial\varphi_\lambda(x_\lambda)\big|\,\mathrm{d}t \le C. \tag{4.58}$$

Now, according to Assumption (C), part (iii), there exist functions $\alpha, \beta \in L^p(0,T)$ and v_h measurable from $[0,T]$ to U such that $\|v_h(t)\| \le \beta(t)$ a.e. $t \in]0,T[$ and $L_\lambda(t, x^*(t) + \rho h, v_h(t)) \le L(t, x^*(t) + \rho h, v_h(t)) \le \alpha(t)$ a.e. $t \in]0,T[$, $|h| = 1$, $\lambda > 0$. Here, we have used the fact that the function $x^* : [0,T] \to E$ is continuous and so, its graph is compact in E. Next, by (4.40), (4.41) and the definition of ∂L_λ, we have

$$\begin{aligned}&\big(q_\lambda(t), x_\lambda(t) - x^*(t) - \rho h\big)\\&\quad + \big\langle (B^* p_\lambda)(t) + \Psi\big(u^*(t) - u_\lambda(t)\big)\big\|u^*(t) - u_\lambda(t)\big\|^{p-2}, u_\lambda(t) - v_h(t)\big\rangle\\&\geq L_\lambda\big(t, x_\lambda(t), u_\lambda(t)\big) - L_\lambda\big(t, x^*(t) + \rho h, v_h(t)\big).\end{aligned}$$

Along with (4.36), the latter yields

$$\begin{aligned}\rho\big|q_\lambda(t)\big| &\leq \big(q_\lambda(t), x_\lambda(t) - x^*(t)\big) + \alpha(t)\\&\quad + \big\|(B^* p_\lambda)(t)\big\|\,\big\|u_\lambda(t) - v_h(t)\big\| + \big\|u^*(t) - u_\lambda(t)\big\|^{p-1}\big\|u_\lambda(t) - v_h(t)\big\|,\end{aligned}$$

because the duality mapping Ψ is demicontinuous on U and $u_\lambda \to u^*$ in $L^p(0, T; U)$.

Recalling that $x_\lambda(t) \to x^*(t)$ uniformly on $[0, T]$ and $u_\lambda \to u^*$ in $L^p(0, T; U)$, for λ sufficiently small we have

$$\big|q_\lambda(t)\big| \leq C\big(\big\|(B^* p_\lambda)(t)\big\| + \big\|u^*(t) - u_\lambda(t)\big\|^{p-1}\big)\big(\big\|u_\lambda(t)\big\| + \beta(t)\big) + \alpha_1(t), \tag{4.59}$$

where $\alpha_1 \in L^1(0, T)$. Then, by (4.44), (4.45), and Lemma 4.8, we have

$$\big|p_\lambda(t)\big| \leq C\left(1 + \int_t^T \big\|(B^* p_\lambda)(s)\big\|\big(\beta(s) + \big\|u_\lambda(s)\big\|\big)\,\mathrm{d}s\right).$$

Finally, by Hölder's inequality

$$\big|p_\lambda(t)\big|^{p'} \leq C\left(1 + \int_t^T \big\|(B^* p_\lambda)(s)\big\|^{p'}\,\mathrm{d}s\right), \quad 0 \leq t \leq T, \tag{4.60}$$

where $\frac{1}{p} + \frac{1}{p'} = 1$. Since B is "causal", the adjoint $B^* : L^{p'}(0, T; E^*) \to L^{p'}(0, T; U^*)$ is "anticausal", that is, $B^*_{\chi(t,T)} = \chi_{[t,T]} B^*_{\chi[t,T]}$ for all $t \in [0, T]$. (Here, $\chi_{[t,T]}$ is the characteristic function of the interval $[t, T]$.) Then, arguing as in the proof of Proposition 1.8, we obtain the inequality

$$\int_t^T \big\|(B^* y)(s)\big\|^{p'}\,\mathrm{d}s \leq C\int_t^T \big|y(s)\big|^{p'}\,\mathrm{d}s, \quad y \in L^{p'}(0, T; E^*),\ 0 \leq t \leq T.$$

The latter, compared with inequality (4.60), implies via Gronwall's Lemma

$$\big|p_\lambda(t)\big| \leq C, \quad \forall t \in [0, T], \tag{4.61}$$

where C is independent of λ. Next, since $\{B^* p_\lambda\}$ is bounded in $L^{p'}(0, T; U^*)$ and $u_\lambda \to u^*$ in $L^p(0, T; U)$ it follows by (4.59) that $\{q_\lambda\}$ is bounded in $L^1(0, T; E^*)$ and the integrals $\{\int_\Omega |q_\lambda(s)|\,\mathrm{d}s; \Omega \subset [0, T]\}$ are uniformly absolutely continuous, that is, for every $\varepsilon > 0$ there exists $\delta(\varepsilon) > 0$ such that $\int_\Omega |q_\lambda(s)|\,\mathrm{d}s \le \varepsilon$ whenever the Lebesgue measure of Ω is $\le \delta(\varepsilon)$. Then, by the Dunford–Pettis criterion in a Banach space (Theorem 1.121), we may conclude that $\{q_\lambda\}$ is weakly compact in $L^1(0,T;E^*)$. Thus, there exists $q \in L^1(0,T;E^*)$ such that on a subsequence $\lambda \to 0$

$$q_\lambda \to q \quad \text{weakly in } L^1(0,T;E^*). \tag{4.62}$$

Similarly, by (4.61), we have

$$p_\lambda \to p^* \quad \text{weakly in } L^1(0,T;E^*) \tag{4.63}$$

and, therefore,

$$B^* p_\lambda \to B^* p^* \quad \text{weakly in } L^p(0,T;U^*).$$

By (4.43), (4.44), and the definition of ∂L_λ, we have

$$\begin{aligned}
&\int_0^T \big(L_\lambda(t,x_\lambda,u_\lambda) - L_\lambda(t,y,v)\big)\,\mathrm{d}t \\
&\quad \le \int_0^T \big(\big\langle B^* p_\lambda + \Psi(u^* - u_\lambda)\|u^* - u_\lambda\|^{p-2}, u_\lambda - v\big\rangle + (q_\lambda, x_\lambda - y)\big)\,\mathrm{d}t,
\end{aligned}$$

for all $(y,v) \in L^\infty(0,T;E^*) \times L^p(0,T;U)$. Remembering that $L_\lambda(t) \le L(t)$ and (4.52)

$$\liminf_{\lambda\to 0} \int_0^T L_\lambda(t,x_\lambda,u_\lambda)\,\mathrm{d}t \le \int_0^T L(t,x^*,u^*)\,\mathrm{d}t,$$

we obtain by (4.43), (4.44), and (4.62)

$$\int_0^T \big(L(t,x^*,u^*) - L(t,y,v)\big)\,\mathrm{d}t \le \int_0^T \big(\langle B^* p^*, u^* - v\rangle + (q, x^* - y)\big)\,\mathrm{d}t,$$

because Ψ is demicontinuous on U and $u_\lambda \to u^*$ strongly in $L^p(0,T;U)$.

Let Ω be any measurable subset of $[0,T]$ and let $y(t)$, $v(t)$ be defined by

$$y(t) = \begin{cases} \tilde{y}, & \text{on } \Omega, \\ x^*(t), & \text{on } [0,T]\setminus\Omega, \end{cases}$$

$$v(t) = \begin{cases} \tilde{v}, & \text{on } \Omega, \\ u^*(t), & \text{on } [0,T]\setminus\Omega, \end{cases}$$

wherein $(\tilde{y},\tilde{v}) \in E \times U$. We have

$$\begin{aligned}
&\int_\Omega \big(L\big(t,x^*(t),u^*(t)\big) - L(t,\tilde{y},\tilde{v})\big)\,\mathrm{d}t \\
&\quad \le \int_0^T \big(\big\langle B^* p^*, u^*(t) - \tilde{v}\big\rangle + \big(q(t), x^*(t) - \tilde{y}\big)\big)\,\mathrm{d}t
\end{aligned}$$

and, since Ω is arbitrary, it follows that

$$\big(q(t), x^*(t) - \tilde{y}\big) + \big\langle (B^* p^*)(t), u^*(t) - \tilde{v}\big\rangle \geq L\big(t, x^*(t), u^*(t)\big) - L(t, \tilde{y}, \tilde{v})$$
$$\text{a.e. } t \in \,]0, T[.$$

Hence,

$$\big(q(t), (B^* p^*)(t)\big) \in \partial L\big(t, x^*(t), u^*(t)\big) \quad \text{a.e. } t \in \,]0, T[. \tag{4.64}$$

Next, by Lemma 4.9 and (4.39), we may infer that on a subsequence again denoted by λ, we have

$$p_\lambda(0) \to p_0^* \quad \text{weakly in } E^*,$$
$$p_\lambda(T) \to p_T^* \quad \text{weakly in } E^*.$$

Since Φ is demicontinuous on E, it follows by (4.42) that

$$\big(p_0^*, x^*(0) - x_1\big) - \big(p_T^*, x^*(T) - x_2\big) \geq \liminf_{\lambda \to 0} \ell_\lambda\big(x_\lambda(0), x_\lambda(T)\big) - \ell(x_1, x_2),$$
$$\forall (x_1, x_2) \in E \times E.$$

Since, as noticed earlier,

$$\liminf_{\lambda \to 0} \ell_\lambda\big(x_\lambda(0), x_\lambda(T)\big) \geq \ell\big(x^*(0), x^*(T)\big),$$

the latter implies

$$(p_0^*, -p_T^*) \in \partial \ell\big(x^*(0), x^*(T)\big). \tag{4.65}$$

We set

$$w_\lambda(t) = \int_0^t \partial\varphi_\lambda\big(x_\lambda(s)\big)\,\mathrm{d}s, \quad 0 \leq t \leq T.$$

By estimate (4.58), we see that Theorem 1.126 is applicable. Thus, there exist a function $w \in BV([0, T]; E^*)$ and a sequence convergent to zero, again denoted by $\{\lambda\}$ such that $w_\lambda(t) \to w(t)$ weakly in E^* for every $t \in [0, T]$ and

$$\lim_{\lambda \to 0} \int_t^T \big(\partial\varphi_\lambda\big(x_\lambda(s)\big), y(s)\big)\,\mathrm{d}s = \int_t^T (\mathrm{d}w, y)$$
$$\text{for all } y \in C\big([t, T]; E\big),\ \forall t \in [0, T]. \tag{4.66}$$

Hence, for all $t \in [0, T]$, we have

$$\int_t^T U^*(s, t)\partial\varphi_\lambda\big(x_\lambda(s)\big)\,\mathrm{d}s \to \int_t^T U^*(s, t)\,\mathrm{d}w \quad \text{weakly in } E^*$$

and, letting λ tend to zero in (4.39), we see by (4.47), (4.62), and (4.63) that p^* satisfies the equation

$$p^*(t) = U^*(T,t)p_T^* - \int_t^T U^*(s,t)q(s)\,\mathrm{d}s - \int_t^T U^*(s,t)\,\mathrm{d}w(s), \quad 0 \le t \le T,$$

and $p^*(0) = p^*$.

Along with (4.64) and (4.65), the latter shows that the functions p^*, w satisfy together with x^* and u^*, (4.17), (4.18), (4.20), and (4.21). As regards (4.19), it follows by (4.66) and the obvious inequality

$$\int_t^T \big(\partial\varphi_\lambda\big(x_\lambda(s)\big), x_\lambda(s) - y(s)\big)\,\mathrm{d}s \ge 0 \quad \text{for all } y \in \mathscr{K}.$$

Assume now that B is given by (4.3), where $B : [0,T] \to L(U,E)$ is strongly measurable and $\|B(t)\|_{L(U,E)} \le \eta(t)$ a.e. $t \in]0,T[$, where $\eta \in L^\infty(0,T)$. Then, $\|B^*(t)\|_{L(E^*,U^*)} \le \eta(t)$ and by (4.61) we see that $\|(B^* p_\lambda)(t)\| \le C$ a.e. $t \in]0,T[$. Since $u_\lambda \to u^*$ in $L^p(0,T;U)$ and $q_\lambda \to q$ weakly in $L^1(0,T;E^*)$, we may conclude by (4.59) that $q \in L^p(0,T;E^*)$. This concludes the proof.

4.1.5 Proof of Theorem 4.6

If w satisfies (4.25) and (4.26), then clearly (4.19) holds.

Assume now that $w \in BV([0,T];E^*)$ satisfies (4.19). To prove the theorem, it suffices to show that w and $\mathrm{d}w_s$ satisfy (4.25) and (4.26), respectively.

Let t_0 be arbitrary but fixed in $]0,T[$. For $y \in K$ and $\varepsilon > 0$, define the function y_ε

$$y_\varepsilon(t) = \begin{cases} x^*(t), & \text{for } |t-t_0| \ge \varepsilon, \\ (1-\varepsilon^{-1}(t_0-t))y + \varepsilon^{-1}(t_0-t)x^*(t_0-\varepsilon), & \text{for } t \in [t_0-\varepsilon, t_0], \\ (1-\varepsilon^{-1}(t-t_0))y + \varepsilon^{-1}(t-t_0)x^*(t_0+\varepsilon), & \text{for } t \in [t_0, t_0+\varepsilon]. \end{cases}$$

Obviously, y_ε is continuous from $[0,T]$ to E and $y_\varepsilon(t) \in K$ for all $t \in [0,T]$. By (4.19), we have

$$\int_0^T \big(\dot{w}(t), x^*(t) - y_\varepsilon(t)\big)\,\mathrm{d}t + \int_0^T (\mathrm{d}w_s, x^* - y_s) \ge 0. \tag{4.67}$$

We set $\rho_\varepsilon(t) = \varepsilon^{-1}(x^*(t) - y_\varepsilon(t))$. If t_0 happens to be a Lebesgue point for the function $\dot{w}$, then, by an elementary calculation involving the definition of y_ε, we get

$$\lim_{\varepsilon\to 0}\int_0^T \big(\dot{w}(t), \rho_\varepsilon(t)\big)\,\mathrm{d}t = \big(\dot{w}(t_0), x^*(t_0) - y\big). \tag{4.68}$$

Inasmuch as $x^* - y_\varepsilon = 0$, outside $[t_0 - \varepsilon, t_0 + \varepsilon]$, we have

$$\int_0^T (\mathrm{d}w_s, x^* - y_\varepsilon) = \int_{t_0-\varepsilon}^{t_0+\varepsilon} (\mathrm{d}w_s, x^* - y_\varepsilon).$$

On the other hand, for each $\eta > 0$, there exist $\{x^*_{i\eta}\}_{i=1}^N \subset E$ and $\alpha_{i\eta} \in C([0, T])$ such that

$$\left| x^*(t) - \sum_{i=1}^N x^*_{i\eta}\alpha_{i\eta}(t) \right| \le \eta \quad \text{for } t \in [0, T].$$

We set

$$z_\eta(t) = x^*(t) - \sum_{i=1}^N x^*_{i\eta}\alpha_{i\eta}(t).$$

We have

$$\left| \int_{t_0-\varepsilon}^{t_0+\varepsilon} (\mathrm{d}w_s, z_\eta) \right| \le \big(V_s(t_0+\varepsilon) - V_s(t_0-\varepsilon)\big) \sup\big\{ |z_\eta(t)|;\ |t - t_0| \le \varepsilon \big\},$$

where $V_s(t)$ is the variation of w_s on the interval $[0, t]$. Since V_s is a.e. differentiable on $]0, T[$, we may assume that

$$\limsup_{\varepsilon\to 0} \varepsilon^{-1} \int_{t_0-\varepsilon}^{t_0+\varepsilon} (\mathrm{d}w_s, z_\eta) \le C\eta, \tag{4.69}$$

where C is independent of η.

Now, we have

$$\begin{aligned}
\left| \int_{t_0-\varepsilon}^{t_0+\varepsilon} \left(\mathrm{d}w_s, \sum_{i=1}^N x^*_{i\eta}\alpha_{i\eta} \right) \right| &\le \sum_{i=1}^N \left| \int_{t_0-\varepsilon}^{t_0+\varepsilon} \alpha_{i\eta}(t)\, \mathrm{d}(w_s, x^*_{i\eta}) \right| \\
&\le \sum_{i=1}^N \big(V_{i\eta}(t_0+\varepsilon) - V_{i\eta}(t_0-\varepsilon)\big)\gamma_{i\eta},
\end{aligned}$$

where $V_{i\eta}(t)$ is the variation of $(w_s, x^*_{i\eta})$ on $[0, t]$ and $\gamma_{i\eta} = \sup|\alpha_{i\eta}(t)|$. Since the weak derivative $\dot{w}_s$ of w_s is zero a.e. on $]0, T[$, we may infer that

$$\frac{\mathrm{d}}{\mathrm{d}t} V_{i\eta}(t) = 0 \quad \text{a.e. } t \in]0, T[$$

and, therefore, we may assume that

$$\lim_{\varepsilon\to 0} \varepsilon^{-1} \int_{t_0-\varepsilon}^{t_0+\varepsilon} \left(\mathrm{d}w_s, \sum_{i=1}^N x^*_{i\eta}\alpha_{i\eta} \right) = 0 \quad \text{for all } \eta > 0. \tag{4.70}$$

Similarly, we see that

$$\lim_{\varepsilon\to 0} \varepsilon^{-1} \int_{t_0-\varepsilon}^{t_0+\varepsilon} \big(\mathrm{d}w_s, y_\varepsilon(s)\big) = 0 \quad \text{a.e. } t_0 \in \,]0, T[,$$

which along with (4.69) and (4.70) yields

$$\lim_{\varepsilon\to 0} \int_0^T (\mathrm{d}w_s, \rho_\varepsilon) = 0,$$

whereupon, by (4.67) and (4.68),

$$\big(\dot{w}(t_0), x^*(t_0) - y\big) \geq 0 \quad \text{a.e. } t_0 \in \,]0, T[.$$

Since y is arbitrary in K, this implies (4.25).

To conclude the proof, it remains to be shown that $\mathrm{d}w_s \in \mathscr{N}(x^*, \mathscr{K})$, that is,

$$\int_0^T (\mathrm{d}w_s, x^* - y) \geq 0 \quad \text{for all } y \in \mathscr{K}. \tag{4.71}$$

Let Ω be the support of the singular measure $\mathrm{d}w_s$. Then, for any $\varepsilon > 0$, there exists an open subset Ω_0 of $]0, T[$ such that $\Omega \subset \Omega_0$ and the Lebesgue measure of Ω_0 is $\leq \varepsilon$. Let $\rho \in C_0^\infty(R)$ be such that $0 \leq \rho \leq 1$, $\rho = 1$ on Ω and $\rho = 0$ on $]0, T[\setminus \Omega_0$.

We set $y^\varepsilon = \rho y + (1-\rho)x^*$, where $y \in \mathscr{K}$ is arbitrary. By (4.19), we have

$$\int_0^T \big(w_s, x^* - y^\varepsilon\big) + \int_0^T \big(\dot{w}, x^* - y^\varepsilon\big)\, \mathrm{d}t \geq 0.$$

Since $x^* - y^\varepsilon = 0$ on $[0, T] \setminus \Omega_0$, we find that

$$\left| \int_0^T \big(\dot{w}, x^* - y^\varepsilon\big)\, \mathrm{d}t \right| \leq C \int_{\Omega_0} |\dot{w}|\, \mathrm{d}t \leq \delta(\varepsilon),$$

where $\lim_{\varepsilon\to 0} \delta(\varepsilon) = 0$. On the other hand, since $\mathrm{d}w_s = 0$ on $[0, T] \setminus \Omega$ and $\rho = 1$ on Ω, we see that

$$\int_0^T (\mathrm{d}w_s, x^* - y_\varepsilon) = \int_0^T (\mathrm{d}w_s, x^* - y),$$

whereupon it follows that

$$\int_0^T (\mathrm{d}w_s, x^* - y) \geq -\delta(\varepsilon).$$

Since ε is arbitrary, we obtain (4.71), as claimed.

Remark 4.10 A little calculation involving the definition of the Stieltjes–Riemann integral reveals that

$$\lim_{\varepsilon\to 0}\int_0^T (\mathrm{d}w, x^* - y_\varepsilon) = \lim_{\varepsilon\to 0}\int_{t_0-\varepsilon}^{t_0+\varepsilon} (\mathrm{d}w, x^* - y_\varepsilon)$$
$$= \big(w(t_0+0) - w(t_0-0), x^*(t_0) - y\big).$$

Hence, for all $y \in K$, we have

$$\big(w(t_0+0) - w(t_0-0), x^*(t_0) - y\big) \geq 0, \quad t_0 \in [0,T],$$

and, therefore, $w(t_0+0) - w(t_0-0) \in N(x^*(t), K)$ for every $t_0 \in [0,T]$. Inasmuch as, by (4.18),

$$p^*(t_0+0) - p^*(t_0-0) = w(t_0+0) - w(t_0-0),$$

we may conclude that

$$p^*(t_0+0) - p^*(t_0-0) \in N\big(x^*(t_0), K\big) \quad \text{for all } t_0 \in [0,T].$$

As noticed earlier, this amounts to saying that the set of all the points t_0, where the dual extremal arc is discontinuous, is contained in the set of t values for which $x^*(t)$ lies on the boundary of K (we recall that $N(x,K) = \{0\}$ for $x \in \operatorname{int} K$).

4.1.6 Further Remarks on Optimality Theorems

Analyzing the proofs of Theorems 4.5 and 4.6, it is apparent that the assumptions imposed on L, ℓ and K are, at least in certain cases, indeed excessive. For the sake of simplicity, we discuss here the simple case when U and E are Hilbert spaces and $p = 2$.

In passing, we observe that, if $A(t)$ is continuous on E or, more generally, if, for every $p_0 \in E$ and $g \in L^2(0,T;E)$, the forward Cauchy problem

$$p'(t) + A^*(t)p(t) = g(t), \quad 0 \leq t \leq T,$$
$$p(0) = p_0,$$

is well posed, then in Hypothesis (E) we may reduce conditions (4.9) and (4.10) to

$$x_0 \in \operatorname{int}\big\{h \in E;\ (h, x_T) \in K_L\big\}$$

and

$$x_0 \in \operatorname{int}\big\{h \in E;\ (h, x_T) \in \operatorname{Dom}(\ell)\big\}.$$

As mentioned earlier, condition (4.9) in Hypothesis (E) can be regarded as a complete controllability assumption for (4.1), which is, in many important infinite-dimensional examples, a very stringent requirement. However, it turns out that this condition can be weakened by replacing the strong interior by the interior relative to a certain linear closed manifold in E. We illustrate this for a control problem with fixed end points. In other words, the function ℓ is defined by

$$\ell(h_1, h_2) = \begin{cases} 0, & \text{if } h_1 = x_0 \text{ and } h_2 = x_T, \\ +\infty, & \text{otherwise,} \end{cases}$$

where x_0 and x_T are fixed in E. We consider only the particular case $K = E$, $f = 0$ and $x_0 = 0$, the general case being obtained by appropriately translating the special case we consider.

If $u \in L^2(0, T; U)$ is given, then we denote by $x(\cdot, u)$ the response function to (4.1) with control u and initial condition $x(0, u) = 0$. Then, the attainable set at the time T is defined by

$$E_T = \{x(T, u);\ u \in L^2(0, T; U)\}.$$

Let $\overline{E}_T$ be the closure of E_T in E. Clearly, $\overline{E}_T$ is a closed linear manifold in E. Finally, let K_T be defined by

$$K_T = \{h \in E;\ (0, h) \in K_L\}.$$

Obviously, K_T is a subset of $\overline{E}_T$ and, in terms of the above notation, condition (4.9) may be expressed as $x_T \in \operatorname{int} K_T$. In general, the interior of $\overline{E}_T$ in E is empty. If the interior of $\overline{E}_T$ is not empty, then we have $\overline{E}_T = E$, because $\overline{E}_T$ is a linear manifold. However, it turns out that, in the special case we are considering, Theorem 4.5 remains valid if this condition is replaced by the following weaker one:

(E$'$) $x_T \in \operatorname{ri} K_T$,

where ri denotes the interior relative to the manifold $\overline{E}_T$.

Here is the argument. Let (x, u) be an optimal pair of the given problem. Clearly, the control optimal problem:

$$\text{Minimize } \int_0^T \left(L\big(t, y(t), v(t)\big) + \frac{1}{2} \left\| v(t) - u(t) \right\|^2 \right) \mathrm{d}t + \ell^\lambda\big(y(0), y(T)\big),$$
$$\text{over all } (y, v) \in \mathscr{H}$$

has a unique solution $(x_\lambda, u_\lambda) \in \mathscr{H}$. Here, $\ell^\lambda : E \times E \to\,]-\infty, +\infty]$ is defined by

$$\ell^\lambda(x_1, x_2) = \begin{cases} \frac{1}{2\lambda} |x_2 - x_T|^2, & \text{if } x_1 = 0, \\ +\infty, & \text{otherwise.} \end{cases}$$

It should be observed that Hypotheses (A)–(E) are trivially satisfied. Thus, by Theorem 4.5, the boundary-value problem

$$x_\lambda' - A(t)x_\lambda = B(t)u_\lambda, \quad t \in [0, T],$$

$$\left\{p_\lambda' + A^*(t)p_\lambda, B^*(t)p_\lambda + u^* - u_\lambda\right\} \in \partial L(t, x_\lambda, u_\lambda) \quad \text{a.e. on }]0,T[,$$
$$x_\lambda(0) = 0, \quad \lambda p_\lambda(T) + x_\lambda(T) = x_T$$

has at least one solution $(x_\lambda, u_\lambda, p_\lambda)$.

The proof continues by the same argument as that used in the proof of Theorem 4.5, except for Lemma 4.9, where Hypothesis (E) was necessary. However, this lemma can be proved under Hypothesis (E′). In fact, let $\Phi_0 : \overline{E}_T \to]-\infty, +\infty]$ be defined by

$$\Phi_0(h) = \inf\left\{G(y, v);\ (y, v) \in \mathscr{H};\ y(0) = 0,\ y(T) = h\right\},$$

where G is defined as in the proof of Lemma 4.9. Obviously, Φ_0 is convex and lower-semicontinuous on $\overline{E}_T$. Moreover, $\mathrm{Dom}(\Phi_0) = K_T$ and Hypothesis (E′) implies that Φ_0 is locally bounded at $h = x_T$. Since $p_\lambda(T) \in E_T$, reasoning as in the general case, one concludes that $\{|p_\lambda(T)|\}$ is bounded.

While various other extensions of Theorem 4.5 could be pursued, we concentrate on the *control of periodic systems*, which are treated in more detail in Sect. 4.5. If the function $\ell : E \times E \to \overline{\mathbb{R}}^*$ is given by

$$\ell(x_1, x_2) = \begin{cases} 0, & \text{if } x_1 = x_2, \\ +\infty, & \text{if } x_1 \neq x_2, \end{cases}$$

then Problem (P) leads to a problem with periodic conditions.

($\mathrm{P_T}$) *Minimize* $\int_0^T L(t, x(t), u(t))\,\mathrm{d}t$ *on the set of all* $(x, u) \in C([0,T]; E) \times L^2(0,T;U)$ *subject to* (4.1) *with periodic conditions* $x(0) = x(T)$.

For the sake of simplicity, we consider here the special case where $K = E$, B is given by formula (4.3) and $A(t)$, $B(t)$ are time-independent and

$$L(t, x, y) = \varphi^0(x) + \psi(u) \quad \text{for all } x \in E,\ u \in U,$$

where $\varphi^0 : E \to \mathbb{R}$ is a continuous, convex function and $\psi : U \to \overline{\mathbb{R}}^*$ is lower-semicontinuous and convex. Further, we assume that *the operator* $(I - e^{AT})^{-1}$ *is well defined and continuous on all of* E. The latter condition implies, by a standard existence result, that (4.1) has a unique periodic solution with period T.

The optimality equations (4.17)–(4.21) become

$$x^{*\prime} = Ax^* + Bu^*, \quad t \in [0,T], \tag{4.17′}$$

$$p^{*\prime} = -A^*p^* + q, \quad t \in [0,T], \tag{4.18′}$$

$$q(t) \in \partial\varphi^0\big(x^*(t)\big), \qquad B^*p^*(t) \in \partial\psi\big(u^*(t)\big), \quad t \in [0,T], \tag{4.19′}$$

$$x^*(0) = x^*(T), \qquad p^*(0) = p^*(T). \tag{4.20′}$$

It must be noticed that Theorem 4.5 is not applicable since Hypothesis (E) is not satisfied in this case. However, by a slight modification of the proof, we see that (4.17′)–(4.20′) *are necessary and sufficient for optimality in Problem* ($\mathrm{P_T}$). Indeed, by (4.41), it follows that, for $\rho > 0$,

$$\big(q_\lambda(t), x_\lambda(t) - x^*(t) - \rho w\big) \ge \varphi_\lambda^0\big(x_\lambda(t)\big) - \varphi_\lambda^0\big(x^*(t) + \rho w\big) \quad \text{for } t \in [0, T],\ |w| = 1.$$

Hence,

$$\big|q_\lambda(t)\big| \le C \quad \text{for all } t \in [0, T] \text{ and } \lambda > 0,$$

because φ^0 is continuous.

Also, notice that, in this case,

$$\ell_\lambda(x_1, x_2) = \frac{|x_1 - x_2|^2}{4\lambda} \quad \text{for all } x_1, x_2 \in E,$$

and therefore (4.42) becomes

$$p_\lambda(0) + x^*(0) - x_\lambda(0) = (2\lambda)^{-1}\big(x_\lambda(0) - x_\lambda(T)\big) = p_\lambda(T).$$

Since the operator $I - e^{A^*T}$ is invertible and, by (4.46),

$$p_\lambda(t) = e^{A^*(T-t)} p_\lambda(T) - \int_t^T e^{A^*(s-t)} q_\lambda(s)\,\mathrm{d}s, \quad 0 \le t \le T,$$

we may conclude that $\{|p_\lambda(T)|\}$ is bounded, as claimed.

Now, consider the case where E and U are Hilbert, $K = E$ and

$$\ell(x_1, x_2) = 0 \quad \text{if } x_1 = x_0, \qquad \ell(x_1, x_2) = +\infty \quad \text{if } x_1 \ne x_0.$$

If $x(t, x_0, u)$ is the solution to (4.1) with the initial condition $x(0) = x_0$, then Problem (P) can be written, in this case, as

$$\inf\big\{J(u);\ u \in L^2(0, T; U)\big\},$$

where $J : L^2(0, T; U) \to\ ]-\infty, +\infty]$ is the convex function

$$J(u) = \int_0^T L\big(t, x(t, x_0, u), u(t)\big)\,\mathrm{d}t.$$

The subdifferential ∂J is given by

$$\partial J(u) = \big\{w \in L^2(0, T; U);\ w(t) \in \partial_u L\big(x(t), u(t)\big) - B^* p(t)\big\},$$
$$\forall u \in L^2(0, T; U),$$

where (x, p) is a solution to the system

$$x' = A(t)x + Bu; \qquad p' \in -A^*(t)p + \partial_s L(x, u), \quad t \in [0, T],$$
$$x(0) = x_0, \qquad p(T) = 0.$$

This follows by Theorem 4.5, noticing that $w \in \partial J(u)$ if and only if u solves the minimization problem

$$\inf\left\{ J(u) - \int_0^T \langle u(t), w(t)\rangle \,\mathrm{d}t \right\}.$$

We may use the formula for ∂J to construct numerical algorithms for Problem (P). For instance, the classical gradient algorithm

$$u_{i+1} = u_i - \rho_i w_i, \quad w_i \in \partial J(u_i),\ \rho_i > 0,$$

reduces to

$$\begin{aligned}
&x_i' = A(t)x_i + Bu_i, \qquad p_i' \in -A^*(t)p_i + \partial_x L(x_i, u_i), \quad t \in [0, T],\\
&x_i(0) = x_0, \qquad p_i(T) = 0,\\
&u_{i+1} = u_i - \rho_i w_i; \quad w_i \in \partial_u L(x_i, u_i) - B^* p_i,\ i = 0, 1, \dots,
\end{aligned}$$

and we have (see problem (2.2))

$$u_i \to u^* \quad \text{weakly in } L^2(0, T; U),$$

where u^* is optimal.

4.1.7 A Finite-Dimensional Version of Problem (P)

We study here Problem (P) in the special case $E = \mathbb{R}^n$, $U = \mathbb{R}^m$ and

$$\ell(x_1, x_2) = \begin{cases} \ell_0(x_1) + \ell_1(x_2), & \text{if } x_1 \in C_0,\ x_2 \in C_1,\\ +\infty, & \text{otherwise.} \end{cases}$$

Namely,

$$\begin{aligned}
&\text{Minimize } \int_0^T L\big(t, x(t), u(t)\big)\,\mathrm{d}t + \ell_0\big(x(0)\big) + \ell_1\big(x(T)\big)\\
&\qquad\qquad \text{on all } (x, u) \in AC\big([0, T]; \mathbb{R}^n\big) \times \mathscr{U} \qquad (4.72)\\
&\text{subject to} \quad x' = A(t)x(t) + B(t)u + f(t) \quad \text{a.e. } t \in (0, T),\\
&\qquad\qquad x(0) \in C_0, \qquad x(T) \in C_1, \qquad (4.73)
\end{aligned}$$

under the following assumptions.

(k) The function $L : (0, T] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is convex and continuous in (x, u) and measurable in T. The Hamiltonian function

$$H(t, x, p) = \sup\big\{ p \cdot u - L(t, x, u);\ u \in U(t) \big\}$$

belongs to $L^1(0,T)$ for each $(x,p) \in \mathbb{R}^n \times \mathbb{R}^n$. For each $t \in [0,T]$, the set $U(t)$ is closed and convex and $\{t \in [0,T];\ C \cap U(t) \neq \emptyset\}$ is measurable for each closed subset C of $\mathbb{R}^m$.

(kk) The functions $\ell_0, \ell_1 : \mathbb{R}^n \to \mathbb{R}$ are convex and everywhere finite. The sets $C_0, C_1 \subset \mathbb{R}^n$ are closed and convex.

(kkk) There is $[x,u] \in AC([0,T];\mathbb{R}^n) \times \mathscr{U}$ satisfying the state system (4.73) such that $L(t,x,u) \in L^1(0,T)$ and either $x(0) \in \operatorname{int} C_0$ or $x(T) \in \operatorname{int} C_1$.

(kkkk) $A \in L^1(0,T;\mathbb{R}^n \times \mathbb{R}^n)$, $B \in L^\infty(0,T;\mathbb{R}^m \times \mathbb{R}^n)$, $f \in L^1(0,T;\mathbb{R}^n)$.

Here, $AC([0,T];\mathbb{R}^n)$ is the space of absolutely continuous function from $[0,T]$ to $\mathbb{R}^n$ and $\mathscr{U}$ is the set of all measurable functions $u : (0,T) \to \mathbb{R}^m$ such that $u(t) \in U(t)$ a.e. $t \in (0,T)$. We denote by $|\cdot|$ the norm in $\mathbb{R}^n$ and $\mathbb{R}^m$.

By (k), it follows that any optimal control u^* to problem (4.72) belongs to $L^1(0,T;\mathbb{R}^m)$. Indeed, since

$$L(t,x,u) = \sup\bigl\{p \cdot u - H(t,x,p);\ p \in \mathbb{R}^m\bigr\}, \quad \forall (x,u) \in \mathbb{R}^n \times U(t), \tag{4.74}$$

we have

$$L\bigl(t,x^*(t),u^*(t)\bigr) \ge \rho\bigl|u^*(t)\bigr| - H\bigl(t,x^*(t),\rho \operatorname{sgn} u^*(t)\bigr) \quad \text{a.e. } t \in (0,T)$$

and the latter implies that $u^* \in L^1(0,T;\mathbb{R}^m)$.

As regards the maximum principle, it has in this case the following form.

Theorem 4.11 *Assume that conditions* (k)–(kkkk) *are satisfied. Then, the pair* (x^*,u^*) *is optimal in problem* (4.72) *if and only if there exists* $p \in AC([0,T];\mathbb{R}^n)$ *which along with* x^* *and* u^* *satisfies the system*

$$p' + A^*(t)p \in \partial_x L(t,x^*,u^*) \quad \text{a.e. } t \in (0,T), \tag{4.75}$$

$$p(0) \in N_{C_0}\bigl(x^*(0)\bigr) + \partial\ell_0\bigl(x^*(0)\bigr), \tag{4.76}$$

$$-p(T) \in N_{C_1}\bigl(x^*(T)\bigr) + \partial\ell_1\bigl(x^*(T)\bigr),$$

$$B^*(t)p(t) \in \partial_u L\bigl(t,x^*(t),u^*(t)\bigr) + N_{U(t)}\bigl(u^*(t)\bigr) \quad \text{a.e. } t \in (0,T). \tag{4.77}$$

Here, $\partial L = [\partial_x L, \partial_u L]$ is the subdifferential of $L(t,\dots)$ and $A^*(t)$, $B^*(t)$ are the adjoint of $A(t) \in L(\mathbb{R}^n,\mathbb{R}^n)$ and of $B(t) \in L(\mathbb{R}^m,\mathbb{R}^n)$, respectively.

Proof Sufficiency. Let x^*, u^*, p satisfy system (4.73), and (4.76)–(4.77). By the definition of the subdifferential, we have, as in the proof of Theorem 4.5,

$$\begin{aligned} L\bigl(t,x^*(t),u^*(t)\bigr) &\le L\bigl(t,x(t),u(t)\bigr) + B^*(t)p(t) \cdot \bigl(u^*(t) - u(t)\bigr) \\ &\quad + \bigl(p'(t) + A^*(t)p(t)\bigr) \cdot \bigl(x^*(t) - x(t)\bigr) \quad \text{a.e. } t \in (0,T), \\ \ell_0\bigl(x(0)\bigr) + \ell_1\bigl(x^*(T)\bigr) &\le \ell_0\bigl(x(0)\bigr) + \ell_1\bigl(x(T)\bigr) \\ &\quad + p(0) \cdot \bigl(x^*(0) - x(0)\bigr) - p(T) \cdot \bigl(x^*(T) - x(T)\bigr), \end{aligned}$$

for any solution $(x, u) \in AC([0, T]; \mathbb{R}^n) \times \mathscr{U}$ to (4.73). This yields

$$\int_0^T L\big(t, x^*(t), u^*(t)\big)\,\mathrm{d}t + \ell_0\big(x^*(0)\big) + \ell_1\big(x^*(T)\big)$$
$$\leq \int_0^T L\big(t, x(t), u(t)\big)\,\mathrm{d}t + \ell_0\big(x(0)\big) + \ell_1\big(x(T)\big),$$

that is, (x^*, u^*) is optimal into problem (4.72).

Necessity. The proof is similar to that of Theorem 4.5. Denote by $\tilde{L}$, $\tilde{\ell}_0$ and $\tilde{\ell}_1$ the functions

$$\tilde{L}(t, x, u) = \begin{cases} L(t, x, u), & \text{if } u \in U(t), \\ +\infty, & \text{if } u \overline{\in} U(t), \end{cases}$$

$$\tilde{\ell}_0(x) = \begin{cases} \ell_0(x), & \text{if } x \in C_0, \\ +\infty, & \text{if } x \overline{\in} C_0, \end{cases}$$

$$\tilde{\ell}_1(x) = \begin{cases} \ell_1(x), & \text{if } x \in C_1, \\ +\infty, & \text{if } x \overline{\in} C_1, \end{cases}$$

and by $\tilde{L}_\lambda$, $(\tilde{\ell}_0)_\lambda$, $(\tilde{\ell}_1)_\lambda$ the regularized of $\tilde{L}$, $\tilde{\ell}_0$ and $\tilde{\ell}_1$, respectively, that is,

$$\tilde{L}_\lambda(t, x, u) = \inf\left\{\frac{|x-y|^2}{2\lambda} + \frac{|u-v|^2}{2\lambda} + L(t, y, v);\ (y, v) \in \mathbb{R}^n \times \mathbb{R}^m\right\},$$

$$\big(\tilde{\ell}_i\big)_\lambda(x) = \inf\left\{\frac{|x-y|^2}{2\lambda} + \ell_i(y),\ y \in \mathbb{R}^n\right\}, \quad i = 0, 1.$$

We recall that $\tilde{L}_\lambda(t, \cdot)$, $(\tilde{\ell}_i)_\lambda$ are convex, continuously differentiable and

$$\tilde{L}_\lambda(t, x, u) = \frac{|[x, u] - (I + \lambda\partial\tilde{L}(t, \cdot))^{-1}(x, u)|^2}{2\lambda} + \tilde{L}\Big(t, \big(I + \lambda\partial\tilde{L}(t, \cdot)\big)^{-1}(x, u)\Big), \quad \forall\lambda > 0, \tag{4.78}$$

$$\big(\tilde{\ell}_i\big)_\lambda(x) = \frac{|x - (I + \lambda\partial\ell_i)^{-1}x|^2}{2\lambda} + \tilde{\ell}_i\Big(\big(I + \lambda\partial\tilde{\ell}_i\big)\Big)^{-1}x, \quad i = 0, 1,\ \lambda > 0. \tag{4.79}$$

Let (x^*, u^*) be optimal in problem (4.72).

Consider the functions $\Phi_\lambda : L^1(0, T; \mathbb{R}^m) \times \mathbb{R}^n \to \mathbb{R}$ defined by

$$\Phi_\lambda(u, x_0) = \int_0^T \tilde{L}_\lambda\big(t, x(t, x_0, u), u(t)\big)\,\mathrm{d}t + \big(\tilde{\ell}_0\big)_\lambda(x_0) + \big(\tilde{\ell}_1\big)_\lambda\big(x(T, x_0, u)\big)$$
$$+ \varepsilon\int_0^T \big|u(t) - u^*(t)\big|\,\mathrm{d}t + 2^{-1}\big|x^*(0) - x_0\big|^2,$$

where $\varepsilon > 0$ is arbitrary but fixed and $x(t, x_0, u)$ is the solution to (4.73) with the initial value condition $x(0) = x_0$.

According to Ekeland's variational principle (Theorem 3.73), for every $\lambda > 0$ there exist $(u_\lambda, x_0^\lambda) \in L^1(0, T; \mathbb{R}^m) \times \mathbb{R}^n$ such that

$$\begin{aligned} \Phi_\lambda\big(u_\lambda, x_0^\lambda\big) = \inf\big\{&\Phi_\lambda(u, x_0) + \lambda^{\frac{1}{2}} \|u_\lambda - u\|_{L^1(0,T;\mathbb{R}^m)} \\ &+ \lambda^{\frac{1}{2}} \big\|x_0 - x_0^\lambda\big\|;\ (u, x_0) \in L^1\big(0, T; \mathbb{R}^m\big) \times \mathbb{R}^n\big\}. \end{aligned} \tag{4.80}$$

□

We set $x_\lambda = x(t, x_0^\lambda, u_\lambda)$, that is, $x_\lambda(0) = x_0^\lambda$.

Lemma 4.12 *We have*

$$B^*(t) p_\lambda(t) = \nabla_u \tilde{L}_\lambda\big(t, x_\lambda(t), u_\lambda(t)\big) + \varepsilon \eta_\lambda(t) + \lambda^{\frac{1}{2}} \xi_\lambda(t), \tag{4.81}$$

where $\eta_\lambda, \xi_\lambda \in L^\infty(0, T; \mathbb{R}^m)$, $|\eta_\lambda(t)|, |\xi_\lambda(t)| \le 1$, *a.e.* $t \in (0, T)$ *and* $p_\lambda \in AC([0, T]; \mathbb{R}^n)$, *satisfies the system*

$$p_\lambda' = -A^*(t) p_\lambda + \nabla_x \tilde{L}_\lambda(t, x_\lambda, u_\lambda) \quad \text{a.e. } t \in (0, T), \tag{4.82}$$

$$\begin{aligned} &p_\lambda(0) = \nabla\big(\tilde{\ell}_0\big)_\lambda\big(x_\lambda(0)\big) + x_\lambda(0) - x^*(0) + \lambda^{\frac{1}{2}} \nu_\lambda, \quad |\nu_\lambda| \le 1, \\ &p_\lambda(T) = -\nabla(\tilde{\ell}_1)_\lambda\big(x_\lambda(T)\big). \end{aligned} \tag{4.83}$$

Proof By (4.80), it follows that

$$\begin{aligned} &\lim_{h \downarrow 0} h^{-1}\big(\Phi_\lambda\big(u_\lambda + hv, x_0^\lambda + hx_0\big) - \Phi_\lambda\big(u_\lambda, x_0^\lambda\big)\big) \\ &\quad + \lambda^{\frac{1}{2}} \|v\|_{L^1(0,T;\mathbb{R}^m)} + \lambda^{\frac{1}{2}} |x_0| \ge 0, \end{aligned} \tag{4.84}$$

for all $(v, x_0) \in L^1(0, T; \mathbb{R}^m) \times \mathbb{R}^n$.

Now, let p_λ be the solution to (4.82) with final value condition $p_\lambda(T) = -\nabla(\tilde{\ell}_1)_\lambda(x_\lambda(T))$. Then, using the fact that $\tilde{L}_\lambda(t, \cdot), (\tilde{\ell}_0)_\lambda$ and $(\tilde{\ell}_1)_\lambda$ are differentiable, it follows that

$$\begin{aligned} &\lim_{h \to 0} h^{-1}\big(\Phi_\lambda\big(u_\lambda + hv, x_0^\lambda + hx_0\big) - \Phi_\lambda\big(u_\lambda, x_0^\lambda\big)\big) \\ &\quad = \int_0^T \big(\nabla_x \tilde{L}_\lambda(t, x_\lambda, u_\lambda) \cdot z(t) + \nabla_u \tilde{L}_\lambda(t, x_\lambda, u_\lambda) \cdot v(t)\big)\, dt \\ &\quad\quad + \nabla\big(\tilde{\ell}_0\big)_\lambda\big(x_0^\lambda\big) \cdot x_0 + \nabla\big(\tilde{\ell}_1\big)_\lambda\big(x_\lambda(T)\big) \cdot z(T) \\ &\quad\quad + \big(x_\lambda(0) - x^*(0)\big) \cdot x_\lambda(0) + \varepsilon \int_0^T \eta_\lambda(t) \cdot v(t)\, dt, \end{aligned}$$

where $\eta_\lambda(t) = \operatorname{sgn}(u_\lambda(t) - u^*(t))$ and

$$\begin{aligned} z' &= A(t)z + B(t)v \quad \text{a.e. } t \in (0,T), \\ z(0) &= x_0. \end{aligned}$$

(Here, $\operatorname{sgn} v = \frac{v}{|v|}$ if $|v| \le 1$, $\operatorname{sgn} 0 = \{w \in \mathbb{R}^m; |w| \le 1\}$.)

Then, by (4.82), we see that

$$\begin{aligned} &\lim_{h\to 0}\big(\Phi_\lambda\big(u_\lambda + hv, x_0^\lambda + hx_0\big) - \Phi_\lambda\big(u_\lambda, x_0^\lambda\big)\big) \\ &\quad = \int_0^T \big(\nabla_u \tilde{L}(t, x_\lambda, u_\lambda) - B^* p_\lambda + \varepsilon\eta_\lambda\big) \cdot v \,\mathrm{d}t \\ &\qquad + \big(\nabla(\tilde{\ell}_0)_\lambda\big(x_\lambda(0)\big) + x_\lambda(0) - x^*(0) - p_\lambda(0)\big) \cdot x_0 \end{aligned}$$

and, by (4.83), it follows that

$$\begin{aligned} &\int_0^T \big(\nabla \tilde{L}_\lambda(t, x_\lambda, u_\lambda) - B^* p_\lambda + \varepsilon\eta_\lambda\big) \cdot v \,\mathrm{d}t + \lambda^{\frac{1}{2}} \int_0^T \big|v(t)\big| \,\mathrm{d}t \\ &\quad + \big(\nabla\big(\tilde{\ell}_0\big)_\lambda\big(x_\lambda(0)\big) + x_\lambda(0) - x^*(0) - p_\lambda(0)\big) \cdot x_0 \\ &\quad + \lambda^{\frac{1}{2}} |x_0| \ge 0, \quad \forall v \in L^1\big(0,T;\mathbb{R}^m\big),\ \forall x_0 \in \mathbb{R}^n, \end{aligned}$$

and this implies (4.81) and the first end-point condition in (4.83). □

Lemma 4.13 *For* $\lambda \to 0$,

$$\begin{aligned} u_\lambda &\to u^* \quad \textit{strongly in } L^1\big(0,T;\mathbb{R}^m\big), \\ x_\lambda &\to x^* \quad \textit{uniformly on } [0,T]. \end{aligned}$$

Proof We have, by (4.80),

$$\begin{aligned} \Phi_\lambda\big(u_\lambda, x_\lambda(0)\big) &\le \Phi_\lambda\big(u^*, x^*(0)\big) + \lambda^{\frac{1}{2}} \int_0^T |u_\lambda - u^*| \,\mathrm{d}t + \lambda^{\frac{1}{2}} \big|x^*(0) - x_\lambda(0)\big| \\ &\le \int_0^T L(t, x^*, u^*) \,\mathrm{d}t + \ell_0\big(x^*(0)\big) + \ell_1\big(x^*(T)\big) \\ &\quad + \lambda^{\frac{1}{2}} \bigg(\int_0^T |u_\lambda - u^*| \,\mathrm{d}t + \big|x^*(0) - x_\lambda(0)\big|\bigg), \end{aligned} \tag{4.85}$$

because $\tilde{L} \le L$ and $(\tilde{\ell}_i)_\lambda \le \ell_i$, $i = 0, 1$.

We note that, by the conjugation formula (4.74), we have

$$L(t,x,u) \ge -H(t,x,0), \quad \forall x \in \mathbb{R}^n,\ u \in U(t),\ t \in [0,T],$$

and that

$$-H(t,0,0) \le -H(t,x,0) - \eta(t)\cdot x, \quad \forall x \in \mathbb{R}^n,\ t \in [0,T], \tag{4.86}$$

where $\eta(t) \in \partial_x(-H(t,0,0))$ and

$$\|\eta(t)\| \le H(t,0,0) + \sup\{-H(t,y,0);\ |y| \le 1\}$$

and so, by Lemma 4.14 below, $\eta \in L^1(0,T;\mathbb{R}^m)$.

We have, therefore,

$$L(t,x,u) \ge \eta(t)\cdot x + H(t,0,0), \quad \forall x \in \mathbb{R}^n,\ u \in U(t),\ t \in [0,T].$$

Then, replacing, if necessary, $\tilde{L}$ by $\tilde{L}(t,x,u) - \eta(t)\cdot x - H(t,0,0)$ and taking

$$\begin{aligned}\Phi_\lambda(u,x_0) = {} & \int_0^T \big(\tilde{L}_\lambda(t,x_0,u) + \eta(t)\cdot x(t,u,x_0) + H(t,0,0)\big)\,\mathrm{d}t + \big(\tilde{\ell}_0\big)_\lambda(x_0)\\ & + \big(\tilde{\ell}_1\big)_\lambda\big(x(T,x_0,u)\big) + \varepsilon\int_0^T |u-u^*|\,\mathrm{d}t + 2^{-1}\big|x_0 - x^*(0)\big|^2,\end{aligned}$$

we may assume that

$$\int_0^T \tilde{L}_\lambda(t,x_\lambda,u_\lambda)\,\mathrm{d}t \ge \int_0^T \beta(t)\,\mathrm{d}t,$$

where $\beta \in L^1(0,T)$. We have $|x_\lambda(t)| \le C(|x_\lambda(0)| + \int_0^T |u_\lambda(t)|\,\mathrm{d}t)$, $t \in (0,T)$. Then, by (4.85), we see that

$$\|u_\lambda - u^*\|_{L^1(0,T)} + \big|x_\lambda(0)\big| \le C, \quad \forall \lambda > 0.$$

We also note that, by (4.78),

$$\begin{aligned}\int_0^T L(t,x_\lambda,u_\lambda)\,\mathrm{d}t &\le \int_0^T \tilde{L}\big(t,\big(I+\lambda\partial\tilde{L}\big)^{-1}(x_\lambda,u_\lambda)\big)\,\mathrm{d}t\\ &\quad + \lambda^{-1}\int_0^T \big|\big(I+\lambda\partial\tilde{L}\big)^{-1}(x_\lambda,u_\lambda) - (x_\lambda,u_\lambda)\big|^2\,\mathrm{d}t\\ &\le 2\int_0^T \tilde{L}_\lambda(t,x_\lambda,u_\lambda)\,\mathrm{d}t \le C, \quad \forall \lambda > 0.\end{aligned}$$

On the other hand, again by (4.74) and Lemma 4.14, we have

$$L(t,x,u) \ge N|u| - H(t,x,N\operatorname{sgn} u) \ge N|u| - \beta_N(t), \quad \forall t \in (0,T),\ |x| \le N,$$

where $\beta_N \in L^1(0,T)$.

Hence, for each measurable subset $E_0 \subset (0, T)$, we have

$$\int_{E_0} \left|u_\lambda(t)\right| \mathrm{d}t \le \frac{1}{N} \int_{E_0} L(t, x_\lambda, u_\lambda)\, \mathrm{d}t + \frac{1}{N} \int_{E_0} \left|\beta_N(s)\right| \mathrm{d}s + \frac{C}{N} + \frac{1}{N} \int_{E_0} \left|\beta_N(s)\right| \mathrm{d}s.$$

Then, by the Dunford–Pettis theorem (Theorem 1.121), we infer that $\{u_\lambda\}$ is weakly compact in $L^1(0, T; \mathbb{R}^m)$, and so, on a subsequence convergent to zero, again denoted λ, we have

$$\begin{aligned} u_\lambda &\to \tilde{u} \quad \text{weakly in } L^1\big(0, T; \mathbb{R}^m\big), \\ x_\lambda &\to \tilde{x} \quad \text{in } C\big([0, T]; \mathbb{R}^n\big), \\ x_\lambda' &\to \tilde{x}' \quad \text{weakly in } L^1\big(0, T; \mathbb{R}^n\big). \end{aligned}$$

Then, by (4.78), it follows that

$$\lim_{\lambda \to 0} \int_0^T \tilde{L}_\lambda(t, x_\lambda, u_\lambda)\, \mathrm{d}t \ge \int_0^T \tilde{L}(t, \tilde{x}, \tilde{u})\, \mathrm{d}t,$$

because, as seen earlier, the function $(x, u) \to \int_0^T L(t, x, u)\, \mathrm{d}t$ is lower-semicontinuous in $L^1(0, T; \mathbb{R}^n) \times L^1(0, T; \mathbb{R}^m)$. Similarly, by (4.79), we have

$$\begin{aligned} \liminf_{\lambda \to 0} (\ell_0)_\lambda\big(x_\lambda(0)\big) &\ge \tilde{\ell}_0\big(\tilde{x}(0)\big), \\ \liminf_{\lambda \to 0} (\ell_1)_\lambda\big(x_\lambda(T)\big) &\ge \tilde{\ell}_1\big(\tilde{x}(T)\big). \end{aligned}$$

Then, letting λ tend to zero in (4.85), we get

$$\begin{aligned} &\int_0^T \tilde{L}(t, \tilde{x}, \tilde{u})\, \mathrm{d}t + \tilde{\ell}_0\big(\tilde{x}(0)\big) + \tilde{\ell}_1\big(\tilde{x}(T)\big) \\ &\quad + \liminf_{\lambda \downarrow 0} \left(\varepsilon \int_0^T |u_\lambda - u^*|\, \mathrm{d}t + 2^{-1} \left|x_\lambda(0) - x^*(0)\right|^2 \right) \le \inf(4.72), \end{aligned}$$

and the conclusion of Lemma 4.14 follows. □

We are going to obtain the optimality system (4.72)–(4.77) by letting λ tend to zero in (4.81)–(4.83). To this purpose, some a priori estimates on p_λ are necessary. Let (x, u) be an admissible pair chosen as in assumption (kkk) ($x(0) \in \operatorname{int} C_0$). By (4.83), we have

$$\begin{aligned} &\big(p_\lambda(0) + x^*(0) - x_\lambda(0) - \lambda^{\frac{1}{2}} \nu_\lambda\big) \cdot \big(x_\lambda(0) - x(0) - \rho w\big) \\ &\quad \ge \big(\tilde{\ell}_0\big)_\lambda\big(x_\lambda(0)\big) - \big(\tilde{\ell}_0\big)_\lambda\big(x(0) + \rho w\big), \end{aligned}$$

for all $\|w\| = 1$ and $\rho > 0$. Since, for ρ sufficiently small,

$$\big(\tilde{\ell}_0\big)_\lambda\big(x(0) + \rho w\big) \le \tilde{\ell}_0\big(x(0) + \rho w\big) \le C, \quad \forall \lambda > 0,$$

we get

$$\rho\left|p_\lambda(0)\right| \le C + p_\lambda(0)\cdot\big(x_\lambda(0)-x(0)\big) \quad \forall\lambda>0.$$

On the other hand, by (4.82) we see that

$$\begin{aligned}&-p_\lambda(0)\cdot\big(x_\lambda(0)-x(0)\big)+p_\lambda(T)\cdot\big(x_\lambda(T)-x(T)\big)\\&\quad=\int_0^T \nabla_x\tilde{L}_\lambda(t,x_\lambda,u_\lambda)\cdot(x_\lambda-x)\,\mathrm{d}t+\int_0^T B(u_\lambda-u)\cdot p_\lambda\,\mathrm{d}t,\end{aligned}$$

because $x'=A(t)x+B(t)u+f(t)$ a.e. $t\in(0,T)$.

Now, using (4.82), we get

$$\begin{aligned}-p_\lambda(0)\cdot\big(x_\lambda(0)-x(0)\big)&\ge\big(\tilde{\ell}_1\big)_\lambda\big(x_\lambda(T)\big)-\big(\tilde{\ell}_1\big)_\lambda\big(x(T)\big)\\&\quad+\int_0^T\big(\tilde{L}_\lambda(t,x_\lambda,u_\lambda)-\tilde{L}_\lambda(t,x,u)\big)\,\mathrm{d}t\\&\quad+\int_0^T(u_\lambda-u)\big(\varepsilon\eta_\lambda(t)+\lambda^{\frac12}\xi_\lambda(t)\big)\,\mathrm{d}t\ge C,\quad\forall\lambda>0,\end{aligned}$$

because $\tilde{L}_\lambda\le L$ and $(\ell_1)_\lambda\le\ell_1,\ \forall\lambda>0$.

Hence, $\{p_\lambda(0)\}$ is bounded in $\mathbb{R}^n$.

For further estimates, we need the following lemma, which was already invoked in the proof of Lemma 4.13.

Lemma 4.14 *For any $r>0$, there exist $\alpha_r,\beta_r\in L^1(0,T)$ such that*

$$-H(t,x,0)\le\alpha_r(t)\quad a.e.\ t\in(0,T),\ |x|\le r, \tag{4.87}$$

$$H(t,x,w)\le\beta_r(t)\quad a.e.\ t\in(0,T),\ |x|\le r,\ |w|\le r. \tag{4.88}$$

Proof The function $H(t,x,p)$ is convex in p and concave in x. Let $x_1,\dots,x_{n+1}$ be such that the n-dimensional simplex generated by these points contains the ball of radius r centered at the origin. Since, by assumption (k), $H(t,x_i,0)\in L^1(0,T)$ for all i, by convexity of $x\to -H(t,x,0)$ we get (4.87). Similarly, if $w_1,\dots,w_{m+1}$ generates an m-dimensional simplex containing $\{w\in\mathbb{R}^m;\ |w|\le r\}$, we have by the convexity of $w\to H(t,x,w)$ that

$$H(t,x,w)\le\sup_i H(t,y,w_i)\le\sup_i\sup_{|y|\le r}H(t,y,w_i)\quad\text{for } |x|,|w|\le r.$$

Since, by hypothesis (k), $\sup_{\|x\|\le r}H(t,y,w_i)=H(t,x_i,w_i)\in L^1(0,T)$, the latter implies (4.88). In particular, it follows by (4.87) and (4.88) that

$$\sup\big\{|v|;\ v\in\partial_p H(t,y,0)\big\}\le\beta(t)\quad\text{a.e. } t\in(0,T), \tag{4.89}$$

where $\beta\in L^1(0,T)$.

Indeed, by the definition of $\partial_p H$, we have

$$H(t,y,0) \le H(t,y,w) - v \cdot w, \quad \forall w \in \mathbb{R}^m,$$

for every $v \in \partial_p H(t,y,0)$. This yields

$$\|v\| \le H(t,y,w) - H(t,y,0) \le \beta(t), \quad \|w\| = 1.$$

Now, coming back to the proof of Theorem 4.11, we note first that, by (4.74), we have

$$\tilde{L}(t,y,v) = -H(t,y,0), \quad \forall v \in \partial_p H(t,y,0).$$

Then, by (4.89), we see that there exist the functions $\alpha, \beta \in L^1(0,T)$ independent of w and $v_w : [0,T] \to \mathbb{R}^m$ measurable such that $|v_w(t)| \le \beta(t)$ a.e. $t \in (0,T)$ and

$$L\big(t, x^*(t) + \rho w, v_w(t)\big) \le \alpha(t), \quad \text{a.e. } t \in (0,T),\ |w| \le 1. \tag{4.90}$$

By (4.81) and (4.82) we see that

$$\begin{aligned}
&\big(p_\lambda' + A^*(t)p_\lambda\big) \cdot (x_\lambda - x^* - \rho w) + \big(B^* p_\lambda - \varepsilon \eta_\lambda - \lambda^{\frac{1}{2}} \xi_\lambda\big) \cdot (u_\lambda - v_w) \\
&\quad \ge \tilde{L}_\lambda(t, x_\lambda, u_\lambda) - \tilde{L}_\lambda(t, x^* + \rho w, v_w) \\
&\quad \ge \tilde{L}_\lambda(t, x_\lambda, u_\lambda) - \alpha(t), \quad \text{a.e. } t \in (0,T).
\end{aligned}$$

Since $x_\lambda \to x^*$ uniformly on $[0,T]$, the latter yields

$$\begin{aligned}
\big|p_\lambda'(t) + A^*(t)p_\lambda(t)\big| &\le C\big(\alpha(t) - \tilde{L}_\lambda\big(t, x_\lambda(t), u_\lambda(t)\big)\big) \\
&\quad + \big(\big|p_\lambda(t)\big| + \varepsilon + \lambda^{\frac{1}{2}}\big)\big(\big|u_\lambda(t)\big| + \beta(t)\big) \\
&\text{a.e. } \in (0,T).
\end{aligned} \tag{4.91}$$

Since, as easily seen by (4.78) and by Lemma 4.14,

$$\int_0^t \tilde{L}_\lambda\big(s, x_\lambda(s), u_\lambda(s)\big)\,ds \ge C, \quad \forall \lambda > 0,\ t \in [0,T],$$

and $\{|u_\lambda|\}$ is weakly compact in $L^1(0,T)$, the previous inequality implies via Gronwall's Lemma that

$$\big|p_\lambda(t)\big| \le C, \quad \forall \lambda > 0,\ t \in [0,T].$$

We have also

$$\|p_\lambda'\|_{L^1(0,T;\mathbb{R}^n)} \le C, \quad \forall \lambda > 0,$$

and so, $\{p_\lambda\}$ is compact in $C([0,T];\mathbb{R}^n)$. By (4.91), we see also via the Dunford–Pettis theorem that $\{p_\lambda'\}$ is weakly compact in $L^1(0,T;\mathbb{R}^n)$. Hence, on a subsequence, we have

$$\begin{aligned} p_\lambda(t) &\to p(t) \quad \text{uniformly on } [0,T],\\ p'_\lambda &\to p' \quad \text{weakly in } L^1\big(0,T;\mathbb{R}^n\big),\\ \eta_\lambda &\to \eta^\varepsilon \quad \text{weak-star in } L^\infty\big(0,T;\mathbb{R}^m\big),\\ \nabla_x \tilde{L}_\lambda(t,x_\lambda,u_\lambda) &\to q_1 \quad \text{weakly in } L^1\big(0,T;\mathbb{R}^n\big),\\ \nabla_u \tilde{L}_\lambda(t,x_\lambda,u_\lambda) &\to q_2 \quad \text{weak-star in } L^\infty\big(0,T;\mathbb{R}^m\big), \end{aligned}$$

where $p = p_\varepsilon$.

Now, from the inequality

$$\begin{aligned} \int_0^T \tilde{L}_\lambda(t,x_\lambda,u_\lambda)\,\mathrm{d}t \le{} & \int_0^T \tilde{L}_\lambda(t,x,u)\,\mathrm{d}t\\ &+ \int_0^T \big(\nabla_x \tilde{L}_\lambda(t,x_\lambda,u_\lambda)\cdot(x_\lambda - x)\\ &+ \nabla_u \tilde{L}_\lambda(t,x_\lambda,u_\lambda)\cdot(u_\lambda - u)\big)\,\mathrm{d}t,\\ \forall (x,u) \in L^\infty\big(0,T;\mathbb{R}^n\big) \times L^1\big(0,T;\mathbb{R}^m\big), \end{aligned}$$

we get

$$\int_0^T \tilde{L}(t,x^*,u^*)\,\mathrm{d}t \le \int_0^T \tilde{L}(t,x,u)\,\mathrm{d}t + \int_0^T \big(q_1\cdot(x^*-x) + q_2\cdot(u^*-u)\big)\,\mathrm{d}t,$$

for all $(x,u) \in L^\infty(0,T;\mathbb{R}^n) \times L^1(0,T;\mathbb{R}^m)$ and this implies, as in the previous case, that

$$\big(q_1(t), q_2(t)\big) \in \partial \tilde{L}\big(t, x^*(t), u^*(t)\big) \quad \text{a.e. } t \in (0,T).$$

Then, letting $\lambda \to 0$ into (4.81) and (4.82), we get

$$p' + A^*(t)p \in \partial_x L(t,x^*,u^*) \quad \text{a.e. } t \in (0,T), \tag{4.92}$$

$$\begin{aligned} &B^*(t)p(t) \in \partial_u L\big(t,x^*(t),u^*(t)\big) + N_{U(t)}\big(u^*(t)\big) + \varepsilon\eta^\varepsilon(t)\\ &\quad \text{a.e. } t \in (0,T). \end{aligned} \tag{4.93}$$

Next, letting $\lambda \to 0$ into system (4.83) and taking into account that $\nabla(\tilde{\ell}_i)_\lambda \in \partial\tilde{\ell}_i((I + \lambda\partial\tilde{\ell}_i)^{-1})$, it follows by Lemma 4.13 and by relations (4.79) that

$$p(0) \in \partial\tilde{\ell}_0\big(x^*(0)\big), \qquad p(T) \in -\partial\tilde{\ell}_1\big(x^*(T)\big).$$

Since $\partial\tilde{\ell}_i(x) = \partial\ell_i(x) + N_{C_i}(x)$, $i = 0,1$, we get

$$\begin{aligned} p(0) &\in \partial\ell_0\big(x^*(0)\big) + N_{C_0}\big(x^*(0)\big);\\ -p(T) &\in \partial\ell_1\big(x^*(T)\big) + N_{C_1}\big(x^*(T)\big). \end{aligned} \tag{4.94}$$

Here, as well as in (4.93), we have used the additivity formula for subdifferentials (Corollary 2.63).

Now, to conclude the proof, we let $\varepsilon \to 0$ into the above equations. Indeed, if we denote by p_ε the solution to (4.92)–(4.94), that is,

$$\bigl(p'_\varepsilon + A^*(t)p_\varepsilon, B^*(t)p_\varepsilon - \varepsilon\eta^\varepsilon\bigr) \in \partial\tilde{L}(t, x^*, u^*) \quad \text{a.e. in } (0, T), \tag{4.95}$$

$$\begin{aligned} &p_\varepsilon(0) \in \partial\ell_0\bigl(x^*(0)\bigr) + N_{C_0}\bigl(x^*(0)\bigr), \\ &-p_\varepsilon(T) \in \partial\ell_1\bigl(x^*(T)\bigr) + N_{C_1}\bigl(x^*(T)\bigr), \end{aligned} \tag{4.96}$$

we have, as above,

$$p_\varepsilon(0) \cdot \bigl(x^*(0) - x(0) - \rho w\bigr) \ge \ell_0\bigl(x^*(0)\bigr) - \ell_0\bigl(x(0) - \rho w\bigr),$$

where (x, u) is as in hypothesis (kkk). This yields

$$\begin{aligned} \rho\bigl\|p_\varepsilon(0)\bigr\| &\le \rho_\varepsilon(0) \cdot \bigl(x_\varepsilon(0) - x(0)\bigr) + C \\ &\le -\ell_1\bigl(x^*(T)\bigr) + \ell_1\bigl(x(T)\bigr) \\ &\quad + \int_0^T \bigl(L(t, x, u) - L(t, x^*, u^*)\bigr)\,\mathrm{d}t + \varepsilon\int_0^T \|u^* - u\|\,\mathrm{d}t. \end{aligned}$$

Hence, $\{p_\varepsilon(0)\}$ is bounded in $\mathbb{R}^n$ and, arguing as above, we find by (4.90) and from the inequality

$$-\bigl(p'_\varepsilon + A^*(t)p_\varepsilon\bigr) \cdot w + \bigl(B^* p_\varepsilon + \varepsilon\eta^\varepsilon\bigr) \cdot (u_\varepsilon - v_0) \ge L(t, x^*, u^*) - \alpha(t)$$

that $\{p_\varepsilon\}$ is compact in $C([0, T]; \mathbb{R}^n)$ and $\{p'_\varepsilon\}$ is weakly compact in $L^1(0, T; \mathbb{R}^n)$. Then, we may pass to the limit to (4.95) and (4.96) to get (4.75)–(4.77). This completes the proof. □

Remark 4.15 It should be emphasized that Theorem 4.11 cannot be deduced by Theorem 4.5, which refer to optimal controllers u in $L^p(0, T; U)$ with $p > 1$. Theorem 4.11 extends to the case of state-constraint optimal control problems of the form (4.5) (see Rockafellar [41]).

4.1.8 The Dual Control Problem

Having in mind the general duality theory developed in Chap. 3, one may speculate that the dual extremal arc p^* is itself the solution to a certain control problem. We see that this is, indeed, the case in a sense which is explained below.

Given the functions $L(t)$ and ℓ defined on $E \times U$ ($R \times E$, respectively) and the closed convex subset K of E, we set

$$M(t, p, q) = \sup\bigl\{\langle p, v\rangle + (q, y) - L(t, y, v);\ y \in K,\ v \in U\bigr\},$$

$$m(p_1, p_2) = \sup\bigl\{(p_1, h_1) - (p_2, h_2) - \ell(h_1, h_2);\ h_1, h_2 \in E\bigr\},$$

for $(p, q) \in U^* \times E^*$ and $(p_1, p_2) \in E^* \times E^*$.

The hypotheses of Sect. 4.1.1 are understood to hold.

In terms of conjugate functions, $M(t): U^* \times E^* \to \overline{\mathbb{R}}^*$ and $m: E^* \times E^* \to \overline{\mathbb{R}}^*$ can be written as

$$\begin{aligned} M(t,p,q) &= \big(L(t)+I_K\big)^*(q,p),\\ m(p_1,p_2) &= \ell^*(p_1,-p_2). \end{aligned}$$

If $\rho \in BV([0,T];E^*)$ is given, we denote, as usual, by $\dot\rho \in L^1(0,T;E^*)$ the weak derivative of ρ, and by $\rho_s(t) = \rho(t) - \int_0^t \dot\rho(s)\,\mathrm{d}s$ the singular part of ρ.

Now, let $G_0: BV([0,T];E^*) \to \overline{\mathbb{R}}^*$ be the convex function

$$G_0(\rho) = \sup\left\{\int_0^T (\mathrm{d}\rho_s, x);\ x \in \mathscr{K}\right\}.$$

It should be remarked that $G_0(\rho) = H_0(d\rho_s)$, where $H_0: M(0,T;E^*) \to \overline{\mathbb{R}}^*$ is the support function of $\mathscr{K}$ and $d\rho_s$ is the Lebesgue–Stieltjes measure defined by ρ_s (see Sect. 1.3.3).

We take as the dual of (P) the following optimization problem:

Minimize

$$(\mathrm{P}^*) \quad \int_0^T \Big(M\big(t,(B^*p)(t),\dot\rho(t)\big) + \big(f(t),p(t)\big)\Big)\,\mathrm{d}t + m\big(p(0),p(T)\big) + G_0(\rho)$$

over all $\rho \in BV\big([0,T];E^*\big)$ and $p:[0,T] \to E^*$

$$\text{subject to} \quad p(t) = U^*(T,t)p(T) - \int_t^T U^*(s,t)\,\mathrm{d}\rho(s), \quad 0 \le t \le T. \tag{4.97}$$

Here, we agree to consider any $\rho \in BV([0,T];E^*)$ as a left continuous function on $]0,T[$. In particular, this implies that $p(T) = p(T-0)$.

If there are no state constraints in the primal problem, that is, $K = E$, then we take the dual problem to be that of *minimizing over the space* $C([0,T];E^*) \times L^p(0,T;U^*)$ *the functional*

$$\begin{aligned} &\int_0^T \Big(M\big(t,(B^*p)(t),v(t)\big) + \big(f(t),p(t)\big)\Big)\,\mathrm{d}t + m\big(p(0),p(T)\big)\\ &\text{subject to} \quad p' + A^*(t)p = v(t), \quad 0 \le t \le T. \end{aligned} \tag{4.98}$$

Problem (P*) should be compared with the general dual problem $\mathscr{P}^*$ introduced in Sect. 3.2.1 and it can be regarded as a control problem with the input $v = \mathrm{d}\rho$ and the state equation

$$p' + A^*(t)p = \mathrm{d}\rho \quad \text{on } [0,T].$$

Now, we discuss the circumstances under which the integral in the dual problem (P*) is meaningful for all $\rho \in BV([0,T];E^*)$. The existence of this integral ought to involve a measurability condition on $M(t)$ of the following type.

$M(t, q(t), v(t))$ is a Lebesgue measurable function of t whenever $q : [0, T] \to U^*$ and $v : [0, T] \to E^*$ are Lebesgue measurable.

As pointed out before, there are several notable cases in which this condition may be derived from Assumption (C), but we omit the details.

By Assumption (C)(iii), we see that, for all $q \in L^{p'}(0, T; U)$ and $v \in L^1(0, T; E^*)$, there exists some $\gamma \in L^1(0, T)$, such that

$$M\big(t, q(t), v(t)\big) \geq \gamma(t) \quad \text{a.e. } t \in]0, T[.$$

On the other hand, it follows by part (ii) of Assumption (C) that

$$M\big(t, s_0(t), r_0(t)\big) \leq g_0(t) \quad \text{a.e. } t \in]0, T[.$$

It follows from the above inequalities that the functional

$$\int_0^T \Big(M\big(t, (B^* p)(t), \dot{\rho}(t)\big) + \big(f(t), p(t)\big)\Big)\,\mathrm{d}t$$

is well defined (that is, equal either to a real number or $+\infty$) and nonidentically $+\infty$.

In passing, we remark that the dual problem (P*), as well as the primal problem (P), involves implicit constraints on the control ρ

$$\dot{\rho}(t) \in W\big(t, p(t)\big) \quad \text{a.e. } t \in]0, T[,$$

where

$$W(t, p) = \big\{v \in E^*;\ M\big(t, (B^* p)(t), v\big) < +\infty\big\}.$$

Similarly, there is the end-point constraint

$$\big(p(0), p(T)\big) \in D(m) = \big\{(p_1, p_2) \in E^* \times E^*;\ m(p_1, p_2) < +\infty\big\}.$$

Now, we are ready to formulate the duality theorem.

Theorem 4.16 *Let the assumptions of Theorem 4.5 be satisfied. Then the pair $(x^*, u^*) \in C([0, T]; E^*) \times L^p(0, T; U^*)$ is optimal for problem* (P) *if and only if the dual problem* (P*) *has a solution p^* such that*

$$\int_0^T L(t, x^*, u^*)\,\mathrm{d}t + \int_0^T \Big(M(t, B^* p^*, \dot{\rho}) + \big(f(t), p^*(t)\big)\Big)\,\mathrm{d}t$$
$$\ell\big(x^*(0), x^*(T)\big) + m\big(p^*(0), p^*(T)\big) + G_0(\rho) = 0. \tag{4.99}$$

Furthermore, the function p^* satisfies, along with $w(t) = \rho(t) - \int_0^T q(s)\,\mathrm{d}s$ and x^*, u^*, (4.17)–(4.21).

Proof Let (x^*, u^*) be optimal in problem (P) and let p^* be a dual extremal arc. By virtue of the conjugacy relation (4.75), the transversality condition (4.21) is satisfied if and only if

$$\ell\big(x^*(0), x^*(T)\big) + m\big(p^*(0), p^*(T)\big) = \big(x^*(0), p^*(0)\big) - \big(x^*(T), p^*(T)\big), \tag{4.100}$$

while, for arbitrary $y \in C([0,T]; E)$ and $\tilde{p} : [0,T] \to E^*$, it would be true that

$$\ell\big(y(0), y(T)\big) + m\big(\tilde{p}(0), \tilde{p}(T)\big) \ge \big(y(0), \tilde{p}(0)\big) - \big(y(T), \tilde{p}(T)\big). \tag{4.101}$$

Similarly, since by (4.20) and (4.25)

$$(q + \dot{w}, B^* p^*) \in \partial\big(L(t, x^*, u^*) + I_K(x^*)\big) \quad \text{a.e. on }]0,T[,$$

it follows by (4.74) that

$$\begin{aligned} &L\big(t, x^*(t), u^*(t)\big) + M\big(t, (B^*p^*)(t), q(t) + \dot{w}(t)\big) \\ &\quad = \big\langle (B^*p^*)(t), u^*(t)\big\rangle + \big(x^*(t), q(t) + \dot{w}(t)\big) \quad \text{a.e. } t \in]0,T[, \end{aligned} \tag{4.102}$$

while, for arbitrary $(y, v) \in \mathscr{H}$, $y \in \mathscr{K}$, $\tilde{p} \in L^q(0,T; E^*)$, $\frac{1}{p} + \frac{1}{q} = 1$ and $z : [0,T] \to E^*$, it would be true that

$$\begin{aligned} &L\big(t, y(t), v(t)\big) + M\big(t, (B^*\tilde{p})(t), z(t)\big) \ge \big\langle (B^*\tilde{p})(t), v(t)\big\rangle + \big(y(t), z(t)\big) \\ &\quad \text{a.e. } t \in]0,T[. \end{aligned} \tag{4.103}$$

Integrating both sides of inequality (4.102) and adding (4.100), since, by (4.17) and (4.18),

$$\begin{aligned} &\int_0^T \big(\big\langle B^*p^*, u^*(t)\big\rangle + \big(x^*(t), q(t)\big)\big)\,\mathrm{d}t + \int_0^T \big(\mathrm{d}w(t), x^*(t)\big) + \int_0^T \big(f(t), p^*(t)\big)\,\mathrm{d}t \\ &\quad = -\big(x^*(0), p^*(0)\big) + \big(x^*(T), p^*(T)\big), \end{aligned}$$

and, by (4.26),

$$G_0(w) = \int_0^T (\mathrm{d}w_s, x^*),$$

we obtain the equality

$$\begin{aligned} &\int_0^T L(t, x^*, u^*)\,\mathrm{d}t + \ell\big(x^*(0), x^*(T)\big) \\ &\quad + \int_0^T \big(M(t, B^*p^*, q + \dot{w}) + (p^*, f)\big)\,\mathrm{d}t + m\big(p^*(0), p^*(T)\big) \\ &\quad + G_0(w) = 0. \end{aligned} \tag{4.104}$$

Now, integrating (4.103), where $(y, v) \in \mathscr{H}$, $y \in \mathscr{K}$, $z = \dot{\rho}$ and $\tilde{p}$ is the solution to (4.97), we obtain, after some calculations involving Inequality (4.101),

$$\int_0^T L(t, y, v)\,\mathrm{d}t + \int_0^T \big(M(t, B^*\tilde{p}, \dot{\rho}) + (f, \tilde{p})\big)\,\mathrm{d}t \\ + \int_0^T (\mathrm{d}\rho_s, y) + \ell\big(y(0), y(T)\big) + m\big(\tilde{p}(0), \tilde{p}(T)\big) \geq 0.$$

Along with (4.104), the latter shows that p^* is a solution to (P*) corresponding to $\rho = w + \int_0^T q\,\mathrm{d}s$ and equality (4.99) follows.

Conversely, if (p^*, ρ) is an optimal pair for (P*) satisfying (4.99), it follows by (4.101) and (4.103) that

$$L\big(t, x^*(t), u^*(t)\big) + M\big(t, (B^*p^*)(t), \dot{\rho}(t)\big) = \big\langle (B^*p^*)(t), u^*(t)\big\rangle + \big(x^*(t), \dot{\rho}(t)\big) \\ \text{a.e. } t \in]0, T[,$$
$$\ell\big(x^*(0), x^*(t)\big) + m\big(p^*(0), p^*(T)\big) = \big(x^*(0), p^*(0)\big) - \big(x^*(T), p^*(T)\big),$$

and

$$G_0(\rho) = \int_0^T (\mathrm{d}\rho_s, x^*).$$

But, as remarked before, these equations are equivalent to (4.20), (4.21), and (4.19), where $\rho(t) = w(t) + \int_0^T q(s)\,\mathrm{d}s$. Thus, the proof of Theorem 4.16 is complete. □

Remark 4.17 Of course, the duality Theorem 4.16 remains true under the conditions of Theorem 4.11.

Remark 4.18 Let us denote by G the functional which occurs in problem (P*). It is obvious that one has always the inequality

$$\inf F(x, u) \geq -\inf G(\rho).$$

The basic question in the duality theory already discussed in Sect. 3.2.1 is whether the equality holds in the above relation. Within the terminology introduced in Sect. 3.2.1 (see Definition 3.40), Theorem 4.16 amounts to saying that problem (P) is *normal*. For finite-dimensional control problems of the form (P), it turns out (Rockafellar [40, 41]) that, if no state constraints are present (that is, $K = E$), then under the assumptions of Theorem 4.5 one has

$$\inf F(x, u) = -\min G(\rho).$$

Along these lines, a sharper duality result has been obtained by the same author [44] by formulating the primal problem (P), as well as the dual (P*), in the space of functions of bounded variation on $[0, T]$.

4.1.9 Some Examples

The following illustrates the kind of problem to which the results of the previous sections can be applied. The examples we have chosen can be extended in several directions, but we do not attempt here maximum generality nor claim to be comprehensive in any sense.

Example 4.19 Consider an optimal control problem of the following type.

$$\text{Minimize } \int_0^T L_0\big(x(t), u(t)\big)\,\mathrm{d}t + \ell_0\big(x(0), x(T)\big)$$

$$\text{in } x \in C\big([0,T]; E\big) \text{ and } u \in L^p(0,T;U) \tag{4.105}$$

$$\text{subject to } \quad x' = A(t)x + B(t)u(t), \quad 0 \le t \le T, \tag{4.106}$$

$$x(0) \in X_0, \qquad x(T) \in X_1, \tag{4.107}$$

$$u(t) \in U_0 \quad \text{a.e. } t \in]0,T[, \tag{4.108}$$

$$x(t) \in K \quad \text{for every } t \in [0,T], \tag{4.109}$$

where L_0 and ℓ_0 are lower-semicontinuous, everywhere defined convex functions on $E \times U$ and $E \times E$, respectively, U_0 is a nonempty, closed convex subset of U and X_0, X_1, K are nonempty closed convex subsets of E. (In particular, X_0 or X_1 may consist of a single point or all of E.) The operators $A(t): E \to E$ and $B(t): U \to E$ $(0 \le t \le T)$ are assumed to satisfy Hypotheses (A) and (B), respectively. The spaces E and U are strictly convex and separable along with their duals.

To formulate this as a problem of type (P), we define

$$L(x,u) = \begin{cases} L_0(x,u), & \text{if } u \in U_0, \\ +\infty, & \text{otherwise,} \end{cases}$$

and

$$\ell(x_1,x_2) = \begin{cases} \ell_0(x_1,x_2), & \text{if } x_1 \in X_0 \text{ and } x_2 \in X_1, \\ +\infty, & \text{otherwise.} \end{cases}$$

In this way, the given optimal control problem is equivalent to minimizing

$$\int_0^T L(x,u)\,\mathrm{d}t + \ell\big(x(0), x(T)\big)$$

over all $x \in C([0,T];E)$ and $u \in L^p(0,T;U)$ subject to the state constraints (4.106)–(4.109).

As pointed out earlier (see (4.11)), Assumption (C) holds if

$$\lim_{\|u\| \to \infty} \frac{L(x,u)}{\|u\|} = +\infty.$$

(If U_0 is bounded, this is trivially satisfied.)

It may be noted that $\mathrm{Dom}(\ell) = X_0 \times X_1$. Thus, Assumption (D) requires the existence of a pair $(x,u) \in C([0,T];E) \times L^1(0,T;U)$ satisfying (4.106)–(4.109) and such that $L_0(x,u) \in L^1(0,T)$, $x(t) \in \operatorname{int} K$ for every $t \in [0,T]$. It is possible to develop general explicit conditions guaranteeing that such a pair (x,u) exists. For brevity, our attention is confined to the unconstrained control case $U_0 = U = E$, $B(t) = I$ and $A(t) = A$ time-independent and dissipative on E. Assume further that $L_0(y,v) \in L^1(0,T)$ for all $y, v \in C([0,T];E)$ and

$$\mathrm{e}^{At} \operatorname{int} K \subset \operatorname{int} K \quad \text{for all } t > 0.$$

Then Hypothesis (D) is implied by the following more easily verified one.

There exist $x_0 \in X_0 \cap \operatorname{int} K$ *such that* $\mathrm{e}^{AT} x_0 \in X_1 \cap \operatorname{int} K$.

Here is the argument. Define

$$x(t) = \mathrm{e}^{At}\left(\frac{t}{T}\mathrm{e}^{AT}x_0 + \left(1 - \frac{t}{T}\right)x_0\right), \quad t \in (0,T).$$

Clearly, $x(t) \in \operatorname{int} K$ for $t \in [0,T]$ and

$$x'(t) = Ax(t) + u(t) \quad \text{for } t \in [0,T],$$

where $u \in C([0,T];E)$. Moreover, $x(0) \in X_0$, $x(T) \in X_1$ and $L_0(x,u) \in L^1(0,T)$.

As regards Assumption (E), it requires in this case the existence of a feasible function $y \in C([0,T];E)$ which satisfies (4.106)–(4.109) and

$$y(T) \in \operatorname{int} X_1 \tag{4.110}$$

or

$$y(T) \in \operatorname{int}\bigl\{h \in E;\ \bigl(y(0), h\bigr) \in K_L\bigr\}, \tag{4.110$'$}$$

where K_L is the set of all attainable pairs for the given optimal control problem.

If $X_1 = E$, then (4.110) automatically holds as long as at least one feasible arc exists. If $U_0 = U = E$, $B(t) = I$ and $L_0(y,v) \in L^1(0,T)$ for all $(y,v) \in C([0,T];E) \times L^p(0,T;E)$, then K_L is just the set of all end-point pairs $(y(0), y(T))$ arising from arcs $y \in C([0,T];E)$ which satisfy the condition

$$y(t) \in K \quad \text{for } t \in [0,T]. \tag{4.111}$$

Hence, as may readily be seen, Hypothesis (E) is satisfied by assuming that $D(A) \cap X_0 \cap K \neq \emptyset$, $D(A) \cap (\operatorname{int} X_1) \cap K \neq \emptyset$. In fact, $y(t) = (1 - \frac{t}{T})y_0 + \frac{t}{T} y_1$, where $y_0 \in D(A) \cap X_0 \cap K$ and $y_1 \in D(A) \cap (\operatorname{int} X_1) \cap K$, satisfies conditions (4.110) and (4.111).

According to the rule of additivity of the subdifferentials (see Corollary 2.63), (4.20) and (4.21) can be written in this case

$$\begin{aligned} &q(t) \in \partial_x L_0\bigl(x(t), u(t)\bigr) \quad \text{a.e. } \in]0,T[,\\ &B^*(t)p(t) \in \partial_u L_0\bigl(x(t), u(t)\bigr) + N\bigl(u(t), U_0\bigr) \quad \text{a.e. } t \in]0,T[, \end{aligned}$$

and

$$\big(p(0), -p(T)\big) \in \partial \ell_0\big(x(0), x(T)\big) + \big(N_0\big(x(0)\big), N_1\big(x(T)\big)\big),$$

where $N(u, U_0)$, $N_0(x(0))$ and $N_1(x(T))$ are the cones of normals at u, $x(0)$ and $x(T)$ to U_0, X_0 and X_1, respectively.

To calculate the dual problem, we assume for simplicity that $\ell_0 \equiv 0$. Then, we have

$$m(p_1, p_2) = \ell^*(p_1, -p_2) = H_0(p_1) + H_1(-p_2),$$

where H_0 and H_1 are the support functions of X_0 and X_1, respectively. Next, we have

$$M(p, q) = \sup\big\{\langle p, u\rangle + (q, x) - L_0(x, u);\ x \in K,\ u \in U_0\big\}$$

and, since $D(L_0) = E \times U$, we may use the Fenchel theorem (Theorem 3.54) to get

$$M(p, q) = \inf\big\{L_0(q - \tilde{q}, p - \tilde{p}) + H_{U_0}(\tilde{p}) + H_K(\tilde{q});\ \tilde{p} \in U^*,\ \tilde{q} \in E^*\big\},$$

where H_{U_0} and H_K are the support functions of U_0 and K, respectively.

Given a closed and convex subset K_0 of E, consider the following problem: *find the control function u subject to constraints* (4.108) *such that $x(t, x_0, u) \in K_0$ for all $t \in [0, T]$*. Here, $x(t, x_0, u)$ is the solution to (4.106) with initial value condition $(x(0) = x_0$. The least square approach to this controllability problem leads to an optimal control problem of the form (4.105)–(4.109), where $X_0 = \{x_0\}$, $X_1 = E$, $\ell_0 \equiv 0$, $K = E$, and

$$L_0(x, u) = \begin{cases} \alpha d^2(x, K_0), & \text{if } u \in U_0, \\ +\infty, & \text{otherwise.} \end{cases} \tag{4.112}$$

Here, $\alpha > 0$ and $d^2(x, K_0)$ is the distance from x to K_0.

The ill-posed problems associated to (4.106) represent another important source of optimal problems of the form (4.105)–(4.109).

Remark 4.20 If K is a closed convex cone of E, $x(0) \in K$, $(I - \varepsilon A)^{-1} K \subset K$ for all $\varepsilon > 0$ and $B(t)u(t) \in K$ a.e. $t \in]0, T[$, for all the control functions u which satisfy condition (4.109), then $x(t) \in K$ for all $t \in [0, T]$ and therefore the state constraints (4.109) become redundant.

Example 4.21 Consider the following distributed optimal control problem.

$$\text{Minimize } \int_Q g\big(y(t,x),u(t,x)\big)\,\mathrm{d}t\,\mathrm{d}x \quad \text{in } y \in L^2\big(0,T;H_0^1(\Omega)\big) \text{ and } u \in L^2(Q),$$

subject to the linear diffusion process described by the heat equation

$$\begin{aligned}
&y_t - \Delta y = u \quad \text{in } Q =]0,T[\times\Omega,\\
&y(t,x) = 0 \quad \text{in } \Sigma =]0,T[\times\Gamma,\\
&\text{with the constraints}\\
&y(0,x) \ge 0 \quad \text{a.e. } x \in \Omega, \qquad \big|u(t,x)\big| \le 1 \quad \text{a.e. } (t,x) \in Q.
\end{aligned} \tag{4.113}$$

Problems of this type are encountered in the control of industrial heating processes in the presence of internal heat sources (see the book of Butkovskiy [20] for such examples).

We assume that the function $g:\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ is convex and everywhere finite. Further, we assume that there exists $v_0 \in L^2(\Omega)$ such that $|v_0(x)| \le 1$ a.e. $x \in \Omega$ and $g(y,v_0) \in L^1(\Omega)$ for every $y \in L^2(\Omega)$. As seen in Example 4.19, this implies that the function $L : L^2(\Omega)\times L^2(\Omega) \to \overline{\mathbb{R}}^*$ defined by

$$L(y,v) = \begin{cases} \int_\Omega g(y,v)\,\mathrm{d}x, & \text{if } |v(x)| \le 1 \text{ a.e. } x \in \Omega,\\ +\infty, & \text{otherwise} \end{cases}$$

satisfies assumptions (C).

We place ourselves in the framework of Example 4.19 (problem (4.105)–(4.109)), where $E = U = L^2(\Omega)$, $B(t) = I$, $K = E$, $U_0 = \{u \in L^2(\Omega); |u(x)| \le 1$ a.e. $x \in \Omega\}$, $X_1 = L^2(\Omega)$, $X_0 = \{y \in L^2(\Omega); y(x) \ge 0$ a.e. $x \in \Omega\}$ and $A(t) = \Delta$ with $D(A(t)) = H_0^1(\Omega)\cap H^2(\Omega)$. As a matter of fact, we are in the situation presented at the end of Sect. 4.1.3, where $V = H_0^1(\Omega)$ and $A(t) = \Delta$.

It is elementary that Assumptions (A), (B), (D) and (E) are satisfied.

Let $\tilde{g} : \mathbb{R}\times\mathbb{R} \to \overline{\mathbb{R}}$ be the extended real-valued function

$$\tilde{g}(y,v) = \begin{cases} g(y,v), & \text{if } |v| \le 1,\\ +\infty, & \text{otherwise.} \end{cases}$$

By Proposition 2.55, we see that

$$\big(\partial L(y,v)\big)(x) = \partial\tilde{g}\big(y(x),v(x)\big) \quad \text{a.e. } x \in \Omega,$$

whereas

$$\partial\tilde{g}(y,v) = \partial g(y,v) + \big\{(0,\lambda v);\ \lambda \ge 0,\ \lambda(1-|v|) = 0\big\}.$$

Next, observe that the cone $N_0(y)$ of normals to X_0 at y is given by

$$N_0(y) = \big\{w \in L^2(\Omega);\ w(x) \le 0,\ w(x)y(x) = 0 \text{ a.e. } x \in \Omega\big\}.$$

Then, by Theorem 4.5 (we give the extremality system in the form (4.23)), *the pair* $y \in L^2(0,T;H_0^1(\Omega))$, $u \in L^2(Q)$ *is optimal in problem* (4.113) *if and only if there exist* $p \in L^2(0,T;H_0^1(\Omega))$ *with* $p_t \in L^2(0,T;H^{-1}(\Omega))$ *and* $\lambda : Q \to \mathbb{R}^+$ *such that*

$$\begin{aligned}
&p_t + \Delta p \in \partial_y g(y,u) \quad \text{on } Q,\\
&p - \lambda u \in \partial_u g(y,u) \quad \text{a.e. on } Q,\\
&\lambda\bigl(1-|u|\bigr) = 0, \qquad |u| \le 1 \quad \text{a.e. on } Q,\\
&y(0,x) \ge 0, \qquad p(0,x) \le 0, \qquad y(0,x)p(0,x) = 0 \quad \text{a.e. } x \in \Omega,\\
&p(T,x) = 0 \quad \text{a.e. } x \in \Omega.
\end{aligned}$$

The dual problem to (4.113) is that of minimizing

$$\int_0^T \tilde{g}^*\bigl(w(t,x), p(t,x)\bigr)\,\mathrm{d}t\,\mathrm{d}x$$

over all $p \in L^2(0,T;H_0^1(\Omega))$ and $w \in L^2(Q)$, subject to the constraints

$$\begin{aligned}
&p_t(t,x) + \Delta p(t,x) = w(t,x) \quad \text{on } Q,\\
&p(0,x) \le 0, \qquad p(T,x) = 0 \quad \text{a.e. on } \Omega.
\end{aligned}$$

To be more specific, let us suppose that $g(y,v) = \alpha|y| + |v|$ for all y and v in $\mathbb{R}$, where α is a nonnegative constant. Then, as is easily verified,

$$g^*(q,p) = \begin{cases} 0, & \text{if } |q| \le \alpha \text{ and } |p| \le 1,\\ +\infty, & \text{if } |q| > \alpha \text{ or } |p| > 1,\end{cases}$$

and, therefore,

$$\tilde{g}^*(q,p) = \begin{cases} \max(|p|-1,0), & \text{if } |q| \le \alpha,\\ +\infty, & \text{if } |q| > \alpha.\end{cases}$$

Thus, in this case, the dual problem becomes

$$\text{Minimize } \int_0^T \int_\Omega \max\bigl\{\bigl|p(t,x)\bigr| - 1, 0\bigr\}\,\mathrm{d}t\,\mathrm{d}x$$

in $p \in L^2(0,T;H_0^1(\Omega))$, subject to the constraints

$$\begin{aligned}
&\bigl|p_t(t,x) + \Delta p(t,x)\bigr| \le \alpha \quad \text{on } Q,\\
&p(0,x) \le 0, \qquad p(T,x) = 0 \quad \text{a.e. } x \in \Omega.
\end{aligned}$$

Example 4.22 In practice, it is usually impossible to apply the control action $u(t,x)$ in order to influence the state $y(t,x)$ of equation at each point of the spatial domain Ω. Usually, what can be expected is that the control can be applied at isolated points within the spatial domain. An important feature of this case is that the control space U is finite-dimensional. As an example, let us briefly discuss the following variant of the problem presented in Example 4.21.

$$\text{Minimize} \int_0^T \int_\Omega g\big(y(t,x),u(t)\big)\,dt\,dx \quad \text{over all } y \in L^2\big(0,T;H_0^1(\Omega)\big)$$

$$\text{and } u=(u_1,\dots,u_N)\in L^2\big(0,T;\mathbb{R}^N\big) \quad \text{subject to the constraints}$$

$$y_t - \Delta y = \sum_{j=1}^N u_j(t)\chi_j(x) \quad \text{on } Q, \qquad y(t,x)=0 \quad \text{on } \Sigma,$$

$$y(0,x)=y_0(x) \quad \text{on } \Omega,$$

$$\big|u_i(t)\big| \le 1 \quad \text{a.e. on }]0,T[,\ i=1,\dots,N.$$

Here, the function g is finite and convex on $\mathbb{R}^{N+1}$ and $\{\chi_j\}_{j=1}^N$ are the characteristic functions of a family of disjoint measurable subsets Ω_j which cover the domain Ω. In this case, the control is provided by N heat sources. As already mentioned, this problem may be written as a problem of the type (P), where $E=L^2(\Omega)$, $U=\mathbb{R}^N$, $K=L^2(\Omega)$, $X_1=L^2(\Omega)$, $X_0=\{y_0\}$, $A(t)=\Delta$ and $B:\mathbb{R}^N\to L^2(\Omega)$ is defined by

$$(Bu)(x)=\sum_{j=1}^N u_i\chi_j(x) \quad \text{a.e. } x\in\Omega.$$

Noticing that $B^*v=\{\int_\Omega \chi_j(x)v(x)\,dx\}_{j=1}^N$, we leave to the reader the calculation of the optimality system in this case.

Remark 4.23 In Examples 4.21 and 4.22, we may consider more general functions $g:\Omega\times R\times U\to\,]-\infty,+\infty]$, which are measurable in $x\in\Omega$, convex and continuous in $(y,u)\in\mathbb{R}\times\mathbb{R}^N$ and such that $g(x,y,v_0)\in L^1(\Omega)$ for all $y\in L^2(\Omega)$ and some v_0 in the control constraint set.

Example 4.24 We consider the following problem.

$$\text{Minimize} \int_Q g\big(y(t,x),u(t,x)\big)\,dx\,dt + \varphi_0\big(y(T)\big)$$

$$\begin{aligned} \text{subject to} \quad & y\in L^2\big(0,T;H_0^1(\Omega)\big), \qquad y_t\in L^2\big(0,T;H^{-1}(\Omega)\big),\\ & u\in L^2(Q) \quad \text{and} \quad y_t-\Delta y=u \quad \text{in } Q, \qquad y=0 \quad \text{in } \Sigma, \\ & y(0,x)=y_0(x), \quad x\in\Omega,\\ & \int_\Omega F\big(x,\nabla_x y(t,x)\big)\,dx \le 0, \quad \forall t\in[0,T]. \end{aligned} \tag{4.114}$$

Here, $\varphi_0 : H_0^1(\Omega) \to \overline{\mathbb{R}}^*$, $g : \mathbb{R} \times \mathbb{R} \to \overline{\mathbb{R}}^*$ are lower-semicontinuous convex functions and $F : \Omega \times \mathbb{R}^n \to \mathbb{R}$ is a normal convex integrand having the property that $F(x, z) \in L^1(\Omega)$ for every $z \in (L^2(\Omega))^n$. In particular, this implies that the function $\psi : H_0^1(\Omega) \to \mathbb{R}$,

$$\psi(y) = \int_\Omega F\big(x, \nabla y(x)\big)\,\mathrm{d}x$$

is convex and continuous.

Problem (4.114) can be written in the form (P), where

$$E = H_0^1(\Omega), \qquad U = L^2(\Omega) \quad \text{and} \quad K = \big\{y \in H_0^1(\Omega);\ \psi(y) \le 0\big\},$$

$$L(y, u) = \int_\Omega g\big(y(x), u(x)\big)\,\mathrm{d}x, \quad \forall y \in H_0^1(\Omega),\ u \in L^2(\Omega),$$

$$\ell(y_1, y_2) = \begin{cases} \varphi_0(y_2), & \text{if } y_1 = y_0, \\ +\infty, & \text{otherwise.} \end{cases}$$

Assumptions (A), (B), (C) and (D) are obviously satisfied if we impose the following two conditions.

(a) There is $u^0 \in L^2(\Omega)$ such that $g(y, u^0) \in L^1(\Omega)$ for every $y \in H_0^1(\Omega)$.
(b) There is at least one feasible function $\tilde{y}$ such that $\tilde{y}(0) = y_0$, φ_0 is bounded on a neighborhood of $\tilde{y}(T) \in H_0^1(\Omega)$ and $\psi(\tilde{y}(t)) < 0$ for $t \in [0, T]$.

Then, according to Theorem 4.5, the pair (y, u) is optimal in problem (4.114) if and only if there exist the functions $q \in L^1(0, T; H^{-1}(\Omega))$, $q_0 \in H^{-1}(\Omega)$, $w \in BV([0, T]; H^{-1}(\Omega))$, $p \in L^2(Q) \cap C([0, T]; H^{-1}(\Omega)) + BV([0, T]; H^{-1}(\Omega))$ satisfying the system

$$\begin{aligned}
&p_t + \Delta p = q + \mathrm{d}w \quad \text{in } [0, T],\\
&p(T - 0) = q_0 + w(T - 0) - w(T),\\
&\big(q(t, x), p(t, x)\big) \in \partial g\big(y(t, x), u(t, x)\big) \quad \text{a.e. } (t, x) \in Q,\\
&q_0 \in -\partial\varphi_0\big(y(T)\big).
\end{aligned}$$

Taking into account the special form of the set

$$\mathscr{K} = \big\{\tilde{y} \in C\big([0, T]; H_0^1(\Omega)\big), \psi\big(y(t)\big) \le 0,\ \forall t \in [0, T]\big\},$$

we see that the measure $\mathrm{d}w \in M([0, T]; H^{-1}(\Omega))$ can be expressed as

$$\mathrm{d}w = \big\{\lambda\partial\Phi(y),\ \lambda \ge 0,\ \lambda\Phi(y) = 0\big\},$$

where $\Phi(y) = \sup\{\psi(y(t)); t \in [0, T]\}$. We may use this formula to express $\mathrm{d}w$ in terms of the gradient of $r(x, \cdot)$ as in the work [35] by Mackenroth.

4.1.10 The Optimal Control Problem in a Duality Pair $V \subset H \subset V'$

Consider problem (P) in the special case $X = V$, $X' = V'$, where V is a reflexive Banach space such that $V \subset H \subset V'$ algebraically and topologically. Here, H is a real Hilbert space with the norm $|\cdot|$ and scalar product $(\cdot,\cdot)$. The norm of V is denoted $\|\cdot\|_V$ and V' is the dual of V with the norm $\|\cdot\|_{V'}$. The duality ${}_{V'}(\cdot,\cdot)_V$, which coincides with the scalar product $(\cdot,\cdot)$ of H on $H \times H$, is again denoted by $(\cdot,\cdot)$. Assume also that V and V' are strictly convex. The family of the operators $\{A(t); 0 \le t \le T\}$ is assumed to satisfy Assumptions (j), (jj) and (jjj) of Proposition 1.149. For simplicity, we take here $L : (0,T) \times V \times U \to \overline{\mathbb{R}}^*$ and $\ell : H \times H \to \overline{\mathbb{R}}^*$ of the form

$$L(t,y,u) = g(t,y) + h(t,u), \quad \forall y \in V,\ u \in U,$$
$$\ell(y_1,y_2) = \varphi_0(y_2) + I_{\{y_0\}}(y_1), \quad \forall y_1, y_2 \in H,$$

where $g : (0,T) \times V \to \overline{\mathbb{R}}$, $h : (0,T) \times U \to R$, and $\varphi_0 : H \to \mathbb{R}$ satisfy the following conditions.

(l) $g(t,\cdot)$ is convex and continuous on V, measurable in t for every $y \in V$ and $\alpha_1 \le g(t,y) \le \alpha_2 \|y\|_V^2 + \alpha_3$, $\forall y \in V$, for $\alpha_i \in \mathbb{R}$, $i = 1,2,3$.
(ll) $h(t,\cdot)$ is convex and lower-semicontinuous on U and measurable in t. There are $p \ge 2$ and $\tilde{\alpha}_2 > 0$, $\tilde{\alpha}_3 \in \mathbb{R}$ such that $h(t,u) \ge \tilde{\alpha}_2 \|u\|_U^2 + \tilde{\alpha}_3$, $\forall u \in U$, $t \in (0,T)$.
(lll) φ_0 is convex and continuous on H.
(llll) $B(t) \in L(U,V')$ a.e. $t \in (0,T)$ and $\|B(t)\|_{L(U,V')} \in L^\infty(0,T)$. Here, U is a real Hilbert space with the norm $\|\cdot\|_U$ and the scalar product $(\cdot,\cdot)_U$.

Then, problem (P) is, in this case, of the form

$$(\mathrm{P}_0) \quad \mathrm{Min}\left\{ \int_0^T \big(g\big(t,y(t)\big) + h\big(t,u(t)\big)\big)\,\mathrm{d}t + \varphi_0\big(y(T)\big) \right\}$$
$$\text{subject to} \quad y' = A(t)y + B(t)u + f(t), \quad t \in (0,T), \qquad y(0) = y_0.$$

Theorem 4.25 *Assume that conditions* (l)–(llll) *hold and that* $f \in L^2(0,T;V')$, $y_0 \in H$. *Then, there is at least one optimal pair* $(y^*,u^*) \in (C([0,T];H) \cap L^2(0,T;V)) \times L^2(0,T;U)$ *in problem* (P_0). *Moreover, any such a pair is a solution to the Euler–Lagrange system*

$$(y^*)' = A(t)y^* + B(t)u^* + f \quad \text{a.e. } t \in (0,T), \tag{4.115}$$
$$p' = -A^*(t)p + \eta(t) \quad \text{a.e. } t \in (0,T),$$
$$y^*(0) = y_0, \quad p(T) \in -\partial\varphi_0\big(y^*(T)\big), \tag{4.116}$$
$$u^*(t) \in \partial h\big(t, B^*p(t)\big) \quad \text{a.e. } t \in (0,T), \tag{4.117}$$

where $\eta \in L^2(0,T;V')$ *and*

$$\eta(t) \in \partial g\big(t, y^*(t)\big) \quad a.e.\ t \in (0, T). \tag{4.118}$$

Equations (4.115)–(4.118) *are also sufficient for optimality.*

Proof *Existence of an optimal pair* follows by a standard device described in Proposition 4.1 and so it is outlined only. Namely, consider a sequence (y_n, u_n) satisfying the equation

$$\begin{aligned} &y_n' = A(t)y_n + B(t)u_n + f \quad \text{a.e. } t \in (0, T), \\ &y_n(0) = y_0, \end{aligned} \tag{4.119}$$

and such that

$$d \le \int_0^T \big(g(t, y_n) + h(t, u_n)\big)\, \mathrm{d}t + \varphi_0\big(y_n(T)\big) \le d + \frac{1}{n}, \tag{4.120}$$

where d is the infimum in problem (P). By estimates (l), (ll) and (lll), we see that $\{u_n\}$ is bounded in $L^p(0, T; U)$ and, by (4.119), it follows that

$$\|y_n\|^2_{L^2(0,T;V)} + \|y_n'\|_{L^2(0,T;V')} \le C.$$

Hence, on a subsequence, again denoted n, we have

$$\begin{aligned} u_n &\to u^* \quad \text{weakly in } L^2(0, T; U), \\ y_n &\to y^* \quad \text{weakly in } L^2(0, T; V), \\ y_n' &\to (y^*)' \quad \text{weakly in } L^2(0, T; V'), \\ y_n(T) &\to y^*(T) \quad \text{weakly in } H, \end{aligned}$$

because, by (4.119) and by assumptions (j)–(jjj) of Proposition 1.149,

$$\begin{aligned} \frac{1}{2}\big(\big|y_n(t)\big|^2 - |y_0|^2\big) &= \int_0^t \big(A(s)y_n, y_n\big)\, \mathrm{d}s + \int_0^t \big(B(s)u_n, y_n\big)\, \mathrm{d}s \\ &\quad + \int_0^t \big(f(s), y_n(s)\big)\, \mathrm{d}s \\ &\le C \int_0^t \|u_n\|_U \|y_n\|_V\, \mathrm{d}s - \omega \int_0^t \big\|y_n(s)\big\|_V^2\, \mathrm{d}s + \alpha \int_0^t |y_n|^2\, \mathrm{d}s. \end{aligned}$$

Since, as seen earlier, the convex functions $y \to \int_0^T \varphi(t, y)\, \mathrm{d}t$, and $u \to \int_0^T h(t, u)\, \mathrm{d}t$ are convex and lower-semicontinuous in $L^2(0, T; V)$ and $L^2(0, T; U)$, respectively, they are weakly lower-semicontinuous and, therefore,

$$\begin{aligned} \liminf_{n\to\infty} \int_0^T g(t, y_n)\, \mathrm{d}t &\ge \int_0^T g(t, y^*)\, \mathrm{d}t, \\ \liminf_{n\to\infty} \int_0^T h(t, u_n)\, \mathrm{d}t &\ge \int_0^T h(t, u^*)\, \mathrm{d}t. \end{aligned}$$

Similarly,

$$\liminf_{n\to\infty} \int_0^T g\big(y_n(T)\big) \geq g\big(y^*(T)\big)\,\mathrm{d}t.$$

Then, by (4.120), we see that (y^*, u^*) is optimal in problem (P).

Necessary conditions of optimality. Let (y^*, u^*) be optimal in problem (P). Then consider, as in the proof of Theorem 4.5, the approximating optimal control problem

$$\begin{aligned} \mathrm{Min}\bigg\{ &\int_0^T \bigg(g_\lambda(t,y) + h_\lambda(t,u) + \frac{1}{2}\,\|u-u^*\|_U^2\bigg)\,\mathrm{d}t + (\varphi_0)_\lambda\big(y(T)\big); \\ &y' = A(t) + B(t)u + f(t),\ \ y(0) = y_0\bigg\}, \end{aligned} \tag{4.121}$$

where $(\varphi_0)_\lambda$ and $h_\lambda(t,\cdot)$, g_λ are, as in the previous cases, regularizations of φ_0, $h(t,\cdot)$ and g, respectively, that is,

$$g_\lambda(t,y) = \inf\bigg\{\frac{1}{2\lambda}\|y-z\|_V^2 + g(t,z);\ z\in V\bigg\},$$

$$h_\lambda(t,u) = \inf\bigg\{\frac{1}{2\lambda}\|u-v\|_U^2 + h(t,v);\ v\in U\bigg\},$$

$$(\varphi_0)_\lambda(y) = \inf\bigg\{\frac{1}{2\lambda}|y-z|^2 + \varphi_0(z);\ z\in H\bigg\}.$$

As seen earlier, problem (4.121) has at least one optimal pair (y_λ, u_λ) and, arguing as in the proof of Theorem 4.5, it follows that (recall that $g_\lambda(t,\cdot)$, $h_\lambda(t,\cdot)$, $g_\lambda(t,\cdot)$ are Gâteaux differentiable)

$$\begin{aligned} &\int_0^T \big(\nabla g_\lambda\big(t,y_\lambda(t)\big), z(t)\big) + \big(\nabla h_\lambda\big(t,u_\lambda(t)\big), v(t)\big) \\ &\quad + \big(\big(u_\lambda(t)-u^*(t), v(t)\big)_U\big)\,\mathrm{d}t + \big(\nabla(\varphi_0)_\lambda\big(y_\lambda(T)\big), z(T)\big) \geq 0, \end{aligned} \tag{4.122}$$

for all $v\in L^2(0,T;U)$ and $z\in C([0,T];H)\cap L^2(0,T;V)$ solution to the system in variations

$$z' = A^*(t)z + B(t)v, \quad t\in(0,T); \qquad z(0) = 0.$$

Then, if we consider $p_\lambda \in C([0,T];H)\cap L^2(0,T;V)$, $\frac{\mathrm{d}p_\lambda}{\mathrm{d}t}\in L^2(0,T;V')$ the solution to the backward dual system

$$\begin{aligned} p_\lambda' &= -A^*(t)p_\lambda + \nabla g_\lambda(t,y_\lambda) \quad \text{a.e. } t\in(0,T), \\ p_\lambda(T) &= -\nabla(\varphi_0)_\lambda\big(y_\lambda(T)\big), \end{aligned} \tag{4.123}$$

which exists in virtue of assumptions (l) $\sim$ (lll), we get by (4.122) that

$$\int_0^T \big(\big(\nabla h_\lambda\big(t,u_\lambda(t)\big) + \big(u_\lambda(t)-u^*(t)\big)\big) - B^*(t)p_\lambda(t)\big)v(t)\,\mathrm{d}t = 0,$$

for all $v \in L^2(0,T;U)$. Hence,

$$B^*(t)p_\lambda(t) = \nabla h_\lambda\big(t, u_\lambda(t)\big) + \big(u_\lambda(t) - u^*(t)\big) \quad \text{a.e. } t \in (0,T). \tag{4.124}$$

On the other hand, as seen in the proof of Theorem 4.5, we have for $\lambda \to 0$

$$u_\lambda \to u^* \quad \text{strongly in } L^2(0,T;U) \tag{4.125}$$

and, therefore,

$$\begin{aligned} y_\lambda &\to y^* \quad \text{strongly in } L^2(0,T;V) \cap C\big([0,T];H\big), \\ \frac{dy_\lambda}{dt} &\to \frac{dy^*}{dt} \quad \text{strongly in } L^2(0,T;V'). \end{aligned} \tag{4.126}$$

Recalling that (Theorem 2.58)

$$\begin{aligned} &\frac{1}{2\lambda}\big|(I+\lambda\varphi_0)^{-1}y_\lambda(T) - y_\lambda(T)\big|^2 + \varphi_0\big((I+\lambda\partial\varphi_0)^{-1}y_\lambda(T)\big) = (\varphi_0)_\lambda\big(y_\lambda(T)\big), \\ &\frac{1}{2\lambda}\big\|\big(\Lambda+\lambda\partial g(t)\big)^{-1}y_\lambda(t) - y_\lambda(t)\big\|_V^2 + g\big(t, \big(\Lambda+\lambda\partial g(t)\big)^{-1}y_\lambda(t)\big) = g_\lambda\big(t, y_\lambda(t)\big), \end{aligned}$$

where $\Lambda : V \to V'$ is the duality mapping of V, by (4.126) we infer that, for $\lambda \to 0$,

$$\big((I+\lambda\partial\varphi_0)^{-1}y_\lambda(T)\big) \to y^*(T) \quad \text{strongly in } H, \tag{4.127}$$

$$\big(\Lambda+\lambda\partial g(t)\big)^{-1}y_\lambda \to y^* \quad \text{strongly in } L^2(0,T;V). \tag{4.128}$$

Recalling that φ_0 is continuous on H and

$$\nabla(\varphi_0)_\lambda\big(y_\lambda(T)\big) \in \partial\varphi_0\big((I+\lambda\partial\varphi_0)^{-1}y_\lambda(T)\big),$$

it follows, by (4.128), that $\{\nabla(\varphi_0)_\lambda(y_\lambda(T))\}$ is bounded in H and so, on a subsequence, again denoted $\{\lambda\}$,

$$\nabla(\varphi_0)_\lambda\big(y_\lambda(T)\big) \to \xi \in \partial\varphi_0\big(y^*(T)\big) \quad \text{weakly in } H. \tag{4.129}$$

On the other hand, by the inequality

$$\big(\nabla g_\lambda\big(t, y_\lambda(t)\big), y_\lambda(t) - \theta\big) \ge g_\lambda\big(t, y_\lambda(t)\big) - g_\lambda(t,\theta), \quad \forall \theta \in V,\ t \in (0,T),$$

we get for $\|\theta\|_V = \rho\|\nabla g_\lambda(t, y_\lambda(t))\|_{V'}$

$$\begin{aligned} \rho\big\|\nabla g_\lambda\big(t, y_\lambda(t)\big)\big\|_{V'} &\le g_\lambda(t,\theta) + \big(\nabla g_\lambda\big(t, y_\lambda(t)\big), y_\lambda(t)\big) \\ &\le C\rho^2\big\|\nabla g_\lambda\big(t, y_\lambda(t)\big)\big\|_{V'} + \big\|\nabla g_\lambda\big(t, y_\lambda(t)\big)\big\|_{V'}\big\|y_\lambda(t)\big\|_V. \end{aligned}$$

This yields

$$\big\|\nabla g_\lambda(t, y_\lambda)\big\|_{V'}^2 \le C_1\|y_\lambda\|_V^2 + C_2$$

and, therefore, by (4.126) and (4.128), we have

$$\nabla g_\lambda(t, y_\lambda) \to \eta \quad \text{weakly in } L^2\big(0, T; V'\big), \qquad \eta(t) \in \partial g\big(t, y(t)\big), \quad \text{a.e. } t \in (0, T).$$

Then letting λ tend to zero in (4.123) and (4.124) we obtain (4.115)–(4.118), as claimed. The fact that (4.115)–(4.118) are sufficient for optimality is immediate, and therefore we omit the proof. □

The Dual Problem We associate with (P_0) the dual problem (see Problem (P^*) in Sect. 4.1.7)

$$(\mathrm{P}_0^*) \quad \mathrm{Min}\bigg\{\int_0^T \big(h^*\big(t, B^*(t)p\big) + g^*(t, v) + \big(f(t), p\big)\big)\,\mathrm{d}t + \big(y_0, p(0)\big)$$
$$+ \varphi_0^*\big(-p(T)\big);\ p' = -A^*(t)p + v\bigg\},$$

where $g^*(t, \cdot) : V' \to \overline{\mathbb{R}}$, $h^*(t, \cdot) : U \to \overline{\mathbb{R}}$ and $\varphi_0^* : H \to \overline{\mathbb{R}}$ are the conjugates of $g(t, \cdot)$, $h(t, \cdot)$ and φ_0, respectively.

We have the following theorem.

Theorem 4.26 *Under assumptions* (l)–(llll), (y^*, u^*) *is optimal in* (P_0) *if and only if Problem* (P_0^*) *has a solution* (p^*, v^*) *and*

$$\min(\mathrm{P}_0) + \min(\mathrm{P}_0^*) = 0.$$

The proof is identical with that of Theorem 4.16 and therefore it will be omitted.

Example 4.27 Theorems 4.25 and 4.26 can be applied neatly to an optimal control problem of the form

$$\text{Minimize} \int_Q g_0(t, y)\,\mathrm{d}x\,\mathrm{d}t + \int_\Sigma h(t, u)\,\mathrm{d}t\,\mathrm{d}x + \int_\Omega \varphi_0\big(y(T, x)\big)\,\mathrm{d}x, \quad u \in L^2(\Sigma)$$

$$\text{subject to} \quad \frac{\partial y}{\partial t} - \Delta y = f \quad \text{in } Q = (0, T) \times \Omega,$$
$$\frac{\partial y}{\partial \nu} = u \quad \text{on } \Sigma = (0, T) \times \partial\Omega,$$
$$y(0, x) = y_0(x) \quad \text{in } \Omega,$$

where $g_0(t, \cdot) : R \to \mathbb{R}$ is a convex continuous function with quadratic growth, that is,

$$g_0(t, r) \le C_1|r|^2 + C_2, \quad \forall r \in \mathbb{R},\ t \in (0, T),$$

and $h(t, \cdot) : R \to \overline{\mathbb{R}}$ is convex, lower-semicontinuous and

$$h(t, u) \ge C_3|u|^2 + C_4, \quad \forall u \in \mathbb{R},\ \text{where } C_3 > 0.$$

In fact, if we take $V = H^1(\Omega)$, $H = L^2(\Omega)$, $U = L^2(\partial\Omega)$, $A(t) : V \to V'$ defined by

$$\big(A(t)y, \psi\big) = \int_\Omega \nabla y \cdot \nabla \psi \,\mathrm{d}x, \quad \forall \psi \in H^1(\Omega)$$

and $B(t) : L^2(\partial\Omega) \to V' = (H^1(\Omega))'$ given by

$$\big(B(t)u, \psi\big) = \int_{\partial\Omega} u\,\psi \,\mathrm{d}x, \quad \forall \psi \in H^1(\Omega).$$

Then the corresponding Euler–Lagrange system is

$$\begin{aligned}
&\frac{\partial p}{\partial t} = -\Delta p + \partial g_0(t, y) \quad \text{in } Q,\\
&\frac{\partial p}{\partial \nu} = 0 \quad \text{on } \Sigma,\\
&p(T, x) = -\partial\varphi_0\big(y(T, x)\big) \quad x \in \Omega,\\
&u^*(t) = \partial h^*\big(t, B^* p(t)\big) = \partial h^*\big(t, p(t)|_{\partial\Omega}\big),
\end{aligned}$$

and the corresponding dual problem is

$$\begin{aligned}
\operatorname{Min}\bigg\{ &\int_Q \big(g_0^*(t, v) - f v\big) \,\mathrm{d}x \,\mathrm{d}t + \int_\Sigma h^*(t, p) \,\mathrm{d}x \,\mathrm{d}t + \int_\Omega \varphi_0^*\big(-p(T, x)\big) \,\mathrm{d}x\\
&+ \int_\Omega y_0 p(0, x) \,\mathrm{d}x;\ \frac{\partial p}{\partial t} = -\Delta p + v \text{ in } Q;\ \frac{\partial p}{\partial \nu} = 0 \text{ on } \Sigma \bigg\}.
\end{aligned}$$

Another example is that of the optimal control problem governed by the equation

$$\begin{aligned}
&\frac{\partial y}{\partial t} - \Delta y = \sum_{i=1}^m u_i(t)\delta(x_i) \quad \text{in } (0, T) \times (a, b),\\
&y(0, x) = y_0(x), \quad x \in (a, b); \qquad y(t, a) = y(t, b) = 0,
\end{aligned}$$

where $\delta(x_i)$ is the Dirac measure concentrated in $x_i \in (a, b)$, $i = 1, \dots, m$. In this case, $U = \mathbb{R}^m$, $V = H_0^1(a, b)$, $V' = H^{-1}(a, b)$, $A = -\Delta$, $Bu = \sum_{i=1}^m u_i\delta(x_i)$. Now, we use the results of this section to indicate a variational approach to the Cauchy problem in a Banach space.

A Variational Approach to Time-Dependent Cauchy Problem Consider the nonlinear Cauchy problem

$$\begin{aligned}
&y'(t) + \partial\varphi\big(t, y(t)\big) \ni f(t), \quad \text{a.e. } t \in (0, T),\\
&y(0) = y_0,
\end{aligned} \tag{4.130}$$

in a reflexive Banach space V with the dual V'. More precisely, assume that V and V' are in the variational position

$$V \subset H \subset V',$$

where H is a real Hilbert space and the above inclusions are continuous and dense. As above, the norms of V, H, V' are denoted $\|\cdot\|_V, |\cdot|_H, \|\cdot\|_{V'}$. As regards φ, the following hypotheses are assumed.

(m) $\varphi : [0,T] \times V \to \mathbb{R}$ *is measurable in* t *and convex, continuous in* y *on* V. *There are* $\alpha_i > 0$, $\gamma_i \in \mathbb{R}$, $i = 1, 2$, *such that*

$$\gamma_1 + \alpha_1 \|u\|_V^{p_1} \le \varphi(t,u) \le \gamma_2 + \alpha_2 \|u\|_V^{p_2}, \quad \forall u \in V,\ t \in (0,T),$$

where $2 \le p_1 \le p_2 < \infty$.

(mm) *There are* $C_1, C_2 \in \mathbb{R}^+$ *such that*

$$\varphi(t,-u) \le C_1 \varphi(t,u) + C_2, \quad \forall u \in V.$$

We have the following theorem.

Theorem 4.28 *Under the above hypotheses, for each* $y_0 \in V$ *and* $f \in L^{p_1'}(0,T;V')$, $\frac{1}{p_i} + \frac{1}{p_i'} = 1$, (4.130) *has a unique solution*

$$y^* \in L^{p_1}(0,T;V) \cap C\big([0,T];H\big) \cap W^{1,p_2'}\big([0,T];V'\big). \tag{4.131}$$

Moreover, y^* *is the solution to the minimizing problem*

$$\begin{aligned} \operatorname{Min}\bigg\{ &\int_0^T \big(\varphi(t,y) + \varphi^*(t, f - y') - (f,y)\big)\,\mathrm{d}t + \frac{1}{2}\big|y(T)\big|_H^2; \\ &y \in L^{p_1}(0,T;V) \cap W^{1,p_2'}\big([0,T];V'\big),\ y(0) = y_0 \bigg\}. \end{aligned} \tag{4.132}$$

A nice feature of this theorem is not only its generality (which is, however, comparable with the standard existence theorem for nonlinear Cauchy problems of the form $y' + A(t)y = f(t)$, where $A(t) : V \to V'$ is monotone, demicontinuous and coercive) (see, e.g., Lions [33] or Barbu [6, 13]), but, first of all, that it reduces the Cauchy problem (4.130) to a convex optimization problem with all the consequences deriving from such an identification.

Proof of Theorem 4.28 Translating y_0 into origin, we may assume that $y_0 = 0$. Recalling the conjugacy formulas from Proposition 2.2, we may, equivalently, write (4.130) as

$$\varphi\big(t, y(t)\big) + \varphi^*\big(t, f(t) - y'(t)\big) = \big(f(t) - y'(t), y(t)\big) \quad \text{a.e. } t \in (0,T),$$

while

$$\varphi\big(t, z(t)\big) + \varphi^*\big(t, f(t) - z'(t)\big) - \big(f(t) - z'(t), z(t)\big) \ge 0 \quad \text{a.e. } t \in (0, T),$$

for all $z \in L^{p_1}(0, T; V) \cap W^{1,p_2'}([0, T]; V')$. (Here, φ^* is the conjugate of φ as function of y.) Therefore, we are lead to the optimization problem

$$\begin{aligned}\text{Min}\bigg\{ \int_0^T \big(\varphi\big(t, y(t)\big) + \varphi^*\big(t, f(t) - y'(t)\big) - \big(f(t) - y'(t), y(t)\big)\big)\,\mathrm{d}t;\\ y \in L^{p_1}(0, T; V) \cap W^{1,p_2'}\big([0, T]; V'\big),\ y(0) = 0\bigg\}.\end{aligned} \tag{4.133}$$

However, since the integral $\int_0^T (y'(t), y(t))\,\mathrm{d}t$ might not be well defined, taking into account that (see Proposition 1.12)

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|y(t)\|_V^2 = \big(y'(t), y(t)\big) \quad \text{a.e. } t \in (0, T),$$

for each $y \in L^{p_1}(0, T; V) \cap W^{1,p_2'}([0, T]; V')$, we shall replace (4.133) by the following convex optimization problem:

$$\begin{aligned}\text{Min}\bigg\{ \int_0^T \big(\varphi\big(t, y(t)\big)\big) + \varphi^*\big(t, f(t) - y'(t)\big) - \big(f(t), y(t)\big)\,\mathrm{d}t + \frac{1}{2}\|y(T)\|_V^2;\\ y \in L^{p_1}(0, T; V) \cap W^{1,p_2'}\big([0, T]; V'\big),\ y(0) = 0, y(T) \in H\bigg\},\end{aligned} \tag{4.134}$$

which is well defined because, as easily follows by hypothesis (m), we have, by virtue of the conjugacy formulas,

$$\overline{\gamma}_1 + \overline{\alpha}_1\|\theta\|_{V'}^{p_2'} \le \varphi^*(t, \theta) \le \overline{\gamma}_2 + \overline{\alpha}_2\|\theta\|_{V'}^{p_1'}, \quad \forall \theta \in V' \text{ a.e. } t \in (0, T). \tag{4.135}$$

We are going to prove now that problem (4.134) has a solution y^*, which is also a solution to (4.130). To this end, we set $d^* = \inf$ (4.134) and choose a sequence

$$\{y_n\} \subset L^{p_1}(0, T; V) \cap W^{1,p_2'}\big([0, T]; V'\big)$$

such that $y_n(0) = 0$ and

$$\begin{aligned}d^* &\le \int_0^T \big(\varphi\big(t, y_n(t)\big) + \varphi^*\big(t, f(t) - y_n'(t)\big) - \big(f(t), y_n(t)\big)\big)\,\mathrm{d}t + \frac{1}{2}\big|y_n(T)\big|_H^2\\ &\le d^* + \frac{1}{n}, \quad \forall n \in \mathbb{N}.\end{aligned} \tag{4.136}$$

By hypothesis (m) and by (4.135), we see that

$$\|y_n\|_{L^{p_1}(0,T;V)} + \|y_n'\|_{L^{p_2'}(0,T;V')} \le C, \quad \forall n \in \mathbb{N},$$

and, therefore, on a subsequence, we have

$$\begin{aligned} y_n &\to y^* \quad \text{weakly in } L^{p_1}(0,T;V), \\ y_n' &\to (y^*)' \quad \text{weakly in } L^{p_2'}(0,T;V'), \\ y_n(T) &\to y^*(T) \quad \text{weakly in } H. \end{aligned} \tag{4.137}$$

Inasmuch as the functions $y \to \int_0^T \varphi(t,y(t))\,\mathrm{d}t,\ z \to \int_0^T \varphi^*(t,f(t)-z'(t))\,\mathrm{d}t$ and $y_1 \to |y_1|_H^2$ are weakly lower-semicontinuous in $L^{p_1}(0,T;V)$, $L^{p_2'}(0,T;V')$ and H, respectively, letting n tend to zero into (4.136), we see that

$$\begin{aligned} &\int_0^T \Big(\varphi\big(t,y^*(t)\big)+\varphi^*\big(t,f(t)-(y^*)'(t)\big)-\big(f(t),y^*(t)\big)\Big)\,\mathrm{d}t \\ &\quad + \frac{1}{2}\big|y^*(T)\big|_H^2 = d^*, \end{aligned} \tag{4.138}$$

that is, y^* is solution to (4.134). Now, we are going to prove that $d^*=0$. To this aim, we invoke the duality Theorem 4.16. Namely, we have

$$d^* + \min(\mathrm{P}_1^*) = 0, \tag{4.139}$$

where (P_1^*) is the dual optimization problem corresponding to (4.134), that is,

$$\begin{aligned} (\mathrm{P}_1^*) \quad \mathrm{Min}\bigg\{ &\int_0^T \Big(\varphi\big(t,-p(t)\big)+\varphi^*\big(t,f(t)+p'(t)\big)+\big(f(t),p(t)\big)\Big)\,\mathrm{d}t \\ &+ \frac{1}{2}\big|p(T)\big|_H^2;\ p \in L^{p_1'}(0,T;V)\cap W^{1,p_2'}(0,T;V')\bigg\}. \end{aligned}$$

Clearly, for $p=-y$, we get $\min(\mathrm{P}_1^*) \le d^*$ and so, by (4.139), we see that

$$\min(\mathrm{P}_1^*) \le 0. \tag{4.140}$$

On the other hand, if $\tilde{p}$ is optimal in (P_1^*), we have

$$(\tilde{p}',\tilde{p}) \in L^1(0,T), \qquad \int_0^T (\tilde{p}',\tilde{p})\,\mathrm{d}t = \frac{1}{2}\bigg(\big|\tilde{p}(T)\big|_H^2 - \frac{1}{2}\big|\tilde{p}(0)\big|_H^2\bigg). \tag{4.141}$$

Indeed, by Proposition 2.2, we have

$$-\big(\tilde{p}'(t),\tilde{p}(t)\big) \le \varphi\big(t,-\tilde{p}(t)\big)+\varphi^*\big(t,f(t)+p'(t)\big)+\big(f(t),\tilde{p}(t)\big) \quad \text{a.e. } t \in [0,T]$$

and

$$\big(\tilde{p}'(t)+f(t),\tilde{p}(t)\big) \le \varphi\big(t,\tilde{p}(t)\big)+\varphi^*\big(t,f(t)+\tilde{p}'(t)\big) \quad \text{a.e. } t \in [0,T].$$

Since $\varphi(t,-\tilde{p}) \in L^1(0,T)$, by hypothesis (mm), it follows that $\varphi(t,\tilde{p}) \in L^1(0,T)$ too, and therefore $(\tilde{p}',\tilde{p}) \in L^1(0,T)$, as claimed.

Now, since

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left|\tilde{p}(t)\right|_H^2 = \left(\tilde{p}'(t), \tilde{p}(t)\right) \quad \text{a.e. } t \in (0, T),$$

we get (4.141), as claimed. This means that

$$\begin{aligned}\min(\mathrm{P}_1^*) = &\int_0^T \left(\varphi\left(t, -\tilde{p}(t)\right) + \varphi^*\left(t, f(t) + \tilde{p}'(t)\right) + \left(f(t) + \tilde{p}'(t), \tilde{p}(t)\right)\right) \mathrm{d}t \\ &+ \frac{1}{2}\|\tilde{p}(0)\|_H^2 \geq 0\end{aligned}$$

by virtue of Proposition 2.2. Then by (4.140), we get $d^* = 0$, as claimed.

The same relation (4.141) follows for y^* and thus

$$\frac{1}{2}\left(\left|y^*(t)\right|_H^2 - \left|y^*(s)\right|_H^2\right) = \int_s^t \left((y^*)'(\tau), y^*(\tau)\right) \mathrm{d}\tau, \quad \forall 0 \leq s \leq t \leq T.$$

This implies that $y \in C([0, T]; H)$ and

$$\frac{1}{2}\left|y^*(T)\right|^2 = \int_0^T \left((y^*)'(\tau), y^*(\tau)\right) \mathrm{d}\tau.$$

Substituting the latter into (4.138), we see that y^* is solution to (4.130) and also that

$$\int_0^T \left(\left(\varphi\left(t, y^*(t)\right) + \varphi^*\left(t, f(t) - (y^*)'(t)\right) - \left(f(t) - (y^*)'(t), y^*(t)\right)\right)\right) \mathrm{d}t = 0.$$

Hence,

$$\varphi\left(t, y^*(t)\right) + \varphi^*\left(t, f(t) - (y^*)'(t)\right) - \left(f(t) - (y^*)'(t), y^*(t)\right) = 0 \quad \text{a.e } t \in (0, T)$$

and, therefore, $(y^*(t))' + \partial\varphi(t, y^*(t)) \ni f(t)$ a.e. $t \in (0, T)$, as claimed.

The uniqueness of a solution y^* satisfying (4.138) is immediate by monotonicity of $u \to \partial\varphi(t, u)$ because, for two such solutions y_1^* and y_2^*, we have therefore

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\|y_1^*(t) - y_2^*(t)\right\|_H^2 \leq 0 \quad \text{a.e. } t \in (0, T)$$

and, since $y_1^* - y_2^*$ is H-valued continuous and $y_1^*(T) - y_2^*(T) = 0$, we infer that $y_1^* - y_2^* \equiv 0$, as claimed. This completes the proof of Theorem 4.28. □

4.2 Synthesis of Optimal Control

Consider the unconstrained problem (P) with fixed initial point, that is, the problem of minimizing

$$\int_0^T L\left(t, x(t), u(t)\right) \mathrm{d}t + \varphi_0\left(x(T)\right) \tag{4.142}$$

in all $x \in C([0,T];E)$ and $u \in L^p(0,T;U)$, $2 \le p < \infty$, subject to

$$x' = A(t)x + B(t)u, \quad 0 \le t \le T,$$
$$x(0) = x_0. \tag{4.143}$$

By definition, a *feedback control* is a function (possibly multivalued) $\Lambda : [0,T] \times E \to U$ having the property that, for all $x_0 \in E$ and $s \in [0,T]$, the Cauchy problem

$$x' = A(t)x + B(t)\Lambda(t,x), \quad s \le t \le T,$$
$$x(s) = x_0,$$

has a solution ("mild"), $x = x(t,s,x_0)$. We call such a feedback control Λ, *optimal feedback control* or *optimal synthesis function* provided that, for all $s \in [0,T]$ and $x_0 \in E$, $u(t) = \Lambda(t, x(t,s,x_0))$ is an optimal control for problem (4.142) and (4.143) on the interval $[s,T]$. The existence and design of optimal feedback controllers is related to the problem of control in real time of differential systems which is a fundamental problem in automatic.

This section is concerned with the existence of optimal feedback controls and the method of dynamic programming, that is, the Hamilton–Jacobi approach to problem (4.142).

Owing to some delicate technical considerations, we restrict our attention to the case where L, B and A are independent of t, without, however, losing the essential features of the general problem.

4.2.1 Optimal Value Function and Existence of Optimal Synthesis

We consider here problem (4.142) and (4.143) where $A(t) \equiv A$ is the infinitesimal generator of a C_0-semigroup e^{At}, $B(t) \equiv B$ is a linear continuous operator from U to E, φ_0 is a convex continuous function on E and $L(t) = L$ is a lower-semicontinuous convex function on $E \times U$.

Further, we assume the following hypothesis.

(C$'$) *The Hamiltonian function H associated to L is everywhere finite on $E \times U^*$. Moreover, there exist $\gamma > 0$, $p > 1$, and the real numbers α, β such that*

$$L(x,u) \ge \gamma \|u\|^p - \beta|x| + \alpha, \quad \forall x \in E,\ u \in U. \tag{4.144}$$

There exists $u_0 \in U$, such that

$$L(x,u_0) < +\infty \quad \textit{for all } x \in E. \tag{4.145}$$

The spaces E and U are assumed reflexive and strictly convex together with their duals.

For every $s \in [0, T]$, define the function $\varphi : [0, T] \times E \to \mathbb{R}$

$$\varphi(s, h) = \inf\left\{ \int_s^T L(x, u)\,\mathrm{d}t + \varphi_0\big(x(T)\big);\ x' = Ax + Bu; x(s) = h, \right.$$
$$\left. u \in L^p(s, T; U) \right\}, \tag{4.146}$$

which is called the *optimal value function* associated with problem (4.142).

Proposition 4.29 *For all* $(s, h) \in [0, T] \times E$, $-\infty < \varphi(s, h) < +\infty$ *and for every* $h \in E$, *the infimum defining* $\varphi(s, h)$ *is attained. For every* $s \in [0, T]$, *the function* $\varphi(s, \cdot) : E \to \mathbb{R}$ *is convex and continuous.*

Proof Let (s, h) be arbitrary but fixed in $[0, T] \times E$. By condition (4.145), wee see that $\varphi(s, h) < +\infty$, while condition (4.144) implies, by virtue of Proposition 4.29 (assumptions (a) and (c) are trivially satisfied here), that the infimum defining $\varphi(s, h)$ is attained. This implies via a standard argument that, for all $s \in [0, T]$, the function $\varphi(s, \cdot)$ is convex and nowhere $-\infty$.

Now, we prove that $\varphi(s, \cdot)$ is lower-semicontinuous on E. To this end, we consider a sequence $\{h_n\} \subset E$, such that $\varphi(s, h_n) \le M$ for all n and $h_n \to h$, as $n \to \infty$. Let $\{x_s^n, u_s^n\} \in C([s, T]; E) \times L^p(s, T; U)$ be such that

$$\varphi(s, h_n) = \int_s^T L\big(x_s^n, u_s^n\big)\,\mathrm{d}t + \varphi_0\big(x_s^n(T)\big) \le M.$$

Then, by assumption (4.144), we deduce via Gronwall's Lemma that $\{u_s^n\}$ remains in a bounded subset (equivalently, weakly compact subset) of $L^p(s, T; U)$. Thus, without loss of generality, we may assume that

$$u_s^n \to u_s \quad \text{weakly in } L^p(s, T; U)$$

and, therefore,

$$x_s^n(t) \to x_s(t) = \mathrm{e}^{(t-s)A} h + \int_s^t \mathrm{e}^{(t-\tau)A} Bu(\tau)\,\mathrm{d}\tau \quad \text{weakly in } E \text{ for every } t \in [s, T].$$

Since the function $(s, u) \to \int_s^T L(x, u)\,\mathrm{d}t$ and φ_0 are weakly lower-semicontinuous on $L^p(s, T; E) \times L^p(s, T; U)$ and E, respectively, we have

$$\varphi(s, h) \le \int_s^T L(x_s, u_s)\,\mathrm{d}t + \varphi_0\big(x_s(T)\big) \le M,$$

as claimed. Since $\varphi(s, \cdot)$ is convex, lower-semicontinuous and everywhere finite on E, we may conclude that it is continuous (see Proposition 2.16). Thus, the proof is complete. □

Let h be arbitrary but fixed in E, and let (x_s, u_s) be an optimal pair in problem (4.26). In other words,

$$\varphi(s,h) = \int_s^T L(x_s,u_s)\,\mathrm{d}t + \varphi_0\big(x_s(T)\big).$$

As mentioned in Sect. 4.1.1, the condition that $-\infty < H < +\infty$ is more than sufficient to ensure that Assumption (C) of Theorem 4.33 is satisfied. Hence, there exists a function $p_s \in C([s,T];E^*)$ (which is not, in general, uniquely determined) and $q_s \in L^p(s,T;E^*)$ such that

$$x_s(t) = \mathrm{e}^{A(t-s)}h + \int_s^t \mathrm{e}^{A(t-\tau)}Bu_s(\tau)\,\mathrm{d}\tau, \quad s \le t \le T, \tag{4.147}$$

$$p_s(t) = \mathrm{e}^{A^*(T-t)}p_s(T) - \int_t^T \mathrm{e}^{A^*(\tau-t)}q_s(\tau)\,\mathrm{d}\tau, \quad s \le t \le T, \tag{4.148}$$

$$p_s(T) \in -\partial\varphi_0\big(x_s(T)\big), \tag{4.149}$$

$$(q_s, B^*p_s) \in \partial L(x_s,u_s) \quad \text{a.e. on }]s,T[. \tag{4.150}$$

Fixing $(y,v) \in C([s,T];E) \times L^p(s,T;U)$ such that $y' = Ay + Bv$ on $[s,T]$ and $y(s) = h$, we have by (4.149) and the definition of ∂L

$$L(x_s,u_s) \le L(y,v) + (q_s, x_s - y) + \big(p_s, B(u_s - v)\big), \quad \text{a.e. } t \in]s,T[.$$

We integrate over $[s,T]$. By a straightforward calculation involving (4.147), (4.148), and Fubini's theorem, we find that

$$\int_s^T L(x_s,u_s)\,\mathrm{d}t \le \int_s^T L(y,v)\,\mathrm{d}t + \big(p_s(T), x_s(T) - y(T)\big) - \big(p_s(s), h - \tilde{h}\big),$$

whereupon by (4.147)–(4.149) we see that

$$\varphi(s,h) \le \varphi\big(s,\tilde{h}\big) - \big(p_s(s), h - \tilde{h}\big), \quad \forall \tilde{h} \in E. \tag{4.151}$$

Thus, we have shown that $h \in D(\partial\varphi(s,\cdot))$ and $-p_s(s) \in \partial\varphi(s,h)$. Let us denote by $\mathscr{M}_s^h$ the set of all the dual extremal arcs $p_s \in C([s,T];E^*)$ corresponding to problem (4.146). We have a quite unexpected relationship between $\mathscr{M}_s^h$ and $\partial\varphi(s)$.

Proposition 4.30 *For all $s \in [0,T]$ and $h \in E$, we have*

$$\partial\varphi(s,h) = \big\{-p_s(s),\ p_s \in \mathscr{M}_s^h\big\}. \tag{4.152}$$

Proof Let $\mathscr{A} : E \to E^*$ be the mapping defined by

$$\mathscr{A}h = \big\{-p_s(s);\ p \in \mathscr{M}_s^h\big\}.$$

We have already seen that $\mathscr{A} \subset \partial\varphi(s, \cdot)$. To prove the converse inclusion relation, it suffices to show that $\mathscr{A}$ is maximal monotone, that is, $R(\Phi + \mathscr{A}) = E^*$. For any $h_0^* \in E^*$, the equation $\Phi(h) + \mathscr{A}h \ni h_0^*$ can be explicitly written as

$$y' = Ay + Bv \quad \text{on } [s, T], \tag{4.153}$$

$$\tilde{p}' = -A^*\tilde{p} + \tilde{q} \quad \text{on } [s, T], \tag{4.154}$$

$$(\tilde{q}, B^*\tilde{p}) \in \partial L(y, v) \quad \text{a.e. on }]s, T[, \tag{4.155}$$

$$\Phi\big(y(s)\big) - \tilde{p}(s) = h_0^*, \quad \tilde{p}(T) \in -\partial\varphi_0\big(y(T)\big). \tag{4.156}$$

(Equations (4.153) and (4.154) must be considered, of course, in the "mild" sense.) □

Again by Theorem 4.5, system (4.153)–(4.156) has a solution if and only if the control problem

$$\inf\left\{\int_s^T L(y, v)\,\mathrm{d}t + \frac{1}{2}\big|y(s)\big|^2 - \big(h_0^*, y(s)\big) + \varphi_0\big(y(T)\big);\right.$$
$$\left. v \in L^p(s, T; U),\ y' = Ay + Bv \text{ on }]s, T[\right\}$$

has solution. But the latter has a solution by virtue of (C′) and of Proposition 4.29. Hence, the equation $\Phi(h) + \mathscr{A}h \ni h_0^*$ has at least one solution h. Equation (4.152) can be used in certain situations to show that the operator $\partial\varphi(s, \dot{)}$ is single-valued. For instance, we have the following proposition.

Proposition 4.31 *Let $U = E$, $B = I$ and let the function L be of the form*

$$L(x, u) = g(x) + \psi(u), \quad \forall x \in E,\ u \in E. \tag{4.157}$$

If either ψ^* is strictly convex or g^* and φ_0^* are both strictly convex, then $\partial\varphi(s, \cdot)$ is single-valued on E.

Proof It suffices to show that, under the above conditions, the dual extremal arc p_s to problem (4.146) is unique. By Theorem 4.16, every such p_s is the solution to the dual control problem

$$\inf\left\{\int_s^T M(p, v)\,\mathrm{d}t + \varphi_0^*\big(-p(T)\big) + \big(p(0), h\big),\ p' + A^*p = v,\right.$$
$$\left. v \in L^p(s, T; E^*)\right\}, \tag{4.158}$$

where

$$M(p, v) = g^*(v) + \psi^*(p).$$

If ψ^* is strictly convex, then clearly the solution p_s to (4.158) is unique. This also happens if φ_0^* and g^* are strictly convex. □

Remark 4.32 In particular, it follows by Proposition 4.31 that $\varphi(s,\cdot)$ is Gâteaux differentiable within E (see Proposition 2.40 and the comments which follow it).

Now, we return to the optimal control problem (4.142) and (4.143).

Let $(x^*,u^*)\in C([0,T];E)\times L^p(0,T;U)$ be an optimal pair. Then, by Theorem 4.33, there is $p^*\in C([0,T];E^*)$ and $q\in L^p(0,T;E^*)$ satisfying

$$x^{*\prime}=Ax^*+Bu^* \quad \text{on } [0,T], \tag{4.159}$$

$$p^{*\prime}=-A^*p^*+q \quad \text{on } [0,T], \tag{4.160}$$

$$(q,B^*p^*)\in\partial L(x^*,u^*) \quad \text{a.e. on }]0,T[, \tag{4.161}$$

$$\begin{aligned}&p^*(T)\in-\partial\varphi_0\big(x^*(T)\big), \qquad u^*(t)\in\partial_p H\big(x^*(t),B^*p^*(t)\big)\\ &\quad \text{a.e. } t\in\,]0,T[.\end{aligned} \tag{4.162}$$

We see that, for every $s\in[0,T]$, (x^*,u^*) is also an optimal pair for problem (4.146) with initial value $h=x^*(s)$. Another way of saying this is that

$$\varphi\big(s,x^*(s)\big)=\int_s^T L\big(x^*(t),u^*(t)\big)\,\mathrm{d}t+\varphi_0\big(x^*(T)\big) \quad \text{for } s\in[0,T], \tag{4.163}$$

so that, by (4.152), we have

$$p^*(s)+\partial\varphi\big(s,x^*(s)\big)\ni 0 \quad \text{for all } s\in[0,T].$$

This means that

$$\Lambda(t,x)=\partial_p H\big(x,-B^*\partial\varphi(t,x)\big)$$

is an optimal synthesis function for problem (4.142), (4.143). In other words, any optimal control $u^*(t)$ is given by the *feedback law*

$$u^*(t)\in\partial_p H\big(x^*(t),-B^*\partial\varphi\big(t,x^*(t)\big)\big), \quad t\in[0,T], \tag{4.164}$$

while the optimal state x^* is the solution to the *closed loop differential system*

$$\begin{aligned}&x'\in Ax+B\partial_p H\big(x,-B^*\partial\varphi(t,x)\big), \quad 0\le t\le T,\\ &x(0)=x_0.\end{aligned} \tag{4.165}$$

In a few words, the result just established amounts to saying that *every optimal control u is a feedback optimal control.*

4.2.2 Hamilton–Jacobi Equations

In this section, we prove that, under certain circumstances, the optimal value function $\varphi:[0,T]\times E\to\mathbb{R}$ is the solution to a certain nonlinear operator equation (see (4.167) below), which generalizes the well-known Hamilton–Jacobi equation from the calculus of variations and classical mechanics.

This equation is known in the literature (see, for instance, Fleming and Rishel [27], Berkovitz [18]) as the *Bellman equation* or the *partial differential equation of dynamic programming*.

We assume hereafter that E and U are real Hilbert spaces, A is the infinitesimal generator of an analytic semigroup of class C_0, and that the assumptions of Sect. 4.1.1 are satisfied where $L(t) \equiv L$, $\varphi_0 = 0$, $B(t) \equiv B$, $p = 2$.

In addition, we assume that Condition (C$'$) is satisfied with $p = 2$ and that we have the following.

For every $k > 0$, there exists $C_k > 0$ such that

$$\sup\big\{|y|;\ y \in \partial_x H(x,q)\big\} \le C_k\big(1 + \|q\|\big) \quad \textit{for } q \in U,\ |x| \le k. \tag{4.166}$$

The main result is the following theorem.

Theorem 4.33 *Under the above assumptions, for every $s \in [0,T]$, the function $h \to \varphi(s,h)$ is convex and lower-semicontinuous on E and, for every $h \in D(A)$, the function $s \to \varphi(s,h)$ is absolutely continuous on $[0,T]$ and satisfies the equation*

$$\varphi_s(s,h) + \big(Ah, \partial\varphi(s,h)\big) - H\big(h, -B^*\partial\varphi(s,h)\big) = 0$$
$$\textit{a.e. } s \in\,]0,T[,\ \forall h \in D(A), \tag{4.167}$$
$$\varphi(T,h) = 0 \quad \textit{for all } h \in E. \tag{4.168}$$

Here, $\varphi_s(s,h)$ stands for the partial derivative $\frac{\mathrm{d}}{\mathrm{d}s}\varphi(s,h)$ which exists a.e. on $]0,T[$. Equation (4.167) must be understood in the following sense: for all $h \in D(A)$, almost all $s \in\,]0,T[$ and every section $\eta(s,h) \subset \partial\varphi(s,h)$,

$$\varphi_s(s,h) + \big(Ah, \eta(s,h)\big) - H\big(h, -B^*\eta(s,h)\big) = 0.$$

Here, $\partial\varphi(s,h)$ denotes, as usual, the subdifferential of φ as a function of h.

Proof Fix $h \in D(A)$ and $s \in [0,T]$. By Proposition 4.29 and Theorem 4.5, there exist functions $x_s, p_s \in C([s,T];E)$, $q_s \in L^2(s,T;E)$, $u_s \in L^2(s,T;U)$ satisfying

$$x_s' = Ax_s + Bu_s \quad \text{a.e. } t \in\,]s,T[, \tag{4.169}$$
$$p_s' = -A^*p_s + q_s \quad \text{a.e. } t \in\,]s,T[, \tag{4.170}$$
$$x_s(s) = h, \qquad p_s(T) = 0, \tag{4.171}$$
$$(q_s, B^*p_s) \in \partial L(x_s,u_s) \quad \text{a.e. on }]s,T[, \tag{4.172}$$

and

$$\varphi(s,h) = \int_0^T L(x_s,u_s)\,\mathrm{d}t.$$

Equations (4.169), (4.170), and (4.172) can be equivalently expressed as

$$x_s' = Ax_s + B\partial_p H(x_s, B^* p_s) \quad \text{a.e. } t \in]s, T[, \tag{4.173}$$

$$p_s' = -A^* p_s - \partial_x H(x_s, B^* p_s) \quad \text{a.e. } t \in]s, T[. \tag{4.174}$$

As noticed earlier (see Proposition 1.148), since A generates an analytic semigroup, the functions x_s and p_s belong to $W^{1,2}([s, T]; E)$ and are strong solutions to (4.169) and (4.170) ((4.172) and (4.173), respectively).

By condition (4.144), we have

$$\gamma \int_s^T \|u_s(t)\|^2 \,\mathrm{d}t + \alpha \le \beta \int_s^T |x_s(t)| \,\mathrm{d}t + \int_s^T L\big(x_s(t), u_s(t)\big) \,\mathrm{d}t$$

and this yields

$$\gamma \int_s^T \|u_s(t)\|^2 \,\mathrm{d}t + \alpha \le \beta \int_s^T |x_s(t)| \,\mathrm{d}t + \int_s^T L\big(x_s^0, u_s^0\big) \,\mathrm{d}t,$$

where $x_s^0(t) = y(t-s)$, $u_s^0(t) = u(t-s)$ and (y, u) is a feasible pair in problem (4.142).

It follows that

$$\int_s^T \|u_s(t)\|^2 \,\mathrm{d}t \le C\left(\int_s^T |x_s(t)| \,\mathrm{d}t + 1\right), \quad s \in [0, T], \tag{4.175}$$

where C is independent of s.

Along with the variation of constant formula

$$x_s(t) = \mathrm{e}^{A(t-s)} h + \int_s^T \mathrm{e}^{A(t-\tau)} B u_s(\tau) \,\mathrm{d}\tau, \quad s \le t \le T,$$

the latter inequality implies via a standard calculation involving Gronwall's Lemma

$$\int_s^T \|u_s(t)\|^2 \,\mathrm{d}t \le C, \quad 0 \le s \le T. \tag{4.176}$$

Then, again using Proposition 1.148, we get

$$\int_s^T |x_s'(t)|^2 \,\mathrm{d}t \le C, \quad s \in [0, T]. \tag{4.177}$$

(In the following, we denote by C several positive constants independent of s.)

Now, condition (4.166) and (4.174) imply

$$|q_s(t)| \le C\big(1 + |p_s(t)|\big) \quad \text{a.e. } t \in]s, T[, \tag{4.178}$$

because, by virtue of (4.177), $|x_s(t)|$ are uniformly bounded on $[s,T]$. Then, using once again the variation of constant formula in (4.170), we get

$$\left|p_s(t)\right| \le C\int_s^T \left|q_s(\tau)\right| \mathrm{d}\tau, \quad t\in[s,T].$$

Substituting (4.168) in the latter, we obtain by Gronwall's Lemma

$$\left|p_s(t)\right| \le C, \quad t\in[s,T] \tag{4.179}$$

and, therefore,

$$\left|q_s(t)\right| \le C, \quad t\in[s,T]. \tag{4.180}$$

By (4.170) and estimates (4.179) and (4.180), it follows that

$$\int_s^T \left|p_s'(t)\right|^2 \mathrm{d}t \le C, \quad s\in[0,T] \tag{4.181}$$

(because A^* generates an analytic semigroup).

Let $\varepsilon>0$ be such that $s+\varepsilon<T$. We note that

$$\varphi(s+\varepsilon,h) \le \int_{s+\varepsilon}^T L\big(x_s(t-\varepsilon),u_s(t-\varepsilon)\big)\,\mathrm{d}t,$$

whereupon

$$\varphi(s+\varepsilon,h)-\varphi(s,h) \le -\int_{T-\varepsilon}^T L\big(x_s(t),u_s(t)\big)\,\mathrm{d}t. \tag{4.182}$$

On the other hand, a glance at relation (4.161) plus a little calculation reveals that

$$\begin{aligned}\varphi(s,h)-\varphi(s+\varepsilon,h) \le{}& \int_s^{s+\varepsilon} L\big(x_s(t),u_s(t)\big)\,\mathrm{d}t\\ &+\big(p_s(s+\varepsilon),x_s(s)-x_s(s+\varepsilon)\big).\end{aligned} \tag{4.183}$$

We claim that the function $t\to H(x_s(y),B^*p_s(t))$ is absolutely continuous on $[s,T]$ and

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(H\big(x_s(t),B^*p_s(t)\big)+\big(Ax_s(t),p_s(t)\big)\Big)=0, \quad \text{a.e. on }]s,T[. \tag{4.184}$$

Postponing for the moment the verification of these properties, we notice that (4.184) implies that

$$H\big(x_s(t),B^*p_s(t)\big)+\big(Ax_s(t),p_s(t)\big)=\delta(s) \quad \text{for } t\in[s,T].$$

On the other hand, (4.172) yields

$$L\big(x_s(t),u_s(t)\big)=\big(Bu_s(t),p_s(t)\big)-H\big(x_s(t),B^*p_s(t)\big) \quad \text{a.e. on }]s,T[,$$

so that

$$L\big(x_s(t), u_s(t)\big) = \big(x_s'(t), p_s(t)\big) - \delta(s) \quad \text{a.e. } t \in]s, T[.$$

Substituting the above equation in (4.182) and (4.183) gives

$$\begin{aligned} &\big|\varphi(s+\varepsilon, h) - \varphi(s, h) - \varepsilon\delta(s)\big| \\ &\quad \le \max\left\{\int_{T-\varepsilon}^{T} \big|x'(t)\big|\,\big|p_s(t)\big|\,\mathrm{d}t, \int_s^{s+\varepsilon} \big|x_s'(t)\big|\,\big|p_s(t) - p_s(s+\varepsilon)\big|\,\mathrm{d}t\right\}. \end{aligned} \tag{4.185}$$

On the other hand, we have

$$\big|p_s(t)\big|^2 \le \varepsilon \int_{T-\varepsilon}^{T} \big|p'(\tau)\big|^2\,\mathrm{d}\tau \quad \text{for } T-\varepsilon \le t \le T,$$

while

$$\big|p_s(t) - p_s(\varepsilon+s)\big| \le \int_s^{s+\varepsilon} \big|p_s'(\tau)\big|\,\mathrm{d}\tau \quad \text{for } s \le t \le s+\varepsilon.$$

Estimates (4.177), (4.181), and (4.185) taken together show that

$$\big|\varphi(s+\varepsilon, h) - \varphi(s, h) - \varepsilon\delta(s)\big| \le C(\varepsilon)\varepsilon, \tag{4.186}$$

where

$$\lim_{\varepsilon\to 0} C(\varepsilon) = 0.$$

Moreover, it is obvious that

$$\delta(s) = H\big(x_s(T), 0\big).$$

Inasmuch as $\{|x_s(T)|\}$ is bounded in E, condition (4.166) implies, in particular, that $\partial_s H(x_s(T), 0)$ and, consequently, $\delta(s)$ are bounded on $[0, T]$. Thus, inequality (4.186) shows that the function $s \to \varphi(s, h)$ is Lipschitz on $[0, T]$. Moreover, it follows from (4.186) that

$$\frac{\mathrm{d}}{\mathrm{d}s}\varphi(s, h) = \delta(s) = H\big(h, B^* p_s(s)\big) + \big(Ah, p_s(s)\big).$$

Recalling that, by Proposition 4.30, $p_s(s) = -\partial\varphi(s, h)$, we obtain the desired equality (4.167).

We complete the proof of Theorem 4.33 by verifying that the function $H(x_s(t), B^* p_s(t))$ has the properties listed above (equation (4.184)). We have already noticed that the condition $-\infty < H(x, p) < +\infty$, for all $(x, p) \in E \times U$, implies that the subdifferential $\partial H = \{-\partial_x H, \partial_p H\}$ of H is locally bounded in $E \times U$ (see Corollary 2.111). In particular, this implies that the function $(x, p) \to H(x, p)$ is locally Lipschitz on $E \times U$. In other words, for every $(x_0, p_0) \in E \times U$ there is a neighborhood V_0 of (x_0, p_0) and a positive constant M such that

$$\big|H(x, p) - H(y, q)\big| \le M\big(|x - y| + \|p - q\|\big)$$

for all (x, y) and (y, q) in V_0. Since the functions x_s and p_s are E-valued continuous on $[s, T]$, the above inequality implies that

$$\left|H\big(x_s(t), B^* p_s(t)\big) - H\big(x_s(\tilde{t}), B^* p_s(\tilde{t})\big)\right| \le M_1\big(\left|x_s(t) - x_s(\tilde{t})\right| + \left|p_s(t) - p_s(\tilde{t})\right|\big)$$

for all t and $\tilde{t}$ in $[s, T]$. Recalling that x_s and p_s are in $W^{1,2}([s, T]; E)$, we may infer that the function $t \to H(x_s(t), B^* p_s(t))$ is absolutely continuous on $[s, T]$ and, therefore, it is almost everywhere differentiable on $]s, T[$ with $\frac{\mathrm{d}}{\mathrm{d}t} H(x_s(t), B^* p_s(t))$ in $L^2(s, T)$. Next, we show that relation (4.184) holds almost everywhere on $]s, T[$. Let t and $h > 0$ be such that $t, t + h \in [s, T]$. We observe from (4.173), (4.174), and the definition of ∂H that

$$H\big(x_s(t), B^* p_s(t)\big) - H\big(x_s(t), B^* p_s(t+h)\big) \le \big\langle u_s(t), B^*\big(p_s(t) - p_s(t+h)\big)\big\rangle$$

while

$$\begin{aligned}&-H\big(x_s(t+h), B^* p_s(t+h)\big) + H\big(x_s(t), B^* p_s(t+h)\big)\\ &\quad \le \big(q_s(t+h), x_s(t+h) - x_s(t)\big),\end{aligned}$$

wherein $q_s(t) = p_s'(t) + A^* p_s(t)$. Combining the two relations above gives

$$\begin{aligned}&H\big(x_s(t), B^* p_s(t)\big) - H\big(x_s(t+h), B^* p_s(t+h)\big)\\ &\qquad - \big(Bu_s(t+h) - q_s(t), x_s(t+h)\big) - \big(q_s(t), x_s(t+h) - x_s(t)\big)\\ &\quad \le \big(q_s(t+h) - q_s(t), x_s(t+h) - x_s(t)\big).\end{aligned}$$

Similarly,

$$\begin{aligned}&H\big(x_s(t+h), B^* p_s(t+h)\big) - H\big(x_s(t), B^* p_s(t)\big)\\ &\qquad - \big(Bu_s(t), p_s(t+h) - p_s(t)\big) - \big(q_s(t), x_s(t) - x_s(t+h)\big)\\ &\quad \le \big(B\big(u_s(t+h) - u_s(t)\big), p_s(t+h) - p_s(t)\big).\end{aligned}$$

Integrating over $[s, T - h]$ yields

$$\begin{aligned}&\frac{1}{h}\int_s^{T-h} \big|H\big(x_s(t), B^* p_s(t)\big) - H\big(x_s(t+h), B^* p_s(t+h)\big)\\ &\qquad - \big(Bu_s(t), p_s(t) - p_s(t+h)\big) - \big(q_s(t), x_s(t+h) - x_s(t)\big)\big|\,\mathrm{d}t\\ &\quad \le \frac{1}{h}\int_s^{T-h} \big(\big|B\big(u_s(t+h) - u_s(t)\big)\big|\,\big|p_s(t+h) - p_s(t)\big|\\ &\qquad + \big|q_s(t+h) - q_s(t)\big|\,\big|x_s(t) - x_s(t+h)\big|\big)\,\mathrm{d}t.\end{aligned}$$

Since x_s, p_s are in $W^{1,2}([s, T]; E)$ and Bu_s, q_s in $L^2(s, T; E)$, we can take the limits for $h \to 0$ and use the Lebesgue dominated convergence theorem to get

$$\frac{\mathrm{d}}{\mathrm{d}t} H\big(x_s(t), B^* p_s(t)\big) - \big(Bu_s(t), p_s'(t)\big) + \big(q_s(t), x_s'(t)\big) = 0 \quad \text{a.e. } t \in\,]s, T[,$$

as claimed. Theorem 4.33 has now been completely proved. □

Now, we prove a variant of Theorem 4.33 under the following stronger assumptions on H and A.

(a) *The Hamiltonian function H satisfies Condition* (C′), *where $p = 2$, and E, U are real Hilbert spaces. The function $p \to H(x, p)$ is Fréchet differentiable and the function $(x, p) \to \partial_p H(x, p)$ is continuous and bounded on every bounded subset of $E \times U$.*

(b) *$L : E \to U \to \mathbb{R}$ is continuous, convex and locally Lipschitz in x, that is, for every $r > 0$, there exists L such that*

$$\bigl|L(x,u) - L(y,u)\bigr| \le L_r|x - y| \quad \textit{for } |x|, |y|, \|u\| \le r.$$

(c) *$A : D(A) \subset E \to E$ is the infinitesimal generator of a C_0-semigroup on E and $\varphi_0 : E \to \mathbb{R}$ is a convex, continuous function which is bounded on every bounded subset.*

Let $\varphi : [0, T] \times E \to \mathbb{R}$ be the optimal value function (4.146).

Theorem 4.34 *Under assumptions* (a), (b), (c) *and* (4.166), *the function φ satisfies the following conditions.*

(i) *For every $x \in D(A)$, $s \to \varphi(s, x)$ is Lipschitz on* $[0, T]$.
(ii) *For every $s \in [0, T]$, $x \to \varphi(s, x)$ is convex and Lipschitz on every bounded sub set of E.*
(iii) *For all $h \in D(A)$ and for almost all $s \in]0, T[$, there exists $\eta(s, h) \in \partial\varphi(s, h)$ such that*

$$\varphi_s(s,h) + \bigl(Ah, \eta(s,h)\bigr) - H\bigl(h, -B^*\eta(s,h)\bigr) = 0 \quad \textit{a.e. } s \in]0,T[, \tag{4.187}$$

$$\varphi(T,h) = \varphi_0(h) \quad \textit{for all } h \in E. \tag{4.188}$$

Proof We denote by $x(t, s, h, u)$ the "mild" solution to the Cauchy problem

$$\begin{aligned} &x' = Ax + Bu, \quad s \le t \le T,\\ &x(s) = h. \end{aligned} \tag{4.189}$$

Let $(s, h) \in [0, T] \times D(A)$ be arbitrary but fixed. Let $(x_s, u_s) \in C([s, T]; E) \times L^2(s, T; U)$ be a solution to (4.189) such that

$$\varphi(s,h) = \int_s^T L(x_s, u_s)\,\mathrm{d}t + \varphi_0\bigl(x_s(T)\bigr) \tag{4.190}$$

and let p_s be a corresponding dual extremal arc. In other words, x_s, u_s, p_s satisfy (4.169), (4.170), and (4.172) (equivalently (4.173) and (4.174)) along with the transversality conditions

$$x_s(s) = h, \qquad p_s(T) + \partial\varphi_0\bigl(x_s(T)\bigr) \ni 0.$$

Let $0 \le s \le s_1 \le T$ and $u_0 \in U$. By (b), $L(x, u_0) < +\infty$, for all $x \in E$. Consider the function $w : [s, T] \to U$ defined by

$$w(t) = u_0 \quad \text{for } s \le t \le s_1, \qquad w(t) = u_{s_1}(t), \quad s_1 \le t \le T.$$

We have

$$\begin{aligned}\varphi(s,h) - \varphi(s_1,h) \le{}& \int_s^{s_1} L\big(x(t,s,h,u_0),u_0\big)\,\mathrm{d}t + \int_{s_1}^T \big(L\big(x(t,s,h,w),w(t)\big) \\ &- L\big(x_{s_1}(t),u_{s_1}(t)\big)\big)\,\mathrm{d}t + \varphi_0\big(x\big(T,s_1,x(s_1,s,h,w),w\big)\big) \\ &- \varphi_0\big(x(T,s_1,h,w)\big). \end{aligned} \tag{4.191}$$

On the other hand, we have

$$\big|x(t,s,h,u_0) - h\big| \le \big|\mathrm{e}^{A(t-s)}h - h\big| + \int_s^t \big|\mathrm{e}^{A(t-\tau)}Bu_0\big|\,\mathrm{d}\tau \le C|t-s|\big(1+|Ah|\big)$$
$$\text{for } s \le t \le s_1,$$

and

$$\big|x(t,s_1,h,w) - x(t,s_1,h_1,w)\big| \le C|h-h_1| \quad \text{for } s_1 \le t \le T,$$

respectively,

$$\big|x(t,s,h,w) - x(t,s_1,h,w)\big| \le C\big|x(s_1,s,h,u_0) - h\big| \le C|s_1 - s|.$$

We notice that, by assumption (c), φ_0 is locally Lipschitz on E (because φ_0 and, consequently, $\partial\varphi_0$ are bounded on bounded subsets). Then, by (4.191), we see that the function $s \to \varphi(s,h)$ is Lipschitz on $[0,T]$.

Next, for all $s \in [0,T]$, $h, \tilde{h} \in E$, we have

$$\big|x(t,s,h,u) - x\big(t,s,\tilde{h},u\big)\big| \le C\big|h - \tilde{h}\big| \tag{4.192}$$

and

$$\big|x(t,s,h,u)\big| \le C\left(|h| + \int_s^t \big\|u(\tau)\big\|\,\mathrm{d}\tau\right), \quad s \le t \le T.$$

Then, by (4.175), we see that for $|h| \le r$, we may restrict problem (4.146) to those $u \in L^2(s,T;U)$ and $x(t,s,h,u)$ which satisfy the inequality (C_1 is independent of s)

$$\int_s^T \big\|u(\tau)\big\|^2\,\mathrm{d}\tau + \big|x(t,s,h,u)\big| \le C_r \quad \text{for } 0 \le t \le T.$$

Since the functions L and φ_0 are locally Lipschitz, it follows, by (4.192), that

$$\big|\varphi(s,h) - \varphi\big(s,\tilde{h}\big)\big| \le L_r\big|h - \tilde{h}\big| \tag{4.193}$$

for all $s \in [0, T]$ and $|h|, |\tilde{h}| \le r$, where L_r is independent of s.

It remains to prove that φ verifies (4.187). To this purpose, we fix $h \in D(A)$, $s \in [0, T]$ and notice that, for all $s \le t \le T$,

$$\varphi\big(t, x_s(t)\big) = \int_t^T L\big(x_s(\tau), u_s(\tau)\big)\,\mathrm{d}\,\mathrm{d}\tau + \varphi_0\big(x_s(T)\big).$$

Recalling that $u_s(t) = \partial_p H(x_s(t), B^* p_s(t))$ for $t \in [s, T]$ and

$$L\big(x_s(t), u_s(t)\big) + H\big(x_s(t), B^* p_s(t)\big) = \big(Bu_s(t), p_s(t)\big), \tag{4.194}$$

we conclude by assumption (a) that the function $t \to L(x_s(t), u_s(t))$ is continuous on $[s, T]$ and therefore

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\varphi\big(t, x_s(t)\big) + L\big(x_s(t), u_s(t)\big) = 0, \quad s \le t \le T.$$

Let $s \in [0, T]$ be such that the function $t \to \varphi(t, h)$ is differentiable at $t = s$. We have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\varphi\big(t, x_s(t)\big)\bigg|_{t=s} &= \lim_{t\to s}\big(\varphi\big(t, x_s(s)\big) - \varphi\big(s, x_s(s)\big)\big)(t-s)^{-1} \\ &\quad + \lim_{t\to s}\big(\varphi\big(t, x_s(t)\big) - \varphi\big(t, x_s(s)\big)\big)(t-s)^{-1}.\end{aligned} \tag{4.195}$$

By the mean value property (see Proposition 2.66), there exist ζ_t on the line segment between $x_s(t)$ and $x_s(s)$ and $\delta_t \in \partial\varphi(t, \zeta_t)$ such that

$$\varphi\big(t, x_s(t)\big) - \varphi\big(t, x_s(s)\big) = \big(\delta_t, x_s(t) - x_s(s)\big).$$

Since, by (4.193), $\{\delta_t\}$ is bounded for $t \to s$ and $\lim_{t\to s}\zeta_t = x_s(s) = h$, we may assume that

$$\delta_t \to \eta(s, h) \quad \text{weakly in } E,$$

where $\eta(s, h) \in \partial\varphi(s, h)$. On the other hand, since the function u_s is continuous on $[s, T]$ and $x_s(s) = h \in D(A)$, it follows, from the variation of constant formula, that

$$x_s(t) = \mathrm{e}^{A(t-s)}h + \int_s^t \mathrm{e}^{A(t-\tau)}Bu_s(\tau)\,\mathrm{d}\tau, \quad s \le t \le T,$$

and that

$$\lim_{t\to s}\big(x_s(t) - x_s(s)\big)(t-s)^{-1} = Ah + Bu_s(s).$$

Along with (4.195), the latter yields

$$\varphi_s(s, h) + \big(\eta(s, h), Ah + Bu_s(s)\big) + L\big(h, u_s(s)\big) = 0$$

and, by (4.194),

$$\varphi_s(s,h) + \big(\eta(s,h), Ah\big) - H\big(h, -B^*\eta(s,h)\big) = 0,$$

because, by virtue of Proposition 4.30, we may take $p_s(s) = -\eta(s,h) \in \partial\varphi(s,h)$.

The proof of Theorem 4.34 is, therefore, complete. □

As regards the uniqueness in (4.142), we have the following theorem.

Theorem 4.35 *Under the assumptions of Theorem 4.33, let $\varphi : [0,T] \times E \to \mathbb{R}$ be a solution to problem (4.167)–(4.168) having the following properties.*

(j) *For every $s \in [0,T]$, $\varphi(s,\cdot)$ is convex and continuous on E; for every $x \in W^{1,2}([0,T];E)$, the function $t \to \varphi(t,x(t))$ is absolutely continuous and the following formula holds:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi\big(t,x(t)\big) = \varphi_t\big(t,x(t)\big) + \big(\eta\big(t,x(t)\big), x'(t)\big) \quad \text{a.e. } t \in\,]0,T[, \tag{4.196}$$

where $\eta(t,x) \in \partial\varphi(t,x)$.

(jj) *For each $x_0 \in D(A)$ and $s \in [0,T]$, the Cauchy problem*

$$\begin{aligned} &x' \in Ax + B\partial_p H\big(x, -B^*\partial\varphi(t,x)\big), \quad s \le t \le T,\\ &x(s) = x_0, \end{aligned} \tag{4.197}$$

has at least one solution $x_s \in W^{1,2}([s,T];E) \cap L^\infty(s,T;D(A))$.

Then φ is the optimal value function of problem (4.142) and

$$\Lambda(t,x) = \partial_x H\big(x, -B^*\partial\varphi(t,x)\big)$$

is an optimal feedback control.

Proof Let $y \in W^{1,2}([s,T];E)$ and $v \in L^2(s,T;U)$ be such that $y(s) = h \in D(A)$ and

$$y' = Ay + Bv \quad \text{a.e. } t \in\,]s,T[.$$

By formula (4.196) and (4.167), it follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi\big(t,y(t)\big) = H\big(y(t), -B^*\eta\big(t,y(t)\big)\big) + \big(Bv(t), \eta\big(t,y(t)\big)\big) \ge L\big(y(t),v(t)\big)$$
$$\text{a.e. } t \in\,]s,T[,$$

and, integrating over $[s,T]$, this yields

$$\varphi(s,h) \le \int_s^T L(y,v)\,\mathrm{d}t. \tag{4.198}$$

Hence,

$$\varphi(s,h) \le \tilde{\varphi}(s,h) \quad \text{for all } h \in D(A), \tag{4.199}$$

where $\tilde{\varphi}$ is the value function of problem (4.142). Since $D(A)$ is dense in E and $\varphi, \tilde{\varphi}$ are continuous, inequality (4.199) extends to all of E. Now, let $\tilde{x}_s$ be the solution to the Cauchy problem

$$\begin{aligned} &x' = Ax + B\Lambda(t,x), \quad s \le t \le T,\\ &x(s) = h, \end{aligned}$$

and let $\tilde{u}_s = \Lambda(t,\tilde{x}_s)$ be the corresponding control. We have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\varphi\big(t,\tilde{x}(t)\big) &= \varphi_t\big(t,\tilde{x}_s(t)\big) + \big(\partial\varphi\big(t,\tilde{x}_s(t)\big),\tilde{x}_s'(t)\big)\\ &= \varphi_t\big(t,\tilde{x}_s(t)\big) + \big(A\tilde{x}_s(t),\partial\varphi\big(t,\tilde{x}_s(t)\big)\big) + \big(\partial\varphi\big(t,\tilde{x}_s(t)\big),B\tilde{u} - s(t)\big)\\ &\quad \text{a.e. } t \in \,]s,T[\end{aligned}$$

and, therefore, by (4.167)

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\varphi\big(t,\tilde{x}_s(t)\big) &= H\big(\tilde{x}_s(t),-B^*\partial\varphi\big(t,\tilde{x}_s(t)\big)\big) + \big(B\partial\varphi\big(t,\tilde{x}_s(t)\big),\tilde{u}_s(t)\big)\\ &= -L\big(\tilde{x}_s(t),\tilde{u}_s(t)\big) \quad \text{a.e. } t \in \,]s,T[, \end{aligned}$$

since, by (4.164), $\tilde{u}_s(t) \in \partial_p H(\tilde{x}_s(t),-B^*\partial\varphi(t,\tilde{x}_s(t)))$. Integrating the latter over $[s,T]$, we get

$$\varphi(s,h) = \int_s^T L(\tilde{x}_s,\tilde{u}_s)\,\mathrm{d}t$$

and, therefore,

$$\varphi(s,h) = \tilde{\varphi}(s,h) \quad \text{for all } s \in [0,T],\ h \in E.$$

Thus, φ is the optimal value function of problem (4.142) and $\tilde{u}_s$ is an optimal control on $[s,T]$. □

Remark 4.36 In general, the optimal value function φ, defined by (4.146), is called the *variational solution* to the Hamilton–Jacobi equation (4.187) and (4.188), and Theorem 4.35 amounts to saying that, under the additional assumption (4.166), this is a strong solution.

Let us now take a brief look at some particular cases.

If $E = \mathbb{R}^n$ and $A \equiv 0$, (4.187) reduces to the classical Hamilton–Jacobi equation

$$\begin{aligned} &\frac{\partial}{\partial t}\varphi(t,x) - H\left(x,-\frac{\partial\varphi}{\partial x}(t,x)\right) = 0, \quad t \in [0,T],\ x \in \mathbb{R}^n,\\ &\varphi(T,x) = \varphi_0(x) \quad \text{for } x \in \mathbb{R}^n. \end{aligned} \tag{4.200}$$

It is instructive to notice that the differential systems of characteristics is just the extremality systems in the Hamiltonian form associated to the corresponding problem of the calculus of variations. We refer the reader to the book [33] of P.L. Lions for other existence results on (4.200) and its implications in control theory.

As another example, consider the case of the control problem with quadratic cost criterion, that is,

$$L(x,u)=\frac{1}{2}\left(|Cx|^2+\langle Nu,u\rangle\right),\qquad \varphi_0(x)=\frac{1}{2}(P_0x,x),\quad x\in E,\ u\in U,$$

where $C\in L(E,E)$, $P_0\in L(E,E)$ is symmetric and positive, and N is a self-adjoint positive definite isomorphism on U. It is readily seen that

$$H(x,p)=\frac{1}{2}\left(\langle N^{-1}p,p\rangle-|Cx|^2\right),\quad \text{for all } x\in E \text{ and } p\in U.$$

Thus, the corresponding Hamilton–Jacobi equation is

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\varphi(t,h)-\frac{1}{2}\langle N^{-1}B^*\partial\varphi(t,h),B^*\partial\varphi(t,h)\rangle+\left(Ah,\partial\varphi(t)\right)+\frac{1}{2}|Ch|^2=0\\&\qquad\text{a.e. on }]0,T[,\\&\varphi(T,h)=\frac{1}{2}(P_0h,h),\quad \text{for every } h\in D(A).\end{aligned}\tag{4.201}$$

It is easily seen that $D(\varphi(t,\cdot))=E$. Furthermore, (4.173) and (4.174) show that the operator $h\to p_t(t)=\partial\varphi(t,h)$ is linear and, therefore, self-adjoint on E (see Proposition 2.51). Moreover, we have

$$\varphi(t,h)=\frac{1}{2}\left(P(t)h,h\right)\quad\text{for all } h\in E,\qquad P(t)=\partial\varphi(t).$$

In terms of $P(t)$, (4.201) may be rewritten as

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\left(P(t)h,h\right)-\langle N^{-1}B^*P(t)h,B^*P(t)h\rangle+2\left(Ah,P(t)h\right)+|Ch|^2=0\\&\qquad\text{a.e. on }]0,T[\text{ and for all } h\in D(A).\end{aligned}\tag{4.202}$$

Thus, differentiating formally (4.202) (in the Fréchet sense), we obtain for P an operator differential equation of the following type (the *Kalman–Riccati equation*):

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}P(t)+A^*P(t)+P(t)A-P(t)BN^{-1}B^*P(t)+CC^*=0,\\&P(T)=P_0,\end{aligned}\tag{4.203}$$

whereas the optimal feedback control $u(t)$ is expressed by (see formula (4.164))

$$u(t)=-N^{-1}B^*P(t)x(t),\quad 0\le t\le T.$$

In this context, (4.203) is equivalent to the synthesis of optimal controller for the given problem.

Remark 4.37 Equation (4.167) can be studied in a more general context than that treated here, namely, that of viscosity solutions (see Crandall and Lions [21–24]). The concept of viscosity solutions for (4.167) is a very general one and within this framework existence and uniqueness follow for quite general Hamiltonian functions H. However, one must assume some growth conditions on H which are hard to verify for Hamilton–Jacobi equations of the form (4.167), arising in the synthesis of optimal control problems.

4.2.3 The Dual Hamilton–Jacobi Equation

The duality theorem for the optimal control problem (P) can be used to express the solution φ to the Hamilton–Jacobi equation (4.187) and (4.188) in function of a "dual" Hamilton–Jacobi equation associated with the dual problem (P*). Namely, the optimal value function φ given by (4.146) is, in virtue of Theorem 4.16, equivalently expressed as

$$\begin{aligned}\varphi(t,x) &= -\inf_q\bigg\{\varphi_0^*(-q)+\inf\bigg\{\int_t^T M\big(B^*p(s),w(s)\big)\,\mathrm{d}s+\big(p(t),x\big);\\ &\qquad w\in L^1(t,T;E),\ p'=-A^*p+w \text{ a.e. } s\in(t,T),\ p(T)=q\bigg\}\bigg\}\\ &= -\inf\big\{\varphi_0^*(-q)+\chi(t,q);\ q\in E\big\},\end{aligned}$$

where M is given as in Sect. 4.1.8 and

$$\begin{aligned}\chi(t,q) &= \inf\bigg\{\int_t^T M\big(B^*p(s),w(s)\big)\,\mathrm{d}s+\big(p(t),x\big);\ w\in L^1(t,T;E);\\ &\qquad p'=-A^*p+w \text{ a.e. in } (t,T);\ p(T)=q\bigg\}\\ &= \inf\bigg\{\int_0^{T-t} M\big(B^*p(s),w(s)\big)\,\mathrm{d}s+\big(p(T-t),x\big);\ p'=A^*p+w\\ &\qquad \text{a.e. in } (0,T-t),\ w\in L^1(0,T-t;E),\ p(0)=q\bigg\}\\ &= \inf\bigg\{\int_t^T M\big(B^*z(s),v(s)\big)\,\mathrm{d}s+\big(z(T),x\big);\ z'=A^*z-v\\ &\qquad \text{a.e. } s\in(t,T);\ z(t)=q,\ v\in L^1(t,T;E)\bigg\}.\end{aligned}$$

In other words, $\chi : [0, T] \to E \to \mathbb{R}$ is the *variational solution* to the Hamilton–Jacobi equation

$$\begin{aligned} &\chi_t(t,q) - \tilde{H}\big(q, \chi_q(t,q)\big) + \big(A^*q, \chi_q(t,q)\big) = 0, \\ &\chi(T,q) = (q,x), \end{aligned} \tag{4.204}$$

where $\tilde{H}$ is the Hamiltonian function

$$\tilde{H}(q,v) = \sup\big\{(v,w) - M(B^*q,w);\ w \in E\big\}, \quad \forall (q,v) \in E \times E.$$

We call (4.204) the *dual Hamilton–Jacobi equation* corresponding to (4.167).

We have proved, therefore, the following representation formula for the variational solutions to the Hamilton–Jacobi equation (4.187) and (4.188) (see Barbu and Da Prato [15]).

Theorem 4.38 *Under the above assumptions, the variational solution φ to the Hamilton–Jacobi equation* (4.187) *and* (4.188) *is given by*

$$\varphi(t,x) = -\inf\big\{\varphi_0^*(-q) + \chi(t,q);\ q \in E\big\}, \quad \forall (t,x) \in [0,T] \times E, \tag{4.205}$$

where χ is the variational solution to (4.204).

Now, we consider some particular cases. Let

$$L(x,u) = h(u), \quad \forall (x,u) \in E \times E,$$

where h is a continuous convex function such that

$$h^*(z) = \sup\big\{\langle z,u\rangle - h(u);\ u \in U\big\} < \infty, \quad \forall z \in U.$$

Then

$$M(z,w) = \begin{cases} h^*(z), & \text{if } w = 0, \\ +\infty, & \text{otherwise}, \end{cases} \quad \forall (z,w) \in U \times E.$$

In this case, the Hamilton–Jacobi equation (4.187) and (4.188) has the following form:

$$\begin{aligned} &\varphi_t(t,x) + \big(Ax, \varphi_x(t,x)\big) - h^*\big(-B^*\varphi_x(t,x)\big) = 0, \\ &\varphi(T,x) = \varphi_0(x), \end{aligned}$$

while the corresponding dual equation (4.204) is

$$\begin{aligned} &\chi_t(t,q) + \big(A^*q, \chi_q(t,q)\big) + h^*(B^*q) = 0, \\ &\chi(T,q) = (x,q). \end{aligned}$$

This is a linear first-order partial differential equation which has the solution given by

$$\chi(t,q) = \left(e^{A^*(T-t)}q, x\right) + \int_0^{T-t} h^*\left(B^* e^{A^* s} q\right) ds, \quad \forall (t,q) \in [0,T] \times E,$$

and therefore by (4.205), we have

$$\varphi(t,x) = -\inf_q \left\{ \varphi_0^*(-q) + \left(e^{A^*(T-t)}q, x\right) + \int_0^{T-t} h^*\left(B^* e^{A^* s} q\right) ds \right\}.$$

In the special case $A = 0$, this yields

$$\begin{aligned}
\varphi(t,x) &= -\inf\left\{\varphi_0^*(-q) + (q,x) + (T-t)h^*(B^*q)\right\} \\
&= \sup\left\{(p,x) - \varphi_0^*(p) - (t-T)h^*(-B^*p);\ p \in E\right\}.
\end{aligned}$$

Using the Fenchel duality theorem (see Theorem 3.54), we may equivalently write φ as

$$\varphi(t,x) = \inf\left\{\varphi_0(p) + (T-t)H^*\left(\frac{x-p}{T-t}\right);\ p \in E\right\}, \tag{4.206}$$

where $H(p) = h^*(-B^*p)$ and H^* is the conjugate of H.

With this notation, the function φ is the variational solution to the Hamilton–Jacobi equation

$$\begin{aligned}
&\varphi_t - H(\varphi_x) = 0 \quad \text{in } (0,T) \times E, \\
&\varphi(T,x) = \varphi_0(x).
\end{aligned}$$

Formula (4.206) is known in literature as the *Lax–Hopf formula*.

Assume now that $L(x,u) = \frac{1}{2}|Cx|^2 + \langle Nu, u\rangle$ and $\varphi_0(x) = \frac{1}{2}(P_0 x, x)$. Then, as seen earlier, the Hamilton–Jacobi equation (4.187) reduces to the Riccati equation (4.203) (equivalently, (4.202)) and so, by (4.205), we have (see Barbu and Da Prato [16])

$$\begin{aligned}
\frac{1}{2}\left(P(t)x, x\right) &= -\inf_{q \in E}\left\{\frac{1}{2}\left(P_0^{-1}q, q\right) + \psi(t,q)\right\} \\
&= -\inf_{q \in E}\left\{\frac{1}{2}\left(P_0^{-1}q, q\right) + \frac{1}{2}\left(Q(t)q, q\right) + \left(r(t), q\right) + s(t)\right\},
\end{aligned}$$

where Q is the solution to the equation

$$Q' + AQ + QA^* - QC^*CQ + BB^* = 0, \quad Q(T) = 0,$$

and

$$r'(\tau) + \left(A - Q(\tau)C^*C\right)r(\tau) = 0, \quad \tau \in (t,T),\ r(T) = x,$$

$$s(t) = -\frac{1}{2}\int_t^T \left|CU(T, T+t-\sigma)x\right|^2 \mathrm{d}\sigma,$$

where $U(t,s)$ is the evolution operator generated by $A - Q(\cdot)C^*C$ (see Definition 1.147).

4.3 Boundary Optimal Control Problems

We present here a general formulation for the so-called "boundary optimal control problem" in Hilbert spaces. There are some notable differences between this formulation and that given in Sect. 4.1, and the main one is that the operator B arising in system (4.1) is, in this case, unbounded from U to E. This more general formulation allows us to include boundary controllers u into specific problems involving partial differential equations.

4.3.1 Abstract Boundary Control Systems

Let E and U be a pair of real Hilbert spaces with the norms denoted $|\cdot|$ and $\|\cdot\|$, respectively. Let A be a linear, closed and densely defined operator in E with domain $D(A) \subset E$ and U a linear continuous operator from U to E. Denote by $(\cdot,\cdot)$ and $\langle\cdot,\cdot\rangle$ the scalar product of H and U, respectively.

We assume that:

(i) *A generates a C_0-semigroup $S(t) = \mathrm{e}^{At}$ on E*
(ii) *$D \in L(U, E)$.*

An abstract boundary control system is of the form

$$\begin{aligned} &\frac{\mathrm{d}y}{\mathrm{d}t}(t) = A\big(y(t) - Du(t)\big) - \lambda Du(t) + f(t), \quad t \in (0,T) \\ &y(0) = y_0, \end{aligned} \tag{4.207}$$

or, equivalently,

$$\begin{aligned} &\frac{\mathrm{d}y}{\mathrm{d}t}(t) = Az(t) - \lambda Du(t) + f(t), \quad t \in (0,T), \\ &z(t) = y(t) - Du(t), \quad t \in (0,T), \\ &y(0) = y_0, \end{aligned} \tag{4.208}$$

where $y_0 \in E,\ u \in L^2(0,T;U),\ f \in L^2(0,T;E),\ \lambda \in \rho(A)$ (the resolvent of A).

Formally, the solution y to (4.208) is given by

$$y(t) = \mathrm{e}^{At}y_0 - \int_0^t A\mathrm{e}^{A(t-s)}Du(s)\,\mathrm{d}s + \int_0^t \mathrm{e}^{A(t-s)}\big(-\lambda Du(s) + f(s)\big)\,\mathrm{d}s. \tag{4.209}$$

However, since in general ADu is not in $L^2(0,T;E)$, Formula (4.209) must be taken in the generalized sense to be defined below.

We denote by $(D(A^*))'$ the completion of the space E in the norm

$$|||x||| = \left\|(\lambda I - A^*)^{-1}x\right\|, \quad \forall x \in H.$$

We have of course $E \subset (D(A^*))'$ in the algebraic and topological sense. Then we consider the extension $\tilde{A}$ of A defined from E to $(D(A^*))'$

$$\tilde{A}y(\psi) = (y, A^*\psi), \quad \forall \psi \in D(A^*). \tag{4.210}$$

Then we mean by "mild" solution to (4.207) a $(D(A^*))'$-valued continuous function $y : [0,T] \to (D(A^*))'$ such that

$$\begin{aligned}
\big(y(t), \psi\big) &= \big(\mathrm{e}^{At}y_0, \psi\big) - \int_0^t \big(\mathrm{e}^{A(t-s)}Du(s), A^*\psi\big)\,\mathrm{d}s \\
&\quad + \int_0^t \big(\mathrm{e}^{A(t-s)}\big(-\lambda Du(s) + f(s)\big), \psi\big)\,\mathrm{d}s, \\
&\forall \psi \in D(A^*),\ \forall t \in [0,T].
\end{aligned} \tag{4.211}$$

Equivalently,

$$\begin{aligned}
&\frac{\mathrm{d}y}{\mathrm{d}t} = \tilde{A}y + \big(\tilde{A} - \lambda I\big)Du + f, \quad t \in (0,T), \\
&y(0) = y_0.
\end{aligned} \tag{4.212}$$

In this way, the boundary control system (4.207) can be written in the form (4.212), that is,

$$\begin{aligned}
&\frac{\mathrm{d}y}{\mathrm{d}t} = \tilde{A}y + Bu + f, \quad t \in (0,T), \\
&y(0) = y_0,
\end{aligned} \tag{4.213}$$

where $B = (\tilde{A} - \lambda I)D \in L(U, (D(A^*))')$. Therefore, we may view an abstract boundary control as a control system of the form (4.1), but with an unbounded operator B.

We present below a few specific examples.

1° Dirichlet Boundary Control Consider the control system

$$\begin{aligned}
&\frac{\partial y}{\partial t}(t,x) - \Delta y(t,x) = f(t,x), \quad (t,x) \in (0,T) \times \Omega = Q_T, \\
&y(0,x) = y_0(x), \quad x \in \Omega, \\
&y(t,x) = u(t,x), \quad (t,x) \in \Sigma_T = (0,T) \times \partial\Omega.
\end{aligned} \tag{4.214}$$

Here, Ω is a bounded and open subset of $\mathbb{R}^N$ with smooth boundary $\partial\Omega$. In system (4.214), the control input u is taken in $L^2(\Sigma_T) = L^2(0,T;L^2(\partial\Omega))$,

$f \in L^2(Q_T) = L^2(0, T; L^2(\Omega))$ and $y_0 \in L^2(\Omega)$. In order to write (4.214) in the form (4.207), we set $E = L^2(\Omega)$, $U = L^2(\partial\Omega)$, $\lambda = 0$, $A = \Delta$, $D(A) = H_0^1(\Omega) \cap H^2(\Omega)$ and $D \in L(U, E)$ is the Dirichlet map associated with (4.214), defined by

$$Du = z, \tag{4.215}$$

where $z \in L^2(\Omega)$ is the solution to the nonhomogeneous Dirichlet problem

$$\Delta z = 0 \quad \text{in } \Omega; \qquad z = u \quad \text{on } \partial\Omega. \tag{4.216}$$

The solution z to (4.216) is defined by

$$\int_\Omega z \Delta\varphi \,\mathrm{d}x = \int_{\partial\Omega} u \frac{\partial\varphi}{\partial\nu} \,\mathrm{d}x, \quad \forall \varphi \in H_0^1(\Omega) \cap H^2(\Omega), \tag{4.217}$$

and it is well known (see Lasiecka and Triggiani [31, 32]) that $D \in L(L^2(\partial\Omega), H^{\frac{1}{2}}(\Omega))$. Then $B = \tilde{A}D \in L(L^2(\partial\Omega), (D(A^*))')$ and therefore system (4.214) can be written as (4.212) (equivalently, (4.207)).

We note for later use that the adjoint $B^* \in L(D(A), L^2(\partial\Omega))$ of the operator B is given by

$$B^* y = \frac{\partial y}{\partial\nu}, \quad \forall y \in D(A) = H_0^1(\Omega) \cap H^2(\Omega), \tag{4.218}$$

where $\frac{\partial}{\partial\nu}$ is, as usual, the normal derivative.

2° Neumann Boundary Control Consider system (4.214) with Neumann boundary control, that is,

$$\begin{aligned} &\frac{\partial y}{\partial t} - \Delta y = f \quad \text{in } Q_T, \\ &y(0, x) = y_0(x) \quad \text{in } \Omega, \\ &\frac{\partial y}{\partial\nu} = u \quad \text{on } \Sigma_T, \end{aligned} \tag{4.219}$$

where $u \in L^2(\Sigma_T)$.

In this case, $U = L^2(\partial\Omega)$, $E = L^2(\Omega)$, $A = \Delta$, $D(A) = \{y \in H^2(\Omega);\ \frac{\partial y}{\partial\nu} = 0$ on $\partial\Omega\}$ and λ is any negative number. Then $Du = z$ is the solution to the Neumann boundary problem

$$\Delta y - \lambda y = 0 \quad \text{in } \Omega; \qquad \frac{\partial y}{\partial\nu} = u \quad \text{on } \partial\Omega$$

and, as is easily seen, we have $D \in L(L^2(\partial\Omega), H^1(\Omega))$ and

$$B^* y = y|_{\partial\Omega}, \quad \forall y \in D(A^*) = D(A).$$

A simpler and more convenient way to represent (4.219) as an abstract boundary control system is to write it as

$$\begin{aligned}&\frac{\mathrm{d}y}{\mathrm{d}t} = A_0 y + Bu + f \quad \text{a.e. } t \in (0,T),\\ &y(0) = y_0,\end{aligned} \tag{4.220}$$

where the operator $A_0 \in L(V, V')$, $V = H^1(\Omega)$, $V' = (H^1(\Omega))'$ is defined by

$$(A_0 y, \psi) = \int_\Omega \nabla y \cdot \nabla \psi \,\mathrm{d}x, \quad \forall \psi \in V,$$

and $B \in L(L^2(\Omega), V')$ is given by $(By, \psi) = \int_{\partial\Omega} u\psi \,\mathrm{d}x$, $\forall \psi \in V$, $y \in V$.

3° The Oseen–Stokes Boundary Control System Consider the linear system

$$\begin{aligned}&\frac{\partial y}{\partial t} - \nu_0 \Delta y + (a \cdot \nabla)y + (y \cdot \nabla)b = \nabla p + f \quad \text{in } (0,T) \times \Omega = Q_T,\\ &\nabla \cdot y = 0 \quad \text{on } Q_T,\\ &y(0,x) = 0 \quad \text{in } \Omega,\\ &y(t,x) = u(t,x) \quad \text{on } \Sigma_T.\end{aligned} \tag{4.221}$$

Here, $y = \{y_1, \dots, y_n\}$, $\nabla \cdot y = \operatorname{div} y$, $a, b \in (H^2(\Omega))^n$, $\nu_0 > 0$ and $\Omega \subset \mathbb{R}^n$ is a bounded and open domain with smooth boundary $\partial\Omega$. System (4.221) describes the dynamics of an incompressible fluid improving the classical Stokes model. In the special case $a = y_e$, $b = y_e$, this system arises by the linearization of the classical Navier–Stokes equation

$$\begin{aligned}&\frac{\partial y}{\partial t} - \nu_0 \Delta y + (y \cdot \nabla)y = \nabla p + f \quad \text{in } Q_T,\\ &\nabla \cdot y = 0 \quad \text{on } Q_T,\\ &y(0,x) = 0 \quad \text{in } \Omega,\\ &y = u \quad \text{on } Q_T,\end{aligned} \tag{4.222}$$

around the stationary solution $y_e \in (H^2(\Omega)) \cap (H_0^1(\Omega))^n$, $\nabla \cdot y_e = 0$.

We set $E = \{y \in (L^2(\Omega))^n; \nabla \cdot y = 0, y \cdot \nu = 0 \text{ on } \partial\Omega\}$ (the space of free divergence vectors on Ω) and denote by A the operator

$$Ay = P\big(\nu_0 \Delta y - (y_e \cdot \nabla)y - (y \cdot \nabla)y_e\big) \tag{4.223}$$

with the domain $D(A) = \{y \in E; y \in (H_0^1(\Omega))^n \cap (H^2(\Omega))^n\}$. Here, P is the Leray projection of E on $(L^2(\Omega))^n$ (see, e.g., [13], p. 251) and ν is, as usual, the normal to $\partial\Omega$.

Consider the operator $D : U \to E$ defined by

$$-\nu_0 \Delta Du + (a \cdot \nabla)Du + (Du \cdot \nabla)b + \lambda Du = \nabla p \quad \text{in } \Omega,$$
$$\nabla \cdot Du = 0, \qquad Du = u \quad \text{on } \partial\Omega,$$

where $U = \{u \in (L^2(\partial\Omega))^n,\ u \cdot \nu = 0 \text{ on } \partial\Omega\}$ and $\lambda > 0$ is sufficiently large. Then we have $D \in L(U, E)$.

System (4.221) can be written as (4.212), where $\tilde{A} : E \to (D(A^*))'$ is given by (4.210), that is,

$$\begin{aligned}
(\tilde{A}y, \psi) &= \int_\Omega y A^* \psi \, \mathrm{d}x \\
&= -\int_\Omega y_j \big(\nu_0 \Delta \psi_j - D_i\big((y_e)_i \psi_j\big) + D_j\big((y_e)_i \psi_i\big)\big) \, \mathrm{d}x, \\
&\forall \psi \in D(A^*) = D(A),
\end{aligned} \tag{4.224}$$

and

$$B = (\tilde{A} - \lambda I) D : L^2(\partial\Omega) \to E.$$

We have

$$B^* y = \nu_0 \frac{\partial y}{\partial \nu}, \quad \forall y \in D(A). \tag{4.225}$$

Now, coming back to the "mild" solution y to system (4.207), by (4.211) we see that, for $u \in L^p(0, T; U)$, $y : [0, T] \to (D(A^*))'$ is in $L^p(0, T; (D(A^*))' \cap C_w([0, T]; (D(A^*))')$, that is, y is weakly $(D(A^*))'$-valued continuous. However, under additional assumptions on A, the "mild" solution is in $L^p(0, T; E)$. This happens, for instance, if besides (i) and (ii) the operator A satisfies also the following assumption.

(iii) *A is infinitesimal generator of a C_0-analytic semigroup e^{At} and there is $\gamma \in L^1(0, T)$ such that*

$$\left\| B^* \mathrm{e}^{A^* t} \right\|_{L(U,E)} \le \gamma(t), \quad \forall t \in (0, T), \tag{4.226}$$

where $B = (\tilde{A} - \lambda I) \in L(U, (D(A^*))')$.

Then we have the following proposition.

Proposition 4.39 *Under assumptions* (i)–(iii), *the mild solution y to* (4.207) *is in $L^p([0, T]; E)$ if $u \in L^p(0, T; U)$. Moreover, if $\gamma \in L^{p'}(0, T)$, $\frac{1}{p} + \frac{1}{p'} = 1$, then $y \in C([0, T]; E)$.*

Proof Consider the function

$$(L_T u)(t) = \int_0^t \mathrm{e}^{A(t-s)} (\tilde{A} - \lambda) Du(s) \, \mathrm{d}s, \quad t \in [0, T]$$

which, for each $u \in L^p(0,T;U)$ is well defined from $[0,T]$ to $(D(A^*))'$ and belongs to $C([0,T];(D(A^*))')$. Moreover, we have

$$\big(L_T u(t), \psi\big) = \int_0^t \big(u(s), B^* e^{A^*(t-s)}\psi\big)\, ds, \quad \forall \psi \in L^{p'}(0,T;E).$$

By (4.226) and the Young inequality, the latter yields

$$\big|\big(L_T u(t), \psi\big)\big|_{L^1(0,T)} \le C_T \|u\|_{L^p(0,T;U)} \|\psi\|_{L^{p'}(0,T;E)}$$

and, therefore,

$$\|L_T u\|_{L^p(0,T;E)} \le C_T \|u\|_{L^p(0,T;U)}.$$

Hence, $L_T u \in L^p(0,T;E)$, as claimed. □

Assume now that $\gamma \in L^{p'}(0,T)$. We have, as above, $L_T u \in L^\infty(0,T;E)$ and

$$\begin{aligned}
&\big|L_T u(t+\varepsilon) - L_T u(t)\big| \\
&\quad \le \left|\int_t^{t+\varepsilon} e^{A(t+\varepsilon-s)} B u(s)\, ds\right| + \left|\big(e^{A\varepsilon} - I\big)\int_0^t e^{A(t-s)} B u(s)\, ds\right| \\
&\quad \le C \int_t^{t+\varepsilon} \big\|u(s)\big\|_U \gamma(t+\varepsilon-s)\, ds + \left|\big(e^{A\varepsilon}-I\big)\int_0^t e^{A(t-s)} B u(s)\right| ds \\
&\quad \le C\left(\int_t^{t+\varepsilon} \big\|u(s)\big\|_U^p\, ds\right)^{\frac{1}{p}} \left(\int_t^{t+\varepsilon} \big(\gamma(t+\varepsilon-s)\big)^{p'}\, ds\right)^{\frac{1}{p'}} \\
&\qquad + \left|\big(e^{A\varepsilon} - I\big)\int_0^t e^{A(t-s)} B u(s)\, ds\right| \to 0,
\end{aligned}$$

as $\varepsilon \to 0$, because $\gamma \in L^{p'}(0,T)$, $u \in L^p(0,T;U)$ and $L_T u \in L^\infty(0,T;E)$. Hence, $L_T u \in C([0,T];E)$. Taking into account that, by assumption (4.226), we have

$$\big\|e^{At} B\big\|_{L(U,E)} \le \gamma(t), \quad \forall t \in (0,T],$$

it follows by Proposition 4.39 that the "mild" solution $y \in L^p(0,T;E)$ to (4.207) can be equivalently expressed as

$$y(t) = e^{At} y_0 + \int_0^t e^{A(t-s)} B u(s)\, ds, \quad \text{a.e. } t \in (0,T). \tag{4.227}$$

Let us check the key assumption (iii) in the examples considered above.

In the case of Dirichlet, for the boundary control system (4.214), by (4.218), we see that (4.226) reduces to

$$\left\|\frac{\partial y}{\partial \nu}(t)\right\|_{L^2(\partial\Omega)} \le \gamma(t)\|y_0\|_{L^2(\Omega)}, \quad \forall y_0 \in L^2(\Omega), \tag{4.228}$$

where $y(t) = e^{At} y_0$ is the solution to the linear equation

$$\frac{\partial y}{\partial t} - \Delta y = 0 \quad \text{in } (0,T) \times \Omega = Q_T,$$
$$y = 0 \quad \text{on } (0,T) \times \partial\Omega = \Sigma_T,$$
$$y(0,x) = y_0(x), \qquad x \in \Omega.$$

Let us check that (4.228) holds with $\gamma(t) = Ct^{-\frac{3}{4}}$. Indeed, by Green's formula, we have for all $\varphi \in H^2(\Omega)$,

$$\int_{\partial\Omega} \frac{\partial}{\partial\nu} y(x,t)\varphi(x)\,dx = \int_{\Omega} \frac{\partial}{\partial t}\, y(x,t)\varphi(x)\,ds = \int_{\Omega} Ay(x,t)\varphi(x)\,dx. \tag{4.229}$$

Let us denote by A^α, $0 < \alpha < 1$, the fractional power of the operator A. A^α is defined by (see, e.g., Yosida [48], p. 260)

$$(-A)^\alpha x = -\frac{\sin\alpha\pi}{\pi} \int_0^\infty \lambda^{\alpha-1}(\lambda I - A)^{-1} Ax\,d\lambda, \quad \forall x \in D(A).$$

Then $H^{\frac{1}{2}}(\Omega) \subset D(A^\alpha)$ for $0 < \alpha < \frac{1}{4}$ (see Lions–Magenes [34], Lasiecka and Triggiani [31]) and by (4.229) we see that

$$\left|\int_{\partial\Omega} \frac{\partial y}{\partial\nu}\,(x,t)\varphi(x)\,dx\right| \le \int_{\Omega} \left|A^{1-\alpha} y(x,t)\right| \cdot \left|A^\alpha \varphi(x)\right| dx$$
$$\le C\left\|A^{1-\alpha} y(t)\right\|_{L^2(\Omega)} \|\varphi\|_{H^{\frac{1}{2}}(\Omega)}, \quad \forall \varphi \in H^2(\Omega).$$

On the other hand, we have the interpolation inequality

$$\left|A^{1-\alpha} y\right| \le C|Ay|^{1-\alpha}|y|^\alpha, \quad \forall y \in D(A).$$

Since $y(t) = e^{At} y_0$ is an analytic semigroup, we have

$$\left|Ay(t)\right| \le \frac{C}{t}, \quad \forall t > 0,$$

and, therefore,

$$\left\|A^{1-\alpha} y(t)\right\|_{L^2(\Omega)} \le Ct^{\alpha-1}\|y_0\|_{L^2(\Omega)}, \quad \forall y_0 \in L^2(\Omega).$$

In virtue of the trace theorem, this yields

$$\left|\int_{\partial\Omega} \frac{\partial y}{\partial\nu}(x,t)u(x)\,dx\right| \le Ct^{-\frac{3}{4}}\|y_0\|_{L^2(\Omega)}\|u\|_{L^2(\partial\Omega)}, \quad \forall u \in L^2(\partial\Omega),$$

which implies the desired inequality (4.228) with $\gamma(t) = Ct^{-\frac{3}{4}}$.

In particular, it follows by Proposition 4.39 that, for each $u \in L^p(0,T;L^2(\partial\Omega))$, the solution y to (4.228) is in $L^p(0,T;L^2(\Omega))$ and, if $u \in L^4(0,T;L^2(\partial\Omega))$, then $y \in C([0,T];L^2(\Omega))$.

Consider now the Neumann boundary control system (4.219). Then Assumption (4.226) is reduced to

$$\left(\int_{\partial\Omega} \left|y(t,x)\right|^2 \mathrm{d}x\right)^{\frac{1}{2}} \le \gamma(t)\|y_0\|_{L^2(\Omega)}, \quad \forall t \in [0,T],$$

where $y(t) = \mathrm{e}^{At} y_0$ is the solution to

$$\begin{aligned}
&\frac{\partial y}{\partial t} - \Delta y = 0 \quad \text{in } (0,T)\times\Omega,\\
&y(0) = u_0 \quad \text{in } \Omega,\\
&\frac{\partial y}{\partial \nu} = 0 \quad \text{on } (0,T)\times\partial\Omega.
\end{aligned}$$

Since, as is easily seen,

$$\int_0^T \|y(t)\|^2_{H^1(\Omega)} \le \int_0^T \int_\Omega |\nabla y|^2 \,\mathrm{d}s\,\mathrm{d}x \le \frac{1}{2}\|y_0\|^2_{L^2(\Omega)},$$

we get by the trace theorem that

$$\left(\int_0^T \int_{\partial\Omega} \left|y(s,x)\right|^2 \mathrm{d}x\right)^{\frac{1}{2}} \le \sqrt{\frac{1}{2}}\|y_0\|_{L^2(\Omega)}$$

and, therefore, (4.226) holds with $\gamma \in L^2(0,T)$. This implies that the solution y to system (4.219) is in $C([0,T];L^2(\Omega))$ for $u \in L^2(0,T;L^2(\partial\Omega))$.

Consider now the Oseen–Stokes equation (4.221). By (4.225), we have, as above, that

$$\|B^*\mathrm{e}^{A^*t}y_0\|_{(L^2(\partial\Omega))^n} = \nu_0 \left\|\frac{\partial}{\partial\nu}\mathrm{e}^{A^*t}y_0\right\|_{(L^2(\partial\Omega))^n},$$

while, by the trace theorem,

$$\begin{aligned}
\left\|\frac{\partial}{\partial\nu}\mathrm{e}^{A^*t}y_0\right\|_{(L^2(\partial\Omega))^n} &\le C\|\mathrm{e}^{A^*t}y_0\|_{(H^{\frac{3}{2}}(\Omega))^n} = C\left|(A^*)^{\frac{3}{4}}\mathrm{e}^{A^*t}y_0\right|\\
&\le Ct^{-\frac{3}{4}}\|y_0\|_{(L^2(\Omega))^n}.
\end{aligned}$$

Hence, condition (4.226) holds with $\gamma(t) = Ct^{-\frac{3}{4}}$ and we have the same conclusion as in the case of the parabolic system (4.214).

Remark 4.40 The abstract formulation (4.212) includes besides linear parabolic boundary control systems of the type presented above also linear systems with singular distributed controllers. For instance, the parabolic control system

$$\frac{\partial y}{\partial t} - \Delta y = u\mu \quad \text{in } (0,T)\times\Omega,$$
$$y = 0 \quad \text{on } (0,T)\times\partial\Omega, \qquad y(0) = y_0 \quad \text{in } \Omega,$$

where $u \in L^2(0,T)$ and $\mu \in (H_0^1(\Omega)\cap H^2(\Omega))'$ can be represented as (4.12), where $A = -\Delta$, $D(A) = H_0^1(\Omega)$, $U = R$. In particular, if $1 \le n \le 3$, one might take $\mu = \delta(x_0)$ (the Dirac distribution concentrated in $x_0 \in \Omega$). The latter is the case of a pointwise controlled system.

4.3.2 The Boundary Optimal Control Problem

We study here the following unconstrained optimal control problem.

$$\begin{aligned}
&\text{Minimize } \int_0^T L\big(t, y(t), u(t)\big)\,\mathrm{d}t + \ell\big(y(0), y(T)\big) \\
&\text{over all } y \in C\big([0,T]; E\big) \text{ and } u \in L^p(0,T;U) \\
&\text{subject to state equation (4.207) (equivalently, (4.227)).}
\end{aligned} \tag{4.230}$$

Here, $p \in [2,\infty[$ and $L : (0,T)\times E\times U \to \overline{\mathbb{R}}$, $\ell : E\times E \to \overline{\mathbb{R}}$ are convex and lower-semicontinuous functions to be made precise below. We assume that (iii) holds with $\gamma \in L^{p'}(0,T)$, where $\frac{1}{p'} = 1 - \frac{1}{p}$ and so, by Proposition 4.39, $y \in C([0,T];E)$. This gives a meaning to $\ell(y(0), y(T))$.

An end-point pair $(y_1, y_2) \in E\times E$ is called *attainable* for problem (4.230) if there exists $y \in C([0,T];E)$ and $u \in L^p(0,T;U)$ satisfying equation (4.207) (in the "mild" sense (4.227)) and such that $L(t,y,u) \in L^1(0,T)$, $y(0) = y_1$, $y(T) = y_2$. The set of all attainable pairs will be denoted by K_L.

We are now ready to formulate the main result of this section.

Theorem 4.41 *Assume that the functions $L(t)$ and ℓ satisfy Hypotheses* (C) *and* (E) *in Sect.* 4.1, *where K_L was defined above. Then a given pair (y^*, u^*) is optimal in problem* (4.230) *if and only if there exist functions $p^* \in C([0,T];E)$ and $q \in L^1(0,T;E)$ satisfying along with y^* and u^* the system*

$$p^{*\prime} = -A^* p^* + q, \quad t \in [0,T], \tag{4.231}$$

$$\big(q(t), B^* p^*(t)\big) \in \partial L\big(t, y^*(t), u^*(t)\big) \quad \textit{a.e. } t \in\,]0,T[, \tag{4.232}$$

$$\big(p^*(0), -p^*(T)\big) \in \partial\ell\big(y^*(0), y^*(T)\big), \tag{4.233}$$

and such that $B^* p^* \in L^{p'}(0,T;U)$.

Equation (4.231) must be, of course, considered in the following "mild" sense:

$$p^*(t) = S^*(T-t)p^*(T) - \int_t^T S^*(s-t)q(s)\,\mathrm{d}s, \quad 0 \le t \le T,$$

where $S^*(t) = (S(t))^* = \mathrm{e}^{A^*t}$, $t \ge 0$.

Let us briefly present a few examples.

Example 4.42 Minimize

$$\int_Q g(x,y)\,\mathrm{d}x\,\mathrm{d}t + \int_\Sigma h(u)\,\mathrm{d}x\,\mathrm{d}t + \frac{1}{2}\int_\Omega \left|y(T,s) - \xi(x)\right|^2 \mathrm{d}x \tag{4.234}$$

in $y \in C([0,T]; L^2(\Omega))$ and $u \in L^p(0,T; L^2(T))$ subject to

$$\begin{aligned} &y_t - \Delta y = 0 \quad \text{in } Q =]0,T[\times \Omega, \\ &y = u \quad \text{on } \Sigma =]0,T[\times \partial\Omega, \\ &y(0,x) = u_0(x), \quad x \in \Omega. \end{aligned} \tag{4.235}$$

Here, Ω is an open domain of $\mathbb{R}^n$ with a smooth boundary $\partial\Omega$ and $\xi \in L^2(\Omega)$ is a given function. The function $g : \Omega \times \mathbb{R} \to \mathbb{R}$ is continuous and convex in y, measurable in x, and satisfies

$$\left|g(x,y)\right| \le C|y|^2 + \zeta(x) \quad \text{a.e. } x \in \Omega,\ y \in \mathbb{R},$$

where $\zeta \in L^1(\Omega)$. As regards the function $h : R \to \overline{\mathbb{R}}^*$, it is assumed convex, lower-semicontinuous and satisfying the growth condition

$$h(u) \ge C_1|u|^2 + C_2 \quad \text{for all } u \in \mathbb{R},$$

where $C_1 > 0$.

Theorem 4.41 is applicable with $E = L^2(\Omega)$, $U = L^2(\partial\Omega)$, $A = \Delta$, $D(A) = H_0^1(\Omega) \cap H^2(\Omega)$, B defined as in Example 4.42 and

$$\begin{aligned} \ell(y_1,y_2) &= \frac{1}{2}\int_\Omega \left|y_2(x) - \xi(x)\right|^2 \mathrm{d}x \quad \text{if } y_1 = y_0 \quad \text{and} \\ &= +\infty \quad \text{if } y_1 \ne y_0, \\ L(t,y,u) &= \int_\Omega g(x,y)\,\mathrm{d}x + \int_{\partial\Omega} h(u)\,\mathrm{d}x. \end{aligned}$$

According to estimate (4.228), we should choose $p > 4$. Thus, recalling (4.218), by Theorem 4.41, *the pair* (y^*, u^*) *is optimal in problem* (4.234) *if and only if there* *exist* $p^* \in C([0,T]; L^2(\Omega))$ *and* $q \in L^1(0,T; L^2(\Omega))$ *satisfying the system*

$$\begin{aligned}
&p_t^* + \Delta p^* = q \quad \text{in } Q,\\
&q(t,x) \in \partial_y g\big(x, y^*(x,t)\big) \quad \text{a.e. } (t,x) \in Q,\\
&p^* = 0 \quad \text{on } \Sigma,\\
&\frac{\partial p^*}{\partial \nu} \in \partial h(u^*) \quad \text{a.e. in } \Sigma,\\
&y^*(0,x) = y_0(x), \qquad p^*(T,x) + y^*(T,x) = \xi(x) \quad \text{a.e. } x \in \Omega.
\end{aligned} \tag{4.236}$$

Example 4.43 We now present an optimal control problem in fluid dynamics:

$$\begin{aligned}
&\operatorname{Min}\left\{\frac{1}{2}\int_0^T \int_\Omega \big|y(x,t)\big|^2\,\mathrm{d}t + \int_0^T \int_{\partial\Omega} \big|u(x,t)\big|^2\,\mathrm{d}x\,\mathrm{d}t\right\} + \frac{1}{2}\int_\Omega \big|y(x,t)\big|^2\,\mathrm{d}x\\
&\text{subject to (4.221).}
\end{aligned} \tag{4.237}$$

In the context of fluid dynamics governed by the Oseen–Stokes system (4.221), problem (4.237) expresses the regulation of the turbulent kinetic energy of the fluid through the boundary control u.

The existence and uniqueness of an optimal pair (y^*, u^*) is immediate. As regards the first-order optimality conditions, by Theorem 4.41 we have

$$u^* = \nu_0 \frac{\partial q^*}{\partial \nu} \quad \text{in } (0,T) \times \Omega, \tag{4.238}$$

where q^* is the solution to the adjoint system

$$\begin{aligned}
&\frac{\partial q^*}{\partial t} - \nu_0 \Delta q^* - (\nabla \cdot a) q^* + \nabla(b \cdot q^*) = \nabla p + y^* \quad \text{in } (0,T) \times \Omega,\\
&\nabla \cdot q^* = 0, \quad \text{in } (0,T) \times \Omega,\\
&q^* = 0 \quad \text{on } (0,T) \times \partial\Omega,\\
&q^*(T,x) = -y^*(T,x) \quad \text{in } \Omega.
\end{aligned} \tag{4.239}$$

4.3.3 *Proof of Theorem 4.41*

Since the proof follows closely that of Theorem 4.5, it is only sketched.

Let $(y^*, u^*) \in C([0,T];E) \times L^p(0,T;U)$ be an optimal pair in problem (4.230). For any $\lambda > 0$, consider the approximating control problem

$$\text{Minimize}\left\{ \int_0^T \left(L_\lambda(t,y,u) + p^{-1}\|u-u^*\|^p\right) \mathrm{d}t + \ell_\lambda\left(y(0), y(T)\right) \right.$$
$$\left. + \frac{1}{2}\left|y(0) - y^*(0)\right|^2 \right\} \tag{4.240}$$

over all $(y,u) \in C([0,T];E) \times L^p(0,T;U)$ subject to (4.207). It follows, as in the proof of Theorem 4.5, that problem (4.240) has a unique optimal solution (y_λ, u_λ) and $\partial L_\lambda(t, y_\lambda, u_\lambda) \in L^p(0,T;E) \times L^p(0,T;U)$. Let $p_\lambda \in C([0,T];E)$ be defined by

$$p_\lambda(t) = S^*(T-t)p_\lambda(T) - \int_t^T S^*(s-t)\partial_y L_\lambda\left(s, y_\lambda(s), u_\lambda(s)\right) \mathrm{d}s. \tag{4.241}$$

Next, since (y_λ, u_λ) is optimal, we have

$$\int_0^T \left(\left(\partial_y L(t,y_\lambda,u_\lambda), y\right) + \left\langle \partial_u L_\lambda(t,y_\lambda,u_\lambda) + \|u_\lambda - u^*\|^{p-2}(u_\lambda - u^*), v\right\rangle\right) \mathrm{d}t$$
$$+ \left(\partial\ell_\lambda\left(y_\lambda(0), y_\lambda(T)\right), \left(y(0), y(T)\right)\right) + \left(y_\lambda(0) - y^*(0), y(0)\right) = 0, \tag{4.242}$$

for all $v \in L^p(0,T;U)$ and $y \in C([0,T];E)$ satisfying (4.207), where $f = 0$. Then, after some calculations involving Fubini's theorem, we get

$$B^* p_\lambda - \|u^* - u_\lambda\|^{p-2}(u_\lambda - u^*) = \partial_u L_\lambda(t, y_\lambda, u_\lambda), \quad \text{a.e. } t \in]0,T[. \tag{4.243}$$

By (4.242) and (4.243), we also have

$$\left(p_\lambda(0) + y^*(0) - y_\lambda(0), -p_\lambda(T)\right) = \partial\ell_\lambda\left(y_\lambda(0), y_\lambda(T)\right). \tag{4.244}$$

We have

$$\int_0^T \left(\dot{L}_\lambda(t,y_\lambda,u_\lambda) + p^{-1}\|u_\lambda - u^*\|^p\right) \mathrm{d}t + \ell_\lambda\left(y_\lambda(0), y_\lambda(T)\right) + \frac{1}{2}\left|y_\lambda(0) - y^*(0)\right|^2$$
$$\le \int_0^T L(t, y^*, u^*)\,\mathrm{d}t + \ell\left(y^*(0), y^*(T)\right),$$

and thus all u_λ remain in a bounded subset of $L^p(0,T;U)$. Then arguing as in the proof of Lemma 4.8, we find that for $\lambda \to 0$

$$u_\lambda \to u^* \quad \text{strongly in } L^p(0,T;U), \tag{4.245}$$
$$y_\lambda \to y^* \quad \text{in } C\left([0,T];E\right). \tag{4.246}$$

Similarly, by the same reasoning as in the proof of Lemma 4.9, we infer that

$$\left|p_\lambda(T)\right| \le C, \quad 0 < \lambda \le 1. \tag{4.247}$$

Next, according to Assumption C, there exist functions $\alpha, \beta \in L^p(0,T)$ and $v_h : [0,T] \to U$ such that $\|v_h(t)\| \le \beta(t)$, a.e. $t \in]0,T[$ and, for all $h \in E$, $|h| = 1$,

$$L_\lambda\big(t, y^*(t) + \rho h, v_j(t)\big) \le L\big(t, y^*(t) + \rho h, v_j(t)\big) \le \alpha(t).$$

This yields

$$\begin{aligned}\big|\partial_y L_\lambda(t, y_\lambda, u_\lambda)\big| \le C\big(\beta(t) + \big\|u_\lambda(t)\big\|\big)\big(\big\|u^*(t) - u_\lambda(t)\big\|^{p-1} \\ + \big\|B^* p_\lambda(t)\big\|\big) + \delta(t) \quad \text{a.e. } t \in]0,T[,\end{aligned} \tag{4.248}$$

where $\delta \in L^p(0,T)$. We set $q_\lambda = \partial_y L_\lambda(t, y_\lambda, u_\lambda)$. By (4.241) and (4.248), we have

$$\begin{aligned}\big\|B^* p_\lambda(t)\big\| \le C\bigg(\zeta(T-t) + \int_t^T \zeta(s-t)\big(\beta(s) + \big\|u_\lambda(s)\big\|\big) \\ \times \big(\big\|B^* p_\lambda(s)\big\| + \big\|u^*(s) - u_\lambda(s)\big\|^{p-1}\big)\,\mathrm{d}s + 1\bigg).\end{aligned} \tag{4.249}$$

Next, by Young's inequality, we have for $\tilde{p}_\lambda(s) = p_\lambda(T-s)$, $\tilde{u}_\lambda(s) = u_\lambda(T-s)$

$$\begin{aligned}&\bigg(\int_0^\nu \bigg(\int_0^t \zeta(s-t)\big\|\tilde{u}_\lambda(s)\big\|\big\|B^* \tilde{p}_\lambda(s)\big\|\,\mathrm{d}s\bigg)^{p'}\mathrm{d}t\bigg)^{\frac{1}{p'}} \\ &\quad \le \bigg(\int_0^\nu \big|\zeta(t)\big|^{p'}\,\mathrm{d}t\bigg)^{\frac{1}{p'}} \int_0^\nu \big\|\tilde{u}_\lambda(s)\big\|\big\|B^* \tilde{p}_\lambda(s)\big\|\,\mathrm{d}s \\ &\quad \le \eta(\nu)\bigg(\int_0^\nu \big\|B^* p_\lambda(t)\big\|^{p'}\,\mathrm{d}t\bigg)^{\frac{1}{p}}, \quad 0 \le \nu \le T,\end{aligned}$$

where $\lim_{t\to 0} \eta(t) = 0$.

We may, therefore, conclude by (4.249) that $\{\int_{T-\nu}^T \|B^* p_\lambda\|^{p'}\,\mathrm{d}t\}$ is bounded for some positive constant ν. Then, by (4.241), we see that $|p_\lambda(t)|$ are uniformly on $[T-\nu, T]$. Now, reasoning as above, with T replaced by $T-\nu$, we find after several steps that $\{B^* p_\lambda\}$ is bounded in $L^{p'}(0,T;U)$ and

$$\big|p_\lambda(t)\big| \le C, \quad \forall t \in [0,T]. \tag{4.250}$$

It should be observed by (4.248) that $\{q_\lambda\} \subset L^1(0,T;E)$ is bounded and $\{\int_\Omega q_\lambda\,\mathrm{d}t; \Omega \subset [0,T]\}$ are uniformly absolutely continuous. Then, according to the Dunford–Pettis criterion, $\{q_\lambda\}$ is a weakly compact subset of $L^1(0,T;E)$. Thus, extracting a subsequence if necessary, we may assume that

$$\begin{aligned} q_\lambda \to q \quad &\text{weakly in } L^1(0,T;E), \\ p_\lambda(T) \to p_T \quad &\text{weakly in } E, \\ p_\lambda \to p \quad &\text{weak-star in } L^\infty(0,T;E), \\ B^* p_\lambda \to B^* p \quad &\text{weakly in } L^{p'}(0,T;U). \end{aligned} \tag{4.251}$$

It follows by (4.248) and (4.251) that $q \in L^p(0,T;E)$, while by (4.241)

$$p_\lambda(t) \to p(t) = S^*(T-t)p_T - \int_t^T S^*(s-t)q(s)\,\mathrm{d}s \quad \text{for } t \in [0,T]$$

in the weak topology of E.

Since $y_\lambda(t) \to y^*(t)$ uniformly on $[0,T]$, we may pass to the limit into (4.244) to get

$$\big(p(0), -p(T)\big) \in \partial\ell\big(y^*(0), y^*(T)\big).$$

Similarly, by (4.243), (4.245), (4.246), and (4.251), it follows that

$$\big(q(t), B^*p(t)\big) \in \partial L\big(t, y^*(t), u^*(t)\big), \quad \text{a.e. } t \in \,]0,T[,$$

as claimed. This concludes the proof of Theorem 4.41.

Remark 4.44 A duality theory for problem (4.224) could be developed following the pattern of Sect. 4.1.8, but the details are left to the reader.

4.4 Optimal Control Problems on Half-Axis

We study here Problem (P) on the half-axis $\mathbb{R}^+ = (0,\infty)$ and its implication in the stabilization of linear systems. It is apparent from the previous development that, in this framework, an existence theory for problem (P), as well as the maximum principle type result, requires some stabilizability assumption on the pair (A,B).

4.4.1 Formulation of the Problem

We are given two real and reflexive Banach spaces E and U which are strictly convex along with their duals E^* and U^*. $A : D(A) \subset E \to E$ is the infinitesimal generator of a C_0-semigroup $\{S(t); t \geq 0\}$ on E. Then the adjoint operator A^* generates the dual semigroup $S^*(\cdot)$.

Now, we consider the linear evolution Cauchy problem

$$\begin{aligned} &x'(t) = Ax(t) + Bu(t), \quad t \geq 0,\\ &x(0) = x_0, \end{aligned} \tag{4.252}$$

where B is a linear continuous operator from U into E and $u : \mathbb{R}^+ \to U$ is a given integrable function. As in the previous sections, by a solution to (4.252) we here mean a "mild" solution, that is,

$$x(t) = S(t)x_0 + \int_0^t S(t-s)Bu(s)\,\mathrm{d}s, \quad t \geq 0. \tag{4.253}$$

The problem to be studied is that of minimizing

$$(\mathrm{P}_\infty) \qquad \int_0^\infty L\big(x(t), u(t)\big)\,\mathrm{d}t$$

in $u \in L^2_{\mathrm{loc}}(\mathbb{R}^+; U)$ and $x \in C(\mathbb{R}^+; E)$ subject to (4.252). Here, x_0 is fixed in E and $L : E \times U \to \mathbb{R}^+$ is a lower-semicontinuous convex function satisfying the condition

$$L(x, u) \geq C_1\|u\|^2 + C_2|x|^2 + C_3 \quad \text{for } (x, u) \in E \times U, \tag{4.254}$$

where $C_1 > 0$ and C_2, C_3 are real constants.

For any $x_0 \in E$ and $u \in L^2_{\mathrm{loc}}(\mathbb{R}^+; U)$, we denote by $x(t, x_0, u)$ the corresponding solution to (4.253). Let $G : L^2_{\mathrm{loc}}(\mathbb{R}^+; U) \times E \to \overline{\mathbb{R}}^*$ be defined by

$$G(u, x_0) = \int_0^\infty L\big(x(t, x_0, u), u(t)\big)\,\mathrm{d}t.$$

Inasmuch as $L \geq 0$, we infer that $G(u, x_0)$ is well defined (unambiguously either a real number or $+\infty$) for each $(u, x_0) \in L^2_{\mathrm{loc}}(\mathbb{R}; U) \times E$. In terms of the functional G, Problem (P_∞) can equivalently be written as

$$\min\big\{G(u, x_0);\ u \in L^2_{\mathrm{loc}}\big(\mathbb{R}^+; U\big)\big\} = \psi(x_0). \tag{4.255}$$

As in the case of optimal control problems on finite intervals, a function u for which the infimum in (4.255) is attained is referred to as an *optimal arc*.

Let $\mathrm{Dom}(\psi)$ be the effective domain of the function $\psi : E \to \mathbb{R}^+$.

Proposition 4.45 *For every $x_0 \in \mathrm{Dom}(\psi)$, the infimum defining $\psi(x_0)$ is attained. Moreover, the function ψ is convex and lower-semicontinuous on E.*

Proof The proof of existence is standard but we reproduce it for the sake of completeness. Let $x_0 \in \mathrm{Dom}(\psi)$ be fixed and let $\{u_n\} \subset L^2_{\mathrm{loc}}(\mathbb{R}^+; U)$ be such that

$$\psi(x_0) \leq G(u_n, x_0) \leq \psi(x_0) + \frac{1}{n}, \quad n = 1, 2, \ldots. \tag{4.256}$$

We set $x_n(t) = x(t, x_0, u_n)$ and fix any $T > 0$. After some calculation involving (4.252) and inequalities (4.254) and (4.256), we find that the $\{u_n\}$ remain in a bounded subset of $L^2(0, T; U)$. Since the space $L^2(0, T; U)$ is reflexive, we may assume that

$$u_n \to u \quad \text{weakly in } L^2(0, T; U). \tag{4.257}$$

Then, by (4.252), it follows that

$$x_n(t) \to x(t) \quad \text{for all } t \geq 0 \text{ weakly in } E$$

and

$$x_n \to x \quad \text{weakly in every } L^2(0, T; E). \tag{4.258}$$

Since the function $(y, v) \to \int_0^T L(y(y), v(t))\,dt$ is convex and lower-semicontinuous and, therefore, weakly lower-semicontinuous on $L^2(0, T; E) \times L^2(0, T; U)$, it follows that

$$\int_0^T L\big(x(t), u(t)\big)\,dt \le \psi(x_0) \quad \text{for all } T > 0.$$

Hence, $\int_0^\infty L(x(t), u(t))\,dt = \psi(x_0)$, as claimed.

Let x_0 and y_0 be arbitrary but fixed in E and let $\lambda \in [0, 1]$.

We set $x_\lambda^0 = \lambda x_0 + (1-\lambda)y_0$ and consider the pairs (x_1, u_1), (y_1, v_1) and (x_λ, u_λ) such that

$$\psi(x_0) = \int_0^\infty L\big(x_1(t), u_1(t)\big)\,dt, \qquad \psi(y_0) = \int_0^\infty L\big(y_1(t), v_1(t)\big)\,dt,$$

and

$$\psi\big(x_\lambda^0\big) = \int_0^\infty L\big(x_\lambda(t), u_\lambda(t)\big)\,dt.$$

Since $x_\lambda(0) = x_\lambda^0$ and L is convex, we have

$$\begin{aligned}\psi\big(x_\lambda^0\big) &\le \int_0^\infty L\big(\lambda x_1 + (1-\lambda)y_1, \lambda u_1 + (1-\lambda)v_1\big)\,dt \\ &\le \lambda \int_0^\infty L(x_1, u_1)\,dt + (1-\lambda)\int_0^\infty L(y_1, v_1)\,dt = \lambda\psi(x_0) + (1-\lambda)\psi(y_0),\end{aligned}$$

and, therefore, ψ is convex too.

To prove that ψ is lower-semicontinuous, consider a sequence $x_0^n \to x_0$ for $n \to \infty$. Let (x_n, u_n) be such that

$$\psi\big(x_0^n\big) = \int_0^\infty L(x_n, u_n)\,dt.$$

Arguing as in the first part of the proof, we infer that $\{u_n\}$ is weakly compact in $L^2_{\text{loc}}(\mathbb{R}^+; U)$. Hence, on some subsequence, again denoted $\{u_n\}$, we have for every $T > 0$

$$\begin{aligned} u_n &\to \tilde{u} \quad \text{weakly in } L^2(0, T; U), \\ x_n(t) &\to \tilde{x}(t) \quad \text{weakly in } E \text{ for every } t \in [0, T].\end{aligned}$$

Hence,

$$\liminf_{n\to\infty} \int_0^T L(x_n, u_n)\,dt \ge \int_0^T L(\tilde{x}, \tilde{u})\,dt \quad \text{for all } T > 0$$

and, therefore,

$$\psi(x_0) \le \liminf_{n\to\infty} \psi\left(x_0^n\right),$$

as claimed. □

4.4.2 Optimal Feedback Controllers for (P_∞)

Let $H : E \times U^* \to \mathbb{R}$ be the Hamiltonian function associated with L, that is,

$$H(y,q) = \sup\left\{\langle q, v\rangle - L(y,v);\ v \in U\right\}$$

and let $\partial H = -(-\partial_y H, \partial_q H)$ be the subdifferential of H. Unless stated otherwise, the following hypotheses are in effect throughout this section.

(i) *The function H is everywhere finite on $E \times U^*$. Furthermore, one has*

$$H(x,0) \le 0, \qquad H(x,q) \le C\left(|x|^2 + \|q\|^2 + 1\right), \quad \forall (x,q) \in E \times U^*.$$

(ii) *There exists $R > 0$, such that*
   (a) $S(0,R) \subset \mathrm{Dom}(\psi)$.
   (b) *For each $x_0 \in S(0,R)$ there exist a sequence $T_n \to +\infty$ and the controllers $u_n \in L^2(0,T_n;U)$ such that $x(T_n,x_0,u_n) \in S(0,R)$.*

Here, $S(0,R)$ is the open ball $\{x \in E; |x| < R\}$.

As seen earlier, the condition that $-\infty < H < +\infty$ on $E \times U^*$ implies that H is continuous and ∂H is locally bounded on $E \times U^*$.

By hypothesis (i), it follows that L satisfies (4.254) and

$$L(y,v) \ge 0 \quad \text{for all } (y,v) \in E \times U. \tag{4.259}$$

The controllability hypothesis (ii) holds in some notable cases. In particular, it is satisfied in the situation described in Lemma 4.46 below.

Lemma 4.46 *Assume that the control system (4.252) is stabilizable and*

$$L(0,0) = 0, \quad (0,0) \in \operatorname{int}\mathrm{Dom}(L). \tag{4.260}$$

Then hypothesis (ii) holds. In addition, if either $\mathrm{Dom}(L) = E \times U$ or the uncontrolled system (4.252) is asymptotically stable and

$$L(y,0) < +\infty, \quad \forall y \in E,$$

then hypothesis (ii) is trivially satisfied with $R = +\infty$.

Here, $\mathrm{Dom}(L)$ denotes, as usual, the effective domain of L; that is,

$$\mathrm{Dom}(L) = \{(y, v) \in E \times U;\ L(y, v) < +\infty\}.$$

Proof By condition (4.260), there is $r > 0$ such that

$$L(y, v) < +\infty \quad \text{for all } |y| < r,\ \|v\| < r. \tag{4.261}$$

The fact that (4.252) is stabilizable means that there is a bounded linear operator $F : E \to U$ such that the *closed loop system* $y' = (A + BF)y$ is *asymptotically stable*, that is, there exist $\gamma > 0$ and $M > 0$ such that

$$\big|y(t)\big| \le M \exp(-\gamma)\big|y(0)\big|, \quad \forall t \le 0.$$

We take $R = \inf\{\frac{r}{M}, \frac{r}{M}\|F\|\}$, where $\|F\|$ is the operator norm of F. By (4.260), we see that

$$\big(y(t), Fy(t)\big) \in \operatorname{int} \mathrm{Dom}(L), \quad \forall t > 0,$$

for each solution y with initial value $x_0 = y(0)$ in $S(0, R)$.

On the other hand, inasmuch as the subdifferential ∂L of L is locally bounded within $\operatorname{int} \mathrm{Dom}(L)$ we may infer that

$$\big|z_1(t)\big| + \big|z_2(t)\big| \le C, \quad t > 0,$$

for all $(z_1(t), z_2(t)) \in \partial L(y(t), Fy(t))$.

The above estimates, along with the hypotheses of Lemma 4.46 yield

$$L\big(y(t), Fy(t)\big) \le \big(z_1(t), y(t)\big) + \big\langle z_2(t), Fy(t)\big\rangle \le C \exp(-\gamma t), \quad \text{for all } t \ge 0.$$

Thus, part (a) of hypothesis (ii) holds with $\mathbb{R}$ defined as above.

The proof of the last part of the lemma is straightforward, so we omit it. □

Theorem 4.47 *Assume that hypotheses* (i) *and* (ii) *are satisfied and let* (x^*, u^*) *be an optimal pair for Problem* (P_∞) *with* $x^*(0) = x_0 \in S(0, R)$. *Then the optimal control* u^* *is given as a function of optimal arc* x^* *by the feedback law*

$$u^*(t) \in \partial_q H\big(x^*(t), -B^*\partial\psi\big(x^*(t)\big)\big) \quad \text{a.e. } t > 0. \tag{4.262}$$

As usual, $\partial\psi : E \to E^*$ denotes the subdifferential of the function ψ.

In particular, Theorem 4.47 implies that each optimal arc to Problem (P_∞) is a solution to the closed loop system

$$x' - Ax \in B\partial_q H\big(x, -B^*\partial\psi(x)\big), \quad t > 0. \tag{4.263}$$

Proof of Theorem 4.47 Let x_0 be a fixed element in $S(0, R)$ and let (x^*, u^*) be an optimal pair for Problem (P_∞) corresponding to $x_0 \in S(0, R)$. (The existence of such a pair is provided by Proposition 4.45.) If ψ is the function defined by (4.255), then it is easy to see that, for each $T > 0$, (x^*, u^*) is also a solution to the following problem:

$$\min\left\{\int_0^T L(y,v)\,\mathrm{d}t + \psi\big(y(T)\big);\ y(t) = S(t)x_0 + \int_0^t S(t-s)Bv(s)\,\mathrm{d}s,\right.$$
$$\left. v \in L^2(0,T;U)\right\}. \tag{4.264}$$

Here is the argument. Let $v \in L^2(0,T;U)$ and y be the corresponding "mild" solution to (4.252) with $y(0) = x_0$. Since, as seen earlier, the infimum defining $\psi(y(T))$ is attained, there exists $w \in L^2_{\mathrm{loc}}(\mathbb{R}^+;U)$ such that

$$\psi\big(y(T)\big) = \int_0^\infty L(z,w)\,\mathrm{d}t; \qquad z' = Az + Bw, \quad z(0) = y(T).$$

It should be remarked that the pair (y_1, v_1) defined by

$$y_1(t) = \begin{cases} y(t), & 0 \le t \le T, \\ z(t-T), & T < t < \infty, \end{cases}$$

$$v_1(t) = \begin{cases} v(t), & 0 \le t \le T, \\ w(t-T), & T < t < \infty \end{cases}$$

satisfies (4.252). We have

$$\psi(x_0) = \int_0^\infty L(x^*, u^*)\,\mathrm{d}t \le \int_0^\infty L(y_1, v_1)\,\mathrm{d}t = \int_0^T L(y,v)\,\mathrm{d}t + \psi\big(y(T)\big).$$

Thus, we may infer that $\psi(x^*(T)) \ge \int_T^\infty L(x^*, u^*)\,\mathrm{d}t$ and, therefore,

$$\psi\big(x^*(T)\big) = \int_T^\infty L\big(x^*(t), u^*(t)\big)\,\mathrm{d}t.$$

Let $\{T_n\}$ be the sequence defined in hypothesis (ii). According to this hypothesis, for a sufficiently large n there exists an admissible control u_n on $[0, T_n]$ such that $x(T_n, x_0, u_n) \in \operatorname{int}\operatorname{Dom}(\psi)$. Thus, Assumption (E) of Theorem 4.5 is satisfied. Since the other assumptions automatically hold, we may apply this theorem to problem (4.264) to infer that for each n there is a continuous function $p_n; [0, T_n] \to E^*$ which satisfies the equation

$$p_n(t) = S^*(T_n - t)p_n(T_n) - \int_t^{T_n} S^*(s-t)q_1^n(s)\,\mathrm{d}s \tag{4.265}$$

on the interval $[0, T_n]$ and the final condition

$$p_n(T_n) \in -\partial\psi\big(x^*(T_n)\big), \tag{4.266}$$

where $q_1^n \in L^2(0, T_n; E^*)$ satisfies the equation

$$\big(q_1^n(t), B^* p_n(t)\big) \in \partial L\big(x^*(t), u^*(t)\big) \quad \text{a.e. } t \in]0, T_n[. \tag{4.267}$$

(As pointed out before, $\partial L : E \times U \to E^* \times U^*$ stands for the subdifferential of L.) It follows, therefore, that

$$u^*(t) \in \partial_q H\big(x^*(t), B^* p_n(t)\big) \quad \text{a.e. } t \in]0, T_n[. \tag{4.268}$$

To conclude the proof, it remains to show that

$$p_n(t) \in -\partial\psi\big(x^*(t)\big) \quad \text{for all } t \in [0, T_n].$$

Here is the argument. Let h be arbitrary in $\mathrm{Dom}(\psi)$ and let $v \in L^2_{\mathrm{loc}}(\mathbb{R}; U)$ and $y(t) = S(t)h + \int_0^T S(t-s)Bv(s)\,\mathrm{d}s$ be such that $\psi(h) = \int_0^\infty L(y, v)\,\mathrm{d}t$. Let t be arbitrary but fixed in the interval $[0, T_n]$ and let $y_t(s) = y(s-t)$, $v_t(s) = v(s-t)$, $t < s < +\infty$. It follows from (4.252) and the definition of ∂L that

$$\begin{aligned} L\big(x^*(s), u^*(s)\big) \le L\big(y_t(s), v_t(s)\big) + \big(x^*(s) - y_t(s), q_1^n(s)\big) \\ + \big\langle u^*(s) - v_t(s), B^* p_n(s)\big\rangle \quad \text{a.e. } s > 0. \end{aligned}$$

We integrate the latter on the interval $[t, T_n]$ to obtain after some calculations

$$\begin{aligned} &-\big(p_n(t), x^*(t) - h\big) + \big(p_n(T_n), x^*(T_n) - y(T_n - t)\big) \\ &\quad \ge \int_t^{T_n} L\big(x^*(s), u^*(s)\big)\,\mathrm{d}s - \int_0^{T_n - t} L\big(y(s), v(s)\big)\,\mathrm{d}s \\ &\quad = \psi\big(x^*(t)\big) - \psi\big(x^*(T_n)\big) - \psi(h) - \psi\big(y(T_n - t)\big). \end{aligned}$$

Combining the latter with (4.266), we get

$$-\big(p_n(t), x^*(t) - h\big) \ge \psi\big(x^*(t)\big) - \psi(h) \quad \text{for } t \in [0, T_n],$$

as claimed. This completes the proof. □

Let us assume now that every "mild" solution to system (4.252) with the initial value $x(0)$ in $D(A)$ is a.e. differentiable. This happens, for example, if A generates an analytic semigroup and if the function $x \to \partial_q H(x, -B^*\partial\psi(x))$ is Fréchet differentiable. Suppose further that $\mathrm{Dom}(\psi) = E$ and set $K = \partial\psi$.

Let (x^*, u^*) be a solution to (P_∞). We have

$$\big(\psi\big(x^*(t)\big)\big)' + L\big(x^*(t), u^*(t)\big) = 0 \quad \text{a.e. } t > 0,$$

which in conjunction with the conjugacy formula defining the Hamiltonian H yields

$$\big(\psi\big(x^*(t)\big)\big)' - \big\langle u^*(t), B^* K x^*(t)\big\rangle = H\big(x^*(t), -B^* K x^*(t)\big), \quad \text{a.e. } t > 0.$$

Using the chain rule differentiation formula

$$\big(\psi\big(x^*(t)\big)\big)' = \big(Kx^*(t), (x^*)'(t)\big), \quad \text{a.e. } t > 0,$$

and keeping in mind that x^* is a solution to system (4.252), we get

$$H\big(x^*(t), -B^*Kx^*(t)\big) - \big(Ax^*(t), Kx^*(t)\big) = 0 \quad \text{for all } t \ge 0,$$

and, therefore, K must satisfy the stationary Hamilton–Jacobi equation

$$H(h, -B^*Kh) - (Ah, Kh) = 0 \quad \text{for all } h \in D(A). \tag{4.269}$$

It should be mentioned that a direct approach to the existence in (4.269) is hard to obtain and also that the uniqueness of a regular solution is improbable. However, it can be studied in the framework of "viscosity solution" theory.

4.4.3 The Hamiltonian System on Half-Axis

We say that a given pair of continuous functions $x : \mathbb{R}^+ \to E$ and $p : \mathbb{R}^+ \to E^*$ is a solution to the Hamiltonian system

$$\begin{aligned} x' - Ax &\in B\partial_q H(x, B^*p), \quad t \ge 0, \\ p' + A^*p &\in -\partial_x H(x, B^*p), \quad t \ge 0, \end{aligned} \tag{4.270}$$

if there exist functions $q_1 \in L^2_{\text{loc}}(\mathbb{R}^+; E^*)$ and $u \in L^2_{\text{loc}}(\mathbb{R}^+; U)$ such that

$$\begin{aligned} x(t) &= S(t)x(0) + \int_0^t S(t-s)Bq_2(s)\,\mathrm{d}s \quad \text{for } t \ge 0, \\ p(t) &= S^*(T-t)p(T) - \int_t^T S^*(s-t)q_1(s)\,\mathrm{d}s \quad \text{for all } 0 \le t \le T, \end{aligned} \tag{4.271}$$

and

$$\begin{aligned} q_2(t) &\in \partial_q H\big(x(t), B^*p(t)\big), \\ q_1(t) &\in -\partial_x H\big(x(t), B^*p(t)\big), \quad \text{a.e. } t > 0, \end{aligned} \tag{4.272}$$

or, equivalently,

$$\big(q_1(t), B^*p(t)\big) \in \partial L\big(x(t), q_2(t)\big), \quad \text{a.e. } t > 0. \tag{4.272$'$}$$

Here, $(-\partial_y H, \partial_q H) = \partial H$. (See (2.159) and (2.160).)

As seen in Proposition 4.45, under hypotheses (i) and (ii), Problem (P_∞) has at least one solution (x, u). Our concern here is to characterize this optimal pair in terms of the Hamiltonian system (4.270). To this aim, besides (i) and (ii), further assumptions are necessary:

(j) $(0,0)$ is a saddle-point of H and $H(0,0)=0$.
(jj) For each $r>0$, $\inf\{-H(y,0);\ |y|=r\}>0$.
(jjj) For each $r>0$, there is a real positive function ω on $\mathbb{R}^+$ such that

$$\lim_{t\downarrow 0}\frac{\omega(t)}{t}=0, \qquad H(y,q)\le \omega\big(\|q\|\big) \quad \text{for all } q\in U^* \text{ and } |y|\le r. \tag{4.273}$$

For the time being, the following consequences of the above assumptions are useful.

Lemma 4.48 *Let H satisfy* (j) *and* (jjj) *and let L be the Lagrangian function associated with H. Then $L\ge 0$ on $E\times U$, $L(0,0)=0$ and there exists a positive function γ defined on $\mathbb{R}^+$ such that $\frac{\gamma(\rho)}{\rho}\to 0$ for $\rho\to 0$ and*

$$L(y,v)\ge \rho\|v\|-\gamma(\rho)\max(1,y) \quad \textit{for all } \rho>0,\ y\in E. \tag{4.274}$$

Proof By the definition of the Lagrangian function L, we have

$$L(y,v)=\sup\big\{\langle q,v\rangle - H(y,q);\ q\in U\big\},$$

which, in virtue of assumption (j), implies that $L(0,0)=0$ and $L\ge 0$. The latter also implies that

$$L(y,v)\ge \rho\|v\|-H(y,\rho w) \quad \text{for all } \rho>0,\ w=\frac{\Psi(v)}{\|v\|}\cdot$$

On the other hand, since the function $y\to H(y,\rho w)$ is concave, we have

$$H(y,\rho w)\le |y|H\left(\frac{y}{|y|},\rho w\right) \quad \text{for } |y|\le 1.$$

Combining this inequality with assumption (jjj), we find (4.274), as claimed (for $|y|\le 1$, inequality (4.274) is a direct consequence of (4.273)). □

Theorem 4.49 *Let (x^*,u^*) be an optimal pair for Problem* (P_∞) *with $|x^*(0)|<R$. Then, under hypotheses* (i), (ii), (j), (jj) *and* (jjj), *there exists a continuous function p satisfying along with x^* and $q_2=u^*$ system* (4.270) *and the conditions*

$$\lim_{t\to\infty}x^*(t)=0, \qquad \big|p(t)\big| \textit{ bounded on } \mathbb{R}^+. \tag{4.275}$$

Moreover, one has

$$p(t)\in -\partial\psi\big(x^*(t)\big) \quad \textit{for every } t\ge 0. \tag{4.276}$$

Conversely, if the pair (x,u) satisfies system (4.271) *(with $q_2=u$) and conditions* (4.275), *then it is optimal in Problem* (P_∞).

Proof As seen in the proof of Theorem 4.47, for each n there is a function $p_n : [0, T_n] \to E^*$ satisfying (4.265)–(4.267). Moreover, we have

$$\psi\big(x^*(t)\big) = \int_t^\infty L(x^*, u^*)\,\mathrm{d}s \quad \text{for } t \geq 0, \tag{4.277}$$

$$p_n(t) \in -\partial\psi\big(x^*(t)\big) \quad \text{for } t \in [0, T_n]. \tag{4.278}$$

Next, by Lemma 4.48, we have

$$\int_t^{t+h} \big\|u^*(s)\big\|\,\mathrm{d}s \leq \frac{1}{\rho}\int_t^{t+h} L(x^*, u^*)\,\mathrm{d}s + \frac{\gamma(\rho)}{\rho}\bigg(h + \int_t^{t+h} |x^*|\,\mathrm{d}s\bigg),$$

for all $t, h > 0$ and $\rho > 0$. Since $L(x^*, u^*) \in L^1(\mathbb{R}^+)$, we infer that

$$\int_t^{t+h} \big\|u^*(s)\big\|\,\mathrm{d}s \leq \theta(t)\bigg(1 + \int_t^{t+h} |x^*|\,\mathrm{d}s\bigg) \quad \text{for all } \ t \geq 0, \tag{4.279}$$

where $\lim_{t\to\infty}\theta(t) = 0$. The latter combined with the obvious equality

$$x^*(t+h) = S(h)x^*(t) + \int_0^h S(h-s)Bu^*(t+s)\,\mathrm{d}s, \quad t, h > 0,$$

yields

$$\big|x^*(t+h)\big| \leq C\big(\big|x^*(t)\big| + \tilde{\theta}(t)\big) \quad \text{for } t \geq 0,\ h \in [0,1], \tag{4.280}$$

where C is independent of t and h and $\lim_{t\to\infty}\tilde{\theta}(t) = 0$. On the other hand, the obvious inequality

$$L(x,u) \geq -H(x,0) \quad \text{for all } x \in E,\ u \in U,$$

implies that $-H(x^*, 0) \in L^1(\mathbb{R}^+)$. Hence,

$$-H\big(x^*(t), 0\big) \leq \delta(T) \quad \text{for } \ t \in [T, +\infty[\setminus E_T,$$

where $\delta(T)$, and the Lebesgue measure of E_T tends to zero for $T \to +\infty$. Along with hypothesis (jj), this implies that

$$\big|x^*(t)\big| \leq \eta(T) \quad \text{for } t \in [T, +\infty[\setminus E_T,$$

where $\eta(T) \to 0$ for $T \to +\infty$. Then, by estimate (4.280), we may conclude that

$$\lim_{t\to\infty} x^*(t) = 0.$$

In particular, we infer that there is $T > 0$ such that $x^*(t) \in \operatorname{int}\operatorname{Dom}(\psi)$ for $t \geq T$. Since $\partial\psi$ is locally bounded on $\operatorname{int}\operatorname{Dom}(\psi)$, it follows by (4.278) that

$$\big|p_n(t)\big| \leq C \quad \text{for } t \geq T, \text{ and } n \text{ sufficiently large.} \tag{4.281}$$

On the other hand, by (4.267) and the definition of ∂L, we have

$$\begin{aligned}&\rho\big(q_1^n(t), w\big) + \big\langle B^* p_n(t), u^*(t) - v(t)\big\rangle \\ &\quad \geq L\big(x^*(t), u^*(t)\big) - L\big(x^*(t) - \rho w, v(t)\big),\end{aligned} \tag{4.282}$$

for all $\rho > 0$ and $w \in E$. If we take $v(t) \in \partial_p H(x^*(t) - \rho w, 0)$, we get

$$L\big(x^*(t) - \rho w, v(t)\big) = -H\big(x^*(t) - \rho w, 0\big)$$

and, therefore,

$$\big\|v(t)\big\| + L\big(x^*(t) - \rho w, v(t)\big) \leq C,$$

for all $t \geq 0$, $|w| = 1$ and ρ sufficiently small. It now follows from (4.282) that

$$\rho\big|q_1^n(t)\big| \leq C\big(\big\|u^*(t)\big\|\big(\big|p_n(t)\big| + 1\big) + 1\big) \quad \text{for } t \geq 0,$$

which, along with the equation

$$p_n(t) = S^*(T - t)p_n(T) - \int_t^T S^*(s - t)q_1^n(s)\,\mathrm{d}s, \quad 0 \leq t \leq T \leq T_n, \tag{4.283}$$

implies that $\{|p_n(t)|\}$ are uniformly bounded on $[0, T]$. Hence, by (4.281), we may infer that

$$\big|p_n(t)\big| \leq C \quad \text{for all } t \geq 0 \text{ and } n = 1, \ldots. \tag{4.284}$$

Next, by the definition of ∂H (see (2.144)), we have

$$\rho\big(q_1^n, w\big) \leq H(x^*, B^* p_n) - H(x^* - \rho w, B^* p_n) \quad \text{for all } \rho > 0,\ w \in E,$$

and

$$-H(x^* - \rho w, B^* p_n) \leq -H(x^* - \rho w, 0) + \big\langle \partial_q H(x^* - \rho w, 0), B^* p_n\big\rangle.$$

Since H and ∂H are locally bounded, we have

$$\big|q_1^n(t)\big| \leq C \quad \text{for all } t > 0 \text{ and all } n.$$

Thus, extending q_1^n by zero outside the interval $[0, T_n]$, we may assume that

$$q_q^n \to q_1 \quad \text{weak-star in } L^\infty\big(\mathbb{R}^+; E^*\big).$$

On the other hand, it follows by (4.284) that there exists an increasing sequence $t_j \to +\infty$ and a subsequence of $\{p_n\}$ (again denoted p_n) such that, for $n \to \infty$,

$$p_n(t_j) \to p^j \quad \text{weakly in } E^* \text{ for all } j.$$

Since (4.283) is satisfied for all $T = t_j < T_n$, we infer that there exists a function $p : \mathbb{R}^+ \to E^*$ such that $p(t_j) = p^j$,

$$p_n(t_j) \to p(t_j) \quad \text{weakly in } E^* \text{ for all } j$$

and

$$p(t) = S^*(t_j - t)p(t_j) - \int_t^{t_j} S^*(s-t)q_1(s)\,\mathrm{d}s, \quad 0 \le t \le t_j,$$

for all $j = 1, 2, \dots$. Letting $n \to \infty$ in (4.278), we get (4.276), as claimed. Note also that by (4.267) we have

$$\big(q_1(t), B^*p(t)\big) \in \partial L\big(x^*(t), u^*(t)\big), \quad \text{a.e. } t > 0,$$

while (4.284) implies that $|p(t)|$ is bounded over $\mathbb{R}^+$. Thus, x^*, p and u^* satisfy all the requirements of the theorem. The sufficiency of condition (4.275) for optimality is immediate. □

Remark 4.50 It follows from Theorem 4.49 that the feedback control (4.262) stabilizes system (4.252) and

$$\big\{(x,p) \in E \times E^*;\ p + \partial\phi(x) \ni 0\big\}$$

is a positively invariant manifold of system (4.270).

Another important consequence is that, by (4.272) and (4.276), the optimal control u^* is expressed in feedback form as

$$u^*(t) = \partial_q H\big(x^*(t), -B^*\partial\psi\big(x^*(t)\big)\big).$$

We present now other qualitative aspects concerning the Hamiltonian system (4.270). The following supplementary assumptions are imposed:

(v) *$H(y,q)$ is strictly convex in q for each $y \in E$.*
(vj) *$H(y,q)$ is strictly concave in y for each $q \in U$ or H is Gâteaux differentiable in q.*
(vjj) *$N(B^*) = \{0\}$ or $H(y,q)$ is Gâteaux differentiable in y and the pair (A,B) is "controllable", that is, $B^*S^*(t)x_0 = 0$ on some interval $[0,T]$ implies $x_0 = 0$.*

Theorem 4.51 *In Theorem 4.49, suppose further that hypotheses (vj) up to (vjj) hold. Then, for each $x_0 \in S(0,R)$, there exists a unique solution (x^*, p) to system (4.270) satisfying (4.275). Moreover, ψ is Gâteaux differentiable on $S(0,R)$ and*

$$\lim_{t\to\infty} p(t) = 0 \quad \textit{weakly in } E^*.$$

Proof We prove first that for each x there exists at most one function p satisfying system (4.270) along with x and u. Assume the contrary and let p_1 be another solution to this system. Then (4.270) implies that $B^*(p - p_1) = 0$ a.e. $t > 0$, because the function $H(x,\cdot)$ is strictly convex. Moreover, by (4.270) we see that

$$(p - p_1)' + A^*(p - p_1) = 0 \quad \text{for } t > 0.$$

Since the pair (A, B) is "controllable", the latter implies that $p - p_1 = 0$, as claimed. On the other hand, as seen in Theorem 4.49, every solution x to system (4.270) and (4.275) is an optimal arc to Problem (P_∞). If (v) holds, then the function L is strictly convex and, therefore, for each x_0, the solution of Problem (P_∞) must be unique. This fact proves the uniqueness of the solution (x, p) to system (4.270).

Denote by $\Gamma \subset E \times E^*$ the set of all the pairs (x_0, p_0) having the property that there exists a solution $(\tilde{x}, \tilde{p})$ to system (4.270) satisfying

$$\tilde{x}(0) = x_0, \qquad \tilde{p}(0) = p_0,$$

and

$$\tilde{p}(T) \in -\partial\psi\big(\tilde{x}(T)\big), \tag{4.285}$$

where T is a positive number with the property that there is a control $u \in L^2(0, T; U)$ such that $y(T, x_0, u) \in S(0, R)$ and $L(y(t, x_0, u), L^1(0, T)$ (by hypothesis (ii), such a number T always exists). As seen in Theorem 4.5, system (4.270) with condition (4.285) and $x(0) = x_0$ is equivalent to the optimization problem

$$\min\bigg\{\int_0^T L(y, v)\,\mathrm{d}t + \psi\big(y(T)\big);$$
$$y(t) = S(t)x_0 + \int_0^t S(t-s)Bv(s)\,\mathrm{d}s,\ v \in L^2(0, T; U)\bigg\}. \tag{4.286}$$

Since problem (4.286) admits at least one solution, we may conclude that $\Gamma x_0 \neq \emptyset$ for each $x_0 \in S(0, R)$. Furthermore, since, as noticed earlier, L is strictly convex, problem (4.286) admits a unique solution which must agree on $[0, T]$ with (x^*, u^*). In other words, $\tilde{x} = x^*$ and $\tilde{p} = p$, where (x^*, p) is the unique solution to system (4.270) satisfying (4.275) and the initial condition $x^*(0) = x_0$. Thus, Γ can be, equivalently, defined as

$$\Gamma x_0 = p(0) \quad \text{for } x_0 \in S(0, R).$$

In particular, we deduce that Γ is single-valued and $-\Gamma x_0 \subset \partial\psi(x_0)$ for all $|x_0| < R$. We prove that $-\Gamma$ agrees with $\partial\psi$ within $S(0, R)$. To this end, it suffices to show that $-\Gamma$ is maximal monotone on $E \times E^*$, that is, $R(\Phi - \Gamma) = E^*$, where Φ is the duality mapping of E. Let y_0^* be a fixed element in E^*. The equation

$$\Phi(x_0) - \Gamma x_0 = y_0^* \tag{4.287}$$

can equivalently be written as

$$\begin{aligned} &\tilde{x}' - A\tilde{x} \in B\partial_q H(x, B^*\tilde{p}), \quad 0 \le t \le T,\\ &\tilde{p}' + A^*\tilde{p} \in -\partial_x H(\tilde{x}, B^*\tilde{p}), \quad 0 \le t \le T, \end{aligned} \tag{4.288}$$

with the two-point boundary-value conditions

$$\Phi\big(\tilde{x}(0)\big) - \tilde{p}(0) = y_0^*, \qquad p(T) \in -\partial\psi\big(\tilde{x}(T)\big). \tag{4.289}$$

It is apparent that (4.288), and (4.289) are just the extremality equations in the Hamiltonian form (see (4.266) and (4.267)) for the control problem

$$\min\biggl\{\int_0^T L(y,v)\,\mathrm{d}t+\psi\bigl(y(T)\bigr)+\frac{1}{2}\bigl|y(0)\bigr|^2-\bigl(y(0),y_0^*\bigr);$$
$$y(t)=S(t)y(0)+\int_0^T S(t-s)Bv(s)\,\mathrm{d}s,\ v\in L^2(0,T;U)\biggr\},$$

which, by hypothesis (i), admits at least one solution.

We may conclude, therefore, that (4.287) has at least one solution $x_0\in E$. Summarizing, we have shown that $-\Gamma=\partial\psi$ on $S(0;R)$. In particular, we infer that $\partial\psi(x_0)$ is singleton for each $|x_0|<R$. This fact implies that ψ is Gâteaux differentiable on $S(0;R)$ and its gradient $\nabla\psi=\partial\psi$. On the other hand, $-\Gamma$ is demicontinuous on $S(0;R)$ (that is, strongly–weakly continuous) because it is single-valued, maximal monotone and bounded within $S(0;R)$ (its domain contains $S(0;R)$. This fact combined with (4.275) yields

$$p(t)\to 0\quad\text{weakly in } E^*\text{ as } t\to\infty,$$

as claimed. □

4.4.4 The Linear Quadratic Regulator Problem

Here, we consider the special case in which L is quadratic, that is,

$$L(y,v)=\frac{1}{2}\bigl(|Cy|^2+\|v\|^2\bigr)\quad\text{for } y\in E,\ v\in U,$$

where C is a linear continuous operator from E into itself and E,U are real Hilbert spaces. As regards the operators A and B, the assumptions are those from Sect. 4.1.

It should be observed that hypotheses (i), (j), (jjj) are trivially satisfied in this special case. We say that system (4.252) is L^2-*controllable* if, for each $x_0\in E$, there exists $v\in L^2(\mathbb{R}^+;U)$ such that $x(t,x_0,v)\in L^2(\mathbb{R}^+;E)$. As an immediate application of Theorem 4.51, we find the following theorem.

Theorem 4.52 *Assume that system* (4.252) *is L^2-controllable. Then Problem* (P_∞) *has a unique optimal pair (x^*,u^*) related by feedback synthesis law*

$$u^*(t)=-B^*Px^*(t)\quad\textit{for } t\ge 0,\tag{4.290}$$

where P is a linear, continuous, self-adjoint and positive operator on E satisfying the algebraic Riccati equation

$$PBB^*P-PA-A^*P=C^*C.\tag{4.291}$$

The minimal cost in (P_∞) *is* $\psi(x_0)=\frac{1}{2}(Px_0,x_0)$.

Proof The existence and uniqueness of the optimal control u^* follow by Proposition 4.1 and the strict convexity of $u \to L(x,u)$. By Theorem 4.47, $u^*(t) \in -B^*\psi(x^*(t))$, where

$$\psi(x_0) = \inf\left\{\frac{1}{2}\int_0^\infty \big(|Cx|^2 + \|u\|^2\big)\,\mathrm{d}t;\, x(t) = S(t)x_0 + \int_0^t S(t-s)Bu(s)\,\mathrm{d}s\right\}.$$

It is immediate that $\psi(\lambda x_0) = \lambda^2\psi(x_0)$ and, for all $x_0, y_0 \in E$, we have the following equality:

$$\psi(x_0 + y_0) + \psi(x_0 - y_0) = 2\big(\psi(x_0) + \psi(y_0)\big).$$

Define the operator $P : E \to E$ by

$$(Px_0, y_0) = \frac{1}{2}\big(\psi(x_0 + y_0) - \psi(x_0 - y_0)\big).$$

Clearly, the operator P is linear, continuous, self-adjoint and positive. In particular, we find that $\psi(x_0) = \frac{1}{2}(Px_0, x_0)$ and $\nabla\psi = P$. Then, differentiating (4.269), we get (4.291), as claimed. □

Theorem 4.53 *In Theorem 4.52, if we assume, in addition, that*

(a) *there is a linear continuous operator $K : E \to E$ such that $A + K(C^*C)^{\frac{1}{2}}$ generates an exponentially stable semigroup,*

then the feedback law (4.290) stabilizes system (4.252), and (4.291) has a unique self-adjoint and positive solution P.

Proof Let P be the operator defined above. We have $u^* = -B^*Px^* \in L^2(\mathbb{R}^+; U)$ and $(C^*C)^{\frac{1}{2}}x^* \in L^2(\mathbb{R}^+; E)$. On the other hand, x^* is the solution to the closed loop system

$$x' = \big(A + KQ^{\frac{1}{2}}\big)x + Bu^* - KQ^{\frac{1}{2}}x, \quad x(0) = x_0,$$

where $Q = C^*C$. Hence, $x^* \in L^2(\mathbb{R}^+; E)$. By Lyapunov's theorem in Hilbert spaces (see Datko [25]), we conclude that $|x^*(t)| \le C\exp(-\alpha t)|x_0|$ for some $\alpha > 0$, as claimed.

Uniqueness. Suppose that P, P_1 both satisfy (4.291). Under the preceding assumption, we find that $A - BD$ generates an exponentially stable semigroup $S_1(t)$, where $D = B^*P$ (P is given by Theorem 4.52). Write $D_1 = B^*P_1$ and notice that

$$\begin{aligned}
&\big(\big(2P_1(A - BD_1) + D_1^*D_1 + C^*C\big)x, x\big) \\
&\quad = \big(\big(2P(A - BD) + D^*D + C^*C\big)x, x\big) \\
&\qquad + \big(2(P_1 - P)(A - BD_1)x, x\big) + \big\|(D - D_1)x\big\|^2.
\end{aligned}$$

Since P and P_1 are solutions to (4.291), we have

$$\big((2P(A-BD)+D^*D+C^*C\big)x,x\big)=0,\quad \forall x\in D(A),$$
$$\big((2P_1(A-BD_1)+D_1^*\big)+D_1^*D_1+C^*C\big)x,x)=0,\quad \forall x\in D(A),$$

and, therefore,

$$2\big((P_1-P)(A-BD_1)x,x\big)+\big\|(D-D_1)x\big\|^2=0,\quad \forall x\in D(A).$$

We set $S_1(t)=\mathrm{e}^{(A-BD_1)t}$. Then the latter yields

$$2\left((P_1-P)\frac{\mathrm{d}}{\mathrm{d}t}S_1(t)x,S_1(t)x\right)+\big\|v(t)\big\|^2=0,\quad \forall t\ge 0,$$

where $v(t)=(D-D_1)S_1(t)x$. Integrating on $\mathbb{R}^+$ and remembering that, for some $\alpha>0$,

$$\big|S_1(t)x\big|\le C\exp(-\alpha t)|x|,\quad t\ge 0,\ \forall x\in E,$$

we get

$$\big((P_1-P)x,x\big)=\int_0^\infty \big\|v(t)\big\|^2\,\mathrm{d}t\ge 0,\quad \forall x\in E.$$

Hence, $P_1\ge P$ and, therefore, $P_1=P$, as claimed. This concludes the proof. □

Remark 4.54 Condition (a) is known in the literature as a "detectability assumption" and it is satisfied in particular if C is a positive definite operator. In this case, we derive from Theorem 4.52 the following simple result.

Corollary 4.55 *The control system* (4.252) *is stabilizable if and only if it is* L^2-*controllable.*

4.5 Optimal Control of Linear Periodic Resonant Systems

In this section, we study the optimal control problem

$$\text{minimize} \int_0^T \Big(g\big(Cy(t)\big)+h\big(u(t)\big)\Big)\,\mathrm{d}t \tag{4.292}$$

subject to $u\in L^2(0,T;U)$ and $y\in C([0,T];H)$ satisfying the state system

$$\begin{aligned}&\frac{\mathrm{d}y}{\mathrm{d}t}=Ay+Bu+f,\quad t\in(0,T),\\ &y(0)=y(T).\end{aligned} \tag{4.293}$$

Here, H, U and Z are real Hilbert spaces, A is the infinitesimal generator of a C_0 semigroup e^{At} on H, $B \in L(U, H)$, $C \in L(H, Z)$, $g : Z \to \overline{\mathbb{R}}^*$, and $h : U \to \overline{\mathbb{R}}^*$ are lower-semicontinuous convex functions. The solution y to the state system (4.292) is considered in the "mild" sense, that is,

$$y(t) = \mathrm{e}^{At} y(T) + \int_0^t \mathrm{e}^{A(t-s)} \big(Bu(s) + f(s)\big)\, \mathrm{d}s, \quad \forall t \in [0, T]. \tag{4.293$'$}$$

It should be said that, if the system is resonant, that is, the null space of the operator $\frac{\mathrm{d}}{\mathrm{d}t} + A$ with periodic conditions is not trivial, then Assumption (E) from Sect. 4.1 does not hold, so Theorem 4.5 is not applicable, because the operator $(I - \mathrm{e}^{AT})$ is not invertible as in the case of Problem P_T. Thus, for the maximum principle, as well as for the existence in problem (4.292), one must assume some stabilization and detectability conditions for the pairs (A, B) and (A, C), respectively, as in the case of Theorem 4.52. As a matter of fact, the analysis of the periodic optimal control problem (4.292) has many similarities with that of the optimal control problem on half-axis presented in Sect. 4.4.

We use the standard notations for the spaces of vector-valued functions on the interval $[0, T]$. The norms and the scalar products of H, U, Z are denoted by $|\cdot|, |\cdot|_U$, $|\cdot|_Z$ and $(\cdot,\cdot)$, $(\cdot,\cdot)_U$, $(\cdot,\cdot)_Z$, respectively. Given the lower-semicontinuous, convex function φ on the Hilbert space X, we denote, as above, by $\partial\varphi$ the subdifferential of φ, and by φ^* the conjugate of φ. Given a linear, densely defined operator W on a Banach space, we denote by $D(W)$ the domain of W, and by $R(W)$ its range. The dual operator is denoted by W^*.

4.5.1 Weak Solutions and the Closed Range Property

Let $\mathscr{A}$ be the linear operator defined in $L^2(0, T; H)$ as

$$\mathscr{A} y = f \tag{4.294}$$

if and only if

$$\int_0^T \Big(\big(y(t), \varphi'(t) + A^*\varphi(t)\big) + \big(f(t), \varphi(t)\big)\Big)\, \mathrm{d}t = 0,$$

for all $\varphi \in W^{1,2}([0, T]; H)$ such that $A^*\varphi \in L^2(0, T; H)$; $\varphi(0) = \varphi(T)$. A function $y \in L^2(0, T; H)$ satisfying (4.294) is called a *weak solution* to the periodic problem

$$\frac{\mathrm{d}y}{\mathrm{d}t} = Ay + f; \quad y(0) = y(T). \tag{4.295}$$

It is readily seen that the operator $\mathscr{A}$ is closed and densely defined in $L^2(0, T; H)$. Moreover, the dual operator $\mathscr{A}^*$ is defined as

$$\mathscr{A}^* z = g \tag{4.296}$$

if and only if

$$\int_0^T \big(\big(z(t), \varphi'(t) - A\varphi(t)\big) - \big(\varphi(t), g(t)\big)\big)\,\mathrm{d}t = 0, \tag{4.297}$$

for all $\varphi \in W^{1,2}([0,T];H)$ such that $A\varphi \in L^2(0,T;H)$, $\varphi(0) = \varphi(T)$.

Let $N(\mathscr{A})$ and $N(\mathscr{A}^*)$ be the null spaces of $\mathscr{A}$ and $\mathscr{A}^*$, respectively. If $R(\mathscr{A})$ (the range of $\mathscr{A}$) is closed in $L^2(0,T;H)$, then, by virtue of the closed range theorem, so is $R(\mathscr{A}^*)$ and we have

$$L^2(0,T;H) = R(\mathscr{A}) \oplus N(\mathscr{A}^*) = R(\mathscr{A}^*) \oplus N(\mathscr{A}). \tag{4.298}$$

This means that, for each $f \in \mathbb{R}(\mathscr{A})$, the solutions y to the equation $\mathscr{A}y = f$ are expressed as $y = y_1 + N(\mathscr{A})$, where $y_1 \in \mathbb{R}(\mathscr{A}^*)$ is uniquely defined. We define $\mathscr{A}^{-1}$ as $\mathscr{A}^{-1}f = y_1$ and note that, by the closed graph theorem, $\mathscr{A}^{-1} \in L(R(\mathscr{A}), L^2(0,T;H))$. The operator $(\mathscr{A}^*)^{-1} \in L(R(\mathscr{A}^*), L^2(0,T;H))$ is similarly defined.

Proposition 4.56 *Assume that, for each $m \in \mathbb{Z}$, the range Y_m of $\mu_m iI - A$ is closed in H and*

$$\sup\big\{\big\|(\mu_m iI - A)^{-1}\big\|_{L(Y_m,H)};\ m \in \mathbb{Z}\big\} < \infty, \tag{4.299}$$

where $\mu_m = \frac{2m\pi}{T}$. Then $R(\mathscr{A})$ is closed in $L^2(0,T;H)$.

Here, we have again denoted by A the realization of the operator A in the complexified space H.

Proof If $f \in \mathbb{R}(\mathscr{A})$, then there is a $y \in L^2(0,T;H)$ such that

$$y(t) = \sum_{m\in\mathbb{Z}} y_m \exp(\mu_m it), \quad t \in (0,T), \tag{4.300}$$

where

$$y_m = (\mu_m i - A)^{-1} f_m, \quad f_m = T^{-1}\int_0^T \exp(-\mu_m it) f(t)\,\mathrm{d}t,$$

and so, by (4.299) and Parseval's identity, we get

$$\|y\|_{L^2(0,T;H)} \le C\|f\|_{L^2(0,T;H)},$$

where $\mathscr{A}y = f$. This implies that $R(\mathscr{A})$ is closed in $L^2(0,T;H)$, as claimed. □

Let $\mathscr{A}_0 : D(\mathscr{A}_0) \subset L^2(0,T;H) \to L^2(0,T;H)$ be the linear operator defined as

$$\mathscr{A}_0 y = f \tag{4.301}$$

if and only if

$$y(t) = e^{At} y(T) + \int_0^t e^{A(t-s)} f(s)\, ds, \quad t \in (0, T).$$

In other words, $\mathscr{A}_0 y = f$ if and only if y is continuous and it is a "mild" periodic solution to (4.295). It is easily seen that $\mathscr{A}_0$ is closed and densely defined in $L^2(0, T; H)$. Moreover, a simple integration by parts shows that $\mathscr{A}_0 \subset \mathscr{A}$. As a matter of fact, we have the following result.

Proposition 4.57 $\mathscr{A}_0 = \mathscr{A}$.

Proof Since, as noticed earlier, the inclusion $\mathscr{A}_0 \subset \mathscr{A}$ is immediate, we confine ourselves to checking that $\mathscr{A} \subset \mathscr{A}_0$. Let $(y, f) \in \mathscr{A}$. We have

$$y(t) = \sum_{m \in \mathbb{Z}} y_m \exp(\mu_m i t) \quad \text{in } L^2(0, T; H); \qquad (\mu_m i - A) y_m = f_m. \tag{4.302}$$

Then the sequence

$$y_N(t) = \sum_{|m| \le N} y_m \exp(i \mu_m t)$$

is convergent to y in $L^2(0, T; H)$, and, for each N, y_N is a "mild" solution to (4.295), where $f = f_N = \sum_{|m| \le N} f_m \exp(i \mu_m t)$. Hence,

$$\begin{aligned} &y_N(t) = e^{A(t-s)} y_N(s) + \int_s^t e^{A(t-s)} f_N(r)\, dr, \quad 0 < s < t < T, \\ &y_N(0) = y_N(T). \end{aligned} \tag{4.303}$$

Since $y_N \to y$ and $f_N \to f$ in $L^2(0, T; H)$ and a.e. on $(0, T)$ (on some subsequence), we infer by (4.303) that $\{y_n(T)\}$ is strongly convergent in H to some y_1 and, therefore, $y_N(t)$ is uniformly convergent to $y(t) \in C([0, T]; H)$ and $\mathscr{A} y = f$, as claimed. □

By Proposition 4.57, we have

$$R(\mathscr{A}) = \left\{ f \in L^2(0, T; H);\ \int_0^T e^{A(T-t)} f(t)\, dt \in R\big(I - e^{AT}\big) \right\}, \tag{4.304}$$

$$N(\mathscr{A}) = \big\{ y \in L^2(0, T; H);\ y(t) = e^{At} y_0,\ \big(I - e^{AT}\big) y_0 = 0 \big\}. \tag{4.305}$$

Moreover, the dual operator $\mathscr{A}^*$ is given by $\mathscr{A}^* z = g$ if and only if

$$z(t) = e^{A^*(T-t)} z(0) + \int_t^T e^{A^*(s-t)} g(s)\, ds, \quad \forall t \in [0, T]. \tag{4.306}$$

Proposition 4.58 *$R(\mathscr{A})$ is closed in $L^2(0,T;H)$ if and only if $R(I-\mathrm{e}^{AT})$ is closed in H.*

Proof If $R(I-\mathrm{e}^{AT})$ is closed in H, then, by (4.304), we see that $R(\mathscr{A})$ is closed in $L^2(0,T;H)$. Assume, now, that $R(\mathscr{A})$ is closed and consider the linear subspace of H,

$$X=\left\{x\in H;\ \mathrm{e}^{At}x\in R(\mathscr{A})\right\}.$$

(Here, we have denoted by $(\mathrm{e}^{At}x)$ the function $t\to \mathrm{e}^{At}x$.) We have

$$X=R\big(I-\mathrm{e}^{AT}\big). \tag{4.307}$$

Here is the argument. If $x\in R(I-\mathrm{e}^{AT})$, then $T\mathrm{e}^{AT}x\in R(I-\mathrm{e}^{AT})$, and therefore the equation

$$\big(I-\mathrm{e}^{AT}\big)y_0=T\mathrm{e}^{AT}x$$

has at least one solution $y_0\in H$. Then the function

$$y(t)=\mathrm{e}^{At}y_0+\int_0^t \mathrm{e}^{A(t-s)}\mathrm{e}^{As}x\,\mathrm{d}s=\mathrm{e}^{At}y_0+t\mathrm{e}^{At}x$$

is a solution to $\mathscr{A}y=\mathrm{e}^{At}x$, that is, $x\in X$.

Now, let x be in X, and let $y(t)=\mathrm{e}^{At}y(0)+t\mathrm{e}^{At}x$ be a solution to $\mathscr{A}y=\mathrm{e}^{At}x$. Since $y(0)=y(T)$, the latter implies that $\mathrm{e}^{AT}x\in R(I-\mathrm{e}^{AT})$, and therefore $x\in R(I-\mathrm{e}^{AT})$. Since X is closed, it follows from (4.307) that so is $R(I-\mathrm{e}^{AT})$. □

Corollary 4.59 *If $R(\mathscr{A})$ is closed in $L^2(0,T;H)$, then $\mathscr{A}^{-1}f\in C([0,T];H)$ for each $f\in R(\mathscr{A})$ and*

$$\|\mathscr{A}^{-1}f\|_{C([0,T];H)}\le C\|f\|_{L^1(0,T;H)},\quad \forall f\in R(\mathscr{A}). \tag{4.308}$$

Proof Since $R(\mathscr{A})$ is closed, so is $R(I-\mathrm{e}^{AT})$, and we have, therefore,

$$\mathscr{A}^{-1}f(t)=\mathrm{e}^{At}\big(I-\mathrm{e}^{AT}\big)^{-1}\int_0^T \mathrm{e}^{A(T-t)}f(t)\,\mathrm{d}t+\int_0^t \mathrm{e}^{A(t-s)}f(s)\,\mathrm{d}s,\quad \forall t\in[0,T].$$

Recalling that $(I-\mathrm{e}^{AT})^{-1}$ is continuous on $R(I-\mathrm{e}^{AT})$, the latter implies (4.308), as desired. □

By Riesz–Fredholm theory, we also have the following corollary.

Corollary 4.60 *If e^{AT} is compact, then $R(\mathscr{A})$ is closed and the spaces $N(\mathscr{A})$, $N(\mathscr{A}^*)$ are finite-dimensional.*

Given $F\in L(H,U)$, we denote by $\mathscr{A}_F$ the operator $\mathscr{A}+BF$ defined from $L^2(0,T;H)$ to itself and we denote by $\mathscr{A}_F^*=\mathscr{A}^*+F^*B^*$ its dual.

Definition 4.61 The pair (A, B) is said to be π-stabilizable if there is an $F \in L(H, U)$ such that $R(\mathscr{A}_F)$ is closed in $L^2(0, T; H)$ and $N(\mathscr{A}_F^*)$ is finite-dimensional.

By virtue of Proposition 4.58 and of (4.305), the pair (A, B) is p-stabilizable if and only if there is an $F \in L(H, U)$ such that $R(I - \mathrm{e}^{(A+BF)T})$ is closed in H and $\dim N(I - \mathrm{e}^{(A^*+F^*B^*)T}) < \infty$. In particular, this happens if either e^{AT} is compact in H, or if the pair (A, B) is stabilizable, that is, there is an $F \in L(H, U)$ such that $A + BF$ generates an exponentially stable semigroup.

Definition 4.62 The pair (A, C) is said to be π-detectable if there is a $K \in L(Z, H)$ such that $R(\mathscr{A}_K)$ is closed in $L^2(0, T; H)$ and $N(\mathscr{A}_K) < \infty$.

Here, $\mathscr{A}_K = \mathscr{A} + KC$.

Throughout this paper, by a solution y to the state equation (4.293), we mean a weak solution, that is, $\mathscr{A} y = Bu + f$.

4.5.2 Existence and the Maximum Principle

We study the existence in problem (4.292) under the following assumptions:

(i) The pair (A, C) is π-detectable.
(ii) $g : Z \to \overline{\mathbb{R}}^*$, $h : U \to \overline{\mathbb{R}}^*$ are convex and lower-semicontinuous and

$$g(z) \geq \alpha |z|_Z + \beta, \quad \forall z \in Z, \tag{4.309}$$

$$h(u) \geq \omega |u|_U^2 + \gamma, \quad \forall u \in U, \tag{4.310}$$

where $\alpha, \omega > 0$ and $\beta, \gamma \in \mathbb{R}$.

Theorem 4.63 *Assume that there is at least one admissible pair* (y, u) *in problem* (4.292). *Then, under hypotheses* (i) *and* (ii), *problem* (4.292) *has at least one solution,* $(y^*, u^*) \in C([0, T]; H) \times L^2(0, T; U)$.

Proof Let $(y_n, u_n) \in C([0, T]; H) \times L^2(0, T; U)$ be such that $Ay_n = Bu_n + f$ and

$$\inf (4.292) = d \leq \int_0^T \big(g\big(Cy_n(t)\big) + h\big(u_n(t)\big)\big)\, \mathrm{d}t \leq d + n^{-1}. \tag{4.311}$$

By (4.309) and (4.310), we have

$$\|Cy_n\|_{L^2(0,T;H)} + \|u_n\|_{L^2(0,T;U)} \leq C_1. \tag{4.312}$$

By (i), there is a $K \in L(Z, H)$ such that $R(\mathscr{A}_K)$ is closed ($\mathscr{A}_K = \mathscr{A} + KC$) and $\dim\ N(\mathscr{A}_K) < \infty$. We have

$$\mathscr{A}_K y_n = Bu_n + KCy_n + f \tag{4.313}$$

and set $y_n = y_n^1 + y_n^2$, where $y_n^1 = \mathscr{A}_K^{-1}(Bu_n + KCy_n + f) \in R(\mathscr{A}_K^*)$ and $y_n^2 \in N(\mathscr{A}_K)$. Then, by (4.308) and (4.312), we have

$$\|y_n^1\|_{C([0,T];H)} \le C_2, \quad \forall n \in \mathbb{N}. \tag{4.314}$$

On the other hand, by the closed range theorem, we know that

$$N(\mathscr{A}_K) = N(\mathscr{C}_K) \oplus R(\mathscr{C}_K^*).$$

We have denoted by $\mathscr{C}_K \in L(N(\mathscr{A}_K), L^2(0,T;Z))$ the operator $y \to Cy$ restricted to $N(\mathscr{A}_K)$. Since $N(\mathscr{A}_K)$ is finite-dimensional, $\mathscr{C}_K$ has closed range in $L^2(0,T;Z)$, and because $\{Cy_n^2\}$ is bounded in $L^2(0,T;Z)$, it is bounded in $L^2(0,T;Z)$, as well. We have, therefore,

$$y_n^2 = z_n^1 + z_n^2,$$

where $\{z_n^1\}$ is bounded in $L^2(0,T;H)$ and $Cz_n^2 = 0$ a.e. in $(0,T)$. We may assume, therefore, that the sequence $\{y_n^1 + z_n^1\}$ is weakly compact in $L^2(0,T;H)$ and, on a subsequence again denoted $\{n\}$, we have

$$\begin{aligned} u_n &\to u^* \quad \text{weakly in } L^2(0,T;U),\\ y_n^1 + z_n^1 &\to y^* \quad \text{weakly in } L^2(0,T;H). \end{aligned}$$

Recalling that $\mathscr{A}(y_n^1 + z_n^1) = Bu_n + f$, we infer that $\mathscr{A}y^* = Bu^* + f$, and, since the convex integrand is weakly lower-semicontinuous, we get

$$d = \int_0^T \big(g\big(Cy^*(t)\big) + h\big(u^*(t)\big)\big)\,dt; \tag{4.315}$$

that is, (y^*, u^*) is optimal in problem (4.292). This completes the proof. □

In order to get the maximum principle for problem (4.292), we use the following assumptions:

(j) *The pair* (A, B) *is* π-stabilizable.
(jj) *The function* $g : Z \to \mathbb{R}$ *is convex and lower-semicontinuous,* $h : U \to \overline{\mathbb{R}}^*$ *is convex and lower-semicontinuous*, $\operatorname{int}\operatorname{Dom}(h) \neq \emptyset$.
(jjj) *The function* f *is in* $C([0,T];H)$*, and one of the following two conditions hold.*

(jjj)$_1$ $\operatorname{Dom}(h) = U$ *and* h *is bounded on every bounded subset of* U.
(jjj)$_2$ $f(t) = Bf_0(t)$*, where* $f_0 \in C([0,T];U)$ *and* $-f_0(t) \in \operatorname{int}\operatorname{Dom}(h)$, $\forall t \in [0,T]$.

Theorem 4.64 *Assume that hypotheses* (j), (jj) *and* (jjj) *hold. Then the pair* $(y^*, u^*) \in C([0,T];H) \times L^2(0,T;U)$ *is optimal in problem* (4.292) *if and only if there are* $p \in C([0,T];H)$ *and* $\eta \in L^\infty(0,T;Z)$ *such that*

$$\frac{dy^*}{dt} = Ay^* + Bu^* + f \quad \text{in } (0,T); \qquad y^*(0) = y^*(T), \tag{4.316}$$

$$\frac{\mathrm{d}p}{\mathrm{d}t} = -A^* p + C^* \eta \quad \mathit{in}\ (0, T); \qquad p(0) = p(T), \tag{4.317}$$

$$\eta(t) \in \partial g\big(Cy^*(t)\big) \quad \mathit{a.e.}\ t \in (0, T), \tag{4.318}$$

$$u^*(t) \in \partial h^*\big(B^* p(t)\big) \quad \mathit{a.e.}\ t \in (0, T). \tag{4.319}$$

System (4.316) and (4.317) is considered, of course, in the weak sense,

$$\mathscr{A} y = Bu^* + f; \qquad \mathscr{A}^* p = -C^* \eta. \tag{4.316$'$}$$

Proof It is readily seen that (4.316)–(4.319) are sufficient for optimality. To prove necessity, we fix an optimal pair (y^*, u^*) and consider the approximation control problem

$$\begin{gathered}\operatorname{Min}\left\{\int_0^T \big(g_\varepsilon(Cy) + h(u) + 2^{-1}\big(|y - y^*|^2 + |u - u^*|_U^2 + \varepsilon^{-1}|v|^2\big)\big)\,\mathrm{d}t\right\}\\ \text{subject to} \\ \mathscr{A} y = Bu + v + f, \quad u \in L^2(0, T; U),\ v \in L^2(0, T; H),\ y \in C\big([0, T]; H\big).\end{gathered} \tag{4.320}$$

Here, $g_\varepsilon \in C^1(Z)$ is defined as in Sect. 2.2.3.

Arguing as above, it is easily seen that problem (4.320) has a unique solution $(y_\varepsilon, u_\varepsilon, v_\varepsilon) \in C([0, T]; H) \times L^2(0, T; U) \times L^2(0, T; H)$, and arguing as in Sect. 4.1, we have for $\varepsilon \to 0$

$$\begin{aligned} u_\varepsilon &\to u^* \quad \text{strongly in } L^2(0, T; U),\\ y_\varepsilon &\to y^* \quad \text{strongly in } L^2(0, T; H). \end{aligned} \tag{4.321}$$

We also have

$$v_\varepsilon \to 0 \quad \text{strongly in } L^2(0, T; H).$$

Next, we have

$$\begin{aligned} &\int_0^T \Big(\big(C^* \nabla g_\varepsilon(Cy_\varepsilon), z\big) + (y_\varepsilon - y^*, z) + (u_\varepsilon - u^*, w)_U \\ &\qquad + h'(u_\varepsilon, w) + \varepsilon^{-1}(v_\varepsilon, v)\Big)\,\mathrm{d}t \ge 0, \end{aligned} \tag{4.322}$$

$\forall (z, w, v) \in C([0, T]; H) \times L^2(0, T; U) \times L^2(0, T; H)$ such that $\mathscr{A} z = Bw + v$. We set $p_\varepsilon = \varepsilon^{-1} v_\varepsilon$. Then (4.322) yields

$$\begin{aligned} &\int_0^T \Big(\big(C^* \nabla g_\varepsilon(Cy_\varepsilon) + y_\varepsilon - y^*, z\big) + (u_\varepsilon - u^*, w)_U \\ &\qquad + h'(u_\varepsilon, w) + (p_\varepsilon, \mathscr{A} z - Bw)\Big)\,\mathrm{d}t \ge 0, \end{aligned} \tag{4.322$'$}$$

$\forall z \in D(\mathscr{A})$, $\forall w \in L^2(0,T;U)$. (Here, h' is the directional derivative of h.) For $w=0$, the latter yields

$$\mathscr{A}^* p_\varepsilon = -C^* \nabla g_\varepsilon(Cy_\varepsilon) + y^* - y_\varepsilon. \tag{4.323}$$

Substituting the latter into (4.322′), we get

$$\int_0^T (B^* p_\varepsilon + u^* - u_\varepsilon, w)_U \, \mathrm{d}t \le \int_0^T h'(u_\varepsilon, w) \, \mathrm{d}t, \quad \forall w \in U.$$

This yields

$$B^* p_\varepsilon \in \partial h(u_\varepsilon) + u_\varepsilon - u^* \quad \text{a.e. } \in (0,T). \tag{4.324}$$

We note also

$$\mathscr{A} y_\varepsilon = Bu_\varepsilon + \varepsilon p_\varepsilon + f. \tag{4.325}$$

We are going to let ε tend to 0 in (4.323) and (4.324) in order to get (4.317)–(4.319). To this aim, some a priori estimates on p_ε are necessary. Assume, first, that condition $(\mathrm{jjj})_2$ holds. Then, by (4.324) and by the definition of ∂h, we have

$$\begin{aligned} &\big(B^* p_\varepsilon(t) + u^*(t) - u_\varepsilon(t), u_\varepsilon(t) + f_0(t) - \rho w\big)_U \\ &\quad \ge h\big(u_\varepsilon(t)\big) - h\big(\rho w - f_0(t)\big) \quad \text{a.e. } t \in (0,T), \end{aligned} \tag{4.326}$$

$\forall w \in U$, $|w|_U = 1$, and ρ positive and sufficiently small. This yields

$$\rho \int_0^T \big|B^* p_\varepsilon(t)\big|_U \, \mathrm{d}t \le \int_0^T \big(p_\varepsilon(t), \mathscr{A} y_\varepsilon(t) - \varepsilon p_\varepsilon(t)\big) \, \mathrm{d}t + C_3.$$

Finally, by (4.316′) we obtain

$$\begin{aligned} &\int_0^T \big(\rho\big|B^* p_\varepsilon(t)\big|_U + \varepsilon\big|p_\varepsilon(t)\big|^2\big) \, \mathrm{d}t \\ &\quad \le -\int_0^T \Big(\big(\nabla g_\varepsilon\big(Cy_\varepsilon(t)\big), Cy_\varepsilon(t)\big)_Z + \big(y_\varepsilon(t) - y^*(t), y_\varepsilon(t)\big)\Big) \, \mathrm{d}t \le C_4, \end{aligned}$$

because ∇g_ε is monotone. On the other hand, it follows that $\{y_\varepsilon\}$ is strongly convergent to y^* in $C([0,T];H)$. Indeed, by (4.325), we see that

$$\begin{aligned} &y_\varepsilon(t) = \mathrm{e}^{-\mathscr{A}_F(t-s)} y_\varepsilon(s) + \int_s^t \mathrm{e}^{-\mathscr{A}_F(t-r)} \big(Bu_\varepsilon(r) + \varepsilon p_\varepsilon(r) + BFy_\varepsilon(r)\big) \, \mathrm{d}r, \\ &0 \le s \le t \le T, \end{aligned}$$

and the conclusion follows by (4.321). (Here and everywhere in the following, $F \in L(H,U)$ is chosen as in Definition 4.61 and $\mathscr{A}_F = A + BF$.) Since ∂g is locally bounded in H and

$$\nabla g_\varepsilon(z) \in \partial g\big((I+\varepsilon\partial g)^{-1}z\big), \quad \forall z \in Z; \qquad \int_0^T g_\varepsilon(Cy_\varepsilon)\,\mathrm{d}t \le C_5,$$

we have

$$\big|\nabla g_\varepsilon\big(Cy_\varepsilon(t)\big)\big|_Z \le C_6, \quad \forall \varepsilon > 0,\ t \in [0,T]. \tag{4.327}$$

We may rewrite (4.323) as

$$\mathscr{A}_F^* p_\varepsilon = -C^*\nabla g_\varepsilon(Cy_\varepsilon) - (y_\varepsilon - y^*) + F^*B^*p_\varepsilon. \tag{4.328}$$

Then, by (4.327) and Corollary 4.59, it follows that

$$\big|p_\varepsilon^1(t)\big| \le C_7, \quad \forall t \in [0,T], \tag{4.329}$$

where $p_\varepsilon = p_\varepsilon^1 + p_\varepsilon^2$ and $p_\varepsilon^1 \in R(\mathscr{A}_F) = N(\mathscr{A}_F^*)^\perp$, $p_\varepsilon^2 \in N(\mathscr{A}_F^*)$.

Denote by $\mathscr{B}_F^*$ the operator $y \to B^*y$ defined from $N(\mathscr{A}_F^*)$ to $L^2(0,T;U)$. Recalling that the space $N(\mathscr{A}_F^*)$ is finite-dimensional, we infer that $\mathscr{B}_F^*$ has a closed range in $L^2(0,T;U)$, so by the closed range theorem, it has a bounded inverse on its range. Since $N(\mathscr{A}_F^*) \subset C([0,T];H)$ and $\{\mathscr{B}_F^* p_\varepsilon^2\}$ is bounded in $L^1(0,T;U)$, it is bounded in $L^2(0,T;U)$ too, and we have $p_\varepsilon^2 = q_\varepsilon^1 + q_\varepsilon^2$, where $\{q_\varepsilon^1\}$ is bounded in $L^2(0,T;H)$ and $B^*q_\varepsilon^2 = 0$ a.e. in $(0,T)$. We conclude, therefore, that the sequence $\{p_\varepsilon^1 + q_\varepsilon^1\}$ is weakly compact in $L^2(0,T;H)$. Moreover, we may write (4.328) as

$$\mathscr{A}_F^*\big(p_\varepsilon^1 + q_\varepsilon^1\big) = -C^*\nabla g_\varepsilon(Cy_\varepsilon) - (y_\varepsilon - y^*) + F^*B^*\big(p_\varepsilon^1 + q_\varepsilon^1\big). \tag{4.328′}$$

Selecting further subsequences, if necessary, we may assume that

$$\begin{aligned} p_\varepsilon^1 + q_\varepsilon^1 &\to p \quad \text{weakly in } L^2(0,T;H),\\ \nabla g_\varepsilon(Cy_\varepsilon) &\to \eta \quad \text{weak-star in } L^\infty(0,T;Z). \end{aligned}$$

Since ∂g and ∂h are maximal monotone (and, therefore, weakly–strongly closed), we may pass to the limit in (4.323) and (4.328′) to get the optimality system (4.316)–(4.319).

Assume, now, that condition (jjj)$_1$ is satisfied. We set $p_\varepsilon = p_\varepsilon^1 + q_\varepsilon^1$. Then, by (4.324), we have

$$h(u_\varepsilon) - h(\rho w) \le (B^*p_\varepsilon + u^*, u_\varepsilon - \rho w)_U,$$

for all $w \in H$, $\rho > 0$. This yields

$$\begin{aligned} \rho\int_0^T \big|B^*p_\varepsilon(t)\big|_U\,\mathrm{d}t &\le C_8 + Th(\rho w) + \int_0^T \big(p_\varepsilon(t), \mathscr{A}y_\varepsilon(t) - f(t)\big)\,\mathrm{d}t\\ &\le Th(\rho w) + C_9\bigg(1 + \int_0^T \big|p_\varepsilon(t)\big|\,\mathrm{d}t\bigg). \end{aligned}$$

Finally,

$$\int_0^T \left|B^* p_\varepsilon(t)\right|_U \mathrm{d}t \le C_\rho + C_{10}\rho^{-1} \int_0^T \left|p_\varepsilon(t)\right| \mathrm{d}t, \tag{4.330}$$

for all $\rho > 0$. Choosing ρ sufficiently large, it follows by (4.328′) and (4.310) that $\{p_\varepsilon^1 + q_\varepsilon^1\}$ is bounded in $L^2(0, T; H)$, so we may conclude the proof as in the previous case. □

The optimal control problem

$$\mathrm{Min}\left\{ \int_0^T \left(g_*\big(v(t)\big) + h^*\big(B^*(t)\big) + \big(f(t), p(t)\big)\right) \mathrm{d}t; \right.$$
$$\left. p \in C\big([0, T]; H\big),\ v \in L^2(0, T; H)\right\} \tag{4.331}$$

subject to

$$\mathscr{A}^* p = -v \tag{4.332}$$

is the dual of (4.292) in the sense of Sect. 4.1.8. Here, h^* is the conjugate of h, and g_* is the conjugate of the function $y \to g(Cy)$.

Theorem 4.65 *Under the assumptions of Theorem 4.64, the pair (y^*, u^*) is optimal in problem (4.292) if and only if the dual problem (4.331) has a solution (p^*, v^*) and*

$$\int_0^T \Big(g\big(Cy^*(t)\big) + h\big(u^*(t)\big) \mathrm{d}t + \int_0^T \Big(g_*\big(v^*(t)\big) + h^*\big(B^* p^*(t)\big)\Big)$$
$$+ \big(f(t), p^*(t)\big)\Big) \mathrm{d}t = 0. \tag{4.333}$$

Proof The argument is standard (see Sect. 4.1) and so the proof is only sketched. If (y^*, u^*) is optimal in (4.292), then, by Theorem 4.64, the optimality system (4.316)–(4.319) has a solution $(p^*, v^* = C^*\eta)$, and, by virtue of the conjugacy relation, we have

$$\begin{aligned} h(u^*) + h^*(B^* p^*) &= (B^* p^*, u^*)_U, \\ g(Cy^*) + g_*(v^*) &= (y^*, v^*). \end{aligned} \tag{4.334}$$

Integrating from 0 to T, we get (4.333). On the other hand, for all $(p, v) \in C([0, T]; H) \times L^2(0, T; H)$, $\mathscr{A}^* p = -v$, we have

$$\begin{aligned} h(u^*) + h^*(B^* p) &\ge (B^* p, u^*)_U, \quad \text{a.e. on } (0, T), \\ g(Cy^*) + g_*(v) &\ge (y^*, v), \quad \text{a.e. on } (0, T), \end{aligned} \tag{4.335}$$

which imply that the pair (p^*, v^*) is optimal in problem (4.331). Conversely, if (4.333) holds, then by (4.334) and (4.335) we see that y^*, p^*,and u^* satisfy the optimality system (4.316)–(4.319), and therefore (y^*, u^*) is optimal in problem (4.292). $\square$

We end this section with a few examples of linear control systems of the form (4.293) for which the previous theorems are applicable.

1. *Parabolic control problems*. Consider the system

$$\begin{aligned} &\frac{\partial y}{\partial t} - \Delta y + b(x)\cdot\nabla y + c(x)y = Bu + f(t,x), \quad (t,x) \in \Omega \times \mathbb{R}, \\ &y = 0, \quad \text{on } \partial\Omega \times \mathbb{R}, \\ &y(t+T,x) = y(t,x), \quad \forall (t,x) \in \Omega \times \mathbb{R}, \end{aligned} \tag{4.336}$$

where $b \in W^{1,\infty}(\Omega;\mathbb{R}^n)$, $c \in L^\infty(\Omega)$, $f \in L^2_{\mathrm{loc}}(\mathbb{R}; L^2(\Omega))$ is T-periodic in t, while $B \in L(L^2(\Omega), L^2(\Omega))$. Here, Ω is a bounded and open subset of $\mathbb{R}^n$ with a sufficiently smooth boundary $\partial\Omega$. We may write (4.336) in the form (4.293), where $H = U = L^2(\Omega)$ and

$$Ay = \Delta y + b\cdot\nabla y - cy, \qquad D(A) = H_0^1(\Omega)\cap H^2(\Omega). \tag{4.337}$$

Since the semigroup e^{At} generated by A on $L^2(\Omega)$ is compact, it follows that $R(\mathscr{A})$ is closed in $L^2(0,T;H)$ and $N(\mathscr{A}^*)$ is finite-dimensional (Corollary 4.60). Hence, the π-stabilizability hypothesis (j) is satisfied in the present situation with $F = 0$. A similar conclusion can be reached for the π-detectability hypothesis (i).

2. *Linear delay control systems*. Consider the control system governed by the delay system

$$\begin{aligned} y'(t) &= A_0 y(t) + A_1 y(t-h) + B_0 u(t) + f(t), \\ y(t) &= y(t+T), \quad \forall t \in \mathbb{R}, \end{aligned} \tag{4.338}$$

where A_0, A_1 are $n\times n$ matrices, B_0 is an $n\times\ell$ matrix, $f \in L^2_{\mathrm{loc}}(\mathbb{R};\mathbb{R}^n)$, $f(t+T) = f(t)$, $u \in L^2(\mathbb{R};\mathbb{R}^\ell)$, and $u(t) = u(t+T)$. It is well known that this system can be written in the form (4.293), where $H = M_2 = \mathbb{R}^n \times L^2(-h,0;\mathbb{R}^n)$, $U = \mathbb{R}^\ell$, $B = (B_0, 0)$, and

$$\begin{aligned} A\big(y_0, y^0\big) &= \left\{A_0 y_0 + A_1 y^0(-h), \frac{\mathrm{d}y^0}{\mathrm{d}s}\right\}, \\ D(A) &= \big\{\big(y_0, y^0\big) \in \mathbb{R}^n \times W^{1,2}\big([-h,0];\mathbb{R}^n\big),\ y_0 = y^0(0)\big\}. \end{aligned}$$

For each $m \in \mathbb{Z}$, we may rewrite the equation $(\mu_m i I - A)y = (f_0, f_1)$ as

$$\begin{aligned} \big(i\mu_m I - A_0 - e^{-i\mu_m h}A_1\big)y_0 &= f_0 + \int_0^{-h} e^{-i\mu_m(h+s)}A_1 f_1(s)\,\mathrm{d}s, \\ y^0(s) &= e^{i\mu_m s}y_0 - \int_0^s e^{i\mu_m(s-t)} f_1(t)\,\mathrm{d}t, \end{aligned} \tag{4.339}$$

where $\mu_m = 2m\pi T^{-1}$. Moreover, after some calculation, we see that

$$N(\mathscr{A}^*) = \left\{ \sum_m \left(y_m \mathrm{e}^{i\mu_m t}, A_1^* y_m \mathrm{e}^{i\mu_m(t-s-h)}\right); \left(i\mu_m I + A_0^* + \mathrm{e}^{i\mu_m h} A_1^*\right) y_m = 0 \right\}. \tag{4.340}$$

By (4.339), we see that $R(i\mu_m I - A)$ is closed and condition (4.298) in Proposition 4.56 is satisfied. We conclude, therefore, that the corresponding operator $\mathscr{A}$ has a closed range in $L^2(0,T;H)$. Moreover, by (4.339) and (4.340), it follows that $N(\mathscr{A})$ and $N(\mathscr{A}^*)$ are finite-dimensional.

3. *First-order hyperbolic systems*. Consider the control system governed by the linear system

$$\begin{aligned}
&y_t(t,x) - z_x(t,x) = u(t,x) + f(t,x), \quad x \in (0,1),\ t \in (0,T),\\
&z_t(t,x) - y_x(t,x) = B_0 v(t,x) + g(t,x), \quad x \in (0,1),\ t \in (0,T),\\
&y(t,0) = y(t,1) = 0; \qquad y(T,x) = y(x,0), \qquad z(T,x) = z(x,0), \quad \forall x \in (0,1).
\end{aligned} \tag{4.341}$$

Here, $B_0 \in L(L^2(0,1), L^2(0,1))$ and $f, g \in C([0,T]; L^2(0,1))$ are given functions. System (4.341) can be written in the form (4.293), where $H = U = L^2(0,1) \times L^2(0,1)$, $B(u,v) = (u, B_0 v)$, and $A(y,z) = (z_x, y_x)$, $D(A) = \{y, z \in H^1(0,1), y(0) = y(1) = 0\}$. Consider the feedback control $F(y,z) = (-y, 0)$. Then it is easily seen that the corresponding operator $\mathscr{A}_F$ has closed range in $L^2(0,T;H)$, $N(\mathscr{A}_F) = N(\mathscr{A}_F^*) = \{(0,C); C \in \mathbb{R}\}$, and therefore the pair (A,B) is π-stabilizable. This simple example extends to linear control hyperbolic systems in $\mathbb{R}^n \times \mathbb{R}^n$, and it is instructive to notice that, in this case, if T is irrational, then $R(\mathscr{A})$ is not closed; thus assumption (j) does not hold with $F = 0$.

4.5.3 The Optimal Control of the Wave Equation

We study here the optimal control problem

$$\text{minimize} \int_0^T \left(2^{-1}\left|Cy(t)\right|_Z^2 + h\big(u(t)\big)\right) \mathrm{d}t \tag{4.342}$$

subject to $u \in L^2(0,T;H)$, $y \in L^2(0,T;U)$,

$$\begin{aligned}
&y'' + Ay = Bu + f, \quad t \in (0,T),\\
&y(0) = y(T), \qquad y'(0) = y'(T),
\end{aligned} \tag{4.343}$$

where A is a self-adjoint, linear, and positively defined operator in H, $B \in L(U,H)$, $C \in L(H,Z)$, and h is a lower-semicontinuous convex function on U. By a *weak solution* to (4.343), we mean a function $y \in L^2(0,T;H)$ such that

$$\int_0^T \big(y(t), \varphi''(t) + A_0\varphi(t)\big)\,\mathrm{d}t = \int_0^T \big(f(t) + Bu(t), \varphi(t)\big)\,\mathrm{d}t, \tag{4.344}$$

for all $\varphi \in Y = \{\varphi \in C^2([0,T];H) \cap C([0,T];D(A));\ \varphi(0) = \varphi(T),\ \varphi'(0) = \varphi'(T)\}$. Equivalently,

$$\mathscr{W} y = Bu + f, \tag{4.343'}$$

where $\mathscr{W} : D(\mathscr{W}) \subset L^2(0,T;H) \to L^2(0,T;H)$ is the linear operator defined by

$$\mathscr{W} y = f \quad \text{iff} \int_0^T \big(y(t), \varphi''(t) + A_0\varphi(t)\big)\,\mathrm{d}t = \int_0^T \big(f(t), \varphi(t)\big)\,\mathrm{d}t, \quad \forall \varphi \in Y. \tag{4.345}$$

It is readily seen that $\mathscr{W}$ is densely defined and closed in $L^2(0,T;H)$.

Writing (4.343) as a first-order differential equation on the product space $D(A^{\frac{1}{2}}) \times H$, we may apply the general results obtained in the previous section to problem (4.342). However, a direct treatment of such a problem requires less restrictive conditions in specific examples. On the other hand, for the sake of simplicity, we do not put the results of this section in the general framework of the p-stabilizability condition; we confine ourselves to assuming that $R(\mathscr{W})$ is closed in $L^2(0,T;H)$. By virtue of the closed range theorem, this assumption implies that

$$L^2(0,T;H) = R(\mathscr{W}) \oplus N(\mathscr{W}); \qquad \mathscr{W}^{-1} \in L\big(R(\mathscr{W}), L^2(0,T;H)\big).$$

Arguing as in the proof of Theorem 4.63, it follows that if $R(\mathscr{W})$ is closed in $L^2(0,T;H)$ and $N(\mathscr{W})$ is finite-dimensional, then problem (4.342) has at least one solution $(y,u) \in L^2(0,T;H) \times L^2(0,T;U)$. As regards the maximum principle, we have the following theorem.

Theorem 4.66 *Assume that $R(\mathscr{W})$ is closed*, $\dim N(\mathscr{W}) < \infty$, *and h, f satisfy hypotheses* (jj), (jjj). *Then the pair $(y^*, u^*) \in L^2(0,T;H) \times L^2(0,T;U)$ is optimal in problem* (4.342) *if and only if there is a $p \in L^2(0,T;H)$ such that*

$$\mathscr{W} p = -C^*Cy, \tag{4.346}$$

$$u^*(t) \in \partial h^*\big(B^*p(t)\big), \quad \text{a.e. } t \in (0,T). \tag{4.347}$$

We omit the proof because it is identical with that of Theorem 4.64. Since in most of the applications the null space $N(\mathscr{W})$ is finite-dimensional (the state equation is highly resonant), we may relax this condition as follows.

(k) *$R(\mathscr{W})$ is closed and the operator $y \to B^*y$ defined from $N(\mathscr{W})$* to $L^2(0,T;H)$ *has closed range.*

Theorem 4.67 *Assume that hypotheses* (jj), (k) *hold*, $f \in C([0,T];H)$, *and that the function h has quadratic growth, i.e.*,

$$h(u) \le \alpha_1|u|_U^2 + \beta_1, \quad \forall u \in U. \tag{4.348}$$

Then the pair $(y^*, u^*) \in L^2(0,T;H) \times L^2(0,T;U)$ is optimal in problem (4.342) *if and only if it satisfies system* (4.346) *and* (4.347).

Proof Let $(y_\varepsilon, u_\varepsilon, v_\varepsilon)$ be the solution to the approximating problem (see (4.320))

$$\operatorname{Min}\left\{ \int_0^T \left(2^{-1}|Cy|_Z^2 + h(u) + 2^{-1}\left(|y-y^*|^2 + |u-u^*|_U^2 + \varepsilon^{-1}|v|^2\right)\right) \mathrm{d}t; \right.$$
$$\left. \mathscr{W} y = Bu + v + f \right\}.$$

As in the proof of Theorem 4.64, we get (4.311) and (see (4.323) and (4.324))

$$\mathscr{W} p_\varepsilon = -C^*Cy_\varepsilon + y^* - y_\varepsilon, \tag{4.349}$$

$$B^* p_\varepsilon \in \partial h(u_\varepsilon) + u_\varepsilon - u^* \quad \text{a.e. in } (0, T). \tag{4.350}$$

By (4.348) and (4.349), we have

$$\|B^* p_\varepsilon\|^2_{L^2(0,T;U)} \le C_1, \quad \forall \varepsilon > 0,$$

and therefore, by virtue of assumption (k), we conclude via the closed range theorem that

$$\left\{p_\varepsilon^1 + p_\varepsilon^3\right\} \quad \text{is bounded in } L^2(0, T; H),$$

where $p_\varepsilon = p_\varepsilon^1 + p_\varepsilon^3 + p_\varepsilon^4$ and $p_\varepsilon^1 \in \mathbb{R}(\mathscr{W})$, $p_\varepsilon^3, p_\varepsilon^4 \in \mathbb{N}(\mathscr{W})$, and $B^* p_\varepsilon^4 = 0$ a.e. $t \in (0, T)$. Hence, we may pass to the limit in (4.349) and (4.350) to get (4.346) and (4.347), as desired. □

The dual problem of (4.342) is (see (4.331) and (4.332))

$$\operatorname{Min}\left\{ \int_0^T \left(g_*(v) + h^*(B^*p) + (f, p)\right) \mathrm{d}t;\ \mathscr{W}^* p = -v;\ v \in L^2(0, T; H)\right\}. \tag{4.351}$$

By using exactly the same argument, it follows that under the assumptions of Theorem 4.66 or Theorem 4.67, the conclusions of the duality Theorem 4.65 remain valid in the present case.

Example 4.68 The one-dimensional wave equation. Consider the control system

$$\begin{aligned} & y_{tt}(t,x) - v^{-1}(x)\left(v(x)y_x(t,x)\right)_x = Bu(t,x) + f(t,x), \quad (t,x) \in (0,\pi) \times \mathbb{R}, \\ & y(0,t) = y(\pi,t) = 0, \quad t \in \mathbb{R}, \\ & y(x,t+T) = y(t,x), \qquad y_t(x,t+T) = y_t(t,x), \quad (t,x) \in (0,\pi) \times \mathbb{R}, \end{aligned} \tag{4.352}$$

where $v \in H^2(0, T)$, $v(x) > 0$, $\forall x \in [0, \pi]$, $B \in L(L^2(0, \pi), L^2(0, \pi))$, and

$$\operatorname{ess\,sup}\left\{\left(v'(x)\right)^2 - 2v''(x)v(x);\ x \in (0, k\pi)\right\} < 0.$$

In this case, $U = L^2(0, \pi)$, $H = L^2(0, \pi)$ is endowed with the scalar product $(y, z) = \int_0^\pi v(x)y(x)z(x)\,\mathrm{d}x$ and

$$A_0 y = -v^{-1}(v y_x)_x, \qquad D(A) = H_0^1(0, \pi) \cap H^2(0, \pi).$$

Equation (4.352) models the forced vibrations of a nonhomogeneous string as well as the propagation of waves in nonisotropic media. (In the latter case, $\nu = (\rho\mu)^{\frac{1}{2}}$ is the acoustic impedance, ρ is the medium density and μ is the elasticity coefficient.) If T is a rational multiple of π, then $R(\mathscr{W})$ is closed, $N(v)$ is finite-dimensional (see Barbu and Pavel [17]), and Theorem 4.67 is applicable.

4.6 Problems

4.1 Find the maximum principle for the optimal control problem

$$\operatorname{Min}\left\{\int_0^T L\big(y(t),u(t)\big)\,\mathrm{d}t;\ y' = A_0y(t) + A_1y(t-b) + Bu(t),\ t \in (-b,T),\right.$$
$$\left. y(0) = y_0,\ y(s) = y^0(s),\ s \in (-b,0)\right\},$$

where $L : \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}^*$ satisfies conditions of Theorem 4.16 (or Theorem 4.5) and $A_0, A_1 \in L(\mathbb{R}^n,\mathbb{R}^n)$, $B \in L(\mathbb{R}^m,\mathbb{R}^n)$, $b > 0$.

Hint. In the space $F = M^2(-b,0;\mathbb{R}^n) = \mathbb{R}^n \times L^2(-b,0;\mathbb{R}^n)$, we rewrite the above delay system as

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = \mathscr{A}Y + Bu, \quad Y(0) = \big(y_0,y^0\big),$$

where $\mathscr{A}(y_0,y^0) = \{A_0y_0 + A_1y^0(-b), \frac{\mathrm{d}y^0}{\mathrm{d}s}\}$, $D(\mathscr{A}) = \{(y_0,y^0) \in \mathbb{R}^n \times W^{1,2}(-b,0;\mathbb{R}^n);\ y^0(0) = y_0\}$ and $Y(t) = \{y(t), y(t+s)\}$. Then we may rewrite the above problem in the form (P) by redefining L and ℓ as

$$L(Y,u) = L(y_0,u), \quad Y = \big(y_0,y^0\big),$$
$$\ell(Y_1,Y_2) = \begin{cases} 0, & \text{if } Y_1 = (y_0,y^0), \\ +\infty, & \text{otherwise.} \end{cases}$$

Then one might apply, under suitable conditions on L, Theorem 4.5. We note that the dual system (4.100) and (4.101) has the form

$$\tilde{p}(t) = \big(p(t), p_1(t,\theta)\big), \quad \theta \in (-b,0),$$
$$p'(t) + A_T^*p(t) = \begin{cases} q(t), & 0 < t < T-b, \\ q(t) - z(t-T), & T-b < t < T, \end{cases}$$
$$q(t) \in \partial_x L\big(y(t),u(t)\big), \qquad B^*p(t) \in \partial_u L\big(y(t),u(t)\big),$$
$$A_T^*p(t) = \begin{cases} A^*p(t) + A_1^*(t-b)p(t+b), & \text{if } t+b < T, \\ 0, & \text{if } t+b > T, \end{cases}$$

$$p_1(0,\theta)=\begin{cases}A_1(\theta+h), & \text{if } -b<\theta<0,\\ 0, & \text{if } -b<\theta<\theta_1,\end{cases}$$
$$z(t)=p_1(T,t).$$

4.2 Show that if $H:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ is a concave–convex continuous function and $\ell_i:\mathbb{R}\to\mathbb{R}$, $i=1,2$, are convex and differentiable, then the hyperbolic first-order system

$$\begin{aligned}&\frac{\partial y}{\partial x}-\partial_p H(y,p)=0, \quad x\in(0,1),\\ &\frac{\partial p}{\partial x}+\partial_y H(y,p)=f(x), \quad x\in(0,1),\\ &p(0)=\nabla\ell_1\big(y(0)\big), \qquad p(1)+\nabla\ell_2\big(y(1)\big)=0,\end{aligned}$$

has a solution y for each $f\in C[0,1]$.

Hint. We associate with the above system the minimization problem

$$\begin{aligned}\operatorname{Min}\bigg\{&\int_0^1 L\big(y(x),y'(x)\big)\,\mathrm{d}x+\int_0^1 f(x)y(x)\,\mathrm{d}x+\ell_1\big(y(0)\big)+\ell_2\big(y(1)\big);\\ &y\in W^{1,1}\big([0,1];\mathbb{R}^n\big)\bigg\}\end{aligned}$$

and apply Theorem 4.5. (See Barbu [5] for a more general result pertaining to the variational treatment of hyperbolic systems of this type.)

4.3 Let A be an infinitesimal generator of a C_0-semigroup in the Banach space E and let $B\in L(U,E)$, where U is another Banach space. Assume that

$$\big\|p(0)\big\|_E^2\le C_T\int_0^T\big\|B^*p(t)\big\|_U^2\,\mathrm{d}t,$$

for all the solutions p to the backward differential system $p'(t)=-A^*p(t)$. Show that there is a controller $u^*\in L^2(0,T;U)$ such that $\int_0^T\|u^*(t)\|_U^2\,\mathrm{d}t\le C_T\|y_0\|_E^2$ and $y'=Ay+Bu^*$, $t\in(0,T)$, $y(0)=y_0$, $y(T)=0$.

Hint. Consider the control problem

$$\operatorname{Min}\bigg\{\int_0^T\frac{1}{2}\big\|u(t)\big\|_U^2\,\mathrm{d}t+\frac{1}{2\varepsilon}\big\|y(T)\big\|_E^2;\ y'=Ay+Bu,\ y(0)=y_0\bigg\},$$

which has an optimal pair $(y_\varepsilon,u_\varepsilon^*)$ satisfying the maximum principle

$$p_\varepsilon'=-A^*p_\varepsilon, \qquad p_\varepsilon(T)=-\frac{1}{\varepsilon}\,y_\varepsilon(T).$$

This yields

$$\int_0^T \|B^* p_\varepsilon(t)\|_U^2 \,\mathrm{d}t + \frac{1}{\varepsilon}\|y_\varepsilon(T)\|_E^2 = {}_E\big(y_0, p_\varepsilon(0)\big)_{E^*}.$$

Letting $\varepsilon \to 0$, we obtain the result.

4.4 Show that the solution to the *eikonal equation*

$$\begin{aligned} &\varphi_t - \rho\|\varphi_x\| = 0, \quad t \ge 0,\ x \in \mathbb{R}^n, \\ &\varphi(0,x) = \varphi_0(x), \quad x \in \mathbb{R}^n, \end{aligned}$$

is given by $\varphi(t,x) = \sup\{\varphi_0(y); |x-y| \le \rho t, y \in \mathbb{R}^n\}$.

Hint. One applies the Lax–Hopf formula (4.206).

4.7 Bibliographical Notes

4.1. In a particular form, the main results of this section, Theorems 4.5, 4.6, 4.16 and 4.26, were given by the first author in [2–4]. In a series of influential works on the convex control problem of Bolza in $\mathbb{R}^n$, Rockafellar [39, 41, 42, 44] has developed a theory of the "maximum principle" in subdifferential form under convexity assumptions which inspired the present work. However, the infinite-dimensional case treated here presents several significant differences. For general results on control problems governed by ordinary differential systems, we refer the reader to the book of Berkovitz [18].

The first studies on validity of the maximum principle for a specific form of distributed parameter optimal control problems were published in the early 1960s. The general theory of optimal Banach spaces has been studied in the works of Balakrishnan [1] and Datko [25], among others. The book of Lions [33] first published in 1968 represents a systematic treatment of quadratic optimal control problems governed by partial differential equations. The results presented here largely encompass the previous one by the absence of differentiability assumptions on the integrand L as well as by the generality of the problem studied. The ideas contained in the present approach of convex control problems were used to develop a theory of necessary conditions for control problems with a nonconvex cost criterion and with nonlinear state equations (see Barbu [10, 11]). In this case, the extremality conditions are expressed in term of Clarke's generalized gradient.

Theorem 4.28 is new in this context, though it was known a long time ago that the Cauchy problem associated with nonlinear operators of subpotential type can be reformulated as convex optimization problems (see Brezis and Ekeland [19]). Recent extensions to this principle as well as applications to the existence theory of PDEs can be found in Visintin [46] and Ghoussoub [28].

4.2. There exists an extensive literature on the synthesis of finite-dimensional control systems and of linear evolution control systems with quadratic cost (see Fleming and Rishel [27], Lions [33] and the references given there), but there appears to be little previous work on optimal feedback controllers for infinite-dimensional control convex problems. We refer to the works [21–24] of Crandall and Lions for a viscosity solution theory of the Hamilton–Jacobi equations associated with infinite-dimensional optimal control problems (see also Barbu and Da Prato [14]).
4.3. The abstract formulation of the boundary control systems given here arises from the works of Fattorini [26], Balakrishnan [1], Washburn [47], Lasiecka and Triggiani [31]. In particular, the book [32] by Lasiecka and Triggiani contains a complete description of linear boundary control systems of parabolic and hyperbolic type. Theorem 4.41 was taken from Barbu [9]. Optimality conditions for linear parabolic boundary control problems with convex cost criterion and state constraints have been derived by several authors including Mackenroth [35] and Tröltzsch [45].
4.4. The results of this section closely follow the work [7, 8] of Barbu.
4.5. The main results presented here are taken from the work [12] of Barbu.

References

1. Balakrishnan AV (1965) Optimal control problems in Banach spaces. SIAM J Control 3:152–180
2. Barbu V (1975) Convex control problem of Bolza in Hilbert space. SIAM J Control 13:754–771
3. Barbu V (1975) On the control problem of Bolza in Hilbert spaces. SIAM J Control 13:1062–1076
4. Barbu V (1976) Constrained control problems with convex cost in Hilbert space. J Math Anal Appl 56:502–528
5. Barbu V (1976) Nonlinear boundary value problems for a class of hyperbolic systems. Rev Roum Math Pures Appl 22:502–522
6. Barbu V (1976) Nonlinear semigroups and evolution equations in Banach spaces. Noordhoff/Ed Acad, Leyden/Bucureşti
7. Barbu V (1978) Convex control problems and Hamiltonian systems on infinite intervals. SIAM J Control 16:687–702
8. Barbu V (1978) On convex problems on infinite intervals. J Math Anal Appl 65:859–911
9. Barbu V (1980) Boundary control problems with convex cost criterion. SIAM J Control 18:227–254
10. Barbu V (1981) Necessary conditions for distributed control problems governed by parabolic variational inequalities. SIAM J Control 19:64–86
11. Barbu V (1984) Optimal control of variational inequalities. Research notes in mathematics, vol 100. Pitman, London
12. Barbu V (1997) Optimal control of linear periodic systems in Hilbert spaces. SIAM J Control Optim 35:2137–2150
13. Barbu V (2010) Nonlinear differential equations of monotone type in Banach spaces. Springer, Berlin
14. Barbu V, Da Prato G (1984) Hamilton–Jacobi equations in Hilbert spaces. Research notes in mathematics, vol 93. Pitman, London

15. Barbu V, Da Prato G (1985) Hamilton–Jacobi equations in Hilbert spaces; variational and semigroup approach. Ann Mat Pura Appl 142:303–349
16. Barbu V, Da Prato G (1992) A representation formula for the solutions to operator Riccati equation. Differ Integral Equ 4:821–829
17. Barbu V, Pavel N (1997) Periodic solutions to nonlinear one dimensional wave equations with x-dependent coefficients. Trans Am Math Soc 349:2035–2048
18. Berkovitz D (1974) Optimal control theory. Springer, Berlin
19. Brezis H, Ekeland I (1976) Un principe variationnel associé à certaines équations paraboliques. Le cas indépendant du temps. C R Acad Sci Paris 282:971–974
20. Butkovskiy AG (1969) Distributed control systems. American Elsevier, New York
21. Crandall MG, Lions PL (1985) Hamilton–Jacobi equations in infinite dimensions, Part I. J Funct Anal 62:379–396
22. Crandall MG, Lions PL (1986) Hamilton–Jacobi equations in infinite dimensions, Part II. J Funct Anal 65:368–405
23. Crandall MG, Lions PL (1986) Hamilton–Jacobi equations in infinite dimensions, Part III. J Funct Anal 68:214–247
24. Crandall MG, Lions PL (1990) Hamilton–Jacobi equations in infinite dimensions, Part IV. J Funct Anal 90:273–283
25. Datko R (1976) Control problems with quadratic cost. J Differ Equ 21:231–262
26. Fattorini HO (1968) Boundary control systems. SIAM J Control 3:349–384
27. Fleming WH, Rishel RW (1975) Deterministic and stochastic optimal control. Springer, Berlin
28. Ghoussoub N (2008) Selfdual partial differential systems and their variational principles. Springer, Berlin
29. Ioffe AD (1976) An existence theorem for a general Bolza problem. SIAM J Control Optim 14:458–466
30. Ioffe AD (1977) On lower semicontinuity of integral functionals. SIAM J Control 15:521–538; 458–466
31. Lasiecka I, Triggiani R (1983) Dirichlet boundary control problems for parabolic equations with quadratic cost. Analyticity and Riccati's feedback synthesis. SIAM J Control Optim 21:41–67
32. Lasiecka I, Triggiani R (2000) Control theory of partial differential equations. Cambridge University Press, Cambridge
33. Lions JL (1968) Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles. Dunod, Gauthier–Villars, Paris
34. Lions JL, Magenes E (1970) Problèmes aux limites non homogènes et applications. Dunod, Gauthier–Villars, Paris
35. Mackenroth U (1982) Convex parabolic boundary control problems with pointwise state constraints. J Math Anal Appl 87:256–277
36. Olech C (1969) Existence theorems for optimal problems with vector–valued cost function. Trans Am Math Soc 136:159–180
37. Olech C (1969) Existence theorems for optimal control problems involving multiple integrals. J Differ Equ 6:512–526
38. Popescu V (1979) Existence for an abstract control problem in Banach spaces. Numer Funct Anal Optim 1:475–479
39. Rockafellar RT (1971) Existence and duality theorems for convex problem of Bolza. Trans Am Math Soc 159:1–40
40. Rockafellar RT (1972) Dual problems of optimal control. In: Balakrishnan AV (ed) Techniques of optimization. Academic Press, San Diego, pp 423–432
41. Rockafellar RT (1972) State constraints in convex control of Bolza. SIAM J Control 10:691–716
42. Rockafellar RT (1974) Conjugate duality and optimization. CBMS lecture notes series, vol 162. SIAM, Philadelphia
43. Rockafellar RT (1975) Existence theorems for general problems of Bolza and Lagrange. Adv Math 15:312–333

44. Rockafellar RT (1976) Dual problems of Lagrange for arcs of bounded variation. In: Russel DL (ed) Optimal control and the calculus of variations. Academic Press, San Diego
45. Tröltzsch F (1981) A generalized bang-bang principle for a heating problem with constraints on the control and the thermal stress. J Integral Equ 3:345–355
46. Visintin (2008) Extension of the Brezis–Ekeland–Nayroles principle to monotone operators. Adv Math Sci Appl 18:633–650
47. Washburn D (1979) A bound on the boundary input map for parabolic equations with applications to time optimal control problems. SIAM J Control Optim 17:652–691
48. Yosida K (1980) Functional analysis. Springer, Berlin

Index

V. Barbu, T. Precupanu, *Convexity and Optimization in Banach Spaces*,
Springer Monographs in Mathematics,
DOI 10.1007/978-94-007-2247-7, © Springer Science+Business Media B.V. 2012