

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 $f : X \rightarrow \overline{\mathbb{R}}$ is called *convex* if the inequality

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \quad (2.1)$$

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 $f(y) < \infty$.

The function $g : X \rightarrow \overline{\mathbb{R}}$ 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

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i), \quad \lambda_i \geq 0, \quad \sum_{i=1}^n \lambda_i = 1$$

holds for all x_1, \dots, x_n in X , for which the right-hand side makes sense.

Another consequence of convexity of $f : X \rightarrow \overline{\mathbb{R}}$ is the convexity of the *level sets*,

$$\{x \in X; f(x) \leq \lambda\},$$

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 $f : X \rightarrow \overline{\mathbb{R}}$, we denote by $\text{Dom}(f)$ (sometimes $\text{dom } f$) the convex set

$$\text{Dom}(f) = \{x \in X; f(x) < +\infty\}. \quad (2.2)$$

Such a set $\text{Dom}(f)$ is called the *effective domain* of f . If f is proper, then $\text{Dom}(f)$ 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

$$I_A(x) = \begin{cases} 0, & \text{if } x \in A, \\ +\infty, & \text{if } x \notin A, \end{cases} \quad (2.3)$$

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 $f : X \rightarrow \overline{\mathbb{R}}$ be any extended real-valued function on X . The set

$$\text{epi } f = \{(x, \alpha); x \in X, \alpha \in \mathbb{R}, f(x) \leq \alpha\} \quad (2.4)$$

is called the *epigraph* of f . The set

$$\text{hypo } f = \{(x, \alpha); x \in X, \alpha \in \mathbb{R}, f(x) \geq \alpha\} \quad (2.5)$$

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 $f : X \rightarrow \overline{\mathbb{R}}$ 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 $(x, \alpha), (y, \beta) \in \text{epi } f$ 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 $(w, t) \in \text{epi } f$. This proves that $\text{epi } f$ is a convex set of $X \times \mathbb{R}$.

Necessity. Suppose that $\text{epi } f$ is convex, but for some $x, y \in X$ and some $\lambda \in [0, 1]$ the inequality

$$f(w) = f((1 - \lambda)x + \lambda y) > (1 - \lambda)f(x) + \lambda f(y)$$

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 $\text{epi } f$. Thus, for each x, y and λ , one has

$$\inf\{(1 - \lambda)\alpha + \lambda\beta; (x, \alpha), (y, \beta) \in \text{epi } f\} = (1 - \lambda)f(x) + \lambda f(y).$$

Since the epigraph of f is convex, we have

$$f(w) = \inf\{t; (w, t) \in \text{epi } f\} \leq (1 - \lambda)f(x) + \lambda f(y) < f(w).$$

The contradiction we arrived at concludes the proof. \square

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 $f : X \rightarrow \overline{\mathbb{R}}$ is called *lower-semicontinuous* (*upper-semicontinuous*) at x_0 if

$$f(x_0) = \liminf_{x \rightarrow x_0} f(x) \quad \left(f(x_0) = \limsup_{x \rightarrow x_0} f(x) \right). \quad (2.6)$$

We recall that

$$\liminf_{x \rightarrow x_0} f(x) = \sup_{V \in \mathcal{V}(x_0)} \inf_{s \in V} f(s) \quad (2.7)$$

and

$$\limsup_{x \rightarrow x_0} f(x) = \inf_{V \in \mathcal{V}(x_0)} \sup_{s \in V} f(s), \quad (2.8)$$

where $\mathcal{V}(x_0)$ 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:

$$\tau_\ell = \{]a, +\infty[; a \in [-\infty, +\infty[\} \cup \{\emptyset, \overline{\mathbb{R}}\}.$$

It is readily seen that the function $f : X \rightarrow \overline{\mathbb{R}}$ is lower-semicontinuous (l.s.c.) at x_0 if and only if $f : X \rightarrow (\overline{\mathbb{R}}, \tau_\ell)$ 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 properties of lower-semicontinuous functions as required for the purpose of the next section.

Proposition 2.5 *Let X be a topological space and let $f : X \rightarrow \overline{\mathbb{R}}$ 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 $\{x \in X; f(x) \leq \lambda\}$, $\lambda \in \mathbb{R}$, 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 $\{x \in X; f(x) \leq \lambda\} = f^{-1}([-\infty, \lambda])$ and (i) is equivalent to the continuity of $f : X \rightarrow (\overline{\mathbb{R}}, \tau_\ell)$, 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 $\varphi : X \times \mathbb{R} \rightarrow \overline{\mathbb{R}}$ 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

$$\text{epi } f - (0, \lambda) = \{(x, t) \in X \times \mathbb{R}; \varphi(x, t) \leq \lambda\},$$

that is, the level sets of the function φ are translates of $\text{epi } f$. 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

$$\left\{x \in X; \sup_{i \in I} f_i(x) \leq \lambda\right\} = \bigcap_{i \in I} \{x \in X; f_i(x) \leq \lambda\}.$$

□

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, $f(x_0) \leq f(x)$ for all x in X , thereby proving Theorem 2.8. \square

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 $f : X \rightarrow \overline{\mathbb{R}}$ 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 $f : X \rightarrow \overline{\mathbb{R}}$ be any convex function on X . The *closure* of the function f , denoted by $\text{cl } f$, is by definition the lower-semicontinuous hull of f , that is, $\text{cl } f = \liminf_{y \rightarrow x} f(y)$ for all $x \in X$ if $\liminf_{y \rightarrow x'} f(y) > -\infty$ for every $x' \in X$ or $\text{cl } f() = -\infty$ for all $x \in X$ if $\liminf_{y \rightarrow x'} f(y) = -\infty$ for some $x' \in X$. The convex function f is said to be *closed* if $\text{cl } f = f$. Particularly, a *proper convex function* is closed if and only if it is lower-semicontinuous.

For every proper closed convex function one has

$$(\text{cl } f)(x) = \liminf_{y \rightarrow x} f(y), \quad \forall x \in X. \quad (2.9)$$

As a consequence of equality (2.9), one obtains

$$\text{epi}(\text{cl } f) = \overline{\text{epi } f}, \quad (2.10)$$

or, more specifically,

$$\{x \in X; (\text{cl } f)(x) \leq \alpha\} = \bigcap_{\lambda > \alpha} \{\overline{x \in X; f(x) \leq \lambda}\}$$

for every $\alpha \in \mathbb{R}$. In particular, it follows from (2.7) that

$$\inf\{f(x); x \in X\} = \inf\{(\text{cl } f)(x); x \in X\}. \quad (2.11)$$

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 $g(x) \leq f(x)$, for every $x \in X$). \square

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 $f : X \rightarrow]-\infty, +\infty]$ 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 $\text{epi } f$ is (strongly) closed if it is weakly closed. This establishes Proposition 2.10. \square

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.) \square

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*

$$\lim_{\substack{\|x\| \rightarrow +\infty \\ x \in M}} f(x) = +\infty. \quad (2.12)$$

In fact, let $x_1 \in \text{Dom}(f)$ and $k > 0$ be such that

$$f(x) > f(x_1) \quad \text{for } \|x\| > k, \quad x \in M.$$

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 $f : X \rightarrow]-\infty, +\infty]$ be a proper convex function on X . Then, the function f is continuous on $\text{int Dom}(f)$ if and only if f is bounded from above on a neighborhood of an interior point of $\text{Dom}(f)$.*

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 $\text{Dom}(f)$. Let $V \in \mathcal{V}(x_0)$ 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 $x \rightarrow f(x + x_0) - f(x_0)$ 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,

$$0 = f(0) \leq \frac{1}{2}(f(x) + f(-x))$$

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 $\text{int Dom}(f)$. Let z be any point in $\text{int Dom}(f)$ and let $\rho > 1$ be such that $z_0 = \rho z \in \text{Dom}(f)$. 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

$$\begin{aligned} f(u) &= f\left(\frac{1}{\rho}z_0 + \left(1 - \frac{1}{\rho}\right)x\right) \leq \frac{1}{\rho}f(z_0) + \left(1 - \frac{1}{\rho}\right)f(x) \\ &\leq \frac{1}{\rho}f(z_0) + \left(1 - \frac{1}{\rho}\right)k \quad \text{for all } u \in V(z). \end{aligned}$$

Hence, f is bounded above on $V(z)$, as claimed. This completes the proof. \square

As a consequence, we obtain the next corollary.

Corollary 2.15 *If a proper convex function $f : X \rightarrow]-\infty, +\infty]$ is upper-semicontinuous at a point which is interior to its effective domain $\text{Dom}(f)$, then f is continuous on $\text{int Dom}(f)$.*

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 $f : X \rightarrow]-\infty, +\infty]$ be a lower-semicontinuous proper convex function. Then f is continuous at every algebraic interior point of its effective domain $\text{Dom}(f)$.*

Proof Without any loss of generality, we may restrict ourselves again to the case in which the origin is an algebraic interior to the effective domain $\text{Dom}(f)$. We choose any real number α such that $\alpha > f(0)$ and set $A = \{x \in X; f(x) \leq \alpha\}$. 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 $x_0 = \rho x \in \text{Dom}(f)$. Here, we have used the fact that the origin is an algebraic interior point of $\text{Dom}(f)$. Since f is convex, we have

$$f(\lambda \rho x) = f(\lambda x_0 + (1 - \lambda)0) \leq \lambda(f(x_0) - f(0)) + f(0),$$

for every $\lambda \in [0, 1]$. Therefore, there exists $\delta > 0$ such that $f(\lambda \rho x) \leq \alpha$ 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. \square

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 linear 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 $\text{Dom}(f)$ of the function f . Let $\{e_i; i = 1, 2, \dots, n\}$ be a basis of the n -dimensional space X , and let a be a sufficiently small positive number such that

$$U = \left\{ x \in X; x = \sum_{i=1}^n x_i e_i, 0 < x_i < \frac{a}{n}, i = 1, 2, \dots, n \right\} \subset \text{Dom}(f).$$

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

$$f(x) \leq \sum_{i=1}^n \frac{x_i}{a} f(a e_i) + \left(1 - \sum_{i=1}^n \frac{x_i}{a}\right) f(0) \leq \frac{1}{n} \sum_{i=1}^n |f(a e_i)| + |f(0)|$$

for every $x \in U$.

Thus, the function f is bounded from above on $U \subset \text{Dom}(f)$. But it is obvious that U is open. This implies, according to Theorem 2.14, that f is continuous on $\text{int Dom}(f)$, which completes the proof. \square

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 $\text{Dom}(f)$ which do not belong to $\text{int Dom}(f)$. In addition to the continuity of f on X , we introduce the concept of continuity on $\text{Dom}(f)$. These two concepts are clearly equivalent on $\text{int Dom}(f)$, but not necessarily on $\text{Dom}(f)$. Also, we notice for later use that

$$\text{int}(\text{epi } f) = \{(x, \alpha) \in X \times \mathbb{R}; x \in \text{int Dom}(f), f(x) < \alpha\}. \quad (2.13)$$

2.1.4 Conjugate Functions

Let X be a real linear locally convex space and let X^* be its conjugate space. Consider any function $f : X \rightarrow \overline{\mathbb{R}}$. The function $f^* : X^* \rightarrow \overline{\mathbb{R}}$ defined by

$$f^*(x^*) = \sup\{(x, x^*) - f(x); x \in X\}, \quad x^* \in X^* \quad (2.14)$$

is called the *conjugate function* of f . The conjugate of f^* , that is, the function f^{**} on X defined by

$$f^{**}(x) = \sup\{(x, x^*) - f^*(x^*); x^* \in X^*\}, \quad x \in X, \quad (2.15)$$

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

$$f(x) + f^*(x^*) \geq (x, x^*) \quad (2.16)$$

and

$$f^*(x^*) + f^{**}(x) \geq (x, x^*), \quad (2.17)$$

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 $\text{Dom}(f)$.

Example 2.18 The conjugate of the indicator function I_A of a subset A of X is given by

$$I_A^*(x^*) = \sup\{(x, x^*); x \in A\}. \quad (2.18)$$

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 $x \rightarrow (x, x^*)$ 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,

$$A^\circ = \{x^* \in X^*; (x, x^*) \leq 1, \forall x \in A\}. \quad (2.19)$$

In terms of I_A^* defined above, the polar of A may be expressed as

$$A^\circ = \{x^* \in X^*; I_A^*(x^*) \leq 1\}. \quad (2.20)$$

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

$$C^\circ = \{x^* \in X^*; (x, x^*) \leq 0, \forall x \in C\} \quad (2.21)$$

and is called the *dual cone* of C .

If $A = Y$ is a linear subspace of X , then

$$Y^\circ = \{x^* \in X^*; (x, x^*) = 0, \forall x \in Y\} \quad (2.22)$$

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

$$p_{A^\circ}(x^*) = \max\{I_A^*(x^*), 0\}, \quad \forall x^* \in X^*. \quad (2.23)$$

Therefore, if $0 \in A$, then $I_A^*(x^*) \geq 0$ and

$$p_{A^\circ} = I_A^*. \quad (2.24)$$

Furthermore,

$$p_A^* = I_{A^\circ} \quad \text{for every } A \subset X \text{ with } 0 \in A. \quad (2.25)$$

Indeed, if $x^* \notin A^\circ$, then there exists $\bar{x} \in A$ such that $(x^*, \bar{x}) > 1$. This implies that

$$\begin{aligned} p_A^*(x^*) &= \sup_{x \in X} \{(x, x^*) - p_A(x)\} \geq \lambda(\bar{x}, x^*) - p_A(\lambda\bar{x}) \\ &= \lambda[(\bar{x}, x^*) - p_A(\bar{x})] \geq \lambda[(\bar{x}, x^*) - 1], \quad \forall \lambda > 0. \end{aligned}$$

Hence, $p_A^*(x^*) = +\infty$ for every $x^* \notin A^\circ$. Now, if $x^* \in A^\circ$, since for every $x \in \text{Dom}(p_A)$, $x \in (p_A(x) + \varepsilon)A$, for all $\varepsilon > 0$, we have

$$\begin{aligned} p_A^*(x^*) &= \sup\{(x, x^*) - p_A(x); x \in \text{Dom}(p_A)\} \\ &\leq \sup_{a \in A} \sup_{x \in \text{Dom}(p_A)} \{(p_A(x) + \varepsilon)(a, x^*) - p_A(x)\} \leq \varepsilon, \quad \forall \varepsilon > 0. \end{aligned}$$

Hence, $p_A^*(x^*) \leq 0$. Because $0 \in \text{Dom}(p_A)$, 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 $f : X \rightarrow \overline{\mathbb{R}}$ 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 $f^{***} \leq f^*$, while part (iv) implies that $f^* \leq f^{***}$, 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 $f : X \rightarrow]-\infty, +\infty]$ be any convex and lower-semicontinuous function on X , $f \not\equiv +\infty$. As already seen, the epigraph $\text{epi } f$ of f is a proper convex and closed subset of product space $X \times \mathbb{R}$. If $x_0 \in \text{Dom}(f)$, then $(x_0, f(x_0) - \varepsilon) \notin \text{epi } f$ for every $\varepsilon > 0$. Thus, using the Hahn–Banach theorem (see Corollary 1.45), there exists $u \in (X \times \mathbb{R})^*$ such that

$$\sup_{(x,t) \in \text{epi } f} u(x, t) < u(x_0, f(x_0) - \varepsilon).$$

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

$$\sup_{(x,t) \in \text{epi } f} \{x_0^*(x) + t\alpha\} < x_0^*(x_0) + \alpha(f(x_0) - \varepsilon).$$

We observe that $\alpha \neq 0$ and must be negative, since $(x_0, f(x_0) + n) \in \text{epi } f$ for every $n \in \mathbb{N}$. On the other hand, $(x, f(x)) \in \text{epi } f$ for every $x \in \text{Dom}(f)$. Thus,

$$x_0^*(x) + \alpha f(x) \leq x_0^*(x_0) + \alpha f(x_0), \quad \forall x \in \text{Dom}(f),$$

or

$$f(x) \geq -\frac{1}{\alpha} x_0^*(x) + \frac{1}{\alpha} x_0^*(x_0) + f(x_0), \quad \forall x \in \text{Dom}(f),$$

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 $f : X \rightarrow]-\infty, +\infty]$ 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 $f : X \rightarrow]-\infty, +\infty]$ 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 $f^{**}(x) \leq f(x)$, 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, $(x_0, f^{**}(x_0)) \notin \text{epi } f$, 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

$$x_0^*(x_0) + \alpha f^{**}(x_0) > \sup\{x_0^*(x) + \alpha t; (x, t) \in \text{epi } f\}. \quad (2.26)$$

Since $(x, t + n) \in \text{epi } f$ for every $n \in \mathbb{N}$ and $(x, t) \in \text{epi } f$, relation (2.26) implies that $\alpha \leq 0$. Furthermore, α must be negative. Indeed, otherwise (that is, $\alpha = 0$), inequality (2.26) implies that

$$x_0^*(x_0) > \sup\{x_0^*(x); x \in \text{Dom}(f)\}. \quad (2.27)$$

Let $h > 0$ and $y_0^* \in \text{Dom}(f^*)$ be arbitrarily chosen. (We recall that $\text{Dom}(f^*) \neq \emptyset$ because f^* is proper.) One obtains

$$\begin{aligned} f^*(y_0^* + hx_0^*) &= \sup\{(x, y_0^*) + h(x, x_0^*) - f(x); x \in \text{Dom}(f)\} \\ &\leq \sup\{(x, y_0^*) - f(x); x \in \text{Dom}(f)\} \\ &\quad + h \sup\{(x_0^*, x); x \in \text{Dom}(f)\} \\ &= f^*(y_0^*) + h \sup\{(x_0^*, x); x \in \text{Dom}(f)\}. \end{aligned}$$

On the other hand, a simple calculation involving the latter expression and inequality (2.17) yields

$$\begin{aligned} f^{**}(x_0) &\geq (y_0^* + hx_0^*, x_0) - f^*(y_0^* + hx_0^*) \\ &\geq (y_0^*, x_0) - f^*(y_0^*) + h[(x_0^*, x_0) - \sup\{(x_0^*, x); x \in \text{Dom}(f)\}]. \end{aligned}$$

Comparing this inequality with (2.27) and letting $h \rightarrow +\infty$, 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

$$\begin{aligned} x_0^*\left(-\frac{x_0}{\alpha}\right) - f^{**}(x_0) &> \sup\left\{x_0^*\left(-\frac{x}{\alpha}\right) - t; (x, t) \in \text{epi } f\right\} \\ &= \sup\left\{\left(-\frac{1}{\alpha}x_0^*, x\right) - f(x); x \in \text{Dom}(f)\right\} = f^*\left(-\frac{1}{\alpha}x_0^*\right). \end{aligned}$$

But this inequality obviously contradicts inequality (2.17). Hence, $f^{**}(x_0) = f(x_0)$ for every $x_0 \in \text{Dom}(f^{**})$. Since $f^{**}(x) = f(x)$, for all $x \in \text{Dom}(f^{**})$, it results that $f^{**}(x) = f(x)$ for all $x \in X$. Thus, the proof is complete. \square

More generally, if f is not lower-semicontinuous, then $f^{**} = \text{cl } f$. Thus, we obtain the following corollary.

Corollary 2.23 *The biconjugate of a convex function f coincides with its closure, that is, $f^{**} = \text{cl } f$.*

Proof It is clear that $\text{cl } f$ is lower-semicontinuous if it is proper and, therefore, $(\text{cl } f)^{**} = \text{cl } f$ as a consequence of Theorem 2.22. But as has already been mentioned, $f^* = (\text{cl } f)^*$, which shows that $f^{**} = \text{cl } f$, as claimed. If $\text{cl } f$ is not proper, the result is immediately clear, since $f^{**} = (\text{cl } f)^{**} = \text{cl } f \equiv -\infty$. \square

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 $x \rightarrow (x, x^*) - f^*(x^*)$ is an affine continuous function for each $x^* \in \text{Dom}(f^*)$, as claimed. The converse is obvious (see Corollary 2.6). \square

There is a close connection between the effective domain $\text{Dom}(f)$ of a lower-semicontinuous convex function $f : X \rightarrow \overline{\mathbb{R}}^*$ and the growth properties of its conjugate $f^* : X^* \rightarrow \overline{\mathbb{R}}^*$.

Proposition 2.25 *Assume that X is a reflexive Banach space. Then the following two conditions are equivalent:*

- (i) $\text{int Dom}(f) \neq \emptyset$.
- (ii) *There are $\rho > 0$ and $C > 0$ such that*

$$f^*(p) \geq \rho \|p\|_{X^*} - C, \quad \forall p \in X. \quad (2.28)$$

Moreover, $\text{Dom}(f) = X$ if and only if

$$\lim_{\|p\| \rightarrow \infty} \frac{f^*(p)}{\|p\|} = +\infty. \quad (2.29)$$

Proof If $\text{int Dom}(f) \neq \emptyset$, then there is a ball $B(x_0, \rho) \subset \text{int Dom}(f)$ 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

$$f(x) = f^{**}(x) \leq \sup_x \{ (x, x^*) - \rho \|x^*\|_{X^*} - C \} \leq \infty \quad \text{for } \|x\|_X \leq \rho$$

and therefore $B(0, \rho) \subset \text{Dom}(f)$, as claimed.

Now, if $\text{Dom}(f) = X$, then by the above argument it follows that (2.28) holds for all $\rho > 0$, that is, for all $\rho > 0$,

$$f^*(p) \geq \rho \|p\|_{X^*} - C_\rho, \quad \forall p \in X^*,$$

which implies that (2.29) holds. Conversely, if (2.29) holds, then, by (2.15), we see that $\text{Dom}(f) = X$, as claimed. \square

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,*

$$A^{\circ\circ} = \overline{\text{conv}(A \cup \{0\})}. \quad (2.30)$$

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, $A = A^{\circ\circ}$, as claimed. \square

Remark 2.27 We notice that the conjugate correspondence $f \rightarrow f^*$ 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 $g : x \rightarrow \overline{\mathbb{R}}$, the function $g^* : X^* \rightarrow \overline{\mathbb{R}}$ defined by

$$g^*(x) = \inf \{ (x, x^*) - g(x); x \in X \}, \quad (2.31)$$

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 $f^* : X^* \rightarrow \overline{\mathbb{R}}$ be the conjugate function of f . Let $(f^*)^* : X^{**} \rightarrow \overline{\mathbb{R}}$ 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 $f : X \rightarrow \overline{\mathbb{R}}$ be any extended real-valued function on X . Then the function f^* on Y defined by

$$f^*(y) = \sup\{(x, y) - f(x); x \in X\}, \quad y \in Y, \quad (2.32)$$

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 $f : X \rightarrow]-\infty, +\infty]$, the *subdifferential* of such a function is the (generally multivalued) mapping $\partial f : X \rightarrow X^*$ defined by

$$\partial f(x) = \{x^* \in X^*; f(x) - f(u) \leq (x - u, x^*), \forall u \in X\}. \quad (2.33)$$

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 $\text{Dom}(f)$. 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,

$$N_K(x) = \{x^* \in X^*; (x^*, x - u) \geq 0 \text{ for all } u \in K\}.$$

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 $x \in \text{int } K$. 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 $F : X \rightarrow X^*$ (see Definition 1.99). Indeed, if $x^* \in F(x)$, then, by the definition of F , one has

$$\begin{aligned} (x - u, x^*) &= \|x\|^2 - (u, x^*) \geq \|x\|^2 - \|u\| \|x\| \\ &\geq \frac{1}{2} (\|x\|^2 - \|u\|^2), \quad \text{for every } u \in X. \end{aligned}$$

In other words, $x^* \in \partial f(x)$. Conversely, suppose that $x^* \in \partial f(x)$. Hence,

$$(x - u, x^*) \geq \frac{1}{2} (\|x\|^2 - \|u\|^2), \quad \forall u \in X.$$

Taking in the latter inequality $u = x + \lambda v$, where $\lambda \in \mathbb{R}^+$ and $v \in X$, we see that

$$-\lambda(v, x^*) \geq -\frac{1}{2} (2\lambda \|x\| \|v\| + \lambda^2 \|v\|^2).$$

Therefore

$$|(v, x^*)| \leq \|v\| \|x\|, \quad \forall v \in X.$$

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 $f : X \rightarrow]-\infty, +\infty]$ be a proper convex function. Then, the following three properties are equivalent:*

- (i) $x^* \in \partial f(x)$.
- (ii) $f(x) + f^*(x^*) \leq (x, x^*)$.
- (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

$$f(u) - f(x) \geq (u - x, x^*), \quad \forall u \in X,$$

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

$$f^*(x^*) + (f^*)^*(x) = (x, x^*), \quad (2.34)$$

where $(f^*)^* : X^{**} \rightarrow]-\infty, +\infty]$ 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, $\partial f^*(0) \cap X^* \neq \emptyset$.

Remark 2.35 If the space X is reflexive, then it follows from Proposition 2.33 that $\partial f^* : X^* \rightarrow X^{**} = X$ is just the inverse of ∂f , in other words,

$$x \in \partial f^*(x^*) \iff x^* \in \partial f(x). \quad (2.35)$$

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 $f : X \rightarrow]-\infty, +\infty]$ 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,

$$H = \{(x, \lambda) \in X \times \mathbb{R}; f(x) \leq \lambda\}.$$

H is a convex subset of $X \times \mathbb{R}$ and $(x_0, f(x_0) + \varepsilon) \in \text{int } H$ 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

$$\alpha_0(f(x_0) - f(x)) \leq (x_0 - x, x_0^*) \quad \text{for every } x \in \text{Dom}(f). \quad (2.36)$$

It should be observed that $\alpha_0 \neq 0$ (that is, the hyperplane is not vertical) because, otherwise, $(x_0 - x, x_0^*) = 0$ for all x in $\text{Dom}(f)$, 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. \square

Remark 2.37 From the above proof, it follows that a proper convex function f is subdifferentiable in an element $x_0 \in \text{Dom}(f)$ 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*

$$\text{int Dom}(f) \subset D(\partial f). \quad (2.37)$$

Proof We have seen in Sect. 2.1.3 (Proposition 2.16) that f is continuous at every interior point of its effective domain $\text{Dom}(f)$. 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 $\lambda \rightarrow \lambda^{-1}(f(x + \lambda h) - f(x))$ is monotonically increasing on $]0, \infty[$. Thus, the *directional derivative at x in the direction h*

$$f'(x, h) = \lim_{\lambda \downarrow 0} \lambda^{-1}(f(x + \lambda h) - f(x)) = \inf_{\lambda > 0} \lambda^{-1}(f(x + \lambda h) - f(x)) \quad (2.38)$$

exists for every $h \in X$. The function $h \rightarrow f'(x, h)$ is called the *directional differential of f at x* . It is immediate from the definition that for fixed $x \in \text{Dom}(f)$, $f'(x, h)$ is a positively homogeneous subadditive function on X . The function f is said to be *weakly* or *Gâteaux differentiable* at x if $h \rightarrow f'(x, h)$ is a linear continuous function on X . In particular, this implies that

$$-f'(x, -h) = f'(x, h) = \lim_{\lambda \rightarrow 0} \lambda^{-1}(f(x + \lambda h) - f(x))$$

for every $h \in X$. If f is weakly differentiable at x , then we denote by $\nabla f(x)$ or $\text{grad } f(x)$ (the *gradient of f at x*) the element of X^* defined by

$$f'(x, h) = (h, \text{grad } f(x)) \quad \text{for every } h \in X.$$

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. \square

Proposition 2.39 *Let $f : X \rightarrow]-\infty, +\infty]$ be a proper convex function. If f is finite and continuous at x_0 , then*

$$f'(x_0, h) = \sup\{(h, x^*); x^* \in \partial f(x_0)\} \quad (2.39)$$

and, in general, one has

$$\partial f(x_0) = \{x^* \in X; (h, x^*) \leq f'(x_0, h), \forall h \in X\}. \quad (2.40)$$

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

$$(h, w) \leq f_0(h) \leq f(x_0 + h) - f(x_0), \quad \forall w \in \partial f(x_0)$$

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 $x \rightarrow \lambda^{-1}(f(x_0 + \lambda x) - f(x_0))$ is just the function $x^* \rightarrow \lambda^{-1}(f^*(x^*) + f(x_0) - (x_0, x^*))$. Therefore,

$$f_0^*(x^*) = \sup_{\lambda > 0} \lambda^{-1}(f(x_0) + f^*(x^*) - (x_0, x^*)),$$

because

$$f_0(h) = \inf_{\lambda > 0} \lambda^{-1}(f(x_0 + \lambda h) - f(x_0)).$$

According to Proposition 2.33, one has

$$\partial f(x_0) = \{x^* \in X^*; f(x_0) + f^*(x^*) - (x_0, x^*) = 0\}$$

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

$$f'(x_0, h) = (h, \partial f(x_0)) \quad \text{for every } h \in X.$$

In particular, this implies that f is Gâteaux differentiable at x_0 and $\text{grad } f(x_0) = \partial f(x_0)$. It follows that the converse result is also true. \square

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 $x_0^* = \text{grad } f(x_0)$. 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 $\text{grad } f(x_0) = \partial f(x_0)$.*

Proof Suppose that f is Gâteaux differentiable at x_0 , that is,

$$(h, \text{grad } f(x_0)) = \lim_{\lambda \rightarrow 0} \lambda^{-1} (f(x_0 + \lambda h) - f(x_0)), \quad \forall h \in X.$$

However,

$$\lambda^{-1} (f(x_0 + \lambda h) - f(x_0)) \leq f(x_0 + h) - f(x_0) \quad \text{for } \lambda \in]0, 1[$$

because f is convex. This implies that

$$f(x_0) - f(x_0 + h) \leq -(h, \text{grad } f(x_0)) \quad \text{for all } h \in X,$$

that is, $\text{grad } f(x_0) \in \partial f(x_0)$. Now, let x_0^* be any element of $\partial f(x_0)$. We have

$$f(x_0) - f(u) \leq (x_0 - u, x_0^*), \quad \forall u \in X,$$

and, therefore,

$$\lambda^{-1} (f(x_0 + \lambda h) - f(x_0)) \geq (h, x_0^*) \quad \text{for every } \lambda > 0.$$

This shows that $(\text{grad } f(x_0) - x_0^*, h) \geq 0$ for all $h \in X$, that is, $x_0^* = \text{grad } f(x_0)$. We conclude the proof by noting that the second part of Proposition 2.40 has already been proven by the above remarks. \square

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 $\partial f : X \rightarrow X^*$ is monotone in $X \times X^*$. In other words,

$$(x_1 - x_2, x_1^* - x_2^*) \geq 0 \quad \text{for } x_i^* \in \partial f(x_i), \quad i = 1, 2. \quad (2.41)$$

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, $F : X \rightarrow X^*$ 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, $f_1 : X \rightarrow]-\infty, +\infty]$ is convex and lower-semicontinuous on X . Moreover, since f is bounded from below by an affine function, we may infer that

$$\lim_{\|x\| \rightarrow +\infty} f_1(x) = +\infty.$$

Thus, using Theorem 2.11 (see Remark 2.13), the infimum of f_1 on X is attained. In other words, there is $x_0 \in \text{Dom}(f)$ such that

$$f_1(x_0) \leq f_1(x) \quad \text{for every } x \in X.$$

We write this inequality in the form

$$f(x_0) - f(x) \leq (x_0 - x, x_0^*) + (x - x_0, F(x)) \quad \text{for every } x \in X$$

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

$$f(x_0) - f(u) \leq (x_0 - u, x_0^*) + (u - x_0, F(tx_0 + (1 - t)u)).$$

Passing to limit $t \rightarrow 1$, we obtain

$$f(x_0) - f(u) \leq (x_0 - u, x_0^*) + (u - x_0, F(x_0))$$

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. \square

Corollary 2.44 *Let $f : X \rightarrow]-\infty, +\infty]$ be a lower-semicontinuous proper and convex function on X . Then $D(\partial f)$ is a dense subset of $\text{Dom}(f)$.*

Proof For simplicity, we assume that X is reflexive. Let x be any element of $\text{Dom}(f)$. 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

$$\|x_\lambda - x\|^2 + \lambda f(x_\lambda) \leq \lambda f(x)$$

and therefore

$$\lim_{\lambda \rightarrow 0} \|x_\lambda - x\| = 0,$$

because f is bounded from below by an affine function. Therefore, $x \in \overline{D(\partial f)}$ and the corollary has been proved. \square

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) $A : X \rightarrow X^*$ is said to be *cyclically monotone* if

$$(x_0 - x_1, x_0^*) + \cdots + (x_{n-1} - x_n, x_{n-1}^*) + (x_n - x_0, x_n^*) \geq 0, \quad (2.43)$$

for every finite set of points in the graph of A , that is, $x_i^* \in Ax_i$ for $i = 0, 1, \dots, n$. 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

$$f(x) = \sup \{ (x - x_n, x_n^*) + \cdots + (x_1 - x_0, x_0^*) \},$$

where $x_i^* \in Ax_i$ for $i = 1, \dots, n$ 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

$$f(x) \geq \lambda + (x - \tilde{x}, \tilde{x}^*) \quad \text{for all } x \in X. \quad (2.44)$$

Let $\lambda < f(\tilde{x})$. Then, by the definition of f there exist the pairs $[x_i, x_i^*] \in A$, $i = 1, \dots, m$, 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). \square

As mentioned earlier (see Theorem 1.143 and Corollary 1.140), if a maximal monotone operator $A : X \rightarrow X^*$ 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 $f : X \rightarrow]-\infty, +\infty]$ is a lower-semicontinuous proper convex function. Then, the following conditions are equivalent.*

$$\lim_{\|x\| \rightarrow +\infty} \frac{f(x)}{\|x\|} = +\infty, \quad (2.45)$$

$$R(A) = X^* \quad \text{and} \quad A^{-1} \text{ is bounded on bounded subsets.} \quad (2.46)$$

Proof 1°. (2.45) \Rightarrow (2.46). Let x_0 be arbitrary, but fixed in $D(A)$. By the definition of ∂f , one has

$$(\partial f(x), x - x_0) \geq f(x) - f(x_0) \quad \text{for any } x \in D(A)$$

and therefore

$$\lim_{\substack{\|x\| \rightarrow \infty \\ [x, y] \in A}} \frac{(x - x_0, y)}{\|x\|} = +\infty.$$

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) \Rightarrow (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 $u \in \text{Dom}(f)$ 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. \square

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 $\text{Dom}(f^*) = \mathbb{R}$ 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 $j : \mathbb{R} \rightarrow]-\infty, +\infty]$ such that $\partial j = \beta$.* Indeed, there exist $-\infty \leq a \leq b \leq +\infty$ such that $]a, b[\subset \text{Dom}(\beta) \subset [a, b]$. Let $\beta^\circ : \text{Dom}(\beta) \rightarrow \mathbb{R}$ 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 $a \in \text{Dom}(\beta)$ and $\beta(b) = [\beta^\circ(b - 0), +\infty[$ if $b \in \text{Dom}(\beta)$ (this is an immediate consequence of the maximality).

Now, let r_0 be fixed in $\text{Dom}(\beta)$ and define the function $j : \mathbb{R} \rightarrow]-\infty, +\infty]$

$$j(r) = \begin{cases} \int_{r_0}^r \beta^\circ(s) ds, & \text{if } r \in [a, b], \\ +\infty, & \text{if } r \notin [a, b]. \end{cases}$$

Then, we have

$$j(r) - j(t) \leq \int_t^r \beta^\circ(s) ds \leq \xi(r - t),$$

for all $r \in \text{Dom}(\beta)$, $t \in \mathbb{R}$ and $\xi \in \beta(r)$. Hence, $\beta(r) \in \partial j(r)$ for all $r \in \text{Dom}(\beta)$. 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*

$$f(x) = \begin{cases} \frac{1}{2}|A^{\frac{1}{2}}x|^2, & \text{if } x \in D(A^{\frac{1}{2}}), \\ +\infty, & \text{otherwise.} \end{cases} \quad (2.47)$$

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

$$\frac{1}{2}|A^{\frac{1}{2}}x|^2 - \frac{1}{2}|A^{\frac{1}{2}}u|^2 \leq (Ax, x - u), \quad \text{for all } u \in D(A^{\frac{1}{2}}),$$

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 $f : H \rightarrow]-\infty, +\infty]$ 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

$$g(t) - g(s) \leq (t - s)t(Au, u) \quad \text{for } t, s \in [0, 1].$$

The last inequality shows that g is absolutely continuous on $[0, 1]$ and $\frac{d}{dt}g(t) = t(Au, u)$ 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. \square

Example 2.52 (Convex integrands and integral functionals) Let Ω be a Lebesgue measurable subset of \mathbb{R}^n and let $L_m^p(\Omega)$, $1 \leq p < \infty$, be the usual Banach space of p -summable functions $y : \Omega \rightarrow \mathbb{R}^m$.

A function $g : \Omega \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}^* =]-\infty, +\infty]$ is said to be a *normal convex integrand* on $\Omega \times \mathbb{R}^m$ if the following conditions are satisfied:

- (i) $g(x, \cdot) : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}^*$ 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 $y : \Omega \rightarrow \mathbb{R}^m$ the function $x \rightarrow g(x, y(x))$ 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 $\text{int } D(g(x, \cdot)) \neq \emptyset$ 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

$$h(x, y) = (y, \alpha(x)) + \beta(x),$$

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

$$I_g(y) = \int_{\Omega} g(x, y(x)) \, dx. \quad (2.48)$$

More precisely, the functional I_g is defined on $L_m^p(\Omega)$ by

$$I_g(y) = \begin{cases} \int_{\Omega} g(x, y(x)) \, dx, & \text{if } g(x, y) \in L_m^1(\Omega), \\ +\infty, & \text{otherwise.} \end{cases}$$

Proposition 2.53 *Let conditions (i), (ii), (iii) and (iv) be satisfied. Then, the function $I_g : L_m^p(\Omega) \rightarrow \mathbb{R}^*$, $1 \leq p < +\infty$, is convex, lower-semicontinuous and $\neq +\infty$. Moreover, for every $y \in L_m^p(\Omega)$, the subdifferential $\partial I_g(y)$ is given by*

$$\partial I_g(y) = \{w \in L_m^{p'}(\Omega); w(x) \in \partial g(x, y(x)) \text{ a.e. } x \in \Omega\}. \quad (2.49)$$

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

$$y_n(x) \rightarrow y(x) \quad \text{a.e. } x \in \Omega$$

and, therefore,

$$g(x, y_n(x)) - (y_n(x), \alpha(x)) - \beta(x) \rightarrow g(x, y(x)) - (y(x), \alpha(x)) - \beta(x) \\ \text{a.e. } x \in \Omega.$$

Then, by the Fatou Lemma

$$\liminf_{n \rightarrow \infty} I_g(y_n) \geq I_g(y)$$

because $\liminf_{n \rightarrow \infty} g(x, y_n(x)) \geq g(x, y(x))(g(x, \cdot))$ is lower-semicontinuous.

Now, let $w \in \partial I_g(y)$. By the definition of $\partial I_g(y)$, we have

$$\int_{\Omega} (g(x, y(x)) - g(x, u(x))) \, dx \leq \int_{\Omega} (w(x), y(x) - u(x)) \, dx$$

for all $u \in L_m^p(\Omega)$. 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

$$\int_E (g(x, y(x)) - g(x, u) - (w(x), y(x) - u)) \, dx \leq 0.$$

Since E is arbitrary, we may conclude that

$$g(x, y(x)) \leq g(x, u) + (w(x), y(x) - u) \quad \text{a.e. } x \in \Omega,$$

and therefore

$$w(x) \in \partial g(x, y(x)) \quad \text{a.e. } x \in \Omega,$$

as claimed. Conversely, it is easy to see that every $w \in L_m^{p'}(\Omega)$ satisfying the latter belongs to $\partial I_g(y)$. \square

Remark 2.54 Under the assumptions of Proposition 2.53, the function I_g is weakly lower-semicontinuous on $L_m^p(\Omega)$ (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 $\varphi : [0, T] \rightarrow \mathbb{R}$ be such that, for every $t \in [0, T]$, the function $x \rightarrow \varphi(t, x)$ is convex, lower-semicontinuous and $\neq +\infty$. 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

$$\varphi(t, x) \geq (\alpha_0(t), x) + \beta(t), \quad (2.50)$$

for all $x \in H$ and $t \in [0, T]$.

Define the function $I_\varphi : L^p(0, T; H) \rightarrow \overline{\mathbb{R}}^*$, $1 \leq p < \infty$,

$$I_\varphi(x) = \begin{cases} \int_0^T \varphi(t, x) dt, & \text{if } \varphi(t, x) \in L^1(0, T), \\ +\infty, & \text{otherwise.} \end{cases} \quad (2.51)$$

Proposition 2.55 *The function I_φ is convex, lower-semicontinuous and $\neq +\infty$ on $L^p(0, T; H)$. The subdifferential ∂I_φ is given by*

$$\partial I_\varphi(x) = \{w \in L^{p'}(0, T; H); w(t) \in \partial \varphi(t, x(t)) \text{ a.e. } t \in]0, T[\}, \quad (2.52)$$

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 $g : \mathbb{R} \rightarrow \overline{\mathbb{R}}^*$ be a lower-semicontinuous convex function and let $\beta = \partial g$ be its subdifferential. Define the function $\varphi : L^2(\Omega) \rightarrow \overline{\mathbb{R}}^* =]-\infty, +\infty]$

$$\varphi(y) = \begin{cases} \frac{1}{2} \int_\Omega |\text{grad } y|^2 dx + \int_\Omega g(y) dx, & \text{if } y \in H_0^1(\Omega) \text{ and } g(y) \in L^1(\Omega), \\ +\infty, & \text{otherwise.} \end{cases}$$

Proposition 2.57 *The function φ is convex, lower-semicontinuous and*

$$\begin{aligned} \partial \varphi(y) &= \{w \in L^2(\Omega); w(x) \in -\Delta y(x) + \partial g(y(x)) \text{ a.e. } x \in \Omega\}, \\ D(\partial \varphi) &= \{y \in H_0^1(\Omega) \cap H^2(\Omega); \exists \tilde{w} \in L^2(\Omega), \tilde{w}(x) \in \partial g(y(x)) \\ &\quad \text{a.e. } x \in \Omega\}. \end{aligned} \quad (2.53)$$

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 $I_\Delta : L^2(\Omega) \rightarrow \overline{\mathbb{R}}^*$,

$$I_\Delta(y) = -\frac{1}{2} \int_\Omega y \Delta y d\xi = \frac{1}{2} \int_\Omega |\nabla y|^2 d\xi, \quad \forall y \in H_0^1(\Omega).$$

This implies that φ is convex and lower-semicontinuous. If we denote by $F : L^2(\Omega) \rightarrow L^2(\Omega)$ the map defined by the right-hand side of (2.53), we see that $Fy \in \partial\varphi(y)$, $\forall y \in D(F) = \{y \in H_0^1(\Omega) \cap H^2(\Omega); \exists \tilde{w} \in L^2(\Omega), \tilde{w}(x) \in \partial g(y(x)) \text{ a.e. } x \in \Omega\}$.

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

$$y - \Delta y + \partial g(y) \ni f \quad \text{in } \Omega; \quad y \in H_0^1(\Omega) \cap H^2(\Omega)$$

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,

$$\int_{\Omega} g((1 + \varepsilon A)^{-1}y) \, dx \leq \int_{\Omega} g(y) \, dx, \quad \forall y \in L^2(\Omega),$$

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 $\tilde{\varphi} : L^2(\Omega) \rightarrow \overline{\mathbb{R}}$, defined by

$$\tilde{\varphi}(y) = \begin{cases} \frac{1}{2} \int_{\Omega} |\text{grad } y|^2 \, dx + \int_{\Gamma} g(y) \, dx, & \text{if } y \in H^1(\Omega), \, g(y) \in L^1(\Gamma), \\ +\infty, & \text{otherwise.} \end{cases} \quad \square$$

Arguing as in the preceding example, we see that φ is convex and lower-semicontinuous. As regards its subdifferential $\partial\varphi : L^2(\Omega) \rightarrow L^2(\Omega)$, it is given by (see Brezis [11, 12])

$$\partial\varphi(y) = -\Delta y, \quad \forall y \in D(\partial g), \quad (2.54)$$

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 $f : X \rightarrow \overline{\mathbb{R}}^*$ be a lower-semicontinuous convex function and let $A = \partial f$. Since $A : X \rightarrow X^*$ is maximal monotone, for every $\lambda > 0$ the equation

$$F(x_{\lambda} - x) + \lambda Ax_{\lambda} \ni 0, \quad (2.55)$$

where $F : X \rightarrow X^*$ is the duality mapping of X , has at least one solution $x_\lambda \in D(A)$ (see Theorem 1.141). The inequality

$$(F(x) - F(y), x - y) \geq (\|x\| - \|y\|)^2 \quad \text{for all } x, y \text{ in } X$$

and the strict convexity of X and X^* then imply that the solution x_λ of (2.55) is unique. We set

$$x_\lambda = J_\lambda x, \quad (2.56)$$

$$A_\lambda x = -\lambda^{-1} F(x_\lambda - x). \quad (2.57)$$

(See Sect. 1.4.1.)

For every $\lambda > 0$, we define

$$f_\lambda(x) = \inf \left\{ \frac{\|x - y\|^2}{2\lambda} + f(y); y \in X \right\}, \quad x \in X. \quad (2.58)$$

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 $\lambda \rightarrow 0$. Theorem 2.58 given below says that this is indeed the case.

Theorem 2.58 *Let $f : X \rightarrow]-\infty, +\infty]$ 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,*

$$f_\lambda(x) = \left(\frac{\lambda}{2} \right) \|A_\lambda x\|^2 + f(J_\lambda x) \quad \text{for every } x \in X, \quad (2.59)$$

$$\lim_{\lambda \rightarrow 0} f_\lambda(x) = f(x) \quad \text{for every } x \in X, \quad (2.60)$$

$$f(J_\lambda x) \leq f_\lambda(x) \leq f(x) \quad \text{for every } x \in X \text{ and } \lambda > 0. \quad (2.61)$$

Proof It is readily seen that the subdifferential of the function $y \rightarrow \frac{\|x-y\|^2}{2\lambda} + f(y)$ is just the operator $y \rightarrow \lambda^{-1} F(y - x) + \partial f(y)$. 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 $x \in \text{Dom}(f)$, then $\lim_{\lambda \rightarrow \infty} J_\lambda x = x$, by using Corollary 1.70 and Proposition 1.146. This fact, combined with the lower-semicontinuity of f and inequality (2.61), shows that $\lim_{\lambda \rightarrow 0} f_\lambda(x) = f(x)$. Now, assume that $f(x) = +\infty$. We must show that $f_\lambda(x) \rightarrow +\infty$ for $\lambda \rightarrow 0$. Suppose that this is not the case, and that, for example,

$$f_{\lambda_n}(x) \leq C \quad \text{where } \lambda_n \rightarrow 0.$$

If equality (2.59) is used again, it would follow that, under the present circumstances, $J_{\lambda_n}x \rightarrow x$ and $f(J_{\lambda_n}x) \leq C$. Then the lower-semicontinuity of f would imply that $f(x) \leq C$, 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

$$f_\lambda(y) - f_\lambda(x) \leq \frac{\lambda}{2} (\|A_\lambda y\|^2 - \|A_\lambda x\|^2) + (A_\lambda y, J_\lambda y - J_\lambda x),$$

that is,

$$\begin{aligned} f_\lambda(y) - f_\lambda(x) &\leq (A_\lambda y, y - x) + (A_\lambda y, J_\lambda y - y) + (A_\lambda y, x - J_\lambda x) \\ &\quad + \frac{\lambda}{2} (\|A_\lambda y\|^2 + \|A_\lambda x\|^2). \end{aligned}$$

Finally,

$$0 \leq f_\lambda(y) - f_\lambda(x) - (A_\lambda x, y - x) \leq (A_\lambda y - A_\lambda x, y - x), \quad (2.62)$$

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

$$\lim_{t \rightarrow 0} \frac{f_\lambda(x + tu) - f_\lambda(x)}{t} = (A_\lambda x, u) \quad \text{for every } x \in X,$$

because A_λ is demicontinuous by Proposition 1.146. Therefore, f_λ is Gâteaux differentiable at any $x \in X$ and $\partial f_\lambda(x) = A_\lambda x$. \square

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

$$|f_\lambda(y) - f_\lambda(x) - (A_\lambda x, y - x)| \leq \frac{\|y - x\|^2}{\lambda} \quad \text{for all } \lambda > 0,$$

which, obviously, implies that f is Fréchet differentiable on H . \square

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 $x \rightarrow \varphi(x) - f(x)$ 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

$$\partial f(x) + \partial\varphi(x) \subset \partial(f + \varphi)(x) \quad \text{for every } x \in D(\partial f) \cap D(\partial\varphi). \quad (2.63)$$

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 $(\text{int } D(A)) \cap D(B) \neq \emptyset$ (or $(D(A) \cap \text{int } D(B)) \neq \emptyset$), 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 $f : X \rightarrow]-\infty, +\infty]$ be a lower-semicontinuous proper and convex function on X . Assume that at least one of the following conditions is satisfied.*

$$D(A) \cap \text{int Dom}(f) \neq \emptyset, \quad (2.64)$$

$$\text{Dom}(f) \cap \text{int } D(A) \neq \emptyset. \quad (2.65)$$

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

$$0 \in D(A) \cap \text{int Dom}(f), \quad (2.66)$$

or

$$0 \in \text{Dom}(f) \cap \text{int } D(A). \quad (2.67)$$

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

$$F(x) + Ax + \partial f(x) \ni y^*. \quad (2.68)$$

To show that equation (2.68) has at least one solution, consider the approximate equation

$$Fx_\lambda + Ax_\lambda + \partial f_\lambda(x) \ni y^*, \quad \lambda > 0, \quad (2.69)$$

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

$$\|x_\lambda\| \leq \|y^*\| \quad \text{for every } \lambda > 0, \quad (2.70)$$

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 $\text{Dom}(f)$, there is $\rho > 0$ such that

$$f_\lambda(\rho w) \leq f(\rho w) \leq C \quad \text{for every } w \in X, \|w\| = 1,$$

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

$$(Fx_\lambda, x_\lambda - \rho w) + (Ax_\lambda, x_\lambda - \rho w) + f_\lambda(x_\lambda) \leq (y^*, x_\lambda - \rho w) + C. \quad (2.71)$$

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

$$\rho\|y_\lambda^*\| \leq C \quad \text{for all } \lambda > 0. \quad (2.72)$$

(We shall denote by C several positive constants independent of λ .) Thus, with the aid of equations (2.69) and (2.70), this yields

$$\|\partial f_\lambda(x_\lambda)\| \leq C \quad \text{for all } \lambda > 0. \quad (2.73)$$

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

$$\sup\{\|z^*\|; z^* \in Ax; \|x\| \leq \rho\} \leq C. \quad (2.74)$$

Let w be any element in X such that $\|w\| = 1$.

Again, multiplying equation (2.69) by $x_\lambda - \rho w$, we obtain

$$(Fx_\lambda, x_\lambda - \rho w) + (\partial f_\lambda(x_\lambda), x_\lambda - \rho w) + (Ax_\lambda, x_\lambda - \rho w) = (y_\lambda^*, x_\lambda - \rho w).$$

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

$$\|\partial f_\lambda(x_\lambda)\| \leq C \quad \text{for every } \lambda > 0.$$

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

$$\begin{aligned} x_\lambda &\rightarrow x \quad \text{weakly in } X, \\ Fx_\lambda + y_\lambda^* &\rightarrow z^* \quad \text{weakly in } X^*. \end{aligned} \quad (2.75)$$

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

$$(Fx_\lambda + Fx_\mu, x_\lambda - x_\mu) + (y_\lambda^* - y_\mu^*, x_\lambda - x_\mu) + (\partial f_\lambda(x_\lambda) - \partial f_\mu(x_\mu), x_\lambda - x_\mu) = 0$$

and therefore

$$\lim_{\lambda, \mu \rightarrow 0} (Fx_\lambda + y_\lambda^* - Fx_\mu - y_\mu^*, x_\lambda - x_\mu) = 0 \quad (2.76)$$

because

$$\begin{aligned}
 & (\partial f_\lambda(x_\lambda) - \partial f_\mu(x_\mu), x_\lambda - x_\mu) \\
 & \geq (\partial f_\lambda(x_\lambda) - \partial f_\mu(x_\mu), x_\lambda - J_\lambda x_\lambda - x_\mu + J_\mu x_\mu u) \\
 & \geq -(\|\partial f_\lambda(x_\lambda)\| + \|\partial f_\mu(x_\mu)\|)(\lambda \|\partial f_\lambda(x_\lambda)\| + \mu \|\partial f_\mu(x_\mu)\|).
 \end{aligned}$$

Here, we have used relations (2.56), (2.57) and the monotonicity of ∂f . Extracting further subsequences, if necessary, we may assume that

$$\lim_{\lambda \rightarrow 0} (F(x_\lambda) + y_\lambda^*, x_\lambda) = \ell.$$

Then, relation (2.75) shows that $(z^*, x) = \ell$. Now, let $[u, v]$ be any element in the graph of $A + F$. We have

$$(Fx_\lambda + y_\lambda^* - v, x_\lambda - u) \geq 0, \quad \forall \lambda > 0.$$

Hence,

$$(z^* - v, x - u) \geq 0, \tag{2.77}$$

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

$$f_\lambda(x_\lambda) \leq f_\lambda(u) + (y^*, x_\lambda - u) - (x_\lambda + y_\lambda^*, x_\lambda - u) \tag{2.78}$$

and therefore

$$\limsup_{\lambda \rightarrow 0} f_\lambda(x_\lambda) \leq f(u) + (y^*, x - u) - (z^*, x - u). \tag{2.79}$$

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

$$\lim_{\lambda \rightarrow 0} (x_\lambda - J_\lambda(x_\lambda)) = 0 \quad \text{strongly in } X.$$

Hence,

$$J_\lambda(x_\lambda) \rightarrow x \quad \text{weakly in } X \text{ as } \lambda \rightarrow 0.$$

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

$$f(x) \leq f(u) + (y^*, x - u) - (z^*, x - u)$$

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. \square

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.*

$$\text{Dom}(f) \cap \text{int Dom}(\varphi) \neq \emptyset. \quad (2.80)$$

Then

$$\partial(f + \varphi) = \partial f + \partial \varphi. \quad (2.81)$$

Proof Since $D(\partial \varphi)$ is a dense subset of $\text{Dom}(\varphi)$ (see Corollary 2.44), condition (2.80) implies that $\text{Dom}(f) \cap \text{int } D(\partial \varphi) \neq \emptyset$. 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. \square

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 $\text{Dom}(f)$ has a nonvalid interior.

Theorem 2.65 *Let $f : H \rightarrow]-\infty, +\infty]$ 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 $C \in \mathbb{R}$,*

$$f((I + \lambda A)^{-1}(x + \lambda h)) \leq f(x) + C\lambda \quad \text{for all } x \in H \text{ and } \lambda > 0. \quad (2.82)$$

Then the operator $A + \partial f$ is maximal monotone and

$$\overline{D(A + \partial f)} = \overline{D(A) \cap D(\partial f)} = \overline{D(A)} \cap \overline{\text{Dom}(f)}. \quad (2.83)$$

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

$$x + Ax + \partial f(x) \ni y. \quad (2.84)$$

To show that this is indeed the case, consider the equation

$$x_\lambda + A_\lambda x_\lambda + \partial f(x_\lambda) \ni y, \quad (2.85)$$

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 $\|A_\lambda x_0\| \leq \|A^0 x_0\|$ 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

$$\begin{aligned} (\partial f(x), A_\lambda(x + \lambda h)) &= \lambda^{-1}(\partial f(x), x + \lambda h - (I + \lambda A)^{-1}(x + \lambda h)) \\ &\geq (\partial f(x), h) + (f(x) - f(I + \lambda A)^{-1}(x + \lambda h))\lambda^{-1} \\ &\geq -C - \|h\| \|\partial f(x)\|. \end{aligned} \quad (2.82')$$

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 $\lambda \rightarrow 0$. We subtract the defining equations for x_λ and x_μ and then multiply by $x_\lambda - x_\mu$; we obtain

$$\|x_\lambda - x_\mu\|^2 + (A_\lambda x_\lambda - A_\mu x_\mu, x_\lambda - x_\mu) \leq 0.$$

Since $A_\lambda x_\lambda \in AJ_\lambda x_\lambda$ and A is monotone, we see that

$$\|x_\lambda - x_\mu\|^2 \rightarrow 0 \quad \text{as } \lambda, \mu \rightarrow 0.$$

Hence, $\lim_{\lambda \rightarrow 0} x_\lambda = 0$ 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

$$A_{\lambda_n} x_{\lambda_n} \rightarrow y_0 \quad \text{in the weak topology of } H.$$

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 $D(A) \cap \text{Dom}(f)$. Then, there exist $x_\varepsilon \in \text{Dom}(f)$ such that $x_\varepsilon \rightarrow x$ strongly in H as $\varepsilon \rightarrow 0$. We set $u_\varepsilon = (I + \varepsilon A)^{-1}(x_\varepsilon + \varepsilon h)$ and observe that

$$\begin{aligned} \|u_\varepsilon - x\| &\leq \|u_\varepsilon - (I + \varepsilon A)^{-1}x\| + \|(I + \varepsilon A)^{-1}x - x\| \\ &\leq \|x_\varepsilon - x\| + \|(I + \varepsilon A)^{-1}x - x\| + \varepsilon \|h\|. \end{aligned}$$

Hence, $u_\varepsilon \rightarrow x$ as $\varepsilon \rightarrow 0$. Moreover, by condition (2.82), $u_\varepsilon \in D(A) \cap \text{Dom}(f)$. Briefly, we have shown that $\overline{D(A) \cap \text{Dom}(f)} \subset \overline{D(A) \cap \text{Dom}(f)}$. Now, we prove that $D(A) \cap \text{Dom}(f) \subset \overline{D(A) \cap D(\partial f)}$. Let u be any element in $D(A) \cap \text{Dom}(f)$ and let $u_\varepsilon \in D(A) \cap D(\partial f)$ be the unique solution to the equation

$$u_\varepsilon + \varepsilon Au_\varepsilon + \varepsilon \partial f(u_\varepsilon) \ni u.$$

We have

$$f(u_\varepsilon) - f(u) \leq \left(\frac{u - u_\varepsilon}{\varepsilon} - Au_\varepsilon, u_\varepsilon - u \right) \leq -\frac{1}{\varepsilon} \|u_\varepsilon - u\|^2 - (Au, u_\varepsilon - u),$$

which implies that $\lim_{\varepsilon \rightarrow 0} u_\varepsilon = 0$. Since u is arbitrary in $D(A) \cap \text{Dom}(f)$, we may infer that $D(A) \cap \text{Dom}(f) \subset \overline{D(A) \cap D(\partial f)}$, as claimed. Since $\overline{D(A) \cap \text{Dom}(f)} \subset \overline{D(A) \cap D(\partial f)}$, Relation (2.83) follows, and this completes the proof. \square

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

$$Ax_\lambda + \partial f_\lambda(x_\lambda) \ni 0.$$

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 $x \rightarrow \frac{1}{2\lambda} \|x - P_K x\|^2$, $x \in H$, 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 $f : X \rightarrow \mathbb{R}$ 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*

$$f(x) - f(y) = (w, x - y). \quad (2.86)$$

Proof Without loss of generality, we may assume that $y = 0$. Define the function $\varphi : \mathbb{R} \rightarrow \mathbb{R}$

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

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 $\varphi_\lambda(1) - \varphi_\lambda(0) = \partial\varphi_\lambda(\theta_\lambda)$. On a sequence $\lambda_n \rightarrow 0$ we have $\theta_{\lambda_n} \rightarrow \theta$ and $\partial\varphi_{\lambda_n}(\theta_{\lambda_n}) \rightarrow \eta \in \partial\varphi(\theta)$. Since $\varphi_\lambda \rightarrow \varphi$ for $\lambda \rightarrow 0$, we infer that $\varphi(1) - \varphi(0) = \eta \in \partial\varphi(\theta)$, as claimed (obviously, $\theta \in]0, 1[$). \square

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 from X to X^* and let K be a closed convex set of X . We say that x satisfies a *variational inequality* if

$$x \in K, \quad (Ax - f, u - x) \geq 0 \quad \text{for all } u \in K, \quad (2.87)$$

where f is given in X^* . In terms of subdifferentials, inequality (2.87) can be written as

$$Ax + \partial I_K(x) \ni f, \quad (2.88)$$

where $I_K : X \rightarrow [0, +\infty]$ 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 $A : X \rightarrow X^*$ 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 $x_0 \in K$,*

$$\lim_{\{\|x\| \rightarrow +\infty, x \in K\}} (Ax, x - x_0) \|x\|^{-1} = +\infty. \quad (2.89)$$

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. \square

We pause, briefly, to point out an important generalization of Theorem 2.67 (see Brezis [10]).

The operator $A : K \rightarrow X^*$ 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 $\limsup_{n \rightarrow \infty} (Au_n, u_n - u) \leq 0$, then $\liminf_{n \rightarrow \infty} (Au_n, u_n - v) \geq (Au, u - v)$ for all $v \in K$.
- (ii) For every $v \in K$, the mapping $u \rightarrow (Au, u - v)$ 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*

$$(I + \lambda A)^{-1}(x + \lambda h) \in K \quad \text{for all } x \in K \text{ and } \lambda > 0, \quad (2.90)$$

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 $A : V \rightarrow V'$ 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

$$a(u, u) \geq \omega \|u\|^2 \quad \text{for all } u \text{ in } V, \quad (2.91)$$

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

$$a(u, v - u) \geq (f, v - u) \quad \text{for all } v \in K. \quad (2.92)$$

In particular, if the bilinear form a is symmetric, problem (2.92) can be equivalently expressed as

$$\min \left\{ \frac{1}{2} a(v, v) - (f, v); v \in K \right\}. \quad (2.93)$$

□

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 $f \rightarrow u$ 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

$$Av = -(a_{ij}(x)v_{x_i})_{x_j}, \quad (2.94)$$

where the coefficients a_{ij} are in $L^\infty(\Omega)$ and satisfy the condition ($\omega > 0$)

$$a_{ij}(x)\xi_i\xi_j \geq \omega|\xi|^2, \quad \forall \xi \in \mathbb{R}^n, \quad \xi = (\xi_1, \dots, \xi_n).$$

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

$$(Au, v) = a(u, v) = \int_{\Omega} a_{ij}(x) u_{x_i} v_{x_j} dx \quad \text{for all } u, v \in H^1(\Omega). \quad (2.94')$$

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

$$-(a_{ij} u_{x_i})_{x_j} = f \quad \text{in } \Omega, \quad u = 0 \quad \text{in } \Gamma$$

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

$$\int_{\Omega} a_{ij}(x) u_{x_i} (v - u)_{x_j} dx \geq \int_{\Omega} f(v - u) dx \quad \text{for all } v \in K. \quad (2.95)$$

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*):

$$-(a_{ij}(x) u_{x_i})_{x_j} = f \quad \text{in } E = \{x \in \Omega; u(x) > \psi(x)\}, \quad (2.96)$$

$$-(a_{ij}(x) u_{x_i})_{x_j} \geq f \quad \text{in } \Omega, \quad (2.97)$$

$$u \geq \psi \quad \text{on } \Omega, \quad u = \psi \quad \text{in } \Omega \setminus E, \quad u = 0 \quad \text{in } \Gamma. \quad (2.98)$$

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

$$\int_{\Omega} a_{ij} u_{x_i} \alpha_{x_j} dx = \int_E f \alpha dx \quad \text{for all } \alpha \in C_0^\infty(E).$$

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

$$\min \left\{ \int_{\Omega} a_{ij}(x) v_{x_i} v_{x_j} dx - \int_{\Omega} f dx; v \in K \right\}.$$

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

$$\frac{1}{2} \int_{\Omega} |\text{grad } v|^2 dx - \int_{\Omega} f v dx$$

has to be minimized on $K = \{v \in H_0^1(\Omega); |\text{grad } v| \leq 1, \text{ a.e. on } \Omega\}$. As seen earlier, this problem can be equivalently expressed as

$$\int_{\Omega} \text{grad } u \text{ grad}(u - v) dx \leq \int_{\Omega} f(u - v) dx \quad \text{for all } v \in K.$$

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

$$\begin{aligned} -\Delta u &= f & \text{on } \Omega_1, & & u &= 0 & \text{on } \Gamma, \\ |\text{grad } u| &= 1 & \text{on } \Omega_2, & \end{aligned}$$

where $\Omega_1 \cap \Omega_2 = \emptyset$ and $\Omega_1 \cup \Omega_2 = \Omega$.

Example 2.72 Let $a : H^1(\Omega) \times H^1(\Omega) \rightarrow \mathbb{R}$ be the bilinear form

$$a(u, v) = \int_{\Omega} \text{grad } u \text{ grad } v dx + \int_{\Omega} uv dx$$

and

$$K = \{u \in H^1(\Omega); u \geq 0 \text{ a.e. on } \Gamma\}.$$

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

$$a(u, v - u) \geq \int_{\Omega} f(v - u) dx, \quad \text{for all } v \in K, \quad (2.99)$$

has a unique solution $u \in K$. Let $v = u \pm \varphi$, where $\varphi \in C_0^\infty(\Omega)$. Then, inequality (2.99) yields

$$a(u, \varphi) - \int_{\Omega} f \varphi \, dx = 0, \quad \text{for all } \varphi \in C_0^\infty(\Omega).$$

Hence,

$$-\Delta u + u = f \quad \text{on } \Omega \quad (2.100)$$

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

$$\int_{\Omega} (\Delta u - u)v \, dx = \int_{\Gamma} v \frac{\partial u}{\partial \nu} \, d\sigma - a(u, v) \quad \text{or all } v \in H^1(\Omega). \quad (2.101)$$

In formula (2.101), we have denoted by $\int_{\Gamma} v \frac{\partial u}{\partial \nu} \, d\sigma$ 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

$$\int_{\Gamma} (v - u) \frac{\partial u}{\partial \nu} \, d\sigma \geq 0 \quad \text{for all } v \in K.$$

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:

$$\begin{aligned} -\Delta u + u &= f \quad \text{on } \Omega, \\ u &\geq 0, \quad \frac{\partial u}{\partial \nu} \geq 0, \quad u \frac{\partial u}{\partial \nu} = 0 \quad \text{on } \Gamma. \end{aligned} \quad (2.102)$$

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 $\varphi : L^2(\Omega) \rightarrow]-\infty, +\infty]$ is given by (see Example 2.56)

$$\varphi(y) = \frac{1}{2} \int_{\Omega} |\text{grad } y|^2 \, dx + \int_{\Gamma} g(y) \, d\sigma$$

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

$$\begin{aligned} -(a_{ij}(x)u_{x_i})_{x_j} - f(x)(u(x) - \psi(x)) &= 0 \quad \text{a.e. } x \in \Omega, \\ u(x) &\geq \psi(x); \quad -(a_{ij}(x)u_{x_i}(x))_{x_j} \geq f(x) \quad \text{a.e. } x \in \Omega. \end{aligned} \quad (2.96')$$

Indeed, by Corollary 2.68, the equation

$$A_H u + \partial I_K(u) \ni f, \quad (2.103)$$

where

$$\begin{aligned} A_H u &= Au \cap H \quad \text{for } u \in D(A_H) = H_0^1(\Omega) \cap H^2(\Omega) \quad \text{and} \\ K &= \{u \in L^2(\Omega); u(x) \geq \psi(x) \text{ a.e. } x \in \Omega\} \end{aligned} \quad (2.104)$$

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,

$$\partial I_K(u) = \{w \in L^2(\Omega); w(x)(u(x) - \psi(x)) = 0, w(x) \geq 0 \text{ a.e. } x \in \Omega\}, \quad (2.105)$$

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*

$$x_0 \in C, \quad -Ax_0 \in C^\circ, \quad (x_0, Ax_0) = 0, \quad (2.106)$$

where C is a given closed, convex cone with the vertex at 0 in X and C° is its polar, that is, $C^\circ = \{x^ \in X^*; (x, x^*) \leq 0 \text{ for all } x \in C\}$.*

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}_+^n the set of nonnegative n -vectors), the above problem takes the familiar form

$$x_0 \geq 0, \quad Ax_0 \geq 0, \quad (x_0, Ax_0) = 0. \quad (2.107)$$

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*

$$(Ax_0, x - x_0) \geq 0 \quad \text{for all } x \in C. \quad (2.108)$$

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 $(Ax_0, y) \geq 0$. Hence, $-Ax_0 \in C^\circ$. Also, taking $x = 2x_0$, we see that $(x_0, Ax_0) \geq 0$, while, for

$x = 0$, (2.108) implies that $(x_0, Ax_0) \leq 0$. Therefore $(x_0, Ax_0) = 0$. This completes the proof. \square

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. \square

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*

$$\lim_{\|x\| \rightarrow +\infty} \frac{(Ax, x - x_0)}{\|x\|} = +\infty, \quad (2.109)$$

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 $x \in \text{Dom}(f)$ 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 $x \in \text{Dom}(f)$, that is, it is possible that $x \notin D(\partial f)$.

But for any $x \in \text{Dom}(f)$ there exists at least one closed hyperplane passing through $(x, f(x) - \varepsilon)$, $\varepsilon > 0$, such that $\text{epi } f$ 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 $\partial_\varepsilon f : X \rightarrow X^*$ defined by

$$\partial_\varepsilon f(x) = \{x^* \in X^*; f(x) - f(u) \leq (x - u, x^*) + \varepsilon, \forall u \in X^*\}, \quad (2.110)$$

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 $D(\partial_\varepsilon f) \subset \text{Dom}(f)$. For $\varepsilon = 0$ we obtain the subdifferential defined by Definition 2.30. Also, we have

$$\partial f(x) = \bigcap_{\varepsilon > 0} \partial_\varepsilon f(x), \quad x \in \text{Dom}(f). \quad (2.111)$$

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 $x \in \text{Dom}(f)$.*

Proof We have $(x, f(x) - \varepsilon) \in \text{epi } f$ for any fixed $\varepsilon > 0$, $x \in \text{Dom}(f)$. By hypothesis, $\text{epi } f$ 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 $\text{epi } f$. This hyperplane is necessarily nonvertical, that is, it can be considered of the form $(x^*, 1)$. Thus, we obtain $x^* \in \partial_\varepsilon f(x)$. \square

Corollary 2.81 *For any proper convex lower-semicontinuous function f we have $D(\partial_\varepsilon f) = \text{Dom}(f)$, 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 $x \in \text{Dom}(f)$.*

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 $\bar{u} \in \text{Dom}(f)$, while $\partial_\varepsilon f(\bar{u}) = \emptyset$. Hence, f must be a proper function. Now, if $x \in \text{Dom}(f)$ and $(x, \alpha) \in \text{epi } f$, then there exists $\varepsilon > 0$ such that $(x, f(x) - \varepsilon) \in \text{epi } f$. But since $\partial_\varepsilon f(x) \neq \emptyset$, we have a closed nonvertical hyperplane passing through $(x, f(x) - \varepsilon)$ such that $\text{epi } f$ is contained in one of the two closed half-spaces determined by that hyperplane. Consequently, $\text{epi } f$ is an intersection of closed half-spaces. Hence, $\text{epi } f$ 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 $f : X \rightarrow]-\infty, +\infty]$ 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 $\partial_\varepsilon f^* : X \rightarrow X$ 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 $x \in \text{Dom}(f)$, then $f(u) \geq f(x) - \varepsilon$ for all $u \in \text{Dom}(f)$ 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 $A : X \rightarrow X^*$ is called ε -monotone if

$$(x - y, x^* - y^*) \geq -2\varepsilon, \quad \text{for all } x^* \in Ax, y^* \in Ay. \quad (2.112)$$

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, $-2\varepsilon \notin \partial_\varepsilon f(1)$, 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 $f : X \rightarrow \mathbb{R}$ 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

$$\partial_\varepsilon f(x) = Ax + b + \{y \in A; \langle A^{-1}y, y \rangle \leq 2\varepsilon\}, \quad \varepsilon \geq 0, x \in X. \quad (2.113)$$

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

$$\frac{1}{2}\langle Ax, x \rangle + \langle v, x \rangle - \frac{1}{2}\langle z - b, A^{-1}(z - b) \rangle + \langle z - b, A^{-1}(z - b) \rangle \leq \langle x, z \rangle + \varepsilon,$$

from which we obtain

$$\langle Ax, x \rangle + 2\langle x, b - z \rangle + \langle A^{-1}(z - b), z - b \rangle \leq 2\varepsilon,$$

and so,

$$\langle x - A^{-1}(z - b), b - z \rangle + \langle x, b - z + Ax \rangle \leq 2\varepsilon.$$

Therefore, if we denote $y = z - Ax - b$, then

$$\langle A^{-1}y, y \rangle \leq 2\varepsilon,$$

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

$$\langle x - u, Ax + b + y - v \rangle \geq -2\varepsilon, \quad \text{for all } x \in X, \quad (2.114)$$

and every $y \in X$ fulfilling the inequality $\langle A^{-1}y, y \rangle \leq 2\varepsilon$. 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

$$\frac{1}{4} \langle A^{-1}(v - y - v) - u, Au + y + b - v \rangle \geq -2\varepsilon,$$

whenever $\langle A^{-1}y, y \rangle \leq 2\varepsilon$.

Taking $z = v - Au - b$, we get

$$\langle A^{-1}(y - z), y - z \rangle \leq 8\varepsilon, \quad \text{if } \langle A^{-1}y, y \rangle \leq 2\varepsilon. \quad (2.115)$$

Therefore it is necessary that $\langle A^{-1}z, z \rangle \leq 2\varepsilon$. 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 $\langle A^{-1}z, z \rangle \leq 2\varepsilon$, 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

$$\langle A^{-1}y, y \rangle = \langle A^{-\frac{1}{2}}y, A^{-\frac{1}{2}}y \rangle = \|A^{-\frac{1}{2}}y\|^2,$$

and so, $\langle A^{-1}y, y \rangle \leq 2\varepsilon$ if and only if $y = \sqrt{2\varepsilon}A^{\frac{1}{2}}u$, where $\|u\| \leq 1$. Consequently, (2.113) can be rewritten in the form

$$\partial_\varepsilon f(x) = Ax + b + \sqrt{2\varepsilon}A^{\frac{1}{2}}(\bar{S}(0; 1)), \quad \varepsilon \geq 0, x \in X.$$

If A is the identity operator, we obtain

$$\partial_\varepsilon \left(\frac{1}{2} \|\cdot\|^2 \right)(x) = x + \sqrt{2\varepsilon} \bar{S}(0; 1), \quad \varepsilon \geq 0, x \in X. \quad (2.116)$$

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 $A : X \rightarrow X^*$ and $\varepsilon \geq 0$, the ε -enlargement of A , denoted by A^ε , is defined by

$$A^\varepsilon x = \{x^* \in X^*; (x - y, x^* - y^*) \geq -2\varepsilon, \text{ for all } y^* \in Ay\}, \quad x \in X. \quad (2.117)$$

Proposition 2.91 Let $A : X \rightarrow X^*$ 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} , $\text{conv } \overline{A}$, $\overline{\text{conv}} A$ and A^{-1} are ε -monotone.
- (iv) $A^{\varepsilon_1} \subset A^{\varepsilon_2}$ if $0 \leq \varepsilon_1 \leq \varepsilon_2$.
- (v) If A is ε -monotone and locally bounded, then \tilde{A} and $\overline{\text{conv}} \tilde{A}$ are ε -monotone, where $\tilde{A} : X \rightarrow X^*$ is defined as closure of $\text{Graph } A$ 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 $x_i \rightarrow x$, $y_i \rightarrow y$, strongly in X and $x_i^* \rightarrow x^*$, $y_i^* \rightarrow y^*$, 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 $\langle x - y, x^* - y^* \rangle \geq -2\varepsilon$, that is, \tilde{A} is ε -monotone. According to property (iii), $\overline{\text{conv}} \tilde{A}$ is also ε -monotone. \square

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. \square

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 $\langle x^* - y^*, x - y \rangle \geq 0$, 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*

$$\partial f(x) = \bigcap_{\varepsilon > 0} (\partial f)^\varepsilon(x), \quad \text{for all } x \in X. \quad (2.118)$$

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*

$$\partial_\varepsilon f(x) = \left\{ x^* \in X^*; (x^*, x_0 - x) + \sum_{i=0}^{m-1} (x_i^*, x_{i+1} - x_i) + (x_m^*, x - x_m) \leq \varepsilon \right. \\ \left. \text{for all } x_i^* \in \partial f(x_i), i = 0, 1, \dots, m \right\}, \quad (2.119)$$

where $x \in \text{Dom}(f)$ 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 $x \in \text{Dom}(f)$. 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

$$f(x) - f(x_0) - \eta \leq (x^*, x - x_0) + \varepsilon, \quad \text{for all } \eta > 0,$$

that is,

$$f(x) - f(x_0) \leq (x^*, x - x_0) + \varepsilon, \quad \text{for every } x_0 \in D(\partial f).$$

Now, since $D(\partial f)$ is a dense subset of $\text{Dom}(f)$ (see Corollary 2.44), by lower-semi-continuity this inequality holds for every $x_0 \in \text{Dom}(f)$, 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. \square

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,

$$N^\lambda(f) = \{x \in X; f(x) \leq \lambda\}. \quad (2.120)$$

Let us consider the following sets:

$$D_\lambda f(x_0) = \{(x^*, \delta) \in X^* \times (0, \infty); x^*(x_0 - x) \geq \delta \text{ whenever } f(x) \leq \lambda\}, \quad (2.121)$$

for every $x_0 \in X$ and $\lambda \in \mathbb{R}$.

It is obvious that, if $D_\lambda f(x_0) \neq \emptyset$, then $f(x_0) > \lambda$. Indeed, if we suppose that $f(x_0) \leq \lambda$, then, for an element $(x^*, \delta) \in D_\lambda f(x_0)$, we have $0 = x^*(x_0 - x_0) \geq \delta$, 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 $F_f : X \times \mathbb{R} \rightarrow \overline{\mathbb{R}}$ defined by

$$F_f(x, t) = f(x) - t, \quad (x, t) \in X \times \mathbb{R}, \quad (2.122)$$

is quasi-convex, since $N^\lambda(F_f) = -(0, \lambda) + \text{epi } f$, for all $\lambda \in \mathbb{R}$. Thus, we have

$$D_\lambda F_f(x_0, t_0) = \{(x^*, \alpha, \delta) \in X^* \times \mathbb{R} \times (0, \infty); x^*(x_0 - x) + \alpha(t_0 - t) \geq \delta, \\ \text{whenever } f(x) - t \leq \lambda\}.$$

By a simple calculation, we find that $(x^*, \alpha, \delta) \in D_\lambda F_f(x_0, t_0)$ if $\alpha = 0$ and $\sup\{(x^*, x); x \in \text{Dom}(f)\} \leq x^*(x_0) - \delta$ 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 $f(x_0) - t_0 = F_f(x_0, t_0) > \lambda$, $\alpha \leq 0$, whenever $D_\lambda F_f(x_0, t_0) \neq \emptyset$.

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 $x^* \in \partial_{\varepsilon_0} f(x_0)$, $\varepsilon_0 = f(x_0) - t_0 - \lambda > 0$, $x_0 \in \text{Dom}(f)$.

Now, we can establish the following characterization of quasi-convex lower-semicontinuous functions.

Theorem 2.98 *A function $f : X \rightarrow \overline{\mathbb{R}}$ 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 $x_0 \in N^\lambda(f)$ 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 $x^*(x) \leq k$ for all $x \in N^\lambda(f)$. Taking $\delta = x^*(x_0) - k > 0$, we obtain $x^*(x - x_0) \leq -\delta$ for all $x \in N^\lambda(f)$, equivalently $(x^*, \delta) \in D_\lambda f(x_0)$. This finishes the proof of Theorem 2.98. \square

Corollary 2.99 *A proper function $f : X \rightarrow \overline{\mathbb{R}}$ 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

$$\partial_q^\lambda f(x_0) = \left\{ x^* \in X^*; \sup_{x \in N^\lambda f} x^*(x - x_0) < 0 \right\}. \quad (2.123)$$

Proposition 2.100 *Let us consider $f : X \rightarrow \overline{\mathbb{R}}$, $x_0 \in X$, $f(x_0) \neq -\infty$, $\varepsilon > 0$. 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, $0 \notin \partial_q^\lambda f(x_0)$. 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, $f(x) \geq f(x_0) - \varepsilon$ for all $x \in X$. Also, (ii) and (iii) are obviously equivalent. \square

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

$$\partial_{\text{GP}}^\lambda f(x_0) = \{x^* \in X^*; x^*(x - x_0) < 0 \text{ if } f(x) < \lambda\}, \quad x_0 \in X, \quad (2.124)$$

$$\begin{aligned} \partial_{\text{M-L}} f(x_0) = \{x^* \in X^*; \text{there exists } k \in K \text{ such that } k \circ x^* \leq f \\ \text{and } k(x^*(x_0)) = f(x_0)\}, \quad x_0 \in X, \end{aligned} \quad (2.125)$$

where K is a given family of functionals $k \in \mathbb{R} \rightarrow \overline{\mathbb{R}}$ 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 $\partial_{\text{GP}} f(x_0) = \partial_K f(x_0)$ if K is the family of all nondecreasing functions.

The λ -quasi-subdifferential associated to the quasi-subdifferential (2.125) is defined as follows:

$$\begin{aligned} \partial_{\text{M-L}}^\lambda f(x_0) = \{ & x^* \in X^*; \text{ there exists } k \in K \text{ such that } k \circ x^* \leq f \\ & \text{and } k((x^*)(x_0)) \geq \lambda \}. \end{aligned} \quad (2.126)$$

Proposition 2.101 *Let K be the family of all nondecreasing functions $k : \mathbb{R} \rightarrow \overline{\mathbb{R}}$. If $f : X \rightarrow \overline{\mathbb{R}}$, $x_0 \in X$ and $\lambda \in \mathbb{R}$, then*

$$\partial_{\text{GP}}^\lambda f(x_0) = \partial_{\text{M-L}}^\lambda f(x_0).$$

Proof From the definition of $\partial_{\text{M-L}}^\lambda$ given by (2.126), we obtain the inclusion $\partial_{\text{M-L}}^\lambda f(x_0) \subset \partial_{\text{GP}}^\lambda f(x_0)$. Conversely, if $x^* \in \partial_{\text{GP}}^\lambda f(x_0)$, taking $k : \mathbb{R} \rightarrow \overline{\mathbb{R}}$ defined by

$$k(t) = \inf\{a; x^*(x) \geq t \text{ if } f(x) < a\},$$

we have $k(x^*(x)) \leq a$ whenever $f(x) < a$. But k is obvious a nondecreasing function, and so $k \circ x^* \leq f$. Also, $k(x^*(x_0)) \geq \lambda$. Hence, $x^* \in \partial_{\text{M-L}}^\lambda f(x_0)$ and the proof is complete. \square

Proposition 2.102 *Let K be the family of all nondecreasing lower-semicontinuous functions. If $f : X \rightarrow \overline{\mathbb{R}}$, $x_0 \in X$, $\lambda_1, \lambda_2 \in \mathbb{R}$ and $\lambda_1 > \lambda_2$, then*

- (i) $\partial_{\text{M-L}}^{\lambda_1} f(x_0) \subset \partial_q^{\lambda_2} f(x_0) \subset \partial_{\text{M-L}}^{\lambda_2} f(x_0)$.
- (ii) $\bigcap_{\lambda < f(x_0)} \partial_q^\lambda f(x_0) = \bigcap_{\lambda < f(x_0)} \partial_{\text{M-L}}^\lambda f(x_0) = \partial_{\text{M-L}} f(x_0)$, if $f(x_0) \in \mathbb{R}$.

Proof Equality (ii) follows by using (i) and the equality

$$\bigcap_{\lambda < f(x_0)} \partial_q^\lambda f(x_0) = \partial_{\text{M-L}} f(x_0).$$

Now, if $x^* \in \partial_q^\lambda f(x_0)$, taking the function k defined in the proof of Proposition 2.101, we notice that k is also lower-semicontinuous. Hence, $k(x^*(x)) \leq a$ if $f(x) < a$, and so, $k \circ x^* \leq f$. Since $\sup_{x \in N^\lambda(f)} x^*(x - x_0) < 0$, it follows that $k(x^*(x_0)) \geq \lambda$. Hence, $\partial_q^\lambda f(x_0) \subset \partial_{\text{M-L}}^\lambda f(x_0)$. On the other hand, if $\partial_{\text{M-L}}^\lambda f(x_0) = \emptyset$ or $N^\lambda(f) = \emptyset$, then the inclusion of the left-hand side of (i) is obvious. Let us suppose that $N^\lambda(f) \neq \emptyset$. Thus, if $x^* \in \partial_{\text{M-L}}^{\lambda_1} f(x_0)$, we have $k(x^*(x)) - k(x^*(x_0)) < \lambda - \lambda_1$, for all x , such that $f(x) \leq \lambda$. 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 $x^*(x_i) \rightarrow \sup_{x \in N^\lambda(f)} x^*(x)$. Since $k(x^*(x_i)) - k(x^*(x_0)) < \lambda - \lambda_1$, by passing to the limit we obtain

$$k(x^*(x_0) + \alpha) - k(x^*(x_0)) \leq \lambda - \lambda_1 < 0.$$

Hence, $\alpha < 0$ and so, $x^* \in \partial_q^\lambda f(x_0)$. Thus, Proposition 2.102 is completely proved. \square

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 $f : \mathbb{R}^n \rightarrow \mathbb{R}$ 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 $\{\lim_{n \rightarrow \infty} \nabla f(x + x_n)\}$, where $x_n \rightarrow 0$ 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,

$$d_C(x) = \inf\{\|x - y\|; y \in C\}.$$

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

$$\{z \in \mathbb{R}^n; \lambda z \in \partial d_C(x) \text{ for some } \lambda > 0\}. \quad (2.127)$$

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

$$\nabla d_C(x) = (x - P_C(x))\|x - P_C(x)\|^{-1}, \quad x \in C,$$

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, $E(f) = \text{epi } f$).

Given the lower-semicontinuous function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, we define the *upper derivative* of f at x with respect to y , as

$$f^\uparrow(x, y) = \lim_{\substack{x' \rightarrow x \\ f(x') \rightarrow f(x) \\ \lambda \downarrow 0}} \inf_{y' \rightarrow y} \frac{f(x' + \lambda y') - f(x')}{\lambda}. \quad (2.128)$$

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*

$$\partial f(x) = \{z \in \mathbb{R}^n; f^\uparrow(x, y) \geq (y, z), \forall y \in \mathbb{R}^n\}. \quad (2.129)$$

If $f^\uparrow(x, 0) = -\infty$, then $\partial f(x)$ is empty, but otherwise $\partial f(x) \neq \emptyset$ and one has

$$f^\uparrow(x, y) = \max\{(y, z); z \in \partial f(x), \forall y \in \mathbb{R}^n\}. \quad (2.130)$$

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 $f : X \rightarrow \mathbb{R}$ 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

$$f^0(x, z) = \limsup_{\substack{x' \rightarrow x \\ \lambda \downarrow 0}} \frac{f(x' + \lambda z) - f(x')}{\lambda}.$$

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

$$f^0(x, z) \geq (z, x^*) \quad \text{for all } z \in X. \quad (2.131)$$

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 $\eta_i \rightarrow \eta$ weak-star in X^* and $x_i \rightarrow x$ 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

$$\sup_{x \in X} \inf_{y \in Y} F(x, y) \leq \inf_{y \in Y} \sup_{x \in X} F(x, y). \quad (2.132)$$

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

$$F(x, \tilde{y}) \leq F(\tilde{x}, \tilde{y}) \leq F(\tilde{x}, y), \quad \forall (x, y) \in X \times Y. \quad (2.133)$$

Definition 2.104 The pair $(\tilde{x}, \tilde{y}) \in X \times Y$ is said to be a *saddle point* for the function $F : X \times Y \rightarrow \mathbb{R}$ 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

$$\inf_{y \in Y} \sup_{x \in X} F(x, y) \leq \sup_{x \in X} F(x, \tilde{y}) \leq \inf_{y \in Y} F(\tilde{x}, y) \leq \sup_{x \in X} \inf_{y \in Y} F(x, y),$$

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 $\text{cl } f$ to denote its closure (see Sect. 2.1.3). For a concave function g , the closure $\text{cl } g$ is defined by

$$\text{cl } g = -\text{cl}(-g).$$

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 $\text{cl}_1 K$ the function obtained by closing $K(x, y)$ as a concave function of x for each y . Similarly, $\text{cl}_2 K$ 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:

$$\text{cl}_1 \text{cl}_2 K = \text{cl}_1 K, \quad \text{cl}_2 \text{cl}_1 K = \text{cl}_2 K. \quad (2.134)$$

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

$$\text{cl}_1 K = \text{cl}_1 K' \quad \text{and} \quad \text{cl}_2 K = \text{cl}_2 K'.$$

In other words, the saddle function K is closed if $\text{cl}_1 K$ and $\text{cl}_2 K$ 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

$$\inf\{K(x, y); y \in Y\} = \inf\{\text{cl}_2 K(x, y); y \in Y\} \quad \text{for every } x \in X, \quad (2.135)$$

and, similarly,

$$\sup\{K(x, y); x \in X\} = \sup\{\text{cl}_1 K(x, y); x \in X\} \quad \text{for every } y \in Y. \quad (2.136)$$

Hence, if (x_0, y_0) is a saddle point of K , that is,

$$K(x, y_0) \leq K(x_0, y_0) \leq K(x_0, y) \quad \text{for all } (x, y) \in X \times Y,$$

we have

$$\sup\{\text{cl}_1 K(x, y_0); x \in X\} = K(x_0, y_0) = \inf\{\text{cl}_2 K(x_0, y); y \in Y\}$$

and therefore for any saddle function K' equivalent with K ,

$$\sup\{K'(x, y_0); x \in X\} = K(x_0, y_0) = \inf\{K'(x_0, y); y \in Y\},$$

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

$$D_1(K) = \{x \in X; K(x, y) > -\infty \text{ for every } y \in Y\}, \quad (2.137)$$

$$D_2(K) = \{y \in Y; K(x, y) < +\infty \text{ for every } x \in X\}. \quad (2.138)$$

It is easy to see that $D_1(K)$ and $D_2(K)$ are convex sets. The set

$$\text{dom } K = D_1(K) \times D_2(K) \quad (2.139)$$

is called the *effective domain* of K . Obviously, K is finite on $\text{dom } K$ and, if K is finite everywhere, one has $\text{dom } K = X \times Y$.

As an example, let A and B be nonempty convex sets in X and Y , respectively, and let

$$K(x, y) = \begin{cases} K_0(x, y), & \text{if } x \in A \text{ and } y \in B, \\ +\infty, & \text{if } x \in A \text{ and } y \notin B, \\ -\infty, & \text{if } x \notin A \text{ and } y \in Y, \end{cases} \quad (2.140)$$

where K_0 is any finite saddle function on $A \times B$. Then, K is a saddle function on $X \times Y$ with

$$\text{dom } K = A \times B.$$

A saddle function $K : X \times Y \rightarrow \overline{\mathbb{R}} = [-\infty, +\infty]$ is called *proper* if $\text{dom } K \neq \emptyset$.

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 $y \in \text{int } D_2(K)$, 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 $y \in \text{int } D_1(K)$, 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 $\text{cl}_1 \text{cl}_2 K = \text{cl}_1 K$. Hence

$$\text{cl}_1 K(x, y) = \lim_{\varepsilon \rightarrow 0} \sup_{\|x-u\| \leq \varepsilon} \text{cl}_2 K(u, y) \quad \text{for every } y \in D_2(K).$$

We set

$$\varphi_\varepsilon(x, y) = \sup_{\|x-u\| \leq \varepsilon} \text{cl}_2 K(u, y).$$

Since $\text{cl}_2 K \leq \text{cl}_1 K$ and the function $x \rightarrow \text{cl}_1 K(x, y)$, $x \in X$, is upper-semicontinuous and concave on X , we may infer that

$$\varphi_\varepsilon(x, y) < +\infty \quad \text{for every } x \in X \text{ and } y \in D_2(K). \quad (2.141)$$

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 $\text{cl}_2 K(u, \cdot)$. Therefore, $\varphi_\varepsilon(x, y)$ is, for any $\varepsilon > 0$, a continuous function of $y \in \text{int } D_2(K)$ (see Proposition 2.16). But this function majorizes the convex function $\text{cl}_1 K(x, \cdot)$, and hence we may conclude that the latter is also continuous on $\text{int } D_2(K)$. Of course, $\text{cl}_1 K \geq K \geq \text{cl}_2 K$, while the closedness of K implies that $\text{cl}_2 K = \text{cl}_2 \text{cl}_1 K$. From the latter relation, we have

$$\text{cl}_1 K(x, y) = \text{cl}_2 K(x, y) \quad \text{for every } x \in X \text{ and } y \in \text{int } D_2(K),$$

hence

$$K(x, y) = \text{cl}_1 K(x, y) \quad \text{for every } x \in X \text{ and } y \in \text{int } D_2(K).$$

Hence, $K(\cdot, y)$ is concave and upper-semicontinuous for every $y \in \text{int } D_2(K)$. 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 $\text{int } D_2(K)$.

Therefore, the convex function $y \rightarrow \text{cl}_2 K(x_0, y)$, $y \in Y$, is not identically $-\infty$ which shows that $\text{cl}_2 K(x_0, y)$ 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 $K : X \times Y \rightarrow \overline{\mathbb{R}}$, 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,

$$\partial_y K(x, y) = \{y^* \in Y^*; K(x, y) \leq K(x, y) + (y - v, y^*), \forall v \in Y\}, \quad (2.142)$$

$$\partial_x K(x, y) = \{x^* \in X^*; K(u, y) \leq K(x, y) + (u - x, x^*), \forall u \in X\}. \quad (2.143)$$

The multivalued operator $\partial K : X \times Y \rightarrow X^* \times Y^*$ defined by

$$\partial K(x, y) = \{-\partial_x K(x, y), \partial_y K(x, y)\}, \quad (x, y) \in X \times Y, \quad (2.144)$$

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

$$(0, 0) \in \partial K(x_0, y_0). \quad (2.145)$$

□

Proposition 2.109 *Let K be a proper saddle function on $X \times Y$. The multivalued mapping $\partial K : X \times Y \rightarrow X^* \times Y^*$ is a monotone operator with*

$$D(\partial K) \subset \text{dom } K. \quad (2.146)$$

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 $(x_1, y_1) \in \text{dom } K$, 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). □

Corollary 2.110 *Let K be a proper closed saddle function on $X \times Y$. Then*

$$\text{int dom } K \subset D(\partial K) \subset \text{dom } K. \quad (2.151)$$

Proof Let $(x, y) \in \text{int dom } K$. Thus, $x \in \text{int } D_1(K)$ and $y \in \text{int } D_2(K)$, so that Theorem 2.108 together with Corollary 2.38 imply that K is subdifferentiable at (x, y) , establishing (2.151). \square

Corollary 2.111 *Let K be a proper and closed saddle function on $X \times Y$. Then K is continuous on $\text{int dom } K$.*

Proof From Theorem 1.144, and Corollary 2.110, it follows that the monotone operator ∂K is locally bounded on $\text{int dom } K \subset \text{int } D(\partial K)$. Let (x_0, y_0) be any element in $\text{int dom } K$. By definition, for all $(x \times Y)$, one has

$$K(x_0, y_0) - K(x, y) \leq (y_0 - y, y_0^*) + (x - x_0, x^*) \quad (2.152)$$

and

$$K(x, y) - K(x_0, y_0) \leq (y - y_0, y^*) + (x_0 - x, x_0^*), \quad (2.153)$$

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

$$|K(x_0, y_0) - K(x, y)| \leq C_1(\|x - x_0\| + \|y - y_0\|),$$

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. \square

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*

$$L(x, y^*) = \sup\{(y, y^*) - K(x, y); y \in Y\}, \quad (2.154)$$

$$K(x, y) = \sup\{(y, y^*) - L(x, y^*); y^* \in Y\} \quad (2.155)$$

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*

$$\text{cl}_2 \text{cl}_1 K = K. \quad (2.156)$$

Moreover, under this correspondence, one has

$$(x^*, y^*) \in \partial K(x, y) \iff (-x^*, y) \in \partial L(x, y^*). \quad (2.157)$$

Proof Let $L : X \times Y^* \rightarrow]-\infty, +\infty]$ be convex, lower-semicontinuous and non-identically $+\infty$ on $X \times Y^*$. Formula (2.155) says that K is the 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 $L^* : X^* \times Y \rightarrow]-\infty, +\infty]$ 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

$$L^*(x^*, y) = \sup\{(x, x^*) + K(x, y); x \in X\}. \quad (2.158)$$

Hence,

$$\text{cl}_1 K(x, y) = -\sup\{(x, x^*) - L^*(x^*, y); x^* \in X^*\}. \quad (2.159)$$

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

$$L(x, y^*) = \sup\{(y, y^*) - \text{cl}_1 K(x, y); y \in Y\},$$

and therefore

$$\text{cl}_2 \text{cl}_1 K(x, y) = \sup\{(y, y^*) - L(x, y^*); y^* \in Y^*\}.$$

Combining this with relation (2.155), we obtain

$$\text{cl}_2 \text{cl}_1 K(x, y) = K(x, y) \quad \text{for every } (x, y) \in X \times Y,$$

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 $\text{dom } K \neq \emptyset$, 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

$$\text{cl } L(x, y^*) = \sup\{(x, x^*) + (y, y^*) - L^*(x^*, y); x^* \in X^*, y \in Y\}.$$

Combining this with equality (2.159), we obtain

$$\begin{aligned} \text{cl } L(x, y^*) &= \sup\{(y, y^*) - \text{cl}_1 K(x, y); y \in Y\} \\ &= \sup\{(y, y^*) - \text{cl}_2 \text{cl}_1 K(x, y); y \in Y\}, \end{aligned}$$

which is equivalent to

$$\text{cl } L(x, y^*) = \sup\{(y, y^*) - K(x, y); y \in Y\} = L(x, y^*)$$

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

$$\begin{aligned} -(x^*, x - x_1) + (y, y^* - y_1^*) &\geq -K(x, y) + K(x_1, y) + (y, y^* - y_1^*) \\ \text{for all } x_1 \in X, y_1^* \in Y^*. \end{aligned} \tag{2.160}$$

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 $y \rightarrow K(x, y)$ 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. \square

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 $K' = \text{cl}_2 K$. 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 $K : X \times Y \rightarrow \overline{\mathbb{R}}$ be a proper, closed saddle function on $X \times Y$. Then the operator $\partial K : X \times Y \rightarrow X^* \times Y^*$ 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 $(x, y) \rightarrow K(x, y) + (x, x_0^*) - (y, y_0^*)$ which is in turn equivalent to $(x, y) \rightarrow K'(x, y) + (x, x_0^*) - (y, y_0^*)$. 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 $\text{cl}_2 K$, 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. \square

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 $K : X \times Y \rightarrow \overline{\mathbb{R}}$ 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 $\text{dom } K$.*

Proof Let (x_0, y_0) be any element of $\text{dom } K$, and let $(x_\lambda, y_\lambda) \in X \times Y$ be such that

$$F_1(x_\lambda - x_0) - \lambda \partial_x K(x_\lambda, y_\lambda) \ni 0, \quad \lambda > 0, \quad (2.164)$$

$$F_2(y_\lambda - y_0) - \lambda \partial_y K(x_\lambda, y_\lambda) \ni 0, \quad \lambda > 0, \quad (2.165)$$

where $F_1 : X \rightarrow X^*$ and $F_2 : Y \rightarrow Y^*$ are duality mappings of X and Y , respectively. Since ∂K is maximal monotone and the operator $(x, y) \rightarrow (F_1(x - x_0), F_2(y - y_0))$ 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

$$\begin{aligned} & (F_1(x_\lambda - x_0), x_\lambda - x_0) + (F_2(y_\lambda - y_0), y_\lambda - y_0) \\ & \leq \lambda(K(x_\lambda, y_0) - K(x_0, y_\lambda)), \quad \text{for all } \lambda > 0. \end{aligned} \quad (2.166)$$

Inasmuch as $(x_0, y_0) \in \text{dom } K$, the functions $x \rightarrow -K(x, y_0)$ and $y \rightarrow K(x_0, y)$ 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

$$\|x_\lambda - x_0\|^2 + \|y_\lambda - y_0\|^2 \leq C\lambda(\|x_\lambda\| + \|y_\lambda\| + 1). \quad (2.167)$$

Therefore $x_\lambda \rightarrow x_0$ and $y_\lambda \rightarrow y_0$ as $\lambda \rightarrow 0$, thereby proving Corollary 2.116. \square

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*

$$\lim_{\substack{\|x\| + \|y\| \rightarrow +\infty \\ x \in A, y \in B}} (K_0(x, y_0) - K_0(x_0, y)) = -\infty. \quad (2.168)$$

Then, the function K_0 has at least one saddle point on $A \times B$.

Proof Let $K : X \times Y \rightarrow [-\infty, +\infty]$ be the closed saddle function defined by (2.140). By Theorem 2.114, the operator $\partial K : X \times Y \rightarrow X^* \times Y^*$ is maximal monotone. Hence, for each $\lambda > 0$ $(x_\lambda, y_\lambda) \in D(\partial K) = A \times B$ such that

$$\lambda F_1(x_\lambda) - \partial_x K(x_\lambda, y_\lambda) \ni 0, \quad (2.169)$$

$$\lambda F_2(y_\lambda) + \partial_y K(x_\lambda, y_\lambda) \ni 0, \quad (2.170)$$

where $F_1 : X \rightarrow X^*$ and $F_2 : Y \rightarrow Y^*$ 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

$$\lambda(F_1(x_\lambda), x_\lambda - x_0) \leq K(x_\lambda, y_\lambda) - K(x_0, y_\lambda),$$

$$\lambda(F_2(y_\lambda), y_\lambda - y_0) \leq K(x_\lambda, y_\lambda) + K(x_\lambda, y_0).$$

Therefore,

$$\lambda(\|x_\lambda\|^2 + \|y_\lambda\|^2) \leq \lambda(\|x_\lambda\| \|x_0\| + \|y_\lambda\| \|y_0\|) + K(x_\lambda, y_0) - K(x_0, y_\lambda).$$

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

$$\begin{aligned} x_\lambda &\rightarrow \tilde{x} \quad \text{weakly in } X, \\ y_\lambda &\rightarrow \tilde{y} \quad \text{weakly in } Y, \end{aligned} \quad (2.171)$$

as $\lambda \rightarrow 0$. If we let $\lambda \rightarrow 0$ in equations (2.169) and (2.170), we may infer that

$$\lim_{\lambda \rightarrow 0} \partial K(x_\lambda, y_\lambda) = (0, 0) \quad \text{strongly in } X^* \times Y^*. \quad (2.172)$$

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

$$(0, 0) \in \partial K(\tilde{x}, \tilde{y}). \quad (2.173)$$

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,

$$K_0(x, \tilde{y}) \leq K_0(\tilde{x}, \tilde{y}) \leq K_0(\tilde{x}, y) \quad \text{for all } x \in A \text{ and } y \in B,$$

and this establishes Corollary 2.118. \square

Let $K^* : X^* \times Y^* \rightarrow \overline{\mathbb{R}}$ 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,

$$(x, y) \in \partial K^*(x^*, y^*) \iff (x^*, y^*) \in \partial K(x, y). \quad (2.174)$$

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 $F : X \times Y \rightarrow \overline{\mathbb{R}}$. An important problem is to establish certain conditions on F, X and Y under which the mini-max equality

$$\max_{x \in X} \min_{y \in Y} F(x, y) = \min_{y \in Y} \max_{x \in X} F(x, y) \quad (2.175)$$

is true or at least a saddle value exists, that is,

$$\sup_{x \in X} \inf_{y \in Y} F(x, y) = \inf_{y \in Y} \sup_{x \in X} F(x, y). \quad (2.176)$$

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 be 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 $y_1, y_2, \dots, y_n \in B$ such that $\alpha > \max_{x \in A} \min_{1 \leq i \leq n} F(x, y_i)$, there is $y_0 \in B$ such that $\alpha > \max_{x \in A} F(x, y_0)$.*
- (b) *F satisfies the equality*

$$\max_{x \in A} \inf_{y \in B} F(x, y) = \inf_{y \in B} \max_{x \in A} F(x, y). \quad (2.177)$$

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 $A_y = \{x \in A; F(x, y) \geq \alpha\}$, 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 $y_1, \dots, y_n \in B$ with $\bigcap_{i=1}^n A_{y_i} = \emptyset$, which implies $\min_{1 \leq i \leq n} F(x, y_i) < \alpha$, for each $x \in X$. Thus, $\max_{x \in A} \min_{1 \leq i \leq n} F(x, y_i) < \alpha$ 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

$$\max_{x \in A} \inf_{y \in B} F(x, y) \geq \inf_{y \in B} \max_{x \in A} F(x, y).$$

Moreover, it follows from (2.132) that equality (2.177) holds. \square

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 $F(x, y_3) \leq F(x, y_1)$ and $F(x, y_3) \leq F(x, y_2)$ 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*

$$\lim_{n \rightarrow \infty} \max_{x \in A} f_n(x) = \max_{x \in A} \lim_{n \rightarrow \infty} f_n(x). \quad (2.178)$$

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). \square

Remark 2.122 The previous theorem is not really a mini-max theorem. If, moreover, B is a compact set and $y \rightarrow F(x, y)$ 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 $F : X \times Y \rightarrow \overline{\mathbb{R}}$ 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

$$tF(x_1, y) + (1 - t)F(x_2, y) \leq F(x_3, y) \quad \text{for all } y \in Y, \quad (2.179)$$

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

$$F(x, y_3) \leq tF(x, y_1) + (1 - t)F(x, y_2) \quad \text{for all } x \in X, \quad (2.180)$$

whenever the right-hand side is well defined.

Definition 2.124 A function $F : X \times Y \rightarrow \overline{\mathbb{R}}$ 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 $t_1, t_2, \dots, t_n \geq 0$ with $\sum_{i=1}^n t_i = 1$, there is an $x_0 \in X$ such that

$$\sum_{i=1}^n t_i F(x_i, y) \leq F(x_0, y) \quad \text{for all } y \in Y, \quad (2.181)$$

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

$$\alpha > \max_{x \in A} \min_{1 \leq i \leq n} F(x, y_i). \quad (2.182)$$

Now, we consider the following convex sets of \mathbb{R}^n :

$$\begin{aligned} C_1 &= \text{conv}\{(F(x, y_1), F(x, y_2), \dots, F(x, y_n)); x \in A\}, \\ C_2 &= \{(u_1, u_2, \dots, u_n); u_i \geq \alpha, i = 1, 2, \dots, n\}. \end{aligned}$$

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 $\alpha_j \geq 0$, $j = 1, 2, \dots, m$, with $\sum_{j=1}^m \alpha_j = 1$, such that $u_i = \sum_{j=1}^m \alpha_j F(x_j, y_i)$ for every $i = 1, 2, \dots, n$. Now, from (i)', there exists a point $x_0 \in A$ such that

$$F(x_0, y) \geq \sum_{j=1}^m \alpha_j F(x_j, y). \quad (2.183)$$

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, $u = (u_1, u_2, \dots, u_n) \notin C_2$. 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

$$\sup_{u \in C_1} \sum_{i=1}^n c_i u_i \leq \inf_{u \in C_2} \sum_{i=1}^n c_i u_i. \quad (2.184)$$

However, the cone C_2 contains all the points $(\alpha, \alpha, \dots, \alpha, \alpha + n, \alpha, \dots, \alpha)$, $n \in \mathbb{N}$, and therefore $c_i \geq 0$; 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 $\sum_{i=1}^n c'_i F(x, y_i) \leq \alpha$ for all $x \in A$. Combining this with property (ii) from Definition 2.123, there is a point $y_0 \in B$ such that $F(x, y_0) \leq \alpha$ for every $x \in A$; hence, $\alpha \geq \max_{x \in A} F(x, y_0)$ 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). \square

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 is 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. \square

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 $\mathcal{F}(u) \subset E$ be a compact set such that the convex hull of every finite subset $\{u_1, u_2, \dots, u_n\} \subset U$ is contained in the corresponding union $\bigcup_{i=1}^n \mathcal{F}(u_i)$. Then, $\bigcap_{u \in U} \mathcal{F}(u) \neq \emptyset$.*

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*

$$\sup_{x \in A} \inf_{y \in B} F(x, y) = \inf_{y \in B} \sup_{x \in A} F(x, y). \quad (2.185)$$

Proof Suppose by contradiction that equality (2.185) is not true. From inequality (2.132), there is $\alpha > \alpha_0$, such that

$$\sup_{x \in A} \inf_{y \in B} F(x, y) < \alpha < \inf_{y \in B} \sup_{x \in A} F(x, y). \quad (2.186)$$

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

$$\bigcap_{y \in B} A_y = \emptyset, \quad \bigcap_{x \in A} B_x = \emptyset.$$

Since A_{y_0} is compact, there are $y_1, \dots, y_n \in B$ such that $\bigcap_{i=1}^n A_{y_i} = \emptyset$. On the other hand, as the convex sets finitely generated are compact, there are $x_1, \dots, x_m \in A$ such that

$$\bigcap_{i=1}^m B_{x_i} \cap \text{conv}\{y_i; i = 1, 2, \dots, n\} = \emptyset.$$

Let $A' = \text{conv}\{x_1, x_2, \dots, x_m\}$ and $B' = \text{conv}\{y_1, y_2, \dots, y_n\}$. Define the multi-valued mapping \mathcal{F} on $A' \times B'$ by

$$\mathcal{F}(u, v) = \{(w, s) \in A' \times B'; F(w, v) \geq \alpha \text{ or } F(u, s) \leq \alpha\}. \quad (2.187)$$

One may easily show that all the conditions of Theorem 2.129 are fulfilled. Indeed, $\mathcal{F}(u, v)$ 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

$$\sum_{i=1}^p \lambda_i(u_i, v_i) \bar{\in} \mathcal{F}(u_j, v_j) \quad \text{for all } j = 1, 2, \dots, p;$$

it follows that

$$F\left(\sum_{i=1}^p \lambda_i u_i, v_j\right) < \alpha \quad \text{and} \quad F\left(u_j, \sum_{i=1}^p \lambda_i v_i\right) > \alpha, \quad j = 1, 2, \dots, p.$$

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,

$$\sum_{i=1}^p \lambda_i(u_i, v_i) \in \bigcup_{i=1}^p \mathcal{F}(u_i, v_i).$$

Thus, according to Theorem 2.129, there is $(x_0, y_0) \in A' \times B'$ such that $(x_0, y_0) \in \mathcal{F}(x, y)$ 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 $x_0 \bar{\in} A_{y_{i_0}}$ and $y_0 \bar{\in} B_{x_{j_0}}$, 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*

$$\lim_{\substack{\|x\| + \|y\| \rightarrow \infty \\ (x, y) \in A \times B}} \{F(x_0, y) - F(x, y_0)\} = \infty, \quad (2.188)$$

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

$$F(x_0, y) - F(x, y_0) > 0. \quad (2.189)$$

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,

$$F(x, y') \leq F(x', y') \leq F(x', y), \quad (2.190)$$

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

$$F(x', y') \leq F(x', \lambda y + (1 - \lambda)y') \leq \lambda F(x', y) + (1 - \lambda)F(x', y'),$$

which leads to

$$F(x', y') \leq F(x', y),$$

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

$$F(x, y') \leq F(x', y'),$$

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). \square

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

$$\lim_{\substack{\|x\| \rightarrow \infty \\ x \in A}} F(x, y_0) = -\infty, \quad (2.191)$$

or, by the symmetric condition

(b)' A is bounded and there is $x_0 \in A$ such that

$$\lim_{\substack{\|y\| \rightarrow \infty \\ y \in B}} F(x_0, y) = +\infty. \quad (2.192)$$

Also, relations (2.191) and (2.192) together are sufficient.

All the results in this section can be applied to functions with values in $\overline{\mathbb{R}}$, 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 $F|_{\text{dom } F} = F_0$ (see (2.140) from Sect. 2.3.2). Moreover, we have

$$\sup_{x \in X} \inf_{y \in Y} F(x, y) = \sup_{x \in A} \inf_{y \in B} F_0(x, y), \quad (2.193)$$

$$\inf_{y \in Y} \sup_{x \in X} F(x, y) = \inf_{y \in B} \sup_{x \in A} F_0(x, y). \quad (2.194)$$

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 $F : X \times Y \rightarrow \overline{\mathbb{R}}$, 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 $A \times B = \text{dom } F$.*

Proof By definition of $A = D_1(F)$, we have

$$\sup_{x \in X} \inf_{y \in Y} F(x, y) = \sup_{x \in X} \inf_{y \in Y} \text{cl}_2 F(x, y) = \sup_{x \in A} \inf_{y \in Y} \text{cl}_2 F(x, y).$$

On the other hand, since F is closed, by definition of $B = D_2(F)$ we have

$$\inf_{y \in Y} \text{cl}_2 F(x, y) = \inf_{y \in Y} \text{cl}_2 \text{cl}_1 F(x, y) = \inf_{y \in Y} \text{cl}_1 F(x, y) = \inf_{y \in B} \text{cl}_1 F(x, y),$$

hence

$$\sup_{x \in X} \inf_{y \in Y} F(x, y) = \sup_{x \in A} \inf_{y \in B} \text{cl}_1 F(x, y) \geq \sup_{x \in A} \inf_{y \in B} F(x, y).$$

Also, the converse inequality holds

$$\sup_{x \in X} \inf_{y \in Y} F(x, y) = \sup_{x \in A} \inf_{y \in Y} \text{cl}_2 F(x, y) = \sup_{x \in A} \inf_{y \in Y} F(x, y) \leq \sup_{x \in A} \inf_{y \in B} F(x, y).$$

Similarly an obtains (2.194). \square

2.4 Problems

2.1 Let $f : I \rightarrow \overline{\mathbb{R}}$ 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 $f(x_n) \rightarrow \alpha$. 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

$$K = \left\{ w - \lim_{n_k \rightarrow \infty} x_{n_k} \right\},$$

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

$$\xi_1 = w - \lim_{n_k \rightarrow \infty} x_{n_k} \quad \text{and} \quad \xi_2 = w - \lim_{n'_k \rightarrow \infty} x_{n'_k},$$

this implies that

$$\begin{aligned} \lim_{n'_k \rightarrow \infty} |x_{n'_k} - \xi_1|^2 &= \lim_{n''_k \rightarrow \infty} |x_{n''_k} - \xi_1|^2, \\ \lim_{n''_k \rightarrow \infty} |x_{n''_k} - \xi_2|^2 &= \lim_{n'_k \rightarrow \infty} |x_{n'_k} - \xi_2|^2 \end{aligned}$$

and therefore $\xi_1 = \xi_2$, as claimed.

2.3 Let K be a closed convex subsets of \mathbb{R}^m and let

$$\mathcal{K} = \{y \in (L^p(\Omega))^m; y(x) \in K, \text{ a.e. } x \in \Omega\},$$

where $1 \leq p < \infty$ and Ω is a measurable sub set of \mathbb{R}^n . Find the normal cone $N_{\mathcal{K}}(y) \subset (L^q(\Omega))^m$ to \mathcal{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 $y \notin K$.

2.4 Find the normal cone $N_{\mathcal{K}}$ for

$$\begin{aligned} \mathcal{K} &= \{y \in L^p(\Omega); a \leq y(x) \leq b, \text{ a.e. } x \in \Omega\}, \\ \mathcal{K} &= \{y \in (L^p(\Omega))^m; \|y(x)\|_m \leq \rho, \text{ a.e. } x \in \Omega\}, \end{aligned}$$

where $\|\cdot\|_m$ is the Euclidean norm in \mathbb{R}^m .

2.5 Find the normal cone $N_{\mathcal{K}}$ to the set $\mathcal{K} = \{y \in L^2(\Omega); a \leq y(x) \leq b, \text{ a.e. } x \in \Omega, \int_{\Omega} y(x) dx = \ell\}$, where $am(\Omega) \leq \ell \leq bm(\Omega)$ (m is the Lebesgue measure).

Hint. We represent $\mathcal{K} = \mathcal{K}_1 \cap \mathcal{K}_2$ where $\mathcal{K}_1 = \{y \in L^2(\Omega); a \leq y(x) \leq b, \text{ a.e. } x \in \Omega\}$, $\mathcal{K}_2 = \{y \in L^2(\Omega); \int_{\Omega} y(x) dx = \ell\}$ and show that

$$N_{\mathcal{K}}(y) = N_{\mathcal{K}_1}(y) + N_{\mathcal{K}_2}(y), \quad \forall y \in \mathcal{K}.$$

Since $N_{\mathcal{K}_1}(y) + N_{\mathcal{K}_2}(y) \subset N_{\mathcal{K}}(y)$, it suffices to show that, for every $f \in L^2(\Omega)$, the equation $y + N_{\mathcal{K}_1}(y) + N_{\mathcal{K}_2}(y) \ni f$ has a solution $y \in \mathcal{K}$. Since $N_{\mathcal{K}_2}(y) = \mathbb{R}$, the above equation reduces to $y = P_{\mathcal{K}_1}(f - \lambda)$, $\lambda \in \mathbb{R}$, where $P_{\mathcal{K}_1}$ is projection on \mathcal{K}_1 .

2.6 Let $g : \mathbb{R} \rightarrow \mathbb{R}$ be a lower-continuous convex function such that $\lim_{|r| \rightarrow \infty} \frac{g(r)}{|r|} = +\infty$ and let $\varphi : H^{-1}(\Omega) \rightarrow \overline{\mathbb{R}}^*$ be defined by

$$\varphi(y) = \begin{cases} \int_{\Omega} g(y(x)) dx, & \text{if } g(y) \in L^1(\Omega), \\ +\infty, & \text{otherwise.} \end{cases}$$

Show that φ is lower-semicontinuous and that

$$\begin{aligned} \partial\varphi(y) &= \{-\Delta w; w \in H_0^1(\Omega), y \in H^{-1}(\Omega) \cap L^1(\Omega), \\ &\quad w(x) \in \partial g(y(x)) \text{ a.e. } x \in \Omega\}. \end{aligned} \quad (2.195)$$

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 $\{f_n\} \subset L^2(\Omega)$, $f_n \rightarrow f$ 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 $j : \mathbb{R} \rightarrow \mathbb{R}$ be a lower-semicontinuous convex function such that

$$\omega_2 |r|^p + c_2 \leq j(r) \leq \omega_1 |r|^p + c_1, \quad \forall r \in \mathbb{R},$$

where $\omega_1, \omega_2 > 0$ and $p > 1$. We set $\beta = \partial j$. Consider the function $\varphi : W_0^{1,p}(\Omega) \rightarrow \overline{\mathbb{R}}^*$ defined by

$$\varphi(y) = \int_{\Omega} j(\nabla y) dx.$$

Show that φ is convex, lower-semicontinuous and its subdifferential $\partial\varphi : W_0^{1,p}(\Omega) \rightarrow W^{-1,p'}(\Omega)$ is given by

$$\partial\varphi(y) = \{w \in W^{-1,p'}(\Omega); w = -\operatorname{div} \eta, \eta(x) \in \partial j(\nabla y(x)), \text{ a.e. } x \in \Omega\}. \quad (2.196)$$

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,

$$\varphi_t(x) = \inf \left\{ \frac{|x - y|^2}{2t} + \varphi(y); y \in H \right\} = S(t)\varphi, \quad t \geq 0.$$

Show that $S(t + s) = S(t)S(s)\varphi$, $\forall t, s > 0$, and

$$\frac{d^+}{dt} \varphi(t, x) + \frac{1}{2} |\nabla_x \varphi(t, x)|^2 = 0, \quad \forall t > 0, x \in H.$$

Remark 2.137 This means that $t \rightarrow S(t)\varphi$ is a continuous semigroup on the space of all continuous convex functions on H with infinitesimal generator $\varphi \rightarrow -\frac{1}{2} |\nabla_x \varphi(x)|^2$.

2.9 Let H be a Hilbert space and let F be a convex and continuously differentiable function on H such that

$$\lim_{|x| \rightarrow \infty} \frac{F(x)}{|x|} = +\infty, \quad \nabla F \text{ is locally Lipschitz,}$$

$$(F'(x) - F'(y), x - y) \geq \omega_r |x - y|^2, \quad \forall x, y, |x|, |y| \leq r.$$

We set

$$(S(t)\varphi)(x) = (\varphi^* + tF)^*(t), \quad t \geq 0, x \in H.$$

Show that:

- (1) $\lim_{t \rightarrow 0} S(t)\varphi(x) = \varphi(x)$.
- (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)

$$\begin{aligned} -y''(x) + y(x) &= f(x) && \text{in } [x \in [0, T]; y(x) > \rho], \\ -y''(x) + y(x) &\leq f(x) && \text{in } [x \in [0, 1]; y(x) = \rho], \\ y(x) &\geq \varphi, \quad \forall x \in [0, 1], && y(0) = y(1) = 0, \end{aligned}$$

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 Brønsted [15] and, subsequently, studied by Brønsted 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 Kararmadian [35], Habelter and Price [24], Eaves [20], Saigal [67], among others. Theorem 2.76 may be compared most closely with some results given by Kararmadian [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 inéquations 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 applications*. 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) Fonctionnelles convexes. Séminaire sur les équations aux dérivées partielles, Collège 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 Șt Univ Ali Cuza Iași, 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, Szegő 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. *Kibernetika* 6:65–70 (in Russian)



<http://www.springer.com/978-94-007-2246-0>

Convexity and Optimization in Banach Spaces

Barbu, V.; Precupanu, T.

2012, XII, 368 p., Hardcover

ISBN: 978-94-007-2246-0