2

Mathematical Preliminaries

In this chapter, we provide the mathematical background as it will be used in later chapters.

2.1 Basic Functional Analysis

The purpose of this section is to provide a survey of basic results from functional analysis that will be used in the sequel. However, we will assume that the reader is familiar with some elementary notions such as metric spaces, Banach spaces, and Hilbert spaces, as well as notions related with the topological structure of these spaces. Unless otherwise indicated, all linear spaces considered in this book are assumed to be defined over the real number field $\mathbb{R}$. The proofs of the results presented in this section can be found in standard textbooks, e.g., [5, 13, 24, 129, 200, 222].

2.1.1 Operators in Normed Linear Spaces

Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed linear spaces, and let

$$A : D(A) \subset X \to Y$$

be an operator with domain $D(A)$ and range denoted by range(A). When $D(A) = X$, we write

$$A : X \to Y.$$

Note that usually we drop the subscripts X and Y in the notation of the norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively, if no ambiguity exists.

Definition 2.1. *Let $A : D(A) \subset X \to Y$.*

(i) *A is continuous at the point $u \in D(A)$ iff for each sequence (u_n) in $D(A)$,*

$$u_n \to u \quad \textit{implies} \quad Au_n \to Au.$$

The operator $A : D(A) \subset X \to Y$ is called continuous iff it is continuous at each point $u \in D(A)$.

(ii) *A is called compact iff A is continuous, and A maps bounded sets into relatively compact sets.*

Note that one sometimes uses the notion *completely continuous* for compact. For compact operators, the following fixed-point theorem from Schauder holds.

Theorem 2.2 (Schauder's Fixed-Point Theorem). *Let X be a Banach space, and let*

$$A : M \to M$$

be a compact operator that maps a nonempty subset M of X into itself. Then A has a fixed point provided M is bounded, closed, and convex.

In finite-dimensional normed linear spaces, Theorem 2.2 reduces to Brouwer's fixed-point theorem.

Corollary 2.3 (Brouwer's Fixed-Point Theorem). *If the operator*

$$A : M \to M$$

is continuous, then A has a fixed point provided M is a compact, convex, nonempty subset in a finite-dimensional normed linear space.

Let

$$A : D(A) \subset X \to Y$$

be a *linear* operator, which means that the domain $D(A)$ of the operator A is a linear subspace of X and A satisfies

$$A(\alpha u + \beta v) = \alpha A u + \beta A v \quad \text{for all } u, v \in D(A),\ \alpha, \beta \in \mathbb{R}.$$

Proposition 2.4. *Let $A : X \to Y$ be a linear operator. Then the following two conditions are equivalent:*

(i) *A is continuous.*

(ii) *A is bounded; i.e., there is a constant $c > 0$ such that*

$$\|Au\| \leq c\|u\| \quad \textit{for all} \ \ u \in X.$$

For a linear continuous operator $A : X \to Y$, the operator norm $\|A\|$ is defined by

$$\|A\| = \sup_{\|u\| \leq 1} \|Au\|,$$

which can easily be shown to be equal to

$$\|A\| = \sup_{\|u\| = 1} \|Au\|.$$

Proposition 2.5. *Let $L(X,Y)$ denote the space of linear continuous operators $A : X \to Y$, where X is a normed linear space and Y is a Banach space. Then $L(X,Y)$ is a Banach space with respect to the operator norm.*

Definition 2.6. *Let*

$$A : D(A) \subset X \to Y$$

be a linear operator. The graph of A denoted by $\mathrm{Gr}(A)$ *is defined by the subset*

$$\mathrm{Gr}(A) = \{(u, Au) : u \in D(A)\}$$

of the product space $X \times Y$. The operator A is called closed (or graph-closed) iff $\mathrm{Gr}(A)$ *is closed in $X \times Y$, which means that for each sequence (u_n) in $D(A)$, it follows from*

$$u_n \to u \quad \text{in } X \quad \text{and} \quad Au_n \to v \quad \text{in } Y$$

that $u \in D(A)$ and $v = Au$. Finally, on $D(A)$, the so-called graph norm $\|\cdot\|_A$ is defined by

$$\|u\|_A = \|u\| + \|Au\| \quad \text{for} \ \ u \in D(A).$$

Corollary 2.7. *If X and Y are Banach spaces and $A : D(A) \subset X \to Y$ is closed, then $D(A)$ equipped with the graph norm, i.e., $(D(A), \|\cdot\|_A)$, is a Banach space.*

Theorem 2.8 (Banach's Closed Graph Theorem). *Let X and Y be Banach spaces. Then each closed linear operator $A : X \to Y$ is continuous.*

For completeness, we shall recall the Uniform Boundedness Theorem and the Open Mapping Theorem, which together with Banach's Closed Graph Theorem are all consequences of Baire's Theorem.

Theorem 2.9 (Uniform Boundedness Theorem). *Let $\mathcal{F}$ be a nonempty set of continuous maps*

$$F : X \to Y,$$

where X is a Banach space and Y is a normed linear space. Assume that

$$\sup_{F \in \mathcal{F}} \|Fu\| < \infty \quad \text{for each} \ \ u \in X.$$

Then a closed ball $\overline{B}$ in X of positive radius exists such that

$$\sup_{u \in \overline{B}} \big(\sup_{F \in \mathcal{F}} \|Fu\| \big) < \infty.$$

Corollary 2.10 (Banach–Steinhaus Theorem). *Let $\mathcal{L} \subset L(X, Y)$ be a nonempty set of linear continuous operators*

$$A : X \to Y,$$

where X is a Banach space and Y is a normed linear space. Assume that

$$\sup_{A \in \mathcal{L}} \|Au\| < \infty \quad \textit{for each} \quad u \in X.$$

Then $\sup_{A \in \mathcal{L}} \|A\| < \infty$.

Theorem 2.11 (Banach's Open Mapping Theorem). *Let X and Y be Banach spaces and $A : X \to Y$ be a linear continuous operator. Then the following two conditions are equivalent:*

(i) *A is surjective.*
(ii) *A is open, which means that A maps open sets onto open sets.*

Corollary 2.12 (Banach's Continuous Inverse Theorem). *Let X and Y be Banach spaces and $A : X \to Y$ be a linear continuous operator. If the inverse operator*

$$A^{-1} : Y \to X$$

exists, then A^{-1} is continuous.

Definition 2.13 (Embedding Operator). *Let X and Y be normed linear spaces with*

$$X \subset Y.$$

The embedding operator $i : X \to Y$ is defined by $i(u) = u$; i.e., i is the identity operator from X into Y.

(i) *The embedding $X \subset Y$ is called continuous iff the embedding operator $i : X \to Y$ is continuous; i.e., a constant $c > 0$ exists such that*

$$\|u\|_Y \le c \, \|u\|_X \quad \textit{for all} \quad u \in X,$$

which is equivalent with

$$u_n \to u \quad \textit{in} \quad X \quad \textit{implies} \quad u_n \to u \quad \textit{in} \quad Y.$$

(ii) *The embedding $X \subset Y$ is called compact iff the embedding operator $i : X \to Y$ is compact; i.e., i is continuous and each bounded sequence (u_n) in X has a subsequence that converges in Y.*

Remark 2.14. More generally, one can define a continuous embedding of a normed linear space X into a normed linear space Y, whenever a linear, continuous, and injective operator $i : X \to Y$ exists. Similarly, X is compactly embedded into Y iff a linear, compact, and injective operator $i : X \to Y$ exists.

2.1.2 Duality in Banach Spaces

Definition 2.15. *Let X be a normed linear space. A linear continuous functional on X is a linear continuous operator*

$$f : X \to \mathbb{R}.$$

The set of all linear continuous functionals on X is called the dual space X^* of X; i.e., $X^* = L(X, \mathbb{R})$. For the image $f(u)$ of the functional f at $u \in X$, we write

$$\langle f, u \rangle = f(u) \quad u \in X, \quad f \in X^*,$$

and $\langle \cdot, \cdot \rangle$ is called the duality pairing.

According to the operator norm defined in Sect. 2.1.1, the norm of f is given through

$$\|f\| = \sup_{\|u\| \leq 1} |\langle f, u \rangle|.$$

As a consequence of Proposition 2.5, we get the following result.

Corollary 2.16. *Let X be a normed linear space. Then the dual space X^* is a Banach space with respect to the norm $\|f\|$ for $f \in X^*$.*

The most important theorem about the structure of linear functionals on normed linear spaces is the Hahn–Banach Theorem. For real linear spaces, the Hahn–Banach Theorem reads as follows (see [24]).

Theorem 2.17 (Hahn–Banach Theorem). *Let $p : E \to \mathbb{R}$ be a function on a real linear space E satisfying*

$$p(\lambda x) = \lambda p(x), \quad \forall\ x \in E,\ \forall\ \ \lambda \geq 0,$$
$$p(x + y) \leq p(x) + p(y), \quad \forall\ \ x, y \in E.$$

Let G be a linear subspace of E, and let $g : G \to \mathbb{R}$ be a linear functional such that

$$g(x) \leq p(x), \quad \forall\ \ x \in G.$$

Then a linear functional $f : E \to \mathbb{R}$ exists with the properties

$$f(x) = g(x), \quad \forall\ \ x \in G$$

and

$$f(x) \leq p(x), \quad \forall\ \ x \in E.$$

As an immediate consequence from Theorem 2.17, we obtain the following theorem, which is the Hahn–Banach Theorem for normed linear spaces.

Theorem 2.18. *Let X be a normed linear space. Assume M is a linear subspace of X, and $F : M \to \mathbb{R}$ is a linear functional such that*

$$|F(u)| \leq c\,\|u\| \quad \textit{for all} \;\; u \in M,$$

where c is some positive constant. Then F can be extended to a linear continuous functional $f : X \to \mathbb{R}$ that satisfies

$$|\langle f, u\rangle| \leq c\,\|u\| \quad \textit{for all} \;\; u \in X.$$

First consequences from the Hahn–Banach Theorem are given in the following corollary.

Corollary 2.19. *Let X be a normed linear space.*

(i) *For each given $u_0 \in X$ with $u_0 \neq 0$, a functional $f \in X^*$ exists such that*

$$\langle f, u_0\rangle = \|u_0\| \quad \textit{and} \quad \|f\| = 1.$$

(ii) *For all $u \in X$, one has*

$$\|u\| = \sup_{f \in X^*,\ \|f\| \leq 1} |\langle f, u\rangle|.$$

(iii) *If for $u \in X$ the condition*

$$\langle f, u\rangle = 0 \quad \textit{for all} \;\; f \in X^*$$

holds, then $u = 0$.

We set

$$X^{**} = (X^*)^*,$$

which is called the bidual space and which consists of all linear continuous functionals $F : X^* \to \mathbb{R}$.

Proposition 2.20. *Let X be a normed linear space. The operator $j : X \to X^{**}$ defined by*

$$j(u)(f) = \langle f, u\rangle \quad \textit{for all} \;\; u \in X,\ f \in X^*$$

has the following properties:

(i) *j is linear and*

$$\|j(u)\| = \|u\| \quad \textit{for all} \;\; u \in X.$$

(ii) *$j(X)$ is a closed subspace of X^{**} if and only if X is a Banach space.*

The operator $j : X \to X^{**}$ is called the canonical embedding of X into X^{**}.

Definition 2.21. *A normed linear space X is called reflexive if the canonical embedding $j : X \to X^{**}$ is surjective; i.e., $j(X) = X^{**}$.*

We readily observe that every reflexive normed linear space X is in fact a Banach space, which is isometrically isomorphic to X^{**}, and thus, we may write $X = X^{**}$.

Corollary 2.22. (i) *Each Hilbert space is reflexive.*

(ii) *Every closed linear subspace of a reflexive Banach space X is again reflexive.*

(iii) *The product of a finite number of reflexive Banach spaces is a reflexive Banach space.*

(iv) *Let X and Y be two isomorphic normed linear spaces. If X is a reflexive Banach space, then Y is also a reflexive Banach space.*

(v) *Let X be a Banach space. Then X is reflexive if and only if X^* is reflexive.*

(vi) *If X is a separable and reflexive Banach space, then X^* is separable.*

Next we define the *dual* or *adjoint operator* of a linear operator $A : D(A) \subset X \to Y$, where X and Y are two Banach spaces.

Definition 2.23. *Assume $D(A)$ is dense in X. Then the dual operator*

$$A^* : D(A^*) \subset Y^* \to X^*$$

is defined by the following relation:

$$\langle A^* v, u\rangle = \langle v, Au\rangle \quad \textit{for all} \ \ v \in D(A^*), \quad u \in D(A),$$

where $v \in Y^*$ belongs to $D(A^*)$ if and only if a $w \in X^*$ exists such that

$$\langle w, u\rangle = \langle v, Au\rangle \quad \textit{for all} \ \ u \in D(A).$$

To verify that A^* is well defined, we note first that according to Definition 2.23, an element $v \in Y^*$ belongs to $D(A^*)$ if and only if a $w \in X^*$ exists such that

$$\langle w, u\rangle = \langle v, Au\rangle \quad \text{for all} \ \ u \in D(A).$$

We set $A^*v = w$. As $D(A)$ is dense in X, the element w is uniquely determined by v, and thus, the operator A^* is well defined. Moreover, one readily observes that A^* is linear and graph-closed. In the special case that $D(A) = X$, we have the following results.

Proposition 2.24. *Let X and Y be two Banach spaces, and let $A : X \to Y$ be a linear and continuous operator. Then the dual operator*

$$A^* : Y^* \to X^*$$

is also linear and continuous, and we have

$$\|A^*\| = \|A\|.$$

Moreover, if the linear operator $A : X \to Y$ is compact, then so is the dual operator $A^* : Y^* \to X^*$.

The following facts about the duality of embeddings are important, e.g., for the understanding of the concept of evolution triple, which will be introduced in Sect. 2.4.3.

Proposition 2.25. *Let X and Y be Banach spaces with $X \subset Y$ such that X is dense in Y, and the embedding*

$$i : X \to Y$$

is continuous. Then the following is true:

(i) *The embedding $Y^* \subset X^*$ is continuous, and the embedding operator $\hat{i} : Y^* \to X^*$ is identical with the dual operator of i; i.e., $\hat{i} = i^*$.*
(ii) *If X is, in addition, reflexive, then Y^* is dense in X^*.*
(iii) *If the embedding $X \subset Y$ is compact, then so is the embedding $Y^* \subset X^*$.*

Proof: As for (i), density arguments show that each element of Y^* can be uniquely identified with an element of X^*, and the continuity of the embedding $Y^* \subset X^*$ follows from the continuity of i. The proof of (ii) makes use of the Hahn–Banach Theorem in connection with the reflexivity of X. (see [222, Chap. 18, 21]), and (iii) follows from Proposition 2.24. □

In finite-dimensional Banach spaces, closed and bounded sets are compact. This result is no longer true for infinite-dimensional Banach spaces because of the following famous theorem due to Riesz.

Theorem 2.26 (Riesz' Lemma). *Let X be a normed linear space. Then, the closed unit ball in X is compact if and only if X is finite-dimensional.*

According to Theorem 2.26, in infinite-dimensional Banach spaces, there are bounded sequences that have no convergent subsequence. This lack of compactness in infinite-dimensional spaces is one of the main reasons for many difficulties in the functional analytical treatment of variational problems. To overcome these difficulties, new concepts of convergence (or new topologies) have been introduced with respect to which the unit ball is compact (respectively, sequentially compact).

Definition 2.27. *Let X be a Banach space. A sequence $(u_n) \subset X$ is called weakly convergent in X to an element $u \in X$ iff*

$$\langle f, u_n \rangle \to \langle f, u \rangle \quad \textit{for all} \quad f \in X^*.$$

The weak convergence is denoted by

$$u_n \rightharpoonup u \quad \textit{as} \quad n \to \infty \qquad \textit{or} \qquad \mathrm{w}-\lim_{n\to\infty} u_n = u.$$

Note, in contrast to the weak convergence, we call the usual convergence with respect to the norm ($u_n \to u$) sometimes the strong convergence. The following theorem provides a compactness result with respect to the topology introduced by the weak convergence.

Theorem 2.28 (Eberlein–Smulian Theorem). *Let X be a reflexive Banach space. Then, each bounded sequence $(u_n) \subset X$ has a weakly convergent subsequence.*

A few properties of weak convergence are summarized in the next proposition.

Proposition 2.29. *Let X be Banach spaces, and $(u_n) \subset X$.*

(i) $u_n \to u$ *implies* $u_n \rightharpoonup u$.
(ii) *If X is finite-dimensional, then strong and weak convergence are equivalent.*
(iii) *If $u_n \rightharpoonup u$, then (u_n) is bounded and*

$$\|u\| \leq \liminf_{n\to\infty} \|u_n\|.$$

(iv) *If $u_n \rightharpoonup u$ in X and $f_n \to f$ in X^*, then it follows that*

$$\langle f_n, u_n \rangle \to \langle f, u \rangle.$$

(v) *If $u_n \to u$ in X and $f_n \rightharpoonup f$ in X^*, then it follows that*

$$\langle f_n, u_n \rangle \to \langle f, u \rangle.$$

The reverse of the Eberlein–Smulian Theorem is also true; i.e, a Banach space is reflexive if and only if every bounded sequence has a weakly convergent subsequence. Thus, the compactness result given by Theorem 2.28 is only valid in reflexive Banach spaces. To deal with nonreflexive Banach spaces, the following so-called weak* convergence has been introduced.

Definition 2.30. *Let X be a Banach space. A sequence $(f_n) \subset X^*$ is called weakly* convergent to an element $f \in X^*$ iff*

$$\langle f_n, u \rangle \to \langle f, u \rangle \quad \textit{for all} \ \ u \in X.$$

The weak convergence is denoted by*

$$f_n \rightharpoonup^* f \quad \textit{as} \ \ n \to \infty, \qquad \textit{or} \ \ \mathrm{w}^*\!-\lim_{n\to\infty} f_n = f.$$

Proposition 2.31. *Let X be a Banach space, and let (f_n) be a sequence in the dual space X^*.*

(i) $f_n \to f$ *in* X^* *implies* $f_n \rightharpoonup^* f$.
(ii) *If $f_n \rightharpoonup^* f$, then (f_n) is bounded in X^* and*

$$\|f\| \leq \liminf_{n\to\infty} \|f_n\|.$$

(iii) *If $u_n \to u$ in X and $f_n \rightharpoonup^* f$ in X^*, then it follows that*

$$\langle f_n, u_n \rangle \to \langle f, u \rangle.$$

(iv) $f_n \rightharpoonup f$ *in* X^* *implies* $f_n \rightharpoonup^* f$.
(v) *If* X *is reflexive, then* $f_n \rightharpoonup^* f$ *is equivalent to* $f_n \rightharpoonup f$.

Definition 2.32. *Let* $A : X \to Y$ *be a linear operator, where* X *and* Y *are Banach spaces.* A *is called weakly sequentially continuous iff*

$$u_n \rightharpoonup u \quad \textit{implies} \quad Au_n \rightharpoonup Au.$$

A *is called strongly continuous iff*

$$u_n \rightharpoonup u \quad \textit{implies} \quad Au_n \to Au.$$

A few simple consequences are provided in the next proposition.

Proposition 2.33. *Let* $A : X \to Y$ *be a linear operator, where* X *and* Y *are Banach spaces.*

(i) *If* A *is continuous, then* A *is weakly sequentially continuous.*
(ii) *If* A *is compact, then* A *is strongly continuous.*
(iii) *If* A *is strongly continuous and* X *is reflexive, then* A *is compact.*

2.1.3 Convex Analysis and Calculus in Banach Spaces

Let X be a normed linear space. A subset K of X is convex iff

$$u,\, v \in K \quad \text{implies} \quad tu + (1-t)v \in K \ \text{ for all } \ 0 \le t \le 1.$$

Theorem 2.34. *Let* H *be a Hilbert space with inner product* $(\cdot,\cdot)$, *and let* K *be a nonempty, closed, and convex subset of* H. *Then to each* $u \in H$, *a uniquely defined* $v \in K$ *closest to* u *exists, that is,*

$$v \in K : \quad \|u - v\| = \inf_{w \in K} \|u - w\|.$$

Equivalently, $v \in K$ *is the uniquely defined solution of the variational inequality*

$$v \in K : \quad (u - v, w - v) \le 0 \quad \textit{for all} \ \ w \in K.$$

Consequences of Theorem 2.34 are the well-known Orthogonal Projection Theorem and the Riesz Representation Theorem of linear continuous functionals on Hilbert spaces. The latter implies that a Hilbert space H is isometrically isomorphic with its dual space H^*. A generalization of the Riesz Representation Theorem is the Lax–Milgram Theorem (see Sect. 2.3).

Important consequences of the Hahn–Banach Theorem are various separation theorems, such as the following one.

Theorem 2.35 (Separation Theorem). *Let* X *be a normed linear space, and let* $K \subset X$ *be a closed and convex subset. If* $u_0 \in X \setminus K$, *then a linear continuous functional* $f \in X^*$ *and an* $\alpha \in \mathbb{R}$ *exists such that*

$$\langle f, u \rangle \le \alpha \quad \textit{for all} \ \ u \in K, \quad \textit{and} \ \ \langle f, u_0 \rangle > \alpha.$$

Definition 2.36. *A subset M of a normed linear space X is called weakly sequentially closed if the limit of every weakly convergent sequence $(u_n) \subset M$ belongs to M; i.e.,*

$$(u_n) \subset M \text{ and } u_n \rightharpoonup u \quad \text{implies} \quad u \in M.$$

Simple examples show that, in general, closed sets of a normed linear space need not be weakly sequentially closed. However, by means of Theorem 2.35, one gets the following equivalence.

Proposition 2.37. *Let M be a convex subset of a normed linear space X. Then, M is closed if and only if M is weakly sequentially closed.*

Next we present some convexity and smoothness properties of the norm in Banach spaces that are important for proving existence results for abstract operator equations involving operators of monotone type (see Theorem 2.156 in Sect. 2.4.4).

Definition 2.38. *A Banach space X is called strictly convex if and only if*

$$\|tu + (1-t)v\| < 1 \quad \text{provided that} \quad \|u\| = \|v\| = 1,\ u \neq v, \quad \text{and } 0 < t < 1.$$

A Banach space X is called locally uniformly convex if and only if for each $\varepsilon \in (0, 2]$, and for each $u \in X$ with $\|u\| = 1$, a $\delta(\varepsilon, u) > 0$ exists such that for all v with $\|v\| = 1$ and $\|u - v\| \geq \varepsilon$, the following holds:

$$\frac{1}{2}\|u + v\| \leq 1 - \delta(\varepsilon, u).$$

A Banach space X is called uniformly convex if and only if X is locally uniformly convex and δ can be chosen to be independent of u.

Obviously we have the following implications:

$$\text{uniformly convex} \implies \text{locally uniformly convex} \implies \text{strictly convex}.$$

Example 2.39. Each Hilbert space is uniformly convex. This readily follows from the parallelogram identity

$$\left\|\frac{1}{2}(u - v)\right\|^2 + \left\|\frac{1}{2}(u + v)\right\|^2 = \frac{1}{2}(\|u\|^2 + \|v\|^2).$$

Example 2.40. Let $1 < p < \infty$ and $\Omega \subset \mathbb{R}^N$ be a domain; then from Clarkson's inequality (see Sect. 2.2.4), it follows that $L^p(\Omega)$ is uniformly convex. By using this result, one readily sees that the Sobolev spaces $W^{m,p}(\Omega)$ are uniformly convex too, for $1 < p < \infty$ and $m = 0, 1, \ldots$.

Furthermore, the following theorems hold.

Theorem 2.41 (Milman–Pettis Theorem). *Each uniformly convex Banach space is reflexive.*

Convexity properties of the norm are closely related with smoothness properties of the norm, i.e., the smoothness of the function $u \mapsto \|u\|$.

Theorem 2.42. *Let X be a reflexive Banach space. Then the following holds:*

(i) *If X^* is strictly convex, then the function $u \mapsto \|u\|$ is Gâteaux-differentiable on $X \setminus \{0\}$.*
(ii) *If X^* is locally uniformly convex, then the function $u \mapsto \|u\|$ is Fréchet-differentiable on $X \setminus \{0\}$.*
(iii) (Troyanski) *In every reflexive Banach space X, an equivalent norm can be introduced so that both X and X^* are locally uniformly convex.*

The notions of Gâteaux and Fréchet derivatives that occur in Theorem 2.42 are natural generalizations of the directional and total derivative of functions $f : \mathbb{R}^n \to \mathbb{R}^m$, respectively, to mappings between Banach spaces. In particular, in the calculus of variations, these notions allow us to generalize the classic criteria in the study of extrema for real-valued functions in $\mathbb{R}^n$ to real-valued functionals $F : D(F) \subset X \to \mathbb{R}$ defined on a subset of a Banach space X.

Definition 2.43 (Gâteaux Derivative). *Let X and Y be Banach spaces, and let $f : U \subset X \to Y$ be a map whose domain $D(f) = U$ is an open subset of X. The directional derivative of f at $u \in U$ in the direction $h \in X$ is given by*

$$\delta f(u; h) = \lim_{t \to 0} \frac{f(u + th) - f(u)}{t}$$

provided this limit exists. If $\delta f(u; h)$ exists for every $h \in X$, and if the mapping $D_G f(u) : X \to Y$ defined by

$$D_G f(u)h = \delta f(u; h)$$

is linear and continuous, then we say that f is Gâteaux-differentiable at u, and we call $D_G f(u)$ the Gâteaux derivative of f at u.

Definition 2.44 (Fréchet Derivative). *Let X and Y be Banach spaces, and let $f : U \subset X \to Y$, where the domain $D(f) = U$ is an open subset of X. Then f is called Fréchet-differentiable at $u \in U$ if and only if a linear and continuous mapping $A : X \to Y$ exists such that*

$$\lim_{\|h\| \to 0} \frac{\|f(u + h) - f(u) - Ah\|}{\|h\|} = 0$$

or equivalently

$$f(u + h) - f(u) = Ah + o(\|h\|), \quad (h \to 0).$$

If such a mapping A exists, then we call $D_F f(u) = A$ (or simply $f'(u) = A$) the Fréchet derivative of f at u.

Corollary 2.45. *Let X and Y be Banach spaces, and let $f : U \subset X \to Y$. Then the following relations between Gâteaux and Fréchet derivative hold:*

(i) *If f is Fréchet-differentiable at $u \in U$, then f is Gâteaux-differentiable at u.*
(ii) *If f is Gâteaux-differentiable in a neighborhood of u_0 and $D_G f$ is continuous at u_0, then f is Fréchet-differentiable at u_0 and $f'(u_0) = D_G f(u_0)$.*

Remark 2.46. If $f : U \subset X \to Y$ is Fréchet-differentiable in U and $f' : U \to L(X,Y)$ is continuous, then we write $f \in C^1(U;Y)$ or simply $f \in C^1(U)$ if $Y = \mathbb{R}$. In a similar way as for mappings from $\mathbb{R}^n$ into $\mathbb{R}^m$, one can prove chain rules for both the Fréchet and the Gâteaux derivative.

Example 2.47. Let $X = L^p(\Omega)$, where $1 < p < \infty$. We will compute the Gâteaux derivative of the pth power L^p-norm, i.e., of the function $f : X \to \mathbb{R}$ defined by

$$f(u) = \|u\|^p_{L^p(\Omega)}.$$

After elementary calculations, we get

$$D_G f(u)h = \delta f(u;h) = p \int_\Omega |u|^{p-2} uh \, dx$$

if we consider real-valued functions $u : \Omega \to \mathbb{R}$. In case the functions are complex-valued, we get

$$\delta f(u;h) = \frac{p}{2} \int_\Omega |u|^{p-2} (\bar{u}h + u\bar{h}) \, dx.$$

We introduce next the notions of convex and semicontinuous functions (or functionals).

Definition 2.48 (Semicontinuous, Convex Functionals). *Let X be a Banach space and $\phi : M \subset X \to [-\infty, \infty]$ with $M = D(\phi)$.*

(i) *The functional ϕ is called sequentially lower semicontinuous at $u \in M$ if and only if*

$$\phi(u) \le \liminf_{n\to\infty} \phi(u_n) \tag{2.1}$$

holds for each sequence $(u_n) \subset M$ such that $u_n \to u$ as $n \to \infty$.

(ii) *The functional ϕ is called lower semicontinuous if and only if the set M_r is closed relative to M for all $r \in \mathbb{R}$, where*

$$M_r = \{u \in M : \phi(u) \le r\}.$$

(iii) *The functional ϕ is called weak sequentially lower semicontinuous at $u \in M$ if and only if (2.1) holds for each weakly convergent sequence (u_n) to u, i.e., $u_n \rightharpoonup u$.*

(iv) *The functional ϕ is called sequentially upper semicontinuous (respectively, weak sequentially upper semicontinuous, upper semicontinuous) if and only if $-\phi$ is sequentially lower semicontinuous (respectively, weak sequentially lower semicontinuous, lower semicontinuous).*

(v) *The functional ϕ is called convex if and only if M is convex and*

$$\phi(tu + (1-t)v) \leq t\phi(u) + (1-t)\phi(v), \quad 0 \leq t \leq 1, \tag{2.2}$$

for all u, $v \in M$ for which the right-hand side of (2.2) is meaningful; ϕ is called strictly convex if and only if for all t with $0 < t < 1$ and for all u, $v \in M$ with $u \neq v$ inequality (2.2) holds strictly; i.e., (2.2) holds with $\leq$ replaced by $<$.

The following proposition provides the connection between the above notions.

Proposition 2.49. *Let X be a Banach space and $\phi : M \subset X \to [-\infty, \infty]$ with $M = D(\phi)$.*

(i) *ϕ is sequentially lower semicontinuous on M if and only if ϕ is lower semicontinuous on M.*

(ii) *Assume $u \in M$ with $\phi(u) \neq \pm\infty$. Then ϕ is sequentially lower semicontinuous at u if and only if, for each $\varepsilon > 0$, a $\delta(\varepsilon) > 0$ exists such that for all $v \in M$ with*

$$\|v - u\| < \delta(\varepsilon) \quad \textit{implies} \quad \phi(u) < \phi(v) + \varepsilon.$$

(iii) *ϕ is continuous if and only if ϕ is both lower and upper semicontinuous.*

(iv) *If, in addition, M is closed and convex, and ϕ is convex, then lower semicontinuous, sequentially lower semicontinuous and weak sequentially lower semicontinuous are mutually equivalent.*

Let X be a Banach space. In what follows we consider only convex functionals $\phi : X \to \mathbb{R} \cup \{+\infty\}$; i.e., we do not allow "$-\infty$" as a value for the convex functional ϕ. The reason is that if $\phi(u_0) = -\infty$ at some point u_0 and if, in addition, ϕ is lower semicontinuous, then ϕ would be nowhere finite. This can readily be seen by the following arguments. Assume there is some $u \in X$ with $\phi(u) \in \mathbb{R}$. Then from the convexity we get for all $t \in (0,1)$, $\phi(tu_0 + (1-t)u) = -\infty$. Taking the limit $t \to 0$, the lower semicontinuity yields $\phi(u) = -\infty$, a contradiction.

Definition 2.50. *Let X be a Banach space and $\phi : X \to \mathbb{R} \cup \{+\infty\}$ be a convex functional.*

(i) *The effective domain of ϕ is the set $\mathrm{dom}(\phi)$ defined by*

$$\mathrm{dom}(\phi) = \{u \in X : \phi(u) < +\infty\}.$$

(ii) *ϕ is said to be proper if $\mathrm{dom}(\phi) \neq \emptyset$.*

(iii) *The epigraph of* ϕ*, denoted by* $\mathrm{epi}(\phi)$*, is given by*

$$\mathrm{epi}(\phi) = \{(u, \lambda) \in X \times \mathbb{R} : \phi(u) \leq \lambda\}.$$

We summarize some elementary properties of convex functionals as follows.

Corollary 2.51. *Let* X *be a Banach space, and let* $\phi, \phi_i : X \to \mathbb{R} \cup \{+\infty\}$*,* $i = 1, 2$*, be convex functionals. Then the following holds:*

(i) $\mathrm{dom}(\phi)$ *is convex.*
(ii) *If* $\lambda \geq 0$*, then* $\lambda\phi$ *is convex.*
(iii) *If* ϕ_1 *and* ϕ_2 *are convex, then* $\phi_1 + \phi_2$ *is convex.*
(iv) ϕ *is convex, proper, and lower semicontinuous if and only if* $\mathrm{epi}(\phi)$ *is, respectively, convex, nonempty, and closed in* $X \times \mathbb{R}$*.*

Proposition 2.52. *Let* X *be a Banach space, and let* $\phi : X \to \mathbb{R} \cup \{+\infty\}$ *be a convex, proper, and lower semicontinuous functional. Then* ϕ *is locally Lipschitz on the interior of* $\mathrm{dom}(\phi)$*.*

Theorem 2.53 (Weierstrass' Theorem). *Let* X *be a reflexive Banach space. If* $\phi : X \to \mathbb{R} \cup \{+\infty\}$ *is a convex, proper, and lower semicontinuous functional satisfying*

$$\lim_{\|u\| \to \infty} \phi(u) = +\infty,$$

then the problem

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

admits at least one solution.

The following notion of subgradient generalizes the classic concept of a derivative.

Definition 2.54 (Subdifferential). *Let* X *be a Banach space, and let* $\phi : X \to \mathbb{R} \cup \{+\infty\}$ *be a convex and proper functional. An element* $u^* \in X^*$ *is called a subgradient of* ϕ *at* $u \in \mathrm{dom}(\phi)$ *if and only if the following inequality holds:*

$$\phi(v) \geq \phi(u) + \langle u^*, v - u \rangle \quad \text{for all } \; v \in X. \tag{2.3}$$

The set of all $u^* \in X^*$ *satisfying (2.3) is called the subdifferential of* ϕ *at* $u \in \mathrm{dom}(\phi)$*, and is denoted by* $\partial\phi(u)$*.*

First properties of the subdifferential are given in the following proposition.

Proposition 2.55. *Let* X *be a Banach space, and let* $\phi : X \to \mathbb{R} \cup \{+\infty\}$ *be a convex and proper functional. Then we have the following properties of* $\partial\phi$*:*

(i) $\partial\phi(u)$ *is convex and weak***-closed.*
(ii) *If* ϕ *is continuous at* $u \in \mathrm{dom}(\phi)$, *then* $\partial\phi(u)$ *is nonempty, convex, bounded, and weak***-compact.*

Note, in (i) of Proposition 2.55 $\partial\phi(u) = \emptyset$ is possible.

Proposition 2.56. *Let* X *be a Banach space, and let* $\phi : X \to \mathbb{R} \cup \{+\infty\}$ *be a convex and proper functional. If* ϕ *is Gâteaux-differentiable at* $u \in \mathrm{int}(\mathrm{dom}(\phi))$, *then* $\partial\phi(u) = \{D_G\phi(u)\}$. *If* ϕ *is continuous at* u *and* $\partial\phi(u)$ *is a singleton, then* ϕ *is Gâteaux-differentiable at* u.

The following sum rule for the subdifferential is due to Moreau and Rockafellar.

Proposition 2.57 (Sum Rule). *Let* X *be a Banach space, and let* $\phi_1, \phi_2 : X \to \mathbb{R} \cup \{+\infty\}$ *be convex functionals. If there is a point* $u_0 \in \mathrm{dom}(\phi_1) \cap \mathrm{dom}(\phi_2)$ *at which* ϕ_1 *is continuous, then the following holds:*

$$\partial(\phi_1 + \phi_2)(u) = \partial\phi_1(u) + \partial\phi_2(u) \quad \textit{for all} \ \ u \in X.$$

Example 2.58. Let $f : \mathbb{R} \to \mathbb{R}$ be a nondecreasing function with its one-sided limits $\underline{f}$ and $\bar{f}$. Define $\phi : \mathbb{R} \to \mathbb{R}$ by

$$\phi(x) = \int_{x_0}^{x} \underline{f}(s)\, ds = \int_{x_0}^{x} \bar{f}(s)\, ds.$$

Note that ϕ is convex and finite on $\mathbb{R}$, i.e., $\mathrm{dom}(\phi) = \mathbb{R}$, and thus ϕ is even locally Lipschitz. Elementary calculations show that the subdifferential is given by

$$\partial\phi(x) = [\underline{f}(x), \bar{f}(x)].$$

Example 2.59. Let $\phi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ be a convex, proper, lower semicontinuous function, and $\Omega \subset \mathbb{R}^N$ a Lebesgue-measurable set such that either $0 = \phi(0) = \min_{s\in\mathbb{R}} \phi(s)$ or the measurable set Ω has finite measure. Define $\Phi : L^p(\Omega) \to \mathbb{R} \cup \{+\infty\}$, $1 < p < \infty$, by

$$\Phi(u) = \int_{\Omega} \phi(u(x))\, dx \quad \text{if} \ \ \phi(u) \in L^1(\Omega), \quad +\infty \ \ \text{otherwise}.$$

Then $\Phi : L^p(\Omega) \to \mathbb{R} \cup \{+\infty\}$ is convex, proper, lower semicontinuous, and $u^* \in \partial\Phi(u)$ if and only if

$$u^* \in L^q(\Omega), \quad \text{and} \ \ u^*(x) \in \partial\phi(u(x)), \quad \text{for a.e. } x \in \Omega,$$

where q is the Hölder conjugate; i.e., $1/p + 1/q = 1$.

2.1.4 Partially Ordered Sets

Definition 2.60 (Partially Ordered Set). *Let P be a nonempty set. We say that a relation $x \le y$ between certain pairs of elements of P is a partial ordering in P, and that $(P, \le)$ is a partially ordered set, if "$\le$" has the following properties:*

(i) *$x \le x$ for all $x \in P$ (reflexivity).*
(ii) *If $x \le y$ and $y \le x$, then $x = y$ (antisymmetry).*
(iii) *If $x \le y$ and $y \le z$, then $x \le z$ (transitivity).*

Note that $x < y$ stands for $x \le y$ and $x \ne y$. Next we define several notions based on the partial ordering introduced above.

Definition 2.61. *Let $(P, \le)$ be a partially ordered set.*

(i) *An element b of P is called an upper bound of a subset A of P if $x \le b$ for each $x \in A$. If $b \in A$, we say that b is the greatest element of A. A lower bound of A and the smallest element of A are defined similarly, replacing $x \le b$ above by $b \le x$.*
(ii) *If the set of all upper bounds of A has the minimum, we call it a least upper bound of A and denote it by* $\sup A$. *The greatest lower bound,* $\inf A$, *of A is defined similarly.*
(iii) *An element $x \in A$ is called a maximal element of $A \subset P$, if there is no $y \ne x$ in A for which $x \le y$. Similarly, a minimal element of A is defined. Obviously, every greatest element of A is a maximal element of A.*
(iv) *We say that a partially ordered set P is a lattice if* $\inf\{x, y\}$ *and* $\sup\{x, y\}$ *exist for all $x, y \in P$.*
(v) *A subset C of P is said to be upward directed if for each pair $x, y \in C$ there is a $z \in C$ such that $x \le z$ and $y \le z$, and C is downward directed if for each pair $x, y \in C$ there is a $w \in C$ such that $w \le x$ and $w \le y$. If C is both upward and downward directed, it is called directed.*
(vi) *A subset C of a partially ordered set P is called a chain if $x \le y$ or $y \le x$ for all $x, y \in C$.*
(vii) *We say that C is well ordered if each nonempty subset of C has a minimum, and inversely well ordered if each nonempty subset of C has a maximum. Obviously, each (inversely) well-ordered set is a chain and each chain is directed.*

Theorem 2.62 (Zorn's Lemma). *If in a partially ordered set P, every chain has an upper bound, then P possesses a maximal element.*

2.2 Sobolev Spaces

In this section, we summarize the main properties of Sobolev spaces. These properties include, e.g., the approximation of Sobolev functions by smooth functions (density theorems), continuity properties and compactness conditions (embedding theorems), the definition of the boundary values of Sobolev functions (trace theorem), and calculus for Sobolev functions (chain rule).

2.2.1 Spaces of Lebesgue Integrable Functions

Let $\mathbb{R}^N$, $N \geq 1$, be equipped with the Lebesgue measure, and let $\Omega \subset \mathbb{R}^N$ be a domain; i.e., Ω is an open and connected subset of $\mathbb{R}^N$. For $1 \leq p < \infty$, we denote by $L^p(\Omega)$ the Banach space of measurable functions $u : \Omega \to \mathbb{R}$ with respect to the norm

$$\|u\|_{L^p(\Omega)} = \left(\int_\Omega |u|^p dx \right)^{1/p} < \infty.$$

For a measurable function u, we put

$$\|u\|_{L^\infty(\Omega)} = \inf\{\alpha \in \mathbb{R} : \ \mathrm{meas}\,(\{x \in \Omega : |u(x)| > \alpha\}) = 0\}.$$

We denote by $L^\infty(\Omega)$ the Banach space of all measurable functions f satisfying $\|u\|_{L^\infty(\Omega)} < \infty$.

We also introduce the *local* L^p-spaces, denoted by $L^p_{\mathrm{loc}}(\Omega)$. A function u belongs to $L^p_{\mathrm{loc}}(\Omega)$ if it is measurable and

$$\int_K |u|^p \, dx < \infty$$

for every compact subset K of Ω.

The following main theorems can be found in standard textbooks on real analysis and measure theory (see [201, 114]).

Theorem 2.63 (Lebesgue's Dominated Convergence Theorem). *Suppose (u_n) is a sequence in $L^1(\Omega)$ such that*

$$u(x) = \lim_{n\to\infty} u_n(x)$$

exists almost everywhere (a.e.) on Ω. If there is a function $g \in L^1(\Omega)$ such that, for a.e. $x \in \Omega$, and for all $n = 1, 2, \dots,$

$$|u_n(x)| \leq g(x)$$

then $u \in L^1(\Omega)$ and

$$\lim_{n\to\infty} \int_\Omega |u_n - u| \, dx = 0.$$

In some sense the following reverse statement of Theorem 2.63 holds.

Theorem 2.64. *Let* $u_n,\, u \in L^1(\Omega)$, $n \in \mathbb{N}$, *be such that*

$$\lim_{n\to\infty} \int_\Omega |u_n - u|\, dx = 0.$$

Then a subsequence (u_{n_k}) *of* (u_n) *exists with*

$$u_{n_k}(x) \to u(x) \quad \textit{for a.e.} \quad x \in \Omega.$$

Theorem 2.65 (Fatou's Lemma). *Let* (u_n) *be a sequence of measurable functions, and let* $g \in L^1(\Omega)$. *If*

$$u_n \geq g \quad \textit{a.e. on} \quad \Omega,$$

then we have

$$\int_\Omega \liminf_{n\to\infty} u_n\, dx \leq \liminf_{n\to\infty} \int_\Omega u_n\, dx.$$

If $\Omega \subset \mathbb{R}^N$ is a measurable subset, we denote its Lebesgue measure by

$$\mathrm{meas}(\Omega) = |\Omega|.$$

Theorem 2.66 (Egorov's Theorem). *Let* (u_n), u *be measurable functions, and*

$$u_n \to u \quad \textit{a.e. on } \Omega,$$

where $\Omega \subset \mathbb{R}^N$ *is measurable with* $|\Omega| < \infty$. *Then for each* $\varepsilon > 0$, *a measurable subset* $E \subset \Omega$ *exists such that*

(i) $|\Omega \setminus E| < \varepsilon$.
(ii) $u_n \to u \quad$ *uniformly on* E.

A characterization of the dual spaces of $L^p(\Omega)$ is given in the next theorem.

Theorem 2.67 (Dual Space). *Let* $\Omega \subset \mathbb{R}^N$ *be a domain, and let* Φ *be a linear continuous functional on* $L^p(\Omega)$, $1 < p < \infty$. *Then a uniquely defined function* $g \in L^q(\Omega)$ *exists with* q *satisfying* $1/p + 1/q = 1$ *such that*

$$\langle \Phi, u\rangle = \int_\Omega g\, u\, dx \quad \textit{for all} \quad u \in L^p(\Omega)$$

and

$$\|\Phi\|_{(L^p(\Omega))^*} = \|g\|_{L^q(\Omega)}.$$

If Φ *is a linear continuous functional on* $L^1(\Omega)$, *then a uniquely defined function* $g \in L^\infty(\Omega)$ *exists such that*

$$\langle \Phi, u\rangle = \int_\Omega g\, u\, dx \quad \textit{for all} \quad u \in L^1(\Omega)$$

and

$$\|\Phi\|_{(L^1(\Omega))^*} = \|g\|_{L^\infty(\Omega)}.$$

In view of Theorem 2.67, the dual space of $L^p(\Omega)$ is isometrically isomorphic to $L^q(\Omega)$ for $1 \le p < \infty$ with $q = \infty$ if $p = 1$.

We summarize some important properties of L^p-spaces in the following theorem.

Theorem 2.68. *Let $\Omega \subset \mathbb{R}^N$ be a domain.*

(i) *For $1 \le p < \infty$, the spaces $L^p(\Omega)$ are separable.*
(ii) *$L^\infty(\Omega)$ is not separable.*
(iii) *For $1 < p < \infty$, the spaces $L^p(\Omega)$ are reflexive.*
(iv) *$L^1(\Omega)$ and $L^\infty(\Omega)$ are not reflexive.*
(v) *For $1 < p < \infty$, the spaces $L^p(\Omega)$ are uniformly convex.*

2.2.2 Definition of Sobolev Spaces

Let $\alpha = (\alpha_1, \dots, \alpha_N)$ with nonnegative integers $\alpha_1, \dots, \alpha_N$ be a multi-index, and denote its order by $|\alpha| = \alpha_1 + \cdots + \alpha_N$. Set $D_i = \partial/\partial x_i$, $i = 1, \dots, N$, and $D^\alpha u = D_1^{\alpha_1} \cdots D_N^{\alpha_N} u$, with $D^0 u = u$. Let Ω be a domain in $\mathbb{R}^N$ with $N \ge 1$. Then $w \in L^1_{\mathrm{loc}}(\Omega)$ is called the α^{th} *weak* or *generalized derivative* of $u \in L^1_{\mathrm{loc}}(\Omega)$ if and only if

$$\int_\Omega u D^\alpha \varphi \, dx = (-1)^{|\alpha|} \int_\Omega w \varphi \, dx, \quad \text{for all } \varphi \in C_0^\infty(\Omega),$$

holds, where $C_0^\infty(\Omega)$ denotes the space of infinitely differentiable functions with compact support in Ω. The generalized derivative w denoted by $w = D^\alpha u$ is unique up to a change of the values of w on a set of Lebesgue measure zero.

Definition 2.69. *Let $1 \le p \le \infty$ and $m = 0, 1, 2, \dots$. The Sobolev space $W^{m,p}(\Omega)$ is the space of all functions $u \in L^p(\Omega)$, which have generalized derivatives up to order m such that $D^\alpha u \in L^p(\Omega)$ for all α: $|\alpha| \le m$. For $m = 0$, we set $W^{0,p}(\Omega) = L^p(\Omega)$.*

With the corresponding norms given by

$$\|u\|_{W^{m,p}(\Omega)} = \left(\sum_{|\alpha| \le m} \|D^\alpha u\|^p_{L^p(\Omega)} \right)^{1/p}, \quad 1 \le p < \infty,$$

$$\|u\|_{W^{m,\infty}(\Omega)} = \max_{|\alpha| \le m} \|D^\alpha u\|_{L^\infty(\Omega)},$$

$W^{m,p}(\Omega)$ becomes a Banach space.

Definition 2.70. *$W_0^{m,p}(\Omega)$ is the closure of $C_0^\infty(\Omega)$ in $W^{m,p}(\Omega)$.*

$W_0^{m,p}(\Omega)$ is a Banach space with the norm $\| \cdot \|_{W^{m,p}(\Omega)}$.

Before we summarize some basic properties of Sobolev spaces, we need to classify the regularity of boundaries.

Definition 2.71. *Let $\Omega \subset \mathbb{R}^N$ be a bounded domain, with boundary $\partial\Omega$. We say that the boundary $\partial\Omega$ is of class $C^{k,\lambda}$, $k \in \mathbb{N}_0$, $\lambda \in (0,1]$, if there are $m \in \mathbb{N}$ Cartesian coordinate systems C_j, $j = 1, \ldots, m$,*

$$C_j = (x_{j,1}, \ldots, x_{j,N-1}, x_{j,N}) = (x'_j, x_{j,N})$$

and real numbers α, $\beta > 0$, as well as m functions a_j with

$$a_j \in C^{k,\lambda}([-\alpha,\alpha]^{N-1}), \quad j = 1, \ldots, m,$$

such that the sets defined by

$$\begin{aligned}
\Lambda^j &= \{(x'_j, x_{j,N}) \in \mathbb{R}^N : |x'_j| \le \alpha,\ x_{j,N} = a_j(x'_j)\},\\
V^j_+ &= \{(x'_j, x_{j,N}) \in \mathbb{R}^N : |x'_j| \le \alpha,\ a_j(x'_j) < x_{j,N} < a_j(x'_j) + \beta\},\\
V^j_- &= \{(x'_j, x_{j,N}) \in \mathbb{R}^N : |x'_j| \le \alpha,\ a_j(x'_j) - \beta < x_{j,N} < a_j(x'_j)\},
\end{aligned}$$

possess the following properties:

$$\Lambda^j \subset \partial\Omega, \quad V^j_+ \subset \Omega, \quad V^j_- \subset \mathbb{R}^N \setminus \Omega, \quad j = 1, \ldots, m,$$

and

$$\bigcup_{j=1}^m \Lambda^j = \partial\Omega.$$

Remark 2.72. If $\partial\Omega \in C^{0,1}$, then we call $\partial\Omega$ a *Lipschitz boundary*, which means that $\partial\Omega$ is locally the graph of a Lipschitz continuous function. In this case, the $(N-1)$-dimensional surface measure is well defined, on the basis of which $L^p(\partial\Omega)$-spaces can be introduced (see [66]). As Lipschitz continuous functions admit a.e. a gradient, the outer unit normal on $\partial\Omega$ exists for a.a. $x \in \partial\Omega$ (see [94]), which allows us to extend the integration by parts formula to Sobolev functions on Lipschitz domains.

Theorem 2.73. *Let $\Omega \subset \mathbb{R}^N$ be a bounded domain, $N \ge 1$. Then we have the following:*

(i) *$W^{m,p}(\Omega)$ is separable for $1 \le p < \infty$.*
(ii) *$W^{m,p}(\Omega)$ is reflexive for $1 < p < \infty$.*
(iii) *Let $1 \le p < \infty$. Then $C^\infty(\Omega) \cap W^{m,p}(\Omega)$ is dense in $W^{m,p}(\Omega)$, and if $\partial\Omega$ is a Lipschitz boundary, then $C^\infty(\overline{\Omega})$ is dense in $W^{m,p}(\Omega)$, where $C^\infty(\Omega)$ and $C^\infty(\overline{\Omega})$ are the spaces of infinitely differentiable functions in Ω and $\overline{\Omega}$, respectively (cf. [99]).*

As for the proofs of these properties we refer to [99].

Now we state some Sobolev embedding theorems. Let X, Y be two normed linear spaces with $X \subset Y$. We recall the operator $i : X \to Y$ defined by $i(u) = u$ for all $u \in X$ is called the embedding operator of X into Y. We say X is continuously (compactly) embedded in Y if $X \subset Y$ and the embedding operator $i : X \to Y$ is continuous (compact).

Theorem 2.74 (Sobolev Embedding Theorem). *Let $\Omega \subset \mathbb{R}^N$, $N \geq 1$, be a bounded domain with Lipschitz boundary $\partial\Omega$. Then the following holds:*

(i) *If $mp < N$, then the space $W^{m,p}(\Omega)$ is continuously embedded in $L^{p^*}(\Omega)$, $p^* = Np/(N - mp)$, and compactly embedded in $L^q(\Omega)$ for any q with $1 \leq q < p^*$.*
(ii) *If $0 \leq k < m - \frac{N}{p} < k + 1$, then the space $W^{m,p}(\Omega)$ is continuously embedded in $C^{k,\lambda}(\overline{\Omega})$, $\lambda = m - \frac{N}{p} - k$, and compactly embedded in $C^{k,\lambda'}(\overline{\Omega})$ for any $\lambda' < \lambda$.*
(iii) *Let $1 \leq p < \infty$, then the embeddings*

$$L^p(\Omega) \supset W^{1,p}(\Omega) \supset W^{2,p}(\Omega) \supset \cdots$$

are compact.

Here $C^{k,\lambda}(\overline{\Omega})$ denotes the *Hölder space*; cf. [99]. As for the proofs we refer to, e.g., [99, 222].

The proper definition of boundary values for Sobolev functions is based on the following theorem.

Theorem 2.75 (Trace Theorem). *Let $\Omega \subset \mathbb{R}^N$ be a bounded domain with Lipschitz ($C^{0,1}$) boundary $\partial\Omega$, $N \geq 1$, and $1 \leq p < \infty$. Then exactly one continuous linear operator exists*

$$\gamma : W^{1,p}(\Omega) \to L^p(\partial\Omega)$$

such that:

(i) $\gamma(u) = u|_{\partial\Omega}$ *if* $u \in C^1(\overline{\Omega})$.
(ii) $\|\gamma(u)\|_{L^p(\partial\Omega)} \leq C\ \|u\|_{W^{1,p}(\Omega)}$ *with C depending only on p and Ω.*
(iii) *If $u \in W^{1,p}(\Omega)$, then $\gamma(u) = 0$ in $L^p(\partial\Omega)$ if and only if $u \in W_0^{1,p}(\Omega)$.*

Definition 2.76 (Trace). *We call $\gamma(u)$ the trace (or generalized boundary function) of u on $\partial\Omega$.*

Remark 2.77. We note that the trace operator

$$\gamma : W^{1,p}(\Omega) \to L^p(\partial\Omega)$$

in Theorem 2.75 is not surjective; i.e., there are functions $\varphi \in L^p(\partial\Omega)$ that are not the traces of functions u from $W^{1,p}(\Omega)$. To describe precisely the range of the trace operator, Sobolev spaces of fractional order, usually referred to as Sobolev–Slobodeckij spaces, have to be taken into account (see [90, 132, 213, 219]). From [132, Theorem 6.8.13, Theorem 6.9.2], we obtain the following result.

Theorem 2.78. *Let $\Omega \subset \mathbb{R}^N$ be a bounded domain with Lipschitz boundary $\partial\Omega$, $N \geq 1$, and $1 < p < \infty$. Then*

$$\gamma(W^{1,p}(\Omega)) = W^{1-\frac{1}{p},p}(\partial\Omega).$$

The following compactness result of the trace operator holds (see [132]).

Theorem 2.79. *Let $\Omega \subset \mathbb{R}^N$ be a bounded domain with Lipschitz boundary $\partial\Omega$, $N \geq 1$.*

(i) *If $1 < p < N$, then*

$$\gamma : W^{1,p}(\Omega) \to L^q(\partial\Omega)$$

is completely continuous for any q with $1 \leq q < (Np - p)/(N - p)$.

(ii) *If $p \geq N$, then for any $q \geq 1$,*

$$\gamma : W^{1,p}(\Omega) \to L^q(\partial\Omega)$$

is completely continuous.

Sobolev–Slobodeckij spaces form a scale of continuous and even compact embeddings with respect to their fractional order of regularity. More precisely, we can deduce the following compact embedding result for the spaces $W^{l,2}(\Omega)$ with $l \in \mathbb{R}_+$ from [219, Theorem 7.9, Theorem 7.10].

Theorem 2.80. *Let $\Omega \subset \mathbb{R}^N$ be a bounded domain with Lipschitz boundary $\partial\Omega$, $N \geq 1$, and let $l_2 < l_1 \leq 1$, where $l_1, l_2 \in \mathbb{R}_+$. Then the embedding*

$$W^{l_1,2}(\Omega) \subset W^{l_2,2}(\Omega)$$

is compact.

If M is a $C^{k,\kappa}$-manifold ($C^{0,1}$ stands for Lipschitz-manifold) and $l_2 < l_1 < k + \kappa$ with $l_1, l_2 \in \mathbb{R}_+$ (for l_1 integer, $l_1 = k + \kappa$ is admissible), then the embedding

$$W^{l_1,2}(M) \subset W^{l_2,2}(M)$$

is compact.

In a similar way as for Sobolev spaces we have the following trace theorem, which can be deduced from [219, Theorem 8.7].

Theorem 2.81 (Trace Theorem). *Let $\Omega \subset \mathbb{R}^N$ be a bounded domain with Lipschitz boundary $\partial\Omega$, $N \geq 1$, and let $1/2 < l \leq 1$ with $l \in \mathbb{R}_+$. Then a uniquely defined continuous linear operator exists*

$$\gamma : W^{l,2}(\Omega) \to W^{l-1/2,2}(\partial\Omega)$$

such that

$$\gamma(u) = u|_{\partial\Omega} \quad \text{if } \ u \in C^1(\overline{\Omega}).$$

Theorem 2.80 and Theorem 2.81 hold likewise in the general case of the spaces $W^{l,p}(\Omega)$ with $l \in \mathbb{R}_+$, $1 < p < \infty$, and can be found, e.g., in [90, 132, 212, 213, 219].

The following extension result is useful in the study of unbounded domain problems.

Lemma 2.82. *Let $\Omega_0 \subset\subset \Omega$, that is, Ω_0 is compactly contained in Ω. Assume $g \in W^{1,p}(\Omega)$, $u \in W^{1,p}(\Omega_0)$, and $u - g \in W_0^{1,p}(\Omega_0)$, $1 \le p < \infty$. Then the function w defined by*

$$w(x) = \begin{cases} u(x) & \text{if} \quad x \in \Omega_0\,, \\ g(x) & \text{if} \quad x \in \Omega \setminus \Omega_0 \end{cases}$$

is in $W^{1,p}(\Omega)$, and its generalized derivative $D_i w = \partial w/\partial x_i$, $i = 1, \ldots, N$, is given by

$$D_i w(x) = \begin{cases} D_i u(x) & \text{if} \quad x \in \Omega_0\,, \\ D_i g(x) & \text{if} \quad x \in \Omega \setminus \Omega_0\,. \end{cases}$$

For the proof of Lemma 2.82, see [120, Lemma 20.14]. Its proof is based on the density property (iii) of Theorem 2.73 and the characterization of the traces of $W_0^{1,p}(\Omega)$ function.

2.2.3 Chain Rule and Lattice Structure

In this section, we assume that $\Omega \subset \mathbb{R}^N$ is a bounded domain with Lipschitz boundary $\partial\Omega$.

Lemma 2.83 (Chain Rule). *Let $f \in C^1(\mathbb{R})$ and $\sup_{s\in\mathbb{R}} |f'(s)| < \infty$. Let $1 \le p < \infty$ and $u \in W^{1,p}(\Omega)$. Then the composite function $f \circ u \in W^{1,p}(\Omega)$, and its generalized derivatives are given by*

$$D_i(f \circ u) = (f' \circ u) D_i u, \quad i = 1, \ldots, N.$$

Lemma 2.84 (Generalized Chain Rule). *Let $f : \mathbb{R} \to \mathbb{R}$ be continuous and piecewise continuously differentiable with $\sup_{s\in\mathbb{R}} |f'(s)| < \infty$, and $u \in W^{1,p}(\Omega)$, $1 \le p < \infty$. Then $f \circ u \in W^{1,p}(\Omega)$, and its generalized derivative is given by*

$$D_i(f \circ u)(x) = \begin{cases} f'(u(x)) D_i u(x) & \text{if } f \text{ is differentiable at } u(x)\,, \\ 0 & \text{otherwise.} \end{cases}$$

The chain rule may further be extended to Lipschitz continuous f; see [99, 222].

Lemma 2.85 (Generalized Chain Rule). *Let $f : \mathbb{R} \to \mathbb{R}$ be a Lipschitz continuous function and $u \in W^{1,p}(\Omega)$, $1 \le p < \infty$. Then $f \circ u \in W^{1,p}(\Omega)$, and its generalized derivative is given by*

$$D_i(f \circ u)(x) = f_B(u(x)) D_i u(x) \quad \text{for a.e.} \;\; x \in \Omega,$$

where $f_B : \mathbb{R} \to \mathbb{R}$ is a Borel-measurable function such that $f_B = f'$ a.e. in $\mathbb{R}$.

The generalized derivative of the following special functions are frequently used in later chapters.

Example 2.86. Let $1 \le p < \infty$ and $u \in W^{1,p}(\Omega)$. Then $u^+ = \max\{u, 0\}$, $u^- = \max\{-u, 0\}$, and $|u|$ are in $W^{1,p}(\Omega)$, and their generalized derivatives are given by

$$(D_i u^+)(x) = \begin{cases} D_i u(x) & \text{if} \quad u(x) > 0\,, \\ 0 & \text{if} \quad u(x) \le 0\,, \end{cases}$$

$$(D_i u^-)(x) = \begin{cases} 0 & \text{if} \quad u(x) \ge 0\,, \\ -D_i u(x) & \text{if} \quad u(x) < 0\,, \end{cases}$$

$$(D_i |u|)(x) = \begin{cases} D_i u(x) & \text{if} \quad u(x) > 0\,, \\ 0 & \text{if} \quad u(x) = 0\,, \\ -D_i u(x) & \text{if} \quad u(x) < 0\,. \end{cases}$$

As for the traces of u^+ and u^-, we have (cf. [66])

$$\gamma(u^+) = (\gamma(u))^+, \quad \gamma(u^-) = (\gamma(u))^-.$$

Lemma 2.87 (Lattice Structure). *Let $u,\ v \in W^{1,p}(\Omega),\ 1 \le p < \infty$. Then* $\max\{u, v\}$ *and* $\min\{u, v\}$ *are in $W^{1,p}(\Omega)$ with generalized derivatives*

$$D_i \max\{u, v\}(x) = \begin{cases} D_i u(x) & \textit{if} \quad u(x) > v(x)\,, \\ D_i v(x) & \textit{if} \quad v(x) \ge u(x)\,, \end{cases}$$

$$D_i \min\{u, v\}(x) = \begin{cases} D_i u(x) & \textit{if} \quad u(x) < v(x)\,, \\ D_i v(x) & \textit{if} \quad v(x) \le u(x)\,. \end{cases}$$

Proof: The assertion follows easily from the above examples and the generalized chain rule by using $\max\{u, v\} = (u-v)^+ + v$ and $\min\{u, v\} = u-(u-v)^+$; see [112, Theorem 1.20]. □

Lemma 2.88. *If $(u_j), (v_j) \subset W^{1,p}(\Omega)$ $(1 \le p < \infty)$ are such that $u_j \to u$ and $v_j \to v$ in $W^{1,p}(\Omega)$, then* $\min\{u_j, v_j\} \to$ *min*$\{u, v\}$ *and* $\max\{u_j, v_j\} \to \max\{u, v\}$ *in $W^{1,p}(\Omega)$ as $j \to \infty$.*

For the proof, see [112, Lemma 1.22]. By means of Lemma 2.88, we readily obtain the following result.

Lemma 2.89. *Let $\underline{u},\ \bar{u} \in W^{1,p}(\Omega)$ satisfy $\underline{u} \le \bar{u}$, and let T be the truncation operator defined by*

$$Tu(x) = \begin{cases} \bar{u}(x) & \textit{if} \quad u(x) > \bar{u}(x)\,, \\ u(x) & \textit{if} \quad \underline{u}(x) \le u(x) \le \bar{u}(x)\,, \\ \underline{u}(x) & \textit{if} \quad u(x) < \underline{u}(x)\,. \end{cases}$$

Then T is a bounded continuous mapping from $W^{1,p}(\Omega)$ [respectively, $L^p(\Omega)$] into itself.

Proof: The truncation operator T can be represented in the form

$$Tu = \max\{u, \underline{u}\} + \min\{u, \bar{u}\} - u.$$

Thus, the assertion easily follows from Lemma 2.88. □

Lemma 2.90 (Lattice Structure). *If* $u,\ v \in W_0^{1,p}(\Omega)$, *then* $\max\{u, v\}$ *and* $\min\{u, v\}$ *are in* $W_0^{1,p}(\Omega)$.

Lemma 2.90 implies that $W_0^{1,p}(\Omega)$ has a lattice structure as well; see [112].

A partial ordering of traces on $\partial\Omega$ is given as follows.

Definition 2.91. *Let* $u \in W^{1,p}(\Omega),\ 1 \le p < \infty$. *Then* $u \le 0$ *on* $\partial\Omega$ *if* $u^+ \in W_0^{1,p}(\Omega)$.

2.2.4 Some Inequalities

In this section, we recall some well-known inequalities that are frequently used and that can be found in standard textbooks; see [93, 132, 222].

Young's Inequality

Let $1 < p, q < \infty$, and $1/p + 1/q = 1$. Then

$$ab \le \frac{a^p}{p} + \frac{b^q}{q} \quad (a, b \ge 0).$$

Proof: For $a,\ b \in \mathbb{R}_+$ satisfying $ab = 0$, the inequality is trivially satisfied. Let $a,\ b > 0$. As the function $x \mapsto e^x$ is convex, it follows that

$$ab = e^{\log a + \log b} = e^{\frac{1}{p}\log a^p + \frac{1}{q}\log b^q} \le \frac{1}{p}e^{\log a^p} + \frac{1}{q}e^{\log b^q} = \frac{a^p}{p} + \frac{b^q}{q}$$

□

Young's Inequality with Epsilon

Let $1 < p, q < \infty$, and $1/p + 1/q = 1$. Then

$$ab \le \varepsilon a^p + C(\varepsilon) b^q \quad (a, b \ge 0,\ \varepsilon > 0)$$

with $C(\varepsilon) = (\varepsilon p)^{-q/p}\frac{1}{q}$.

Proof: Again we only need to consider the case where $a,\ b > 0$. In this case, we set $ab = ((\varepsilon p)^{1/p} a)(\frac{b}{(\varepsilon p)^{1/p}})$ and apply Young's inequality. □

Equivalent Norms

Let $1 \leq s < \infty$, and $\xi_i \in \mathbb{R},\ \xi_i \geq 0,\ i = 1, \dots, N$, then we have the following inequality:

$$a \left(\sum_{i=1}^{N} \xi_i^s \right)^{1/s} \leq \sum_{i=1}^{N} \xi_i \leq b \left(\sum_{i=1}^{N} \xi_i^s \right)^{1/s},$$

where a and b are some positive constants depending only on N and s.

Proof: The inequality is an immediate consequence of the fact that all norms in $\mathbb{R}^N$ are equivalent to each other. $\square$

Monotonicity Inequality

Let $1 < p < \infty$. Consider the vector-valued function $a : \mathbb{R}^N \to \mathbb{R}^N$ defined by

$$a(\xi) = |\xi|^{p-2}\xi \quad \text{for} \quad \xi \neq 0, \quad a(0) = 0.$$

If $1 < p < 2$, then we have

$$(a(\xi) - a(\xi')) \cdot (\xi - \xi') > 0 \quad \text{for all} \quad \xi, \xi' \in \mathbb{R}^N,\ \xi \neq \xi'.$$

If $2 \leq p < \infty$, then a constant $c > 0$ exists such that

$$(a(\xi) - a(\xi')) \cdot (\xi - \xi') \geq c\, |\xi - \xi'|^p \quad \text{for all} \quad \xi \in \mathbb{R}^N.$$

Hölder's Inequality

Let $1 \leq p, q \leq \infty,\ \frac{1}{p} + \frac{1}{q} = 1$. If $u \in L^p(\Omega),\ v \in L^q(\Omega)$, then one has

$$\int_\Omega |uv|\ dx \leq \|u\|_{L^p(\Omega)} \|v\|_{L^q(\Omega)}.$$

Minkowski's Inequality

Let $1 \leq p \leq \infty$ and $u, v \in L^p(\Omega)$; then

$$\|u + v\|_{L^p(\Omega)} \leq \|u\|_{L^p(\Omega)} + \|v\|_{L^p(\Omega)}.$$

Clarkson's Inequalities

Let u, $v \in L^p(\Omega)$. If $2 \leq p < \infty$, then

$$\|u+v\|^p_{L^p(\Omega)} + \|u-v\|^p_{L^p(\Omega)} \leq 2^{p-1}\Big(\|u\|^p_{L^p(\Omega)} + \|v\|^p_{L^p(\Omega)}\Big).$$

If $1 < p < 2$, then

$$\|u+v\|^p_{L^p(\Omega)} + \|u-v\|^p_{L^p(\Omega)} \leq 2\Big(\|u\|^p_{L^p(\Omega)} + \|v\|^p_{L^p(\Omega)}\Big).$$

Proof: Use the function $\varphi : [0,1] \to \mathbb{R}$ defined by

$$\varphi(t) = \frac{(1+t)^p + (1-t)^p}{1+t^p}, \quad t \in [0,1].$$

□

Remark 2.92. It follows immediately from Clarkson's inequalities that the spaces $L^p(\Omega)$ and the Sobolev spaces $W^{m,p}(\Omega)$ are uniformly convex for $1 < p < \infty$, and $m = 0, 1, \ldots,$.

Poincaré–Friedrichs Inequality

Let $\Omega \subset \mathbb{R}^N$ be a bounded domain, $1 \leq p < \infty$, and $u \in W^{1,p}_0(\Omega)$. Then we have the estimate

$$\|u\|_{L^p(\Omega)} \leq C\ \|\nabla u\|_{L^p(\Omega)},$$

where the constant C only depends on p, N, and Ω.

Remark 2.93. The Poincaré–Friedrichs inequality implies that

$$\|u\|_{W^{1,p}_0(\Omega)} = \|\nabla u\|_{L^p(\Omega)}$$

defines an equivalent norm on $W^{1,p}_0(\Omega)$. Equivalent norms on $W^{1,p}(\Omega)$ play an important role in the treatment of boundary value problems. The following general result provides a tool to identify equivalent norms on $W^{1,p}(\Omega)$.

Proposition 2.94. *Let $\Omega \subset \mathbb{R}^N$, $N \geq 1$, be a bounded domain with Lipschitz boundary $\partial\Omega$. Assume $\varphi : W^{1,p}(\Omega) \to \mathbb{R}_+$, $1 \leq p < \infty$, is a seminorm that satisfies the following conditions:*

(i) *A positive constant d exists such that*

$$\varphi(u) \leq d\,\|u\|_{W^{1,p}(\Omega)} \quad \textit{for all} \ \ u \in W^{1,p}(\Omega).$$

(ii) *If $u =$ constant, then $\varphi(u) = 0$ implies $u = 0$.*

Then $\|\cdot\|_\sim$ *defined by*

$$\|u\|_\sim = \left(\|\nabla u\|^p_{L^p(\Omega)} + \varphi(u)^p\right)^{\frac{1}{p}}$$

defines an equivalent norm in $W^{1,p}(\Omega)$.

As an application of Proposition 2.94, we obtain, e.g., an equivalent norm on the closed subspace V_Γ of $W^{1,p}(\Omega)$ defined by

$$V_\Gamma = \{u \in W^{1,p}(\Omega) : \gamma(u) = 0 \text{ on } \Gamma\},$$

where $\Gamma \subset \partial\Omega$ is some part of the boundary $\partial\Omega$ with strictly positive surface measure $|\Gamma| > 0$. To this end, define φ by

$$\varphi(u) = \left(\int_\Gamma |\gamma(u)|^p \, d\Gamma\right)^{\frac{1}{p}} \quad \text{for all } u \in W^{1,p}(\Omega),$$

where γ is the trace operator. We observe that (i) and (ii) of Proposition 2.94 are satisfied, and thus $\|\cdot\|_\sim$ defined above gives an equivalent norm on $W^{1,p}(\Omega)$. As $\varphi(u) = 0$ for $u \in V_\Gamma$, we see that

$$\|u\|_\sim = \|\nabla u\|_{L^p(\Omega)} \quad \text{for all } u \in V_\Gamma$$

is an equivalent norm on the subspace V_Γ.

2.3 Operators of Monotone Type

In this section, we provide the basic results on pseudomonotone operators from a Banach space X into its dual space X^*.

2.3.1 Main Theorem on Pseudomonotone Operators

Let X be a real, reflexive Banach space with norm $\|\cdot\|$, X^* its dual space, and denote by $\langle\cdot,\cdot\rangle$ the duality pairing between them. The norm convergence in X and X^* is denoted by "$\to$" and the weak convergence by "$\rightharpoonup$".

Definition 2.95. *Let* $A : X \to X^*$; *then* A *is called*

(i) *continuous (respectively, weakly continuous) iff* $u_n \to u$ *implies* $Au_n \to Au$ *(respectively,* $u_n \rightharpoonup u$ *implies* $Au_n \rightharpoonup Au$*)*
(ii) *demicontinuous iff* $u_n \to u$ *implies* $Au_n \rightharpoonup Au$
(iii) *hemicontinuous iff the real function* $t \to \langle A(u + tv), w\rangle$ *is continuous on* $[0, 1]$ *for all* u, v, $w \in X$
(iv) *strongly continuous or completely continuous iff* $u_n \rightharpoonup u$ *implies* $Au_n \to Au$

(v) *bounded iff* A *maps bounded sets into bounded sets*
(vi) *coercive iff* $\lim_{\|u\|\to\infty} \frac{\langle Au,u\rangle}{\|u\|} = +\infty$

Definition 2.96 (Operators of Monotone Type). *Let* $A : X \to X^*$*; then* A *is called*

(i) *monotone (respectively, strictly monotone) iff* $\langle Au-Av, u-v\rangle \geq$ *(respectively,* $>$*)* 0 *for all* $u, v \in X$ *with* $u \neq v$
(ii) *strongly monotone iff there is a constant* $c > 0$ *such that* $\langle Au-Av, u-v\rangle \geq c\|u-v\|^2$ *for all* $u, v \in X$
(iii) *uniformly monotone iff* $\langle Au - Av, u - v\rangle \geq a(\|u - v\|)\|u - v\|$ *for all* $u, v \in X$ *where* $a : [0,\infty) \to [0,\infty)$ *is strictly increasing with* $a(0) = 0$ *and* $a(s) \to +\infty$ *as* $s \to \infty$
(iv) *pseudomonotone iff* $u_n \rightharpoonup u$ *and* $\limsup_{n\to\infty}\langle Au_n, u_n - u\rangle \leq 0$ *implies* $\langle Au, u - w\rangle \leq \liminf_{n\to\infty}\langle Au_n, u_n - w\rangle$ *for all* $w \in X$
(v) *to satisfy* (S_+)*-condition iff* $u_n \rightharpoonup u$ *and* $\limsup_{n\to\infty}\langle Au_n, u_n - u\rangle \leq 0$ *imply* $u_n \to u$

We can show (cf. [18]) that the pseudomonotonicity according to (iv) of Definition 2.96 is equivalent to the following definition.

Definition 2.97. *The operator* $A : X \to X^*$ *is pseudomonotone iff* $u_n \rightharpoonup u$ *and* $\limsup_{n\to\infty}\langle Au_n, u_n - u\rangle \leq 0$ *implies* $Au_n \rightharpoonup Au$ *and* $\langle Au_n, u_n\rangle \to \langle Au, u\rangle$.

For the following result, see [222, Proposition 27.6].

Lemma 2.98. *Let* $A, B : X \to X^*$ *be operators on the real reflexive Banach space* X*. Then the following implications hold:*

(i) *If* A *is monotone and hemicontinuous, then* A *is pseudomonotone.*
(ii) *If* A *is strongly continuous, then* A *is pseudomonotone.*
(iii) *If* A *and* B *are pseudomonotone, then* $A + B$ *is pseudomonotone.*

The main theorem on pseudomonotone operators due to Brézis is given by the next theorem (see [222, Theorem 27.A]).

Theorem 2.99 (Main Theorem on Pseudomonotone Operators). *Let* X *be a real, reflexive Banach space, and let* $A : X \to X^*$ *be a pseudomonotone, bounded, and coercive operator, and* $b \in X^*$*. Then a solution of the equation* $Au = b$ *exists.*

Remark 2.100. Theorem 2.99 contains several important surjectivity results as special cases, such as Lax–Milgram's theorem and the Main Theorem on Monotone Operators, which will be formulated in the following corollaries.

Corollary 2.101 (Main Theorem on Monotone Operators). *Let X be a real, reflexive Banach space, and let $A : X \to X^*$ be a monotone, hemicontinuous, bounded, and coercive operator, and $b \in X^*$. Then a solution of the equation $Au = b$ exists.*

For the proof of Corollary 2.101, we have only to mention that in view of Lemma 2.98, a monotone and hemicontinuous operator is pseudomonotone.

Corollary 2.102 (Lax–Milgram's Theorem). *Let X be a real Hilbert space, and let $a :\ X \times X \to \mathbb{R}$ be a bilinear form. Assume that*

(i) *a is bounded; i.e., there is a $C > 0$ such that*

$$|a(x, y)| \le C\|x\|\|y\| \qquad \text{for } x, y \in X.$$

(ii) *a is coercive, i.e., there is a $C_0 > 0$ such that*

$$a(x, x) \ge C_0\|x\|^2 \qquad \text{for } x \in H.$$

Then, for each f in X^*, there is a unique element u in X such that

$$a(u, v) = \langle f, v\rangle \qquad \text{for } v \in X.$$

The mapping $f \mapsto u$ is one-to-one, continuous, and linear from X^* onto X.

As for the proof, note that the bilinear form a of Corollary 2.102 defines a linear, bounded, and strongly monotone operator $A : X \to X^*$ acccording to

$$\langle Au, v\rangle = a(u, v) \quad \text{for all } \ u, v \in X,$$

and thus the equation $a(u, v) = \langle f, v\rangle$ of Corollary 2.102 is equivalent with the operator equation $Au = f$ in X^*. The existence result for the latter follows immediately from Corollary 2.101, because A is strongly monotone and continuous and therefore, in particular, also coercive. The uniqueness is a consequence of the strong monotonicity of A.

2.3.2 Leray–Lions Operators

An important class of operators of monotone type is the so-called Leray–Lions operators (see [215, 152]). These kinds of operators occur in the functional analytical treatment of nonlinear elliptic and parabolic problems.

Definition 2.103 (Leray–Lions Operator). *Let X be a real, reflexive Banach space. We say that $A : X \to X^*$ is a Leray–Lions operator if it is bounded and satisfies*

$$Au = \mathcal{A}(u, u), \quad \textit{for} \ \ u \in X,$$

where $\mathcal{A} : X \times X \to X^*$ has the following properties:

(i) *For any* $u \in X$, *the mapping* $v \mapsto \mathcal{A}(u,v)$ *is bounded and hemicontinuous from* X *to its dual* X^*, *with*

$$\langle \mathcal{A}(u,u) - \mathcal{A}(u,v), u - v\rangle \ge 0 \quad \text{for} \;\; v \in X.$$

(ii) *For any* $v \in X$, *the mapping* $u \mapsto \mathcal{A}(u,v)$ *is bounded and hemicontinuous from* X *to its dual* X^*.

(iii) *For any* $v \in X$, $\mathcal{A}(u_n,v)$ *converges weakly to* $\mathcal{A}(u,v)$ *in* X^* *if* $(u_n) \subset X$ *is such that* $u_n \rightharpoonup u$ *in* X *and*

$$\langle \mathcal{A}(u_n,u_n) - \mathcal{A}(u_n,u), u_n - u\rangle \to 0.$$

(iv) *For any* $v \in X$, $\langle \mathcal{A}(u_n,v), u_n\rangle$ *converges to* $\langle F, u\rangle$ *if* $(u_n) \subset V$ *is such that* $u_n \rightharpoonup u$ *in* X, *and* $\mathcal{A}(u_n,v) \rightharpoonup F$ *in* X^*.

As for the proof of the next theorem, see [215].

Theorem 2.104. *Every Leray–Lions operator* $A : X \to X^*$ *is pseudomonotone.*

Next we will see that quasilinear elliptic operators satisfying certain structure and growth conditions represent Leray–Lions operators. To this end, we need to study first the mapping properties of superposition operators, which are also called Nemytskij operators.

Definition 2.105 (Nemytskij Operator). *Let* $\Omega \subset \mathbb{R}^N$, $N \ge 1$, *be a nonempty measurable set, and let* $f : \Omega \times \mathbb{R}^m \to \mathbb{R}$, $m \ge 1$, *and* $u : \Omega \to \mathbb{R}^m$ *be a given function. Then the superposition or Nemytskij operator* F *assigns* $u \mapsto f \circ u$; *i.e.,* F *is given by*

$$Fu(x) = (f \circ u)(x) = f(x, u(x)) \;\; \text{for } x \in \Omega.$$

Definition 2.106 (Carathéodory Function). *Let* $\Omega \subset \mathbb{R}^N$, $N \ge 1$, *be a nonempty measurable set, and let* $f : \Omega \times \mathbb{R}^m \to \mathbb{R}$, $m \ge 1$. *The function* f *is called a Carathéodory function if the following two conditions are satisfied:*

(i) $x \mapsto f(x,s)$ *is measurable in* Ω *for all* $s \in \mathbb{R}^m$.
(ii) $s \mapsto f(x,s)$ *is continuous on* $\mathbb{R}^m$ *for a.e.* $x \in \Omega$.

Lemma 2.107. *Let* $f : \Omega \times \mathbb{R}^m \to \mathbb{R}$, $m \ge 1$, *be a Carathéodory function that satisfies a growth condition of the form*

$$|f(x,s)| \le k(x) + c\sum_{i=1}^{m} |s_i|^{p_i/q}, \quad \forall\; s = (s_1, \dots, s_m) \in \mathbb{R}^m, \;\text{a.e. } x \in \Omega,$$

for some positive constant c *and some* $k \in L^q(\Omega)$, *and* $1 \le q, p_i < \infty$ *for all* $i = 1, \dots, m$. *Then the Nemytskij operator* F *defined by*

$$Fu(x) = f(x, u_1(x), \ldots, u_m(x))$$

is continuous and bounded from $L^{p_1}(\Omega) \times \cdots \times L^{p_m}(\Omega)$ *into* $L^q(\Omega)$. *Here* u *denotes the vector function* $u = (u_1, \ldots, u_m)$. *Furthermore,*

$$\|Fu\|_{L^q(\Omega)} \le c \left(\|k\|_{L^q(\Omega)} + \sum_{i=1}^{m} \|u_i\|_{L^{p_i}(\Omega)}^{p_i/q} \right).$$

Definition 2.108. *Let* $\Omega \subset \mathbb{R}^N$, $N \ge 1$, *be a nonempty measurable set. A function* $f : \Omega \times \mathbb{R}^m \to \mathbb{R}$, $m \ge 1$, *is called superpositionally measurable (or sup-measurable) if the function* $x \mapsto Fu(x)$ *is measurable in* Ω *whenever the component functions* $u_i : \Omega \to \mathbb{R}$ *of* $u = (u_1, \ldots, u_m)$ *are measurable.*

Now let $\Omega \subset \mathbb{R}^N$ be a bounded domain with Lipschitz boundary $\partial\Omega$, let A_1 be the second-order quasilinear differential operator in divergence form given by

$$A_1 u(x) = -\sum_{i=1}^{N} \frac{\partial}{\partial x_i} a_i(x, u(x), \nabla u(x)),$$

and let A_0 denote the operator

$$A_0 u(x) = a_0(x, u(x), \nabla u(x)).$$

Let $1 < p < \infty$, $1/p + 1/q = 1$, and assume for the coefficients $a_i : \ \Omega \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$, $i = 0, 1, \ldots, N$ the following conditions.

(H1) Carathéodory and Growth Condition: Each $a_i(x, s, \xi)$ satisfies Carathéodory conditions, i.e., is measurable in $x \in \Omega$ for all $(s, \xi) \in \mathbb{R} \times \mathbb{R}^N$ and continuous in (s, ξ) for a.e. $x \in \Omega$. A constant $c_0 > 0$ and a function $k_0 \in L^q(\Omega)$ exist so that

$$|a_i(x, s, \xi)| \le k_0(x) + c_0(|s|^{p-1} + |\xi|^{p-1})$$

for a.e. $x \in \Omega$ and for all $(s, \xi) \in \mathbb{R} \times \mathbb{R}^N$, with $|\xi|$ denoting the Euclidian norm of the vector ξ.

(H2) Monotonicity Type Condition: The coefficients a_i satisfy a monotonicity condition with respect to ξ in the form

$$\sum_{i=1}^{N} (a_i(x, s, \xi) - a_i(x, s, \xi'))(\xi_i - \xi_i') > 0$$

for a.e. $x \in \Omega$, for all $s \in \mathbb{R}$, and for all $\xi, \xi' \in \mathbb{R}^N$ with $\xi \ne \xi'$.

(H3) Coercivity Type Condition:

$$\sum_{i=1}^{N} a_i(x, s, \xi)\xi_i \ge \nu|\xi|^p - k(x)$$

for a.e. $x \in \Omega$, for all $s \in \mathbb{R}$, and for all $\xi \in \mathbb{R}^N$ with some constant $\nu > 0$ and some function $k \in L^1(\Omega)$.

Let V be a closed subspace of $W^{1,p}(\Omega)$ such that $W_0^{1,p}(\Omega) \subset V \subset W^{1,p}(\Omega)$, then under condition (H1) the differential operators A_1 and A_0 generate mappings from V into its dual space (again denoted by A_1 and A_0, respectively) defined by

$$\langle A_1 u, \varphi\rangle = \sum_{i=1}^{N} \int_\Omega a_i(x,u,\nabla u) \frac{\partial \varphi}{\partial x_i}\, dx\,, \quad \langle A_0 u, \varphi\rangle = \int_\Omega a_0(x,u,\nabla u)\ \varphi\, dx\,.$$

Theorem 2.109. *Set $A = A_1 + A_0$. Then the operators A, A_0, and A_1 have the following properties:*

(i) *If (H1) is satisfied, then the mappings $A, A_1, A_0 : V \to V^*$ are continuous and bounded.*
(ii) *If (H1) and (H2) are satisfied, then $A : V \to V^*$ is pseudomonotone.*
(iii) *If (H1), (H2), and (H3) are satisfied, then A has the (S_+)-property.*

Conditions (H1) and (H2) are the so-called Leray–Lions conditions that guarantee that A is pseudomonotone. In their original paper, Leray and Lions [149] showed the pseudomonotonicity under conditions (H1), (H2), and the following additional condition.

(H4) $\limsup_{|\xi|\to\infty,\ s\in B} \sum_{i=1}^{N} \frac{a_i(x,s,\xi)\xi_i}{|\xi|+|\xi|^{p-1}} = +\infty$, for a.e. $x \in \Omega$ and all bounded sets B.

However, Landes and Mustonen have shown in [136] that condition (H4) is redundant for the pseudomonotonicity of A. As for the proof of the results stated in Theorem 2.109 as well as on existence theorems involving pseudomonotone operators, we refer to [17, 18] and [23, 27, 105, 152, 208, 222].

Example 2.110. Let $\Omega \subset \mathbb{R}^N$ be a bounded domain. A prototype of a monotone elliptic operator in Ω is the negative of the p-Laplacian Δ_p, $1 < p < \infty$, defined by

$$\Delta_p u = \operatorname{div}(|\nabla u|^{p-2}\nabla u) \quad \text{where} \quad \nabla u = (\partial u/\partial x_1, \dots, \partial u/\partial x_N).$$

This operator coincides with the Laplacian Δ if $p = 2$, and is of the form A_1 with the coeffients a_i, $i = 1, \dots, N$, given by

$$a_i(x,s,\xi) = |\xi|^{p-2}\xi_i.$$

Thus, hypothesis (H1) is satisfied with $k_0 = 0$, $c_0 = 1$, and $a_0 = 0$. Hypothesis (H2) follows from the inequalities satisfied by the vector-valued function $\xi \mapsto |\xi|^{p-2}\xi$, (see Sect. 2.2.4) and (H3) is obviously true with $\nu = 1$ and $k = 0$ due to

$$\sum_{i=1}^{N} a_i(x,s,\xi)\xi_i = \sum_{i=1}^{N} |\xi|^{p-2}\xi_i\,\xi_i = |\xi|^p.$$

Therefore, hypotheses (H1)–(H3) are satisfied by the negative p-Laplacian, and in view of Theorem 2.109, we see that $-\Delta_p : V \to V^*$ is continuous, bounded, pseudomonotone, and has the (S_+)-property. Moreover, from the inequality

$$\langle -\Delta_p u - (-\Delta_p v), u - v\rangle = \int_\Omega (|\nabla u|^{p-2}\nabla u - |\nabla v|^{p-2}\nabla v)(\nabla u - \nabla v)\, dx \geq 0,$$

for all u, $v \in V$, we infer that $-\Delta_p : V \to V^*$ is, in particular, also a monotone operator. Depending on the domain of definition of $-\Delta_p$, we can say even more. For example, let $V = W_0^{1,p}(\Omega)$. According to Sect. 2.2.4,

$$\|u\|_V = \left(\int_\Omega |\nabla u|^p\, dx\right)^{1/p}$$

defines an equivalent norm in V. From the inequalities for the function $\xi \mapsto |\xi|^{p-2}\xi$, we see that the operator $-\Delta_p : W_0^{1,p}(\Omega) \to (W_0^{1,p}(\Omega))^*$ has the mapping properties given in the following lemma.

Lemma 2.111. *Let V be a closed subspace of $W^{1,p}(\Omega)$ such that $W_0^{1,p}(\Omega) \subset V \subset W^{1,p}(\Omega)$. Then one has:*

(i) $-\Delta_p : V \to V^*$ *is continuous, bounded, pseudomonotone, and has the (S_+)-property.*
(ii) $-\Delta_p : W_0^{1,p}(\Omega) \to (W_0^{1,p}(\Omega))^*$ *is*
  (a) *strictly monotone if* $1 < p < \infty$.
  (b) *strongly monotone if* $p = 2$ *(Laplacian).*
  (c) *uniformly monotone if* $2 < p < \infty$.

2.3.3 Multivalued Pseudomonotone Operators

In this section, we briefly recall the main results of the theory of pseudomonotone multivalued operators developed by Browder and Hess to the extent it will be needed in the study of variational and hemivariational inequalities. For the proofs and a more detailed presentation, we refer to the monographs [222, 177].

First we present basic results about the continuity of multivalued functions (multifunctions) and provide useful equivalent descriptions of these notions. Even though these notions can be defined in a much more general context, we confine ourselves to mappings between Banach spaces, which is sufficient for our purpose.

Definition 2.112 (Semicontinuous Multifunctions). *Let X, Y be Banach spaces and $A : X \to 2^Y$ be a multifunction.*

(i) *A is called upper semicontinuous at x_0, if for every open subset $V \subset Y$ with $A(x_0) \subset V$, a neighborhood $U(x_0)$ exists such that $A(U(x_0)) \subset V$. If A is upper semicontinuous at every $x_0 \in X$, we call A upper semicontinuous in X.*

(ii) *A is called lower semicontinuous at x_0 if for every neighborhood $V(y)$ of every $y \in A(x_0)$, a neighborhood $U(x_0)$ exists such that*

$$A(u) \cap V(y) \neq \emptyset \quad \text{for all} \;\; u \in U(x_0).$$

If A is lower semicontinuous at every $x_0 \in X$, we call A lower semicontinuous in X.

(iii) *A is called continuous at x_0 if A is both upper and lower semicontinuous at x_0. If A is continuous at every $x_0 \in X$, we call A continuous in X.*

Alternative equivalent continuity criteria are given in the following propositions. To this end, we introduce the *preimage* of a multifunction.

Definition 2.113 (Preimage). *Let $M \subset Y$ and $A : X \to 2^Y$ be a multifunction. The preimage $A^{-1}(M)$ is defined by*

$$A^{-1}(M) = \{x \in X : A(x) \cap M \neq \emptyset\}.$$

Proposition 2.114. *Let X, Y be Banach spaces and $A : X \to 2^Y$ be a multifunction. Then the following statements are equivalent:*

(i) *A is upper semicontinuous.*
(ii) *For all closed sets $C \subset Y$, the preimage $A^{-1}(C)$ is closed.*
(iii) *If $x \in X$, (x_n) is a sequence in X with $x_n \to x$ as $n \to \infty$, and V is an open set in Y such that $A(x) \subset V$, then $n_0 \in \mathbb{N}$ exists depending on V such that for all $n \geq n_0$, we have $A(x_n) \subset V$.*

Proposition 2.115. *Let X, Y be Banach spaces and $A : X \to 2^Y$ be a multifunction. Then the following statements are equivalent:*

(i) *A is lower semicontinuous.*
(ii) *For all open sets $O \subset Y$, the preimage $A^{-1}(O)$ is open.*
(iii) *If $x \in X$, (x_n) is a sequence in X with $x_n \to x$ as $n \to \infty$, and $y \in A(x)$, then for every $n \in \mathbb{N}$, we can find a $y_n \in A(x_n)$, such that $y_n \to y$, as $n \to \infty$.*

Remark 2.116. For a single-valued operator $A : X \to Y$, upper semicontinuous and lower semicontinuous in the multivalued setting is identical with continuous. For $A : M \to 2^N$ having the same corresponding properties, where M and N are subsets of the Banach spaces X and Y, respectively, then M and N have to be equipped with the induced topology.

Next we introduce the notion of multivalued monotone and pseudomonotone operators from a real, reflexive Banach space X into its dual space and formulate the main surjectivity result for these kinds of operators.

Definition 2.117 (Graph). *Let X be a real Banach space, and let $A : X \to 2^{X^*}$ be a multivalued mapping; i.e., to each $u \in X$, there is assigned a subset*

$A(u)$ of X^*, which may be empty if $u \notin D(A)$, where $D(A)$ is the domain of A given by

$$D(A) = \{u \in X : A(u) \neq \emptyset\}.$$

The graph of A denoted by $\mathrm{Gr}(A)$ is given by

$$\mathrm{Gr}(A) = \{(u, u^*) \in X \times X^* : u^* \in A(u)\}.$$

Definition 2.118 (Monotone Operator). *The mapping $A : X \to 2^{X^*}$ is called*

(i) *monotone iff*

$$\langle u^* - v^*, u - v\rangle \geq 0 \quad \textit{for all} \;\; (u, u^*),\, (v, v^*) \in \mathrm{Gr}(A)$$

(ii) *strictly monotone iff*

$$\langle u^* - v^*, u - v\rangle > 0 \quad \textit{for all} \;\; (u, u^*),\, (v, v^*) \in \mathrm{Gr}(A), \;\; u \neq v$$

(iii) *maximal monotone iff A is monotone and there is no monotone mapping $\tilde{A} : X \to 2^{X^*}$ such that $\mathrm{Gr}(A)$ is a proper subset of $\mathrm{Gr}(\tilde{A})$, which is equivalent to the following implication:*

$$(u, u^*) \in X \times X^* : \quad \langle u^* - v^*, u - v\rangle \geq 0 \;\; \textit{for all} \;\; (v, v^*) \in \mathrm{Gr}(A)$$

implies $(u, u^*) \in \mathrm{Gr}(A)$

The notions of strongly and uniformly monotone multivalued operators are defined in a similar way as for single-valued operators.

Example 2.119. If $X = \mathbb{R}$, then a maximal monotone mapping $\beta : \mathbb{R} \to 2^{\mathbb{R}}$ is called *maximal monotone graph* in $\mathbb{R}^2$. For example, an increasing function $f : \mathbb{R} \to \mathbb{R}$ generates a maximal monotone graph β in $\mathbb{R}^2$ given by

$$\beta(s) := [f(s-0), f(s+0)],$$

where $f(s \pm 0)$ are the one-sided limits of f in s.

A single-valued operator

$$A : D(A) \subset X \to X^*$$

is to be understood as a multivalued operator $A : X \to X^*$ by setting $Au = \{Au\}$ if $u \in D(A)$ and $Au = \emptyset$ otherwise. Thus, A is monotone iff

$$\langle Au - Av, u - v\rangle \geq 0 \quad \text{for all} \;\; u, v \in D(A),$$

and $A : D(A) \subset X \to X^*$ is maximal monotone iff A is monotone and the condition

$$(u, u^*) \in X \times X^* : \quad \langle u^* - Av, u - v\rangle \geq 0 \;\; \text{for all} \;\; v \in D(A)$$

implies $u \in D(A)$ and $u^* = Au$.

Definition 2.120 (Pseudomonotone Operator). *Let X be a real reflexive Banach space. The operator $A : X \to 2^{X^*}$ is called pseudomonotone if the following conditions hold:*

(i) *The set $A(u)$ is nonempty, bounded, closed, and convex for all $u \in X$.*
(ii) *A is upper semicontinuous from each finite-dimensional subspace of X to the weak topology on X^*.*
(iii) *If $(u_n) \subset X$ with $u_n \rightharpoonup u$, and if $u_n^* \in A(u_n)$ is such that*

$$\limsup \langle u_n^*, u_n - u\rangle \le 0,$$

then to each element $v \in X$, $u^*(v) \in A(u)$ exists with

$$\liminf \langle u_n^*, u_n - v\rangle \ge \langle u^*(v), u - v\rangle.$$

Definition 2.121 (Generalized Pseudomonotone Operator). *Let X be a real reflexive Banach space. The operator $A : X \to 2^{X^*}$ is called generalized pseudomonotone if the following holds:*
Let $(u_n) \subset X$ and $(u_n^*) \subset X^*$ with $u_n^* \in A(u_n)$. If $u_n \rightharpoonup u$ in X and $u_n^* \rightharpoonup u^*$ in X^* and if $\limsup\langle u_n^*, u_n - u\rangle \le 0$, then the element u^* lies in $A(u)$ and

$$\langle u_n^*, u_n\rangle \to \langle u^*, u\rangle.$$

The next two propositions provide the relation between pseudomonotone and generalized pseudomontone operators.

Proposition 2.122. *Let X be a real reflexive Banach space. If the operator $A : X \to 2^{X^*}$ is pseudomonotone, then A is generalized pseudomonotone.*

Under the additional assumption of boundedness, the following converse of Proposition 2.122 is true.

Proposition 2.123. *Let X be a real reflexive Banach space, and assume that $A : X \to 2^{X^*}$ satisfies the following conditions:*

(i) *For each $u \in X$, we have that $A(u)$ is a nonempty, closed, and convex subset of X^*.*
(ii) *$A : X \to 2^{X^*}$ is bounded.*
(iii) *If $u_n \rightharpoonup u$ in X and $u_n^* \rightharpoonup u^*$ in X^* with $u_n^* \in A(u_n)$ and if $\limsup\langle u_n^*, u_n - u\rangle \le 0$, then $u^* \in A(u)$ and $\langle u_n^*, u_n\rangle \to \langle u^*, u\rangle$.*

Then the operator $A : X \to 2^{X^*}$ is pseudomonotone.

As for the proof of Proposition 2.123 we refer to [177, Chap. 2]. Note that the notion of boundedness of a multivalued operator is exactly the same as for single-valued operators; i.e., the image of a bounded set is again bounded.

The relation between maximal monotone and pseudomonotone operators as well as the invariance of pseudomonotonicity under addition is given in the following theorem.

Theorem 2.124. *Let X be a real reflexive Banach space, and let A, $A_i : X \to 2^{X^*}$, $i = 1, 2$.*

(i) *If A is maximal monotone with $D(A) = X$, then A is pseudomonotone.*
(ii) *If A_1 and A_2 are two pseudomonotone operators, then the sum $A_1 + A_2 : X \to 2^{X^*}$ is pseudomonotone.*

The main theorem on pseudomonotone multivalued operators is formulated in the next theorem.

Theorem 2.125. *Let X be a real reflexive Banach space, and let $A : X \to 2^{X^*}$ be a pseudomonotone and a bounded operator, which is coercive in the sense that a real-valued function $c : \mathbb{R}_+ \to \mathbb{R}$ exists with*

$$c(r) \to +\infty, \quad \textit{as} \;\; r \to +\infty$$

such that for all $(u, u^*) \in Gr(A)$, we have

$$\langle u^*, u - u_0 \rangle \geq c(\|u\|_X)\|u\|_X$$

for some $u_0 \in X$. Then A is surjective; i.e., range$(A) = X$.

Remark 2.126. We remark that the boundedness condition supposed in Theorem 2.125 can be dropped (see [177, Theorem 2.6]). This is because by definition of a multivalued pseudomonotone operator A according to Definition 2.120 the operator A has to be upper semicontinuous from each finite-dimensional subspace X_n of X to the weak topology on X^*. This latter condition along with the coercivity and the properties of the images allows us to get a surjectivity result on finite-dimensional subspaces X_n.

Theorem 2.127. *Let X be a real reflexive Banach space, $\Phi : X \to 2^{X^*}$ a maximal monotone operator, and $u_0 \in D(\Phi)$. Let $A : X \to 2^{X^*}$ be a pseudomonotone operator, and assume that either A_{u_0} is quasi-bounded or Φ_{u_0} is strongly quasi-bounded. Assume further that $A : X \to 2^{X^*}$ is u_0-coercive; i.e., a real-valued function $c : \mathbb{R}_+ \to \mathbb{R}$ exists with $c(r) \to +\infty$ as $r \to +\infty$ such that for all $(u, u^*) \in \mathrm{Gr}\,(A)$, we have $\langle u^*, u - u_0 \rangle \geq c(\|u\|_X)\|u\|_X$. Then $A + \Phi$ is surjective; i.e., range$(A + \Phi) = X^*$.*

The operators A_{u_0} and Φ_{u_0} that appear in Theorem 2.127 are defined by $A_{u_0}(v) := A(u_0 + v)$ and similarly for Φ_{u_0}. As for the notion of *quasi-bounded* and *strongly quasi-bounded,* we refer to [177, p. 51]. In particular, one has that any bounded operator is quasi-bounded and strongly quasi-bounded.

2.4 First-Order Evolution Equations

In this section we present the basic functional analytic tools needed in the study of first-order single- and multivalued evolution equations in the form

$$u \in X,\ u' \in X^* : \quad u' + Au \ni f \quad \text{in} \quad X^*, \quad u(0) = u_0, \tag{2.4}$$

where $X = L^p(0, \tau; V)$, $1 < p < \infty$, with $\tau > 0$ is the L^p-space of vector-valued functions $u : (0, \tau) \to V$ defined on the interval $(0, \tau)$ with values in some Banach space V, and u' is the generalized or distributional derivative of the function $t \mapsto u(t)$ with respect to $t \in (0, \tau)$. The right-hand side $f \in X^*$ is given, and $A : X \to 2^{X^*}$ is some (in general) multivalued operator. The initial values u_0 are taken from some Hilbert space H such that the embedding $V \subset H$ is continuous and dense. Problem (2.4) provides an abstract framework for the functional analytic treatment of initial-boundary value problems for parabolic differential equations and inclusions.

2.4.1 Motivation

To give a motivation for the study of the abstract problem (2.4), let us consider the classic initial-boundary value problem for the heat equation.

Let $\Omega \subset \mathbb{R}^N$ be a bounded domain with smooth boundary $\partial\Omega$, and denote $Q = \Omega \times (0, \tau)$ and $\Gamma = \partial\Omega \times (0, \tau)$ for some $\tau > 0$. We are looking for a function $(x, t) \mapsto u(x, t)$ defined in $\overline{\Omega} \times [0, \tau)$ such that

$$\begin{aligned} u_t - \Delta u &= f && \text{in } Q, \\ u &= 0 && \text{on } \Gamma, \\ u(\cdot, 0) &= u_0(\cdot) && \text{in } \Omega, \end{aligned} \tag{2.5}$$

where the right-hand side $f : Q \to \mathbb{R}$ and the initial values $u_0 : \Omega \to \mathbb{R}$ are given functions. A classic (or strong) solution of (2.5) is a function that satisfies all equations of (2.5) pointwise in the usual sense. This, however, requires sufficient smoothness assumptions on the data f and u_0 as well as on the domain Ω. To be able to deal with (2.5) under relaxed regularity assumptions on the data, one tries instead to consider an appropriate generalized problem corresponding to (2.5), which in turn leads to the notion of weak solutions. To make plausible the definition of weak solutions of (2.5), we temporarily suppose that u is in fact a smooth solution of (2.5). In a similar way as in the explanation of weak solutions of the Dirichlet problem for elliptic equations (see Chap. 1), we formally multiply the heat equation by $v \in C_0^\infty(\Omega)$ and subsequently integrate by parts, which yields

$$\frac{d}{dt} \int_\Omega u(x, t) v(x)\, dx + \int_\Omega \nabla u(x, t) \nabla v(x)\, dx = \int_\Omega f(x, t) v(x)\, dx, \tag{2.6}$$

for all $v \in C_0^\infty(\Omega)$. As $V = W_0^{1,2}(\Omega)$ is the closure of $C_0^\infty(\Omega)$ in $W_0^{1,2}(\Omega)$, we see that (2.6) makes perfect sense for $v \in V$ and $u(\cdot, t) \in V$ with $f \in L^2(Q)$. Now we change our viewpoint concerning the function u in that we deal with the space variable x and the time variable t in different ways. We associate with $u = u(x, t)$ a mapping (again denoted by u) $u : [0, \tau) \to V$ defined by

$$(u(t))(x) = u(x, t), \quad x \in \Omega,\ t \in [0, \tau),$$

which means that we are going to consider u not as a function of x and t together, but as a mapping $u : [0, \tau) \to V$. Let $H = L^2(\Omega)$, and denote by $(\cdot, \cdot)$ the inner product in H. Similarly as for u, we interpret the right-hand side function f as a mapping $f : [0, \tau) \to H$ according to

$$(f(t))(x) = f(x, t), \quad x \in \Omega, \ t \in [0, \tau).$$

By means of the bilinear form $a : V \times V \to \mathbb{R}$ defined by

$$a(w, v) = \int_\Omega \nabla w \nabla v \, dx,$$

we can rewrite (2.6) in the form

$$\frac{d}{dt}(u(t), v) + a(u(t), v) = (f(t), v) \quad \text{for all} \ \ v \in V, \tag{2.7}$$

which together with $u(0) = u_0 \in H$ represents a weak formulation of the initial-boundary value problem (2.5) [note that the homogeneous boundary values are taken into account by $u(t) \in V$]. For fixed t, let us consider the mapping $v \mapsto (f(t), v)$. Apparently this mapping is linear, and in view of the continuous embedding $V \subset H$, we have

$$|(f(t), v)| \leq \|f(t)\|_H \|v\|_H \leq c\|f(t)\|_H \|v\|_V,$$

which shows that the mapping is bounded. Thus, the mapping $v \mapsto (f(t), v)$ belongs to V^*; i.e., there is a $b \in V^*$ such that

$$\langle b, v \rangle = (f(t), v), \quad \text{for all} \ \ v \in V,$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between V and V^*. Next we will see that the functional b is defined in a unique way. Assume there is another $h \in H$ that generates the same functional b. It yields

$$(f(t) - h, v) = 0 \quad \text{for all} \ \ v \in V,$$

and thus $f(t) = h$, because $V \subset H$ is densely embedded. It allows us to identify b with $f(t)$. In this way, the element $f(t) \in H$ has to be considered as an element of V^*, and thus, we have

$$\langle f(t), v \rangle = (f(t), v) \quad \text{for all} \ \ v \in V. \tag{2.8}$$

The bilinear form a defined above, which can easily be seen to be bounded, generates a linear and bounded operator $A : V \to V^*$ through

$$\langle Aw, v \rangle = a(w, v) \quad \text{for all} \ \ w, v \in V. \tag{2.9}$$

Thus, by (2.8) and (2.9), we can rewrite (2.7) in the form

$$\frac{d}{dt}(u(t),v)+\langle Au(t),v\rangle=\langle f(t),v\rangle \quad \text{for all} \ \ v\in V. \tag{2.10}$$

We will later introduce the generalized or distributional derivative of a vector-valued function $t \mapsto u(t)$, whose derivative $u'(t)$ turns out to have the property

$$\langle u'(t),v\rangle=\frac{d}{dt}(u(t),v) \quad \text{for all} \ \ v\in V, \tag{2.11}$$

where d/dt is the generalized derivative of the real-valued function $t \mapsto (u(t),v)$ on $(0,\tau)$. Equations (2.10) and (2.11) result in the operator equation

$$u'(t)+Au(t)=f(t) \quad \text{in} \ \ V^*, \tag{2.12}$$

which is only required to be satisfied for a.e. $t\in(0,\tau)$.

Let $X=L^2(0,\tau;V)$, and denote by X^* its dual space, which is given by $X^*=L^2(0,\tau;V^*)$ (see Sect. 2.4.2). Furthermore, by means of the operator $A:V\to V^*$, we define an operator $\hat{A}:X\to X^*$ by

$$(\hat{A}u)(t)=Au(t), \quad t\in(0,\tau).$$

Thus, in view of (2.12) and the definition of $\hat{A}$, a generalized formulation of the initial-boundary value problem (2.5) reads as follows: For given $u_0\in H$ and $f\in X^*$, we seek a function $u\in X$ such that $u'\in X^*$ and

$$u'+\hat{A}u=f \quad \text{in} \ \ X^*, \quad u(0)=u_0, \tag{2.13}$$

which is of the abstract form of the (single-valued) evolution equation (2.4).

We observe a few particularities that are typical in the functional analytic setting of parabolic problems.

(i) The space and time variables x and t are treated differently, and the function $(x,t)\mapsto u(x,t)$ is considered as a vector-valued function.
(ii) The formulation of the given initial-boundary value problem as an abstract operator equation of the form (2.13) requires the use of two spaces H and V with the need that $V\subset H$ is densely and continuously embedded. It leads to the concept of *evolution triple*: $V\subset H\subset V^*$.
(iii) The solution space for the operator equation (2.13) is given by

$$W=\{u\in X : u'\in X^*\},$$

where u' has to be understood as the distributional derivative of the vector-valued function u.

In the following subsections, we will give the basic notions and existence results for the abstract evolution equation (2.4).

2.4.2 Vector-Valued Functions

Let B be a Banach space with norm $\|\cdot\|$, B^* its dual space, and $0 < \tau < \infty$. We consider vector-valued functions $u : [0,\tau] \to B$ and explain first some notions such as measurability and integrability. Most of the material in this subsection can be found in [93, 208, 222].

Definition 2.128. *Let u and s be vector-valued functions.*

(i) *$s : [0,\tau] \to B$ is called simple (or step function) if it is of the form*

$$s(t) = \sum_{i=1}^{m} \chi_{E_i}(t) u_i, \quad 0 \le t \le \tau,$$

where each E_i is a Lebesgue measurable subset of the interval $[0,\tau]$, $u_i \in B$ $(i = 1,\dots,m)$, and χ_{E_i} is the characteristic function of E_i.

(ii) *$u : [0,\tau] \to B$ is strongly measurable if a sequence (s_k) of simple functions $s_k : [0,\tau] \to B$ exists such that $s_k(t) \to u(t)$ as $k \to \infty$, for a.e. $t \in [0,\tau]$.*

(iii) *$u : [0,\tau] \to B$ is weakly measurable if for each $u^* \in B^*$ the mapping $t \to \langle u^*, u(t)\rangle$ is Lebesgue measurable.*

(iv) *$u : [0,\tau] \to B$ is almost separably valued if a subset $N \subset [0,\tau]$ of zero measure exists such that the set $\{u(t) : t \in [0,\tau] \setminus N\}$ is a separable subset of B.*

Theorem 2.129 (Pettis). *The function $u : [0,\tau] \to B$ is strongly measurable if and only if u is weakly measurable and almost separably valued.*

Definition 2.130. *The integral of vector-valued functions is defined as follows:*

(i) *The integral of the simple function $s(t) = \sum_{i=1}^{m} \chi_{E_i}(t) u_i$ is defined by*

$$\int_0^\tau s(t)\,dt = \sum_{i=1}^{m} \mathrm{meas}(E_i)\ u_i.$$

(ii) *The vector-valued function $u : [0,\tau] \to B$ is called integrable if a sequence (s_k) of simple functions exists such that*

$$\int_0^\tau \|s_k(t) - u(t)\|\ dt \to 0 \quad \text{as } k \to \infty.$$

(iii) *If $u : [0,\tau] \to B$ is integrable, its integral is defined by*

$$\int_0^\tau u(t)\,dt = \lim_{k\to\infty} \int_0^\tau s_k(t)\,dt.$$

Theorem 2.131. *The function $u : [0, \tau] \to B$ is integrable if and only if u is strongly measurable and $t \to \|u(t)\|$ is integrable. Furthermore, one has*

$$\left\| \int_0^\tau u(t)\ dt \right\| \leq \int_0^\tau \|u(t)\|\ dt, \quad \textit{and} \quad \left\langle u^*, \int_0^\tau u(t)\ dt \right\rangle = \int_0^\tau \langle u^*, u(t)\rangle\ dt,$$

for each $u^* \in B^*$.

Definition 2.132. *Let $1 \leq p \leq \infty$. We denote by $L^p(0, \tau; B)$ the space of (equivalent classes of) measurable functions $u : [0, \tau] \to B$ such that $\|u(\cdot)\|$ belongs to $L^p(0, \tau; \mathbb{R})$ with*

$$\|u\|_{L^p(0,\tau;B)} = \left(\int_0^\tau \|u(t)\|^p\, dt \right)^{1/p} \quad \textit{for} \ \ 1 \leq p < \infty,$$

$$\|u\|_{L^\infty(0,\tau;B)} = \operatorname*{ess\,sup}_{0 \leq t \leq \tau} \|u(t)\| < \infty.$$

The space $C([0, \tau]; B)$ comprises of all continuous functions $u : [0, \tau] \to B$ with

$$\|u\|_{C([0,\tau];B)} = \max_{0 \leq t \leq \tau} \|u(t)\| < \infty.$$

Theorem 2.133. *Let B and Y be Banach spaces. Then we have the following results:*

(i) *$L^p(0, \tau; B)$ with $1 \leq p \leq \infty$ and the norm given by Definition 2.132 is a Banach space.*
(ii) *$C([0, \tau]; B)$ is dense in $L^p(0, \tau; B)$ for $1 \leq p < \infty$, and the embedding $C([0, \tau]; B) \subset L^p(0, \tau; B)$ is continuous.*
(iii) *If B is a Hilbert space with scalar product $(\cdot, \cdot)_B$, then $L^2(0, \tau; B)$ is also a Hilbert space with the scalar product*

$$(u, v) = \int_0^\tau (u(t), v(t))_B\ dt.$$

(iv) *$L^p(0, \tau; B)$ is separable if B is separable and $1 \leq p < \infty$.*
(v) *$L^p(0, \tau; B)$ is uniformly (strictly) convex in the case where B is uniformly (strictly) convex and $1 < p < \infty$.*
(vi) *If the embedding $B \subset Y$ is continuous, then the embedding*

$$L^r(0, \tau; B) \subset L^q(0, \tau; Y), \quad 1 \leq q \leq r \leq \infty,$$

is also continuous.
(vii) *Let B be a reflexive and separable Banach space, and let $1 < p < \infty$, $1/p + 1/q = 1$. Then $X = L^p(0, \tau; B)$ is also reflexive and separable, and its dual space X^* is norm-isomorphic to $L^q(0, \tau; B^*)$. Therefore, X^* and $L^q(0, \tau; B^*)$ may be identified. The duality pairing $\langle \cdot, \cdot \rangle_X$ between X and its dual X^* can be written as*

$$\langle v, u \rangle_X = \int_0^\tau \langle v(t), u(t) \rangle_B\, dt \quad \textit{for all} \ \ u \in X,\ v \in X^*.$$

Remark 2.134. We usually drop the subscripts X and B in $\langle v, u\rangle_X$ and $\langle v(t), u(t)\rangle_B$, respectively, because from the context, the type of duality pairing is clear.

2.4.3 Evolution Triple and Generalized Derivative

The material of this subsection is mainly taken from [208, 222].

Definition 2.135 (Evolution Triple). *A triple (V, H, V^*) is called an evolution triple if the following properties hold:*

(i) *V is a real, separable, and reflexive Banach space, and H is a real, separable Hilbert space endowed with the scalar product $(\cdot, \cdot)$.*
(ii) *The embedding $V \subset H$ is continuous, and V is dense in H.*
(iii) *Identifying H with its dual H^* by the Riesz map, we then have $H \subset V^*$ with the equation*

$$\langle h, v\rangle_V = (h, v) \quad \textit{for} \ \ h \in H \subset V^*, \ \ v \in V.$$

Remark 2.136. As V is reflexive and V is dense in H, the space H^* is dense in V^*, and hence, H is dense in V^*. It is a simple consequence of Proposition 2.25 in Sect. 2.1.2 applied to the embedding operator $i : V \to H$.

Example 2.137. Let $\Omega \subset \mathbb{R}^N$ be a bounded domain with Lipschitz boundary $\partial\Omega$, and let V be a closed subspace of $W^{1,p}(\Omega)$ with $2 \le p < \infty$ such that $W_0^{1,p}(\Omega) \subset V \subset W^{1,p}(\Omega)$. Then (V, H, V^*) with $H = L^2(\Omega)$ is an evolution triple with all embeddings being, in addition, compact.

Definition 2.138. *Let Y, Z be Banach spaces, and $u \in L^1(0, \tau; Y)$ and $w \in L^1(0, \tau; Z)$. Then, the function w is called the generalized derivative of the function u in $(0, \tau)$ iff the following relation holds:*

$$\int_0^\tau \varphi'(t)u(t)\, dt = -\int_0^\tau \varphi(t)w(t)\, dt \quad \textit{for all} \ \ \varphi \in C_0^\infty(0, \tau).$$

We write $w = u'$.

Theorem 2.139. *Let $V \subset H \subset V^*$ be an evolution triple, and let $1 \le p, q \le \infty$, $0 < \tau < \infty$. Let $u \in L^p(0, \tau; V)$; then the generalized derivative $u' \in L^q(0, \tau; V^*)$ exists iff there is a function $w \in L^q(0, \tau; V^*)$ such that*

$$\int_0^\tau (u(t), v)_H \varphi'(t)\ dt = -\int_0^\tau \langle w(t), v\rangle_V \varphi(t)\ dt$$

for all $v \in V$ and all $\varphi \in C_0^\infty(0, \tau)$. The generalized derivative u' is uniquely defined and $u' = w$.

Definition 2.140. *Let V be a real, separable, and reflexive Banach space, and let $X = L^p(0,\tau;V)$, $1 < p < \infty$. A space W is defined by*

$$W = \{u \in X : u' \in X^*\},$$

where u' is the generalized derivative, and $X^* = L^q(0,\tau;V^*)$, $1/p + 1/q = 1$.

Theorem 2.141 (Lions–Aubin). *Let B_0, B, B_1 be reflexive Banach spaces with $B_0 \subset B \subset B_1$, and assume $B_0 \subset B$ is compactly and $B \subset B_1$ is continuously embedded. Let $1 < p < \infty$, $1 < q < \infty$, and define $\mathcal{W}$ by*

$$\mathcal{W} = \{u \in L^p(0,\tau;B_0) : u' \in L^q(0,\tau;B_1)\}.$$

Then $\mathcal{W} \subset L^p(0,\tau;B)$ is compactly embedded.

Example 2.142. Let $\Omega \subset \mathbb{R}^N$ be a bounded domain with Lipschitz boundary $\partial\Omega$. As $W^{1,p}(\Omega) \subset L^p(\Omega)$ is compactly embedded, and $L^p(\Omega) \subset W^{1,p}(\Omega)^*$ is continuously embedded for $2 \le p < \infty$, Theorem 2.141 can be applied by setting $B_0 = W^{1,p}(\Omega)$, $B = L^p(\Omega)$ and $B_1 = W^{1,p}(\Omega)^*$, $2 \le p < \infty$. Thus, W defined in Definition 2.140, i.e.,

$$W = \{u \in L^p(0,\tau;W^{1,p}(\Omega)) : u' \in L^q(0,\tau;W^{1,p}(\Omega)^*)\},$$

is compactly embedded in $L^p(0,\tau;L^p(\Omega)) \equiv L^p(Q)$, where $Q = \Omega \times (0,\tau)$.

Let $\Omega \subset \mathbb{R}^N$ be as in Example 2.142, and $\Gamma = \partial\Omega \times (0,\tau)$. If $u \in X = L^p(0,\tau;W^{1,p}(\Omega))$, then for a.e. $t \in (0,\tau)$ the function $t \mapsto \gamma u(t) \in L^p(\partial\Omega)$ is well defined, where $\gamma : W^{1,p}(\Omega) \to L^p(\partial\Omega)$ denotes the trace operator (see Theorem 2.75). In view of Theorem 2.133 (vi) and the continuity of $\gamma : W^{1,p}(\Omega) \to L^p(\partial\Omega)$, we get that $t \mapsto \gamma u(t)$ belongs to $L^p(0,\tau;L^p(\partial\Omega)) \equiv L^p(\Gamma)$. If we denote the mapping that assigns $u \in X$ to the vector-valued function $t \mapsto \gamma u(t)$ again by γ, then it follows that $\gamma : X \to L^p(\Gamma)$ is linear and continuous. Moreover, as the trace operator $\gamma : W^{1,p}(\Omega) \to L^p(\partial\Omega)$ is even compact, we obtain the following result.

Proposition 2.143. *Let $\Omega \subset \mathbb{R}^N$ be a bounded domain with Lipschitz boundary $\partial\Omega$, and let $X = L^p(0,\tau;W^{1,p}(\Omega))$ with $2 \le p < \infty$. Then the trace operator $\gamma : W \to L^p(\Gamma)$ is compact.*

Proof: We apply Theorem 2.141. To this end, let $B_0 = W^{1,p}(\Omega)$, $B = W^{1-\varepsilon,p}(\Omega)$, and $B_1 = B_0^*$. As $B_0 \subset B$ is compactly embedded for any $\varepsilon \in (0,1)$, and $B \subset B_1$ is continuously embedded, from Theorem 2.141, it follows that $W \subset L^p(0,\tau;W^{1-\varepsilon,p}(\Omega))$ is compactly embedded. If we select ε such that $0 < \varepsilon < 1 - 1/p$, then $\gamma : W^{1-\varepsilon,p}(\Omega) \to W^{1-\varepsilon-1/p,p}(\partial\Omega)$ is linear and continuous, and thus $\gamma : L^p(0,\tau;W^{1-\varepsilon,p}(\Omega)) \to L^p(0,\tau;W^{1-\varepsilon-1/p,p}(\Omega)) \subset L^p(\Gamma)$ is linear and continuous, which due to the compact embedding of $W \subset L^p(0,\tau;W^{1-\varepsilon,p}(\Omega))$ completes the proof. $\square$

Theorem 2.144. *Let $V \subset H \subset V^*$ be an evolution triple, and let $1 < p < \infty$, $1/p + 1/q = 1$, $0 < \tau < \infty$. Then the following hold:*

(i) *The space W defined in Definition 2.140 is a real, separable, and reflexive Banach space with the norm*

$$\|u\|_W = \|u\|_X + \|u'\|_{X^*}.$$

(ii) *The embedding $W \subset C([0,\tau]; H)$ is continuous.*
(iii) *For all $u, v \in W$ and arbitrary t, s with $0 \le s \le t \le \tau$, the following generalized integration by parts formula holds:*

$$(u(t), v(t))_H - (u(s), v(s))_H = \int_s^t \langle u'(\zeta), v(\zeta)\rangle_V + \langle v'(\zeta), u(\zeta)\rangle_V \, d\zeta.$$

Remark 2.145. The integration by parts formula is equivalent to

$$\frac{d}{dt}(u(t), v(t))_H = \langle u'(t), v(t)\rangle_V + \langle v'(t), u(t)\rangle_V \quad \text{for a.e. } t \in (0,\tau).$$

In particular, for $u = v$, we obtain

$$\frac{d}{dt}\|u(t)\|_H^2 = 2\langle u'(t), u(t)\rangle_V,$$

which implies

$$\int_s^t \langle u'(\zeta), u(\zeta)\rangle_V \, d\zeta = \frac{1}{2}(\|u(t)\|_H^2 - \|u(s)\|_H^2). \tag{2.14}$$

In case that $V = W^{1,p}(\Omega)$, $2 \le p < \infty$, and $H = L^2(\Omega)$, we obtain the following generalization of formula (2.14), which will be useful for obtaining comparison principles in evolutionary problems.

Lemma 2.146. *Let $X = L^p(0,\tau; W^{1,p}(\Omega))$ with $2 \le p < \infty$ and $W = \{u \in X : u' \in X^*\}$, where $\Omega \subset \mathbb{R}^N$ is a bounded domain with Lipschitz boundary $\partial\Omega$. Let $\theta : \mathbb{R} \to \mathbb{R}$ be continuous and piecewise continuously differentiable with $\theta' \in L^\infty(\mathbb{R})$, and $\theta(0) = 0$, and let Θ denote the primitive of θ defined by*

$$\Theta(r) = \int_0^r \theta(s)\, ds.$$

Then, for $w \in W$, the following formula holds:

$$\int_r^s \langle w'(t), \theta(w(t))\rangle \, dt = \int_\Omega \Theta(w(s))\, dx - \int_\Omega \Theta(w(r))\, dx, \tag{2.15}$$

for a.e. $0 \le r < s \le \tau$.

Proof: The proof makes use of density arguments and the generalized chain rule for Sobolev functions (see Lemma 2.84). Note first that in view of the assumptions on θ and Lemma 2.84, the composed function $\theta(w)$ is in X for $w \in W$. The space $C^1([0,\tau];C^1(\overline{\Omega}))$ of smooth functions is dense in W (cf. [222, Chap. 23]). Let $w \in W$ be given. Then there is a sequence $(w_n) \subset C^1([0,\tau];C^1(\overline{\Omega}))$ with $w_n \to w$ as $n \to \infty$. For the smooth functions w_n, we have

$$\begin{aligned}\int_r^s \langle w_n'(t), \theta(w_n(t))\rangle\, dt &= \int_r^s \int_\Omega w_n'(x,t)\theta(w_n(x,t))\, dxdt \\ &= \int_r^s \int_\Omega \frac{\partial}{\partial t}\Big(\Theta(w_n(x,t))\Big)\, dxdt \\ &= \int_\Omega \Big(\Theta(w_n(x,s)) - \Theta(w_n(x,r))\Big)\, dx. \qquad (2.16)\end{aligned}$$

The assumptions on θ imply that θ is Lipschitz continuous, and thus, it follows that for some subsequence of (w_n) (again denoted by (w_n)),

$$\theta(w_n) \to \theta(w) \quad \text{in} \quad X, \qquad (2.17)$$

and due to the continuous embedding $W \subset C([0,\tau];L^2(\Omega))$, one gets for all $t \in [0,\tau]$

$$\Theta(w_n(t)) \to \Theta(w(t)) \quad \text{in} \quad L^2(\Omega). \qquad (2.18)$$

By using (2.17), (2.18), we may pass to the limit in (2.16) for some subsequence, which completes the proof. □

Example 2.147. Let $\theta(s) = s$. Then θ trivially satisfies all assumptions of Lemma 2.146, and the primitive Θ is given by $\Theta(s) = (1/2)s^2$, and thus, formula (2.15) becomes

$$\begin{aligned}\int_r^s \langle w'(t), w(t)\rangle\, dt &= \frac{1}{2}\int_\Omega (w(s))^2\, dx - \frac{1}{2}\int_\Omega (w(r))^2\, dx \\ &= \frac{1}{2}(\|w(s)\|_H^2 - \|w(r)\|_H^2), \qquad (2.19)\end{aligned}$$

for all $0 \le r < s \le \tau$, where $H = L^2(\Omega)$, which is formula (2.14.)

The following example will play a crucial rule in obtaining comparison results.

Example 2.148. If $\theta(s) = s^+ = \max\{s, 0\}$, then its primitive can easily be seen to be $\Theta(s) = (1/2)(s^+)^2$, and thus, for $w \in W$, we get the formula

$$\int_r^s \langle w'(t), (w(t))^+\rangle\, dt = \frac{1}{2}(\|(w(s))^+\|_H^2 - \|(w(r))^+\|_H^2). \qquad (2.20)$$

2.4.4 Existence Results for Evolution Equations

The material of this subsection is mainly based on results obtained in [19, 20, 208]; see also [152, 222].

Let $V \subset H \subset V^*$ be an evolution triple, and let $X = L^p(0,\tau;V)$, X^* and W be the spaces of vector-valued functions as defined in Sect. 2.4.3 with $1 < p < \infty$, $1/p + 1/q = 1$, and $0 < \tau < \infty$. We provide an existence result for the evolution equation

$$u \in W: \quad u'(t) + A(t)u(t) = f(t), \ 0 < t < \tau, \quad u(0) = 0, \tag{2.21}$$

where $f \in X^*$ is given and $A(t) : V \to V^*$ is some operator specified later. Without loss of generality, homogeneous initial values have been assumed, because inhomogeneous initial values can be transformed to homogeneous ones by translation. The generalized derivative $Lu = u'$ restricted to the subset

$$D(L) = \{u \in X : u' \in X^* \ \text{ and } \ u(0) = 0\} = \{u \in W : u(0) = 0\}$$

defines a linear operator $L : D(L) \to X^*$ given by

$$\langle Lu, v \rangle = \int_0^\tau \langle u'(t), v(t) \rangle \, dt \quad \text{for all} \ \ v \in X.$$

The operator L has the following properties.

Lemma 2.149. *Let $V \subset H \subset V^*$ be an evolution triple, and let $X = L^p(0,\tau;V)$, where $1 < p < \infty$. Then the operator $L : D(L) \subset X \to X^*$ is densely defined, closed, and maximal monotone.*

Proof: First we note that the set M defined by

$$M = \{u \in C^1([0,\tau];V) : u(0) = 0\}$$

satisfies $M \subset D(L)$ and $\overline{M} = X$, which shows that $\overline{D(L)} = X$, and thus, L is densely defined. Due to the continuous embedding $W \subset C([0,\tau];H)$, it follows that $D(L)$ is closed in W, and thus, L is closed. From formula (2.19), we get

$$\langle Lu, u \rangle = \int_0^\tau \langle Lu(t), u(t) \rangle \, dt = \frac{1}{2}(\|u(\tau)\|_H^2 - \|u(0)\|_H^2) = \frac{1}{2}\|u(\tau)\|_H^2 \geq 0, \tag{2.22}$$

which shows that L is monotone. To prove that L is maximal monotone, we make use of the characterization of single-valued maximal monotone operators (see Sect. 2.3.3). To this end, suppose $(v,w) \in X \times X^*$ and

$$\langle w - Lu, v - u \rangle \geq 0 \quad \text{for all} \ \ u \in D(L). \tag{2.23}$$

We need to show that $v \in D(L)$ and $w = Lv = v'$. Let u be chosen as

$$u = \varphi z, \quad \textit{with } \varphi \in C_0^\infty(0,\tau) \text{ and } z \in V.$$

Obviously, $u \in D(L)$ with $u' = \varphi' z$, and $\langle Lu, u\rangle = 0$ in view of (2.22). Due to inequality (2.23), we then have

$$0 \le \langle w, v\rangle - \int_0^\tau \langle \varphi'(t)v(t) + \varphi(t)w(t), z\rangle_V \, dt \text{ for all } z \in V, \tag{2.24}$$

which implies

$$\int_0^\tau \langle \varphi'(t)v(t) + \varphi(t)w(t), z\rangle_V \, dt = 0 \text{ for all } \varphi \in C_0^\infty(0,\tau),$$

and thus, $v' = w$. It remains to show that $v \in D(L)$. Again by applying formula (2.22) with u replaced by $v - u$, we obtain

$$0 \le \langle v' - u', v - u\rangle = \frac{1}{2}(\|v(\tau) - u(\tau)\|_H^2 - \|v(0) - u(0)\|_H^2). \tag{2.25}$$

To complete the proof, we only need to show that $v(0) = 0$. To this end, choose a sequence $(v_n) \subset V$ with $\tau v_n \to v(\tau)$ in H (note that V is dense in H) and specialize $u(t) = t v_n$. Then $u \in D(L)$, and from (2.25), one obtains

$$0 \le \frac{1}{2}(\|v(\tau) - \tau v_n\|_H^2 - \|v(0)\|_H^2),$$

which by passing to the limit as $n \to \infty$ results in $v(0) = 0$. □

Remark 2.150. With only slight modifications one can prove that $L : D(L) \subset X \to X^*$ defined by

$$Lu = u' : \quad D(L) = \{u \in X : u' \in X^* \text{ and } u(0) = u(\tau)\}$$

is a densely defined, closed, and maximal monotone operator (cf. [222, Proposition 32.10]).

Now we state the following conditions on the time-dependent operators $A(t) : V \to V^*$:

(H1) $\|A(t)u\|_{V^*} \le c_0\Big(\|u\|_V^{p-1} + k_0(t)\Big)$ for all $u \in V$ and $t \in [0,\tau]$ with some positive constant c_0 and $k_0 \in L^q(0,\tau)$.
(H2) $A(t) : V \to V^*$ is demicontinuous for each $t \in [0,\tau]$.
(H3) The function $t \to \langle A(t)u, v\rangle$ is measurable on $(0,\tau)$ for all $u, v \in V$.
(H4) $\langle A(t)u, u\rangle \ge c_1(\|u\|_V^p - k_1(t))$ for all $u \in V$ and $t \in [0,\tau]$ with some constant $c_1 > 0$ and some function $k_1 \in L^1(0,\tau)$.

Define an operator $\hat{A}$ related with $A(t)$ by

$$\hat{A}(u)(t) = A(t)u(t), \quad t \in [0,\tau], \tag{2.26}$$

which may be considered as the associated Nemytskij operator generated by the operator-valued function $t \mapsto A(t)$. Thus, problem (2.21) corresponds to the following one:

$$u \in D(L) : Lu + \hat{A}(u) = f \quad \text{in} \quad X^*. \tag{2.27}$$

Definition 2.151. *Let $D(L)$ be equipped with the graph norm; that is,*

$$\|u\|_L = \|u\|_X + \|Lu\|_{X^*}.$$

The operator $\hat{A} : X \to X^*$ is called pseudomonotone with respect to the graph norm topology of $D(L)$ (or pseudomonotone w.r.t. $D(L)$ for short); if for any sequence $(u_n) \in D(L)$ satisfying

$$u_n \rightharpoonup u \text{ in } X, \ Lu_n \rightharpoonup Lu \text{ in } X^*, \text{ and } \limsup_{n\to\infty} \langle \hat{A}(u_n), u_n - u\rangle \le 0,$$

it follows that

$$\hat{A}(u_n) \rightharpoonup \hat{A}(u) \text{ in } X^* \text{ and } \langle \hat{A}(u_n), u_n\rangle \to \langle \hat{A}(u), u\rangle.$$

In an obvious similar way, the (S_+)-condition with respect to $D(L)$ is defined.

For the following surjectivity result, which yields the existence for problem (2.27), we refer to [19, 152].

Theorem 2.152. *Let $L : D(L) \subset X \to X^*$ be as given above, and let $\hat{A} : X \to X^*$ defined by (2.26) be bounded, demicontinuous, and pseudomonotone w.r.t. $D(L)$. If $\hat{A}$ is coercive, then $(L + \hat{A})(D(L)) = X^*$; that is, $L + \hat{A}$ is surjective.*

The next result shows that certain properties of the operators $A(t)$ are transfered to its Nemytskij operator $\hat{A}$; cf. [20].

Theorem 2.153. *Let hypotheses (H1)–(H4) be satisfied. Then we have the following results:*

(i) *If $A(t) : V \to V^*$ is pseudomonotone for all $t \in [0, \tau]$, then $\hat{A} : X \to X^*$ is pseudomonotone with respect to $D(L)$ according to Definition 2.151.*
(ii) *If $A(t) : V \to V^*$ has the (S_+)-property for all $t \in [0, \tau]$, then $\hat{A} : X \to X^*$ has the (S_+)-property with respect to $D(L)$.*
(iii) *Hypotheses (H1) and (H3) imply that $\hat{A} : X \to X^*$ is bounded.*
(iv) *Hypotheses (H1)–(H3) imply that $\hat{A} : X \to X^*$ is demicontinuous.*
(v) *Hypothesis (H4) implies that $\hat{A} : X \to X^*$ is coercive.*

2.4.5 Multivalued Evolution Equations

In this section, we briefly recall a general surjectivity result for multivalued operators in a real reflexive Banach space X, which allows us to deal with multivalued evolution equations in the form

$$u \in X: \quad u' + A(u) \ni f \ \text{ in } \ X^*, \quad u(0) = u_0. \tag{2.28}$$

To this end, we introduce first the notion of a multivalued pseudomonotone operator with respect to the graph norm topology of the domain $D(L)$ (w.r.t. $D(L)$ for short) of some linear, closed, densely defined, and maximal monotone operator $L : D(L) \subset X \to X^*$.

Definition 2.154. *Let $L : D(L) \subset X \to X^*$ be a linear, closed, densely defined, and maximal monotone operator. The operator $A : X \to 2^{X^*}$ is called pseudomonotone w.r.t. D(L) if the following conditions are satisfied:*

(i) *The set $A(u)$ is nonempty, bounded, closed, and convex for all $u \in X$.*
(ii) *A is upper semicontinuous from each finite-dimensional subspace of X to the weak topology of X^*.*
(iii) *If $(u_n) \subset D(L)$ with $u_n \rightharpoonup u$ in X, $Lu_n \rightharpoonup Lu$ in X^*, $u_n^* \in A(u_n)$ with $u_n^* \rightharpoonup u^*$ in X^*, and $\limsup\langle u_n^*, u_n - u\rangle \le 0$, then $u^* \in A(u)$ and $\langle u_n^*, u_n\rangle \to \langle u^*, u\rangle$.*

Definition 2.155. *The operator $A : X \to 2^{X^*}$ is called coercive iff either the domain $D(A)$ of A is bounded or $D(A)$ is unbounded and*

$$\frac{\inf\{\langle v^*, v\rangle : v^* \in A(v)\}}{\|v\|_X} \to +\infty \quad \textit{as} \quad \|v\|_X \to \infty, \ v \in D(A).$$

The following surjectivity result can be found in [79, Theorem 1.3.73, p. 62].

Theorem 2.156. *Let X be a real reflexive, strictly convex Banach space with dual space X^*, and let $L : D(L) \subset X \to X^*$ be a linear, closed, densely defined, and maximal monotone operator. If the multivalued operator $A : X \to 2^{X^*}$ is pseudomonotone w.r.t. $D(L)$, bounded, and coercive, then $L + A$ is surjective; i.e., $(L + A)(D(L)) = X^*$.*

Consider the multivalued evolution equation

$$u \in X: \quad u' + A(u) \ni f \ \text{ in } \ X^*, \quad u(0) = 0, \tag{2.29}$$

where

$$X = L^p(0, \tau; V), \quad 1 < p < \infty,$$

and $V \subset H \subset V^*$ is an evolution triple with V being strictly convex. As earlier, we define the operator L by

$$Lu = u', \quad \text{with } D(L) = \{u \in W : u(0) = 0\} \tag{2.30}$$

(for W see Definition 2.140). Thus, problem (2.29) can equivalently be written in the form

$$u \in D(L) : \quad u' + A(u) \ni f \ \text{ in } \ X^*. \tag{2.31}$$

Corollary 2.157. *If the multivalued operator $A : X \to 2^{X^*}$ in (2.29) is pseudomonotone w.r.t. $D(L)$, bounded, and coercive, then problem (2.29) has at least one solution.*

Proof: In view of Theorem 2.133, the Banach space X is reflexive and strictly convex. The operator $L : D(L) \subset X \to X^*$ given by (2.30) is densely defined, linear, closed, and maximal monotone (see Lemma 2.149). Thus, the assertion follows from Theorem 2.156. □

2.5 Nonsmooth Analysis

The area of nonsmooth analysis is closely related with the development of a critical point theory for nondifferentiable functions, in particular, for locally Lipschitz continuous functions based on Clarke's generalized gradient. It provides an appropriate mathematical framework to extend the classic critical point theory for C^1-functionals in a natural way, and to meet specific needs in applications, such as in nonsmooth mechanics and engineering. In this section, we provide basic facts and results of nonsmooth analysis to such an extent as it will be needed in the study of the problems we shall be investigating in this book.

2.5.1 Clarke's Generalized Gradient

Throughout this section, X stands for a real Banach space endowed with the norm $\|\cdot\|$. The dual space of X is denoted X^*, and the notation $\langle \cdot, \cdot \rangle$ means the duality pairing between X^* and X.

We recall the following well-known definition.

Definition 2.158. *A functional $f : X \to \mathbb{R}$ is said to be locally Lipschitz if for every point $x \in X$ a neighborhood V of x in X and a constant $K > 0$ exist such that*

$$|f(y) - f(z)| \le K\|y - z\|, \quad \forall\, y, z \in V.$$

Example 2.159. A convex and continuous function $f : X \to \mathbb{R}$ is locally Lipschitz. More generally, a convex function $f : X \to \mathbb{R}$, which is bounded above on a neighborhood of some point is locally Lipschitz (see [68, p. 34]).

Example 2.160. A functional $f : X \to \mathbb{R}$, which is Lipschitz continuous on bounded subsets of X is locally Lipschitz. The converse assertion is not generally true. For instance, consider the next situation (given through [203]). On the Hilbert space ℓ^2, let the function $f : \ell^2 \to \mathbb{R}$ be defined by

$$f(x) = \sup_{n \geq 0}(2n|x_n| - n), \quad \forall\, x \in \ell^2,$$

where x_n are the components of x. The function f is convex, continuous, and not bounded on the bounded sets. Indeed, f is defined on ℓ^2 because for any $x \in \ell^2$, the set

$$\{n : 2n|x_n| - n \geq 0\} = \left\{n : |x_n| \geq \frac{1}{2}\right\}$$

is finite. The function f is convex because it is the upper hull of the convex functions f_n on ℓ^2 given by $f_n(x) = 2n|x_n| - n$. We note that f is zero on the ball centered at 0 and radius $\frac{1}{2}$ because $0 = f_0(x) \leq f(x)$ and $2|x_n| \leq 1$ if $\|x\| < \frac{1}{2}$. Being bounded on a nonempty open set, the function f is continuous. Finally, it is seen that $f(e_n) = n$, where e_n is the n-th vector of the canonical basis of ℓ^2. It turns out that the function f is not bounded from above on the unit sphere in ℓ^2. Consequently, the function f is not Lipschitz continuous on bounded subsets, but as pointed out in Example 2.159, f is locally Lipschitz.

The classic theory of differentiability does not work in the case of locally Lipschitz functions. However, a suitable subdifferential calculus approach has been developed by Clarke [68]. Here we give a brief introduction. Further details can be found in [68, 43, 79, 103, 173].

Definition 2.161. *Let $f : X \to \mathbb{R}$ be a locally Lipschitz function, and fix two points $u, v \in X$. The generalized directional derivative of f at u in the direction v is defined as follows:*

$$f^o(u; v) = \limsup_{\substack{x \to u \\ t \downarrow 0}} \frac{f(x + tv) - f(x)}{t}.$$

As f is locally Lipschitz, it is clear that $f^o(u; v) \in \mathbb{R}$.

Proposition 2.162. *If $f : X \to \mathbb{R}$ is a locally Lipschitz function, then the following holds:*

(i) *The function $f^o(u; \cdot) : X \to \mathbb{R}$ is subadditive, positively homogeneous, and satisfies the inequality*

$$|f^o(u; v)| \leq K\|v\|, \quad \forall\, v \in X,$$

where $K > 0$ is the Lipschitz constant of f near the point $u \in X$.

(ii) $f^o(u; -v) = (-f)^o(u; v), \quad \forall\, v \in X$.

(iii) *The function $(u, v) \in X \times X \mapsto f^o(u; v) \in \mathbb{R}$ is upper semicontinuous.*

Proof: The result follows directly from Definition 2.161. □

The next definition focuses on the case where $f^o(u;v)$ reduces to the usual directional derivative

$$f'(u;v) = \lim_{t\downarrow 0} \frac{f(u+tv) - f(u)}{t}.$$

Definition 2.163. *A locally Lipschitz function $f : X \to \mathbb{R}$ is said to be regular at a point $u \in X$ if*

(i) *the directional derivative $f'(u;v)$ exists, for every $v \in X$.*
(ii) *$f^o(u;v) = f'(u;v)$, $\forall\ v \in X$.*

Significant classes of regular functions are given in the following examples.

Example 2.164. If the function $f : X \to \mathbb{R}$ is strictly differentiable, that is, for all $u \in X$, $f'(u) \in X^*$ exists such that

$$\lim_{\substack{w\to u \\ t\downarrow 0}} \frac{f(w+tv) - f(w)}{t} = \langle f'(u), v\rangle, \quad \forall\ v \in X,$$

where the convergence is uniform for v in compact sets, then f is locally Lipschitz and regular in the sense of Definition 2.163. In particular, if $f : X \to \mathbb{R}$ is a continuously differentiable function, then f is strictly differentiable, so it is locally Lipschitz and regular.

Example 2.165. A convex and continuous function $f : X \to \mathbb{R}$ is regular.

On the basis of Definition 2.161, one introduces the main notion in this section.

Definition 2.166. *The generalized gradient of a locally Lipschitz functional $f : X \to \mathbb{R}$ at a point $u \in X$ is the subset of X^* defined by*

$$\partial f(u) = \{\zeta \in X^* \ :\ f^o(u;v) \geq \langle \zeta, v\rangle,\ \forall\ v \in X\}.$$

By using the Hahn–Banach theorem (see [24, p. 1]), it follows $\partial f(u) \neq \emptyset$.

Example 2.167. If $f : X \to \mathbb{R}$ is a locally Lipschitz function that is Gâteaux differentiable and regular at the point $u \in X$, then one has $\partial f(u) = \{D_G f(u)\}$, where $D_G f(u)$ denotes the Gâteaux differential of f at u. Indeed, as f is Gâteaux differentiable and regular at u, we may write

$$\langle D_G f(u), v\rangle = f'(u;v) = f^o(u;v), \quad \forall\ v \in X,$$

that implies $D_G f(u) \in \partial f(u)$. Conversely, if $\zeta \in \partial f(u)$, from Definitions 2.166 and 2.163 in conjunction with the assumption that f is Gâteaux differentiable at u, it turns out that

$$\langle \zeta, v\rangle \leq f^o(u;v) = f'(u;v) = \langle D_G f(u), v\rangle, \quad \forall\ v \in X,$$

so $\zeta = D_G f(u)$.

Example 2.168. If $f : X \to \mathbb{R}$ is continuously differentiable, then $\partial f(u) = \{f'(u)\}$ for all $u \in X$, where $f'(u)$ denotes the Fréchet differential of f at u. It is a direct consequence of Example 2.167.

Example 2.169. If $f : X \to \mathbb{R}$ is convex and continuous, then the generalized gradient $\partial f(u)$ coincides with the subdifferential of f at u in the sense of convex analysis. It follows from Examples 2.159 and 2.165.

Remark 2.170. It is seen from Definition 2.166, Example 2.169, and Proposition 2.162(i) that the generalized gradient of a locally Lipschitz functional $f : X \to \mathbb{R}$ at a point $u \in X$ is given by

$$\partial f(u) = \partial(f^o(u;\cdot))(0),$$

where in the right-hand side, the subdifferential in the sense of convex analysis is written.

The next proposition presents some important properties of generalized gradients.

Proposition 2.171. *Let $f : X \to \mathbb{R}$ be a locally Lipschitz function. Then for any $u \in X$, the following properties hold:*

(i) *$\partial f(u)$ is a convex, weak*-compact subset of X^* and*

$$\|\zeta\|_{X^*} \le K\,, \quad \forall\, \zeta \in \partial f(u),$$

where $K > 0$ is the Lipschitz constant of f near u.

(ii) *$f^o(u;v) = \max\{\langle \zeta, v\rangle : \zeta \in \partial f(u)\}$, $\forall\, v \in X$.*

(iii) *The mapping $u \mapsto \partial f(u)$ is weak*-closed from X into X^*.*

(iv) *The mapping $u \mapsto \partial f(u)$ is upper semicontinuous from X into X^*, where X^* is equipped with the weak*-topology.*

Proof: As for (i) and (ii), one applies Definitions 2.161 and 2.166, and as for (iv), see [68]. To see (iii), let $(u_n) \subset X$ satisfy $u_n \to u$ in X, and let $\zeta_n \in \partial f(u_n)$ with $\zeta_n \rightharpoonup^* \zeta$ in X^*. We need to show that $\zeta \in \partial f(u)$. By Definition 2.166, we have $\langle \zeta_n, v\rangle \le f^o(u_n;v)$ for all $v \in X$, which from the weak*-convergence of (ζ_n) and the upper semicontinuity of the function $x \mapsto f^o(x;v)$ according to Proposition 2.162(iii) implies

$$\langle \zeta, v\rangle \le \limsup_{n\to\infty} f^o(u_n;v) \le f^o(u;v) \ \text{ for all } v \in X,$$

and thus, $\zeta \in \partial f(u)$. $\square$

Remark 2.172. The definitions and results given here are applicable to a locally Lipschitz function $f : U \to \mathbb{R}$ on a nonempty, open subset U of the Banach space X.

In the final part of this subsection, we present, for the sake of providing more information on the development of generalized differentiation theory in variational analysis, some basic elements of another subdifferential calculus for nonsmooth functionals, namely the one introduced by Mordukhovich ([164], [165]). The subsequent chapters of the book will not make use of this theory because we need calculus rules and specific properties related to Clarke's concept of generalized gradient for locally Lipschitz functionals in the sense of Definition 2.166 as will be given in the next subsection devoted to calculus, but we consider that it is useful to outline here the subdifferentiation approach in [164], [165] (see also [22]).

Given a nonempty subset S of a Banach space X and a point $u \in S$, it is introduced, for every number $\varepsilon \geq 0$, the set of ε- normals to S at u as the subset of X^* equal to

$$\hat{N}_\varepsilon(u; S) = \{\zeta \in X^* : \limsup_{\substack{w \to u \\ w \in S}} \frac{\langle \zeta, w-u \rangle}{\|w-u\|} \leq \varepsilon\}.$$

The basic normal cone $N(u; S)$ to S at u is defined by

$$N(u; S) = \limsup_{\substack{w \to u, w \in S \\ \varepsilon \downarrow 0}} \hat{N}_\varepsilon(w; S),$$

where in the right-hand side, the sequential Painlevé–Kuratowski upper limit is written. Explicitly, this means that

$$N(u; S) = \{\zeta \in X^* : \text{ there are sequences } w_k \to u \text{ with } w_k \in S, \varepsilon_k \downarrow 0,$$
$$\zeta_k \rightharpoonup^* \zeta \text{ with } \zeta_k \in \hat{N}_{\varepsilon_k}(w_k; S) \text{ for all } k\}.$$

It is shown in [165, Theorem 2.9] that in the case where X is an Asplund space (i.e., every separable subspace of X has a separable dual), the formula of $N(u; S)$ results in

$$N(u; S) = \limsup_{w \to u, w \in S} \hat{N}_0(w; S).$$

Now we are in a position to introduce the notion of subdifferential of an extended real-valued function $f : X \to [-\infty, +\infty]$ at $u \in X$ with $f(u) \in \mathbb{R}$ as follows:

$$\tilde{\partial} f(u) := \{\zeta \in X^* : (\zeta, -1) \in N((u, f(u)); \mathrm{epi}(f))\}.$$

If $f : X \to \mathbb{R}$ is locally Lipschitz and X is an Asplund space, the relationship between the above subdifferential $\tilde{\partial} f(u)$ and Clarke's generalized gradient in the sense of Definition 2.166 is expressed by the following formula:

$$\partial f(u) = \mathrm{cl}^* \mathrm{co}\, \tilde{\partial} f(u), \text{ for all } u \in X$$

(see [165, Theorem 8.11]), where the notation cl*co stands for the convex closure in the weak* topology on the space X^*.

2.5.2 Some Calculus

This subsection is devoted to the basic calculus rules with generalized gradients.

Proposition 2.173. *Let $f : X \to \mathbb{R}$ be a locally Lipschitz function, let $\lambda \in \mathbb{R}$, and let $u \in X$. Then the following formula holds:*

$$\partial(\lambda f)(u) = \lambda \partial f(u).$$

In particular, one has

$$\partial(-f)(u) = -\partial f(u).$$

Proof: If $\lambda = 0$, the property is obvious. If $\lambda > 0$, we have

$$\begin{aligned}\zeta \in \partial(\lambda f)(u) &\iff \langle \frac{1}{\lambda}\,\zeta, v\rangle \leq \frac{1}{\lambda}\,(\lambda f)^o(u; v) = f^o(u; v), \quad \forall\ v \in X\\ &\iff \zeta \in \lambda \partial f(u).\end{aligned}$$

If $\lambda < 0$, we have $\zeta \in \partial(\lambda f)(u) \iff$

$$\begin{aligned}\langle \frac{1}{\lambda}\,\zeta, v\rangle &= -\frac{1}{\lambda}\langle \zeta, -v\rangle \leq -\frac{1}{\lambda}\,(\lambda f)^o(u; -v) = -\frac{1}{\lambda}\,(-\lambda f)^o(u; v)\\ &= f^o(u; v), \quad \forall\ v \in X\end{aligned}$$

$\iff \zeta \in \lambda\partial f(u)$, where Proposition 2.162(ii) has been used. □

Proposition 2.174. *Let $f, g : X \to \mathbb{R}$ be locally Lipschitz functions. Then for every $u \in X$, the following inclusion holds:*

$$\partial(f + g)(u) \subset \partial f(u) + \partial g(u).$$

If, in addition, the functions f and g are regular at the point $u \in X$, then the above inclusion becomes an equality, and $f + g$ is regular at u.

Proof: Let $\zeta \in \partial(f + g)(u)$. Definition 2.166 ensures

$$\langle \zeta, v\rangle \leq f^o(u; v) + g^o(u; v), \quad \forall\ v \in X. \tag{2.32}$$

Arguing by contradiction, let us admit that $\zeta \notin \partial f(u) + \partial g(u)$. Then, by separation in the space X^* endowed with the w^*-topology, $w \in X$ exists such that

$$\begin{aligned}\langle \zeta, w\rangle &> \max\{\langle z, w\rangle\ :\ z \in \partial f(u) + \partial g(u)\}\\ &= \max\{\langle z_1, w\rangle\ :\ z_1 \in \partial f(u)\} + \max\{\langle z_2, w\rangle\ :\ z_2 \in \partial g(u)\}\\ &= f^o(u; w) + g^o(u; w),\end{aligned}$$

where Proposition 2.171(ii) has been employed. It contradicts (2.32), which proves the first assertion in Proposition 2.174.

Suppose now that f and g are regular at u in X. Then, by means of Definition 2.166 for every $\zeta \in \partial f(u) + \partial g(u)$, we have

$$\begin{aligned}\langle \zeta, v\rangle &\leq f^o(u;v) + g^o(u;v)\\ &= f'(u;v) + g'(u;v)\\ &= (f+g)'(u;v)\\ &\leq (f+g)^o(u;v)\end{aligned}$$

for all $v \in X$. We conclude $\zeta \in \partial(f+g)(u)$, which completes the proof. □

Remark 2.175. The inclusion of Proposition 2.174 becomes an equality also when at least one of the two locally Lipschitz functions is strictly differentiable.

We state a useful necessary condition of optimality in the case of locally Lipschitz functions.

Proposition 2.176. *If $u \in X$ is a local minimum or maximum point for the locally Lipschitz function $f : X \to \mathbb{R}$, then $0 \in \partial f(u)$.*

Proof: We may assume that u is a local minimum (if u is a local maximum, we can argue with $-f$). Then we obtain that $f^o(u;v) \geq 0$, $\forall\, v \in X$, which is equivalent to $0 \in \partial f(u)$. □

The result below presents the mean value property for locally Lipschitz functionals due to Lebourg [148].

Theorem 2.177. *Let $f : X \to \mathbb{R}$ be a locally Lipschitz function. Then for all $x, y \in X$, $u = x + t_0(y-x)$ with $0 < t_0 < 1$, and $\zeta \in \partial f(u)$ exist, such that*

$$f(y) - f(x) = \langle \zeta, y - x\rangle.$$

Proof: Consider the function $\theta : [0,1] \to \mathbb{R}$ defined by

$$\theta(t) = f(x + t(y-x)) + t[f(x) - f(y)], \quad \forall\, t \in [0,1].$$

The continuity of θ combined with the equalities $\theta(0) = \theta(1) = f(x)$ yields a point $t_0 \in (0,1)$ where θ assumes the minimum or maximum. By Proposition 2.176, we find that

$$0 \in \partial\theta(t_0) \subset \langle \partial f(x + t_0(y-x)), y - x\rangle + [f(x) - f(y)].$$

The conclusion of Theorem 2.177 follows. □

Another important result in the calculus with generalized gradients is the chain rule.

Theorem 2.178. *Let $F : X \to Y$ be a continuously differentiable mapping between the Banach spaces X, Y, and let $g : Y \to \mathbb{R}$ be a locally Lipschitz* *function. Then the function $g \circ F : X \to \mathbb{R}$ is locally Lipschitz, and for any point $u \in X$, the formula holds:*

$$\partial(g \circ F)(u) \subset \partial g(F(u)) \circ DF(u), \tag{2.33}$$

in the sense that every element $z \in \partial(g \circ F)(u)$ can be expressed as

$$z = DF(u)^*\zeta, \quad \textit{for some } \zeta \in \partial g(F(u)),$$

where $DF(u)^*$ denotes the adjoint operator associated with the Fréchet differential $DF(u)$ of F at u. If, in addition, F maps every neighborhood of u onto a dense subset of a neighborhood of $F(u)$, then (2.33) is satisfied with equality.

Proof: The mean value theorem for the continuously differentiable mapping F readily yields that $g \circ F$ is locally Lipschitz. According to Proposition 2.171 (ii), inclusion (2.33) is equivalent to the inequality

$$\begin{aligned}(g \circ F)^o(u; v) &\le \max\{\langle z, DF(u)v\rangle \,:\, z \in \partial g(F(u))\rangle\} \\ &= g^o(F(u); DF(u)v), \quad \forall\, v \in X\,. \end{aligned} \tag{2.34}$$

Fix $w, v \in X$ and $t > 0$. Applying Theorem 2.177 ensures the existence of $t_0, t_1 \in (0, 1)$ and $\zeta \in \partial g(F(w) + t_0(F(w + tv) - F(w)))$ such that

$$g \circ F(w + tv) - g \circ F(w) = \langle \zeta, F(w + tv) - F(w)\rangle = t\langle \zeta, DF(w + t_1tv)v\rangle.$$

Dividing by t, then letting $w \to u$ in X and $t \to 0$, and taking into account that the multifunction ∂g is upper semicontinuous from X to X^* endowed with the w^*-topology [cf. Proposition 2.171(iv)], we obtain (2.34). Assuming now that F maps an arbitrary neighborhood of u onto a dense subset of a neighborhood of $F(u)$ implies

$$\begin{aligned} g^o(F(u); DF(u)v) &= \limsup_{\substack{x \to u \\ t \downarrow 0}} \frac{g(F(x) + tDF(u)v) - g(F(x))}{t} \\ &= \limsup_{\substack{x \to u \\ t \downarrow 0}} \frac{g(F(x + tv)) - g(F(x))}{t} \\ &= (g \circ F)^o(u; v), \quad \forall\, v \in X. \end{aligned}$$

Therefore, (2.34) holds with equality, so the same is true for (2.33). □

Corollary 2.179. *Under the assumptions of the first part of Theorem 2.178, if g (or $-g$) is regular at $F(u)$, then $g \circ F$ (or $-g \circ F$) is regular at u and equality holds in (2.33).*

Proof: As $\partial(-g)(F(u)) = -\partial g(F(u))$, it is sufficient to suppose that g is regular at $F(u)$. It turns out that

$$
\begin{aligned}
g^o(F(u); DF(u)v) &= g'(F(u); DF(u)v) \\
&= \lim_{t\downarrow 0} \frac{g(F(u) + tDF(u)v) - g(F(u))}{t} \\
&= \lim_{t\downarrow 0} \left[\frac{g(F(u) + tDF(u)v) - g(F(u+tv))}{t} \right. \\
&\qquad \left. + \frac{g(F(u+tv)) - g(F(u))}{t} \right] \\
&= (g \circ F)'(u; v) \le (g \circ F)^o(u; v), \quad \forall\, v \in X.
\end{aligned}
$$

Consequently, we have equality in (2.34), so equality holds in (2.33). □

Corollary 2.180. *If a linear continuous embedding $i : X \to Y$ of the Banach space X into a Banach space Y exists, then for every locally Lipschitz function $g : Y \to \mathbb{R}$, we have*

$$\partial(g \circ i)(u) \subset i^* \partial g(i(u)), \quad \forall\, u \in X.$$

If, in addition, $i(X)$ is dense in Y, then

$$\partial(g \circ i)(u) = i^* \partial g(i(u)), \quad \forall\, u \in X.$$

Proof: One applies Theorem 2.178 for $F = i$. □

Finally, we give Aubin–Clarke's Theorem [9] of subdifferentiation under the integral sign.

Let numbers $m \ge 1$, $1 < p < +\infty$, and let T be a positive complete measure space with $|T| < \infty$, where $|T|$ stands for the measure of T. Let $j : T \times \mathbb{R}^m \to \mathbb{R}$ be a function such that $j(\cdot, y) : T \to \mathbb{R}$ is measurable whenever $y \in \mathbb{R}^m$, and satisfies either

$$|j(x, y_1) - j(x, y_2)| \le k(x)|y_1 - y_2|, \quad \text{a.a. } x \in T,\ \forall\, y_1, y_2 \in \mathbb{R}^m, \tag{2.35}$$

with a function $k \in L^q(T)$ and $1/p + 1/q = 1$, or, $j(x, \cdot) : \mathbb{R}^m \to \mathbb{R}$ is locally Lipschitz for almost all $x \in T$ and there are a constant $c > 0$ and a function $h \in L^q(T)$ such that

$$|z| \le h(x) + c|y|^{p-1}, \quad \text{a.a. } x \in T, \quad \forall\, y \in \mathbb{R}^m,\ \forall\, z \in \partial_y j(x, y). \tag{2.36}$$

The notation $\partial_y j(x, y)$ in (2.36) means the generalized gradient of j with respect to the second variable $y \in \mathbb{R}^m$; i.e., $\partial_y j(x, y) = \partial j(x, \cdot)(y)$. We introduce the functional $J : L^p(T; \mathbb{R}^m) \to \mathbb{R}$ by

$$J(v) = \int_T j(x, v(x))dx\,, \quad \forall\, v \in L^p(T; \mathbb{R}^m). \tag{2.37}$$

Theorem 2.181 (Aubin-Clarke's Theorem). *Under assumption (2.35) or (2.36), one has that the functional $J : L^p(T;\mathbb{R}^m) \to \mathbb{R}$ in (2.37) is Lipschitz continuous on the bounded subsets of $L^p(T;\mathbb{R}^m)$ and its generalized gradient satisfies*

$$\partial J(u) \subset \{w \in L^q(T;\mathbb{R}^m) : \ w(x) \in \partial_y j(x, u(x)) \ \text{ for a.e. } x \in T\}. \tag{2.38}$$

Moreover, if $j(x,\cdot)$ is regular at $u(x)$ for almost all $x \in T$, then J is regular at u and (2.38) holds with equality.

Proof: Using Hölder's inequality in conjunction with (2.35) or (2.36), one verifies easily that J is Lipschitz continuous on bounded subsets of $L^p(T;\mathbb{R}^m)$. Definition 2.161 ensures that the map $x \mapsto j^o_y(x, u(x); v(x))$ is measurable on T [see the arguments given in the proof of Theorem 2.7.2 in Clarke [68] related with the superpositional measurability of $s \mapsto j^o_y(\cdot, s; 1)$], where the subscript y indicates that the generalized directional derivative j^o is taken with respect to the second variable. Furthermore, by assumption (2.35) or (2.36), it is known that this function is integrable. Let us check the inequality

$$J^o(u; v) \leq \int_T j^o_y(x, u(x); v(x))dx\,, \quad \forall\, u, v \in L^p(T;\mathbb{R}^m). \tag{2.39}$$

If (2.35) is assumed, then Fatou's lemma leads directly to (2.39). In the case where (2.36) is admitted, Theorem 2.177 enables us to write

$$\frac{j(x, u(x) + \lambda v(x)) - j(x, u(x))}{\lambda} = \langle \zeta_x, v(x)\rangle,$$

with $\zeta_x \in \partial j(x, u^*(x))$ for some $u^*(x)$ lying on the open segment in $\mathbb{R}^m$ with endpoints $u(x)$ and $u(x) + \lambda v(x)$. Then Fatou's lemma implies (2.39). Notice that the application of Fatou's lemma is possible because of the growth condition in (2.36). In view of (2.39), any $z \in \partial J(u)$ belongs to the subdifferential at 0 of the convex function on $L^p(T;\mathbb{R}^m)$ given by

$$v \in L^p(T;\mathbb{R}^m) \mapsto \int_T j^o_y(x, u(x); v(x))dx \in \mathbb{R}.$$

The subdifferentiation under the integral for the convex integrands (see [79]) and Remark 2.170 allow us to conclude that (2.38) holds. Finally, assume further that $j(x,\cdot)$ is regular at $u(x)$ for almost all $x \in T$. Then, under either assumption (2.35) or (2.36), we may apply Fatou's lemma to get

$$\begin{aligned}\liminf_{\lambda\downarrow 0} \frac{1}{\lambda}(J(u+\lambda v) - J(u)) &\geq \int_T j'_y(x, u(x); v(x))dx \\ &= \int_T j^0_y(x, u(x); v(x))dx\,, \quad \forall\, v \in L^p(T;\mathbb{R}^m).\end{aligned}$$

Combining with (2.39), it follows that the directional derivative $J'(u; v)$ exists and $J'(u; v) = J^o(u; v)$ for every $v \in L^p(T;\mathbb{R}^m)$, thus we obtain the regularity of J at u, as well as the equality

$$J^o(u;v) = J'(u;v) = \int_T j_y'(x,u(x);v(x))dx\,, \quad \forall\, v \in L^p(T;\mathbb{R}^m).$$

Thereby, due to the regularity assumption for $j(x,\cdot)$, it is seen that (2.38) becomes an equality. □

2.5.3 Critical Point Theory

In this subsection, we present basic elements of a general critical point theory for nonsmooth functionals $I : X \to \mathbb{R} \cup \{+\infty\}$ on a real Banach space X verifying the structural hypothesis

(H) $I = \Phi + \Psi$, with $\Phi : X \to \mathbb{R}$ locally Lipschitz and $\Psi : X \to \mathbb{R} \cup \{+\infty\}$ convex, lower semicontinuous, and proper (i.e., $\not\equiv +\infty$).

For more details and developments, we refer to the works [102], [103, Chap. 4], [156], [171, Chap. 3], [173, Chap. 2].

Definition 2.182. *An element $u \in X$ is called a critical point of the functional $I : X \to \mathbb{R} \cup \{+\infty\}$ satisfying* (H) *if*

$$\Phi^o(u;v-u) + \Psi(v) - \Psi(u) \geq 0 \quad \forall\, v \in X, \tag{2.40}$$

where the notation $\Phi^o(u;\cdot)$ means the generalized directional derivative of Φ at u (see Definition 2.161).

Definition 2.182 can be expressed equivalently as follows.

Proposition 2.183. *An element $u \in X$ is a critical point of the functional $I : X \to \mathbb{R} \cup \{+\infty\}$ satisfying* (H) *if and only if $u \in D(\partial\Psi)$ and*

$$0 \in \partial\Phi(u) + \partial\Psi(u), \tag{2.41}$$

where the notations $\partial\Phi(u)$ and $\partial\Psi(u)$ stand for the generalized gradient of Φ at u and the subdifferential (in the sense of convex analysis) of Ψ at u, respectively, whereas $D(\partial\Psi)$ denotes the domain of the subdifferential $\partial\Psi$; i.e., $D(\partial\Psi) = \{x \in X : \ \partial\Psi(x) \neq \emptyset\}$.

Proof: Assume that $u \in X$ satisfies relation (2.40), or equivalently,

$$\Phi^o(u;w) + \Psi(w+u) - \Psi(u) \geq 0 \quad \forall\, w \in X.$$

It follows that 0 is a minimum point of the convex function

$$w \mapsto \Phi^o(u;w) + \Psi(w+u) - \Psi(u),$$

so $u \in D(\partial\Psi)$, and by using the subdifferential calculus for convex functions,

$$0 \in \partial(\Phi^o(u;\cdot) + \Psi(\cdot + u) - \Psi(u))(0) = \partial(\Phi^o(u;\cdot))(0) + \partial\Psi(u) = \partial\Phi(u) + \partial\Psi(u)$$

(see the last part of the statement of Proposition 2.174). Conversely, if (2.41) is satisfied, $\xi \in \partial\Phi(u)$ and $\eta \in \partial\Psi(u)$ exist such that $0 = \xi + \eta$ in X^*. Taking into account Definition 2.166 and because $\eta \in \partial\Psi(u)$, we derive

$$\Phi^o(u; v-u) + \Psi(v) - \Psi(u) \geq \langle \xi, v-u\rangle + \langle \eta, v-u\rangle = \langle \xi+\eta, v-u\rangle = 0$$

for all $v \in X$. □

Corollary 2.184. *Let $\Phi : X \to \mathbb{R}$ be a locally Lipschitz function, and let K be a nonempty, closed, convex subset of X. Denote by $I_K : X \to \mathbb{R} \cup \{+\infty\}$ the indicator function of K; i.e., $I_K(x) = 0$ whenever $x \in K$ and $I_K = +\infty$ otherwise. Then $u \in X$ is a critical point of $\Phi + I_K$ if and only if $u \in K$ and $0 \in \partial\Phi(u) + N_K(u)$, where $N_K(u) = \{\eta \in X^* : \ \langle \eta, v-u\rangle \leq 0, \ \forall v \in K\}$ is the normal cone of K at u.*

Proof: One applies Proposition 2.183 for $\Psi = I_K$. □

The examples below illustrate the concept of critical point introduced in Definition 2.182.

Example 2.185. Every local minimum $u \in X$ of a nonsmooth functional $I : X \to \mathbb{R} \cup \{+\infty\}$ satisfying (H) with $I(u) < +\infty$ is a critical point in the sense of Definition 2.182. Indeed, if $u \in X$ with $I(u) < +\infty$ is a local minimum of I, then, by convexity of Ψ, for any $v \in X$ and a small $t > 0$, we have

$$0 \leq I(u + t(v-u)) - I(u) \leq \Phi(u + t(v-u)) - \Phi(u) + t(\Psi(v) - \Psi(u)).$$

Dividing by t and letting $t \to 0^+$, we deduce that u fulfills Definition 2.182.

Example 2.186. Every minimum $u \in X$ of $\Phi|_K$ with $\Phi : X \to \mathbb{R}$ locally Lipschitz and a nonempty, closed, convex subset $K \subset X$ is a critical point of $\Phi + I_K$ in the sense of Definition 2.182. Indeed, if u is a minimum of Φ on K, then $u \in K$ and

$$\inf_X (\Phi + I_K) = (\Phi + I_K)(u) = \Phi(u),$$

and Example 2.185 leads to the desired conclusion.

Example 2.187. Every local maximum $u \in X$ of a nonsmooth functional $I : X \to \mathbb{R} \cup \{+\infty\}$ satisfying (H) with $I(u) < +\infty$ is a critical point in the sense of Definition 2.182. Indeed, under the given hypotheses, u is in the interior of the effective domain of Ψ, and thus, Ψ is Lipschitz continuous near u. Actually, $I = \Phi + \Psi$ is Lipschitz continuous near u and Proposition 2.174 yields

$$0 \in \partial I(u) = \partial(\Phi + \Psi)(u) \subset \partial\Phi(u) + \partial\Psi(u),$$

where $\partial\Phi(u)$ is the generalized gradient of Φ and $\partial\Psi(u)$ is the subdifferential of Ψ in the sense of convex analysis (see Example 2.169). According to Proposition 2.183, u is a critical point of I.

Example 2.188. Let $\Phi : X \to \mathbb{R}$ be a locally Lipschitz function. Setting $\Psi = 0$ in (H), we see by Definition 2.182 that $u \in X$ is a critical point of Φ if and only if $0 \in \partial\Phi(u)$. Therefore, in this case, Definition 2.182 reduces to the definition of Chang [64] for a critical point of a locally Lipschitz function. In particular, if $\Phi \in C^1(X)$ and $\Psi = 0$ in (H), one obtains the notion of critical point in the smooth critical point theory.

Example 2.189. Consider in assumption (H) that $\Phi \in C^1(X)$ and $\Psi : X \to \mathbb{R} \cup \{+\infty\}$ is convex, lower semicontinuous, and proper. Notice that the functional $I = \Phi + \Psi : X \to \mathbb{R} \cup \{+\infty\}$ complies with hypothesis (H). Then, according to Definition 2.182, $u \in X$ is a critical point of $I = \Phi + \Psi$ if and only if

$$\langle \Phi'(u), v - u\rangle + \Psi(v) - \Psi(u) \geq 0 \quad \forall\, v \in X;$$

i.e., $-\Phi'(u) \in \partial\Psi(u)$. Consequently, in this case, Definition 2.182 reduces to the definition of critical point as given by Szulkin [211].

We present the Palais–Smale condition for the class of nonsmooth functionals satisfying the structural hypothesis (H).

Definition 2.190. *The functional $I = \Phi + \Psi : X \to \mathbb{R} \cup \{+\infty\}$ in* (H) *is said to satisfy the Palais–Smale condition (for short,* (PS)*) if every sequence $(u_n) \subset X$ such that $(I(u_n))$ is bounded in $\mathbb{R}$ and*

$$\Phi^o(u_n; v - u_n) + \Psi(v) - \Psi(u_n) \geq -\varepsilon_n \|v - u_n\|, \quad \forall\, v \in X,$$

for a sequence (ε_n) with $\varepsilon_n \downarrow 0$, contains a strongly convergent subsequence.

Example 2.191. If $\Phi \in C^1(X)$ and $\Psi : X \to \mathbb{R} \cup \{+\infty\}$ is convex, lower semicontinuous, and proper, then Definition 2.190 coincides with the (PS) condition in the sense of Szulkin [211]. In particular, if $\Phi \in C^1(X)$ and $\Psi = 0$, then Definition 2.190 is the usual smooth (PS) condition.

We need the following result from [211].

Lemma 2.192. *Let $\chi : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous, convex function with $\chi(0) = 0$. If*

$$\chi(x) \geq -\|x\|, \quad \forall\, x \in X,$$

then $z \in X^*$ exists such that $\|z\|_{X^*} \leq 1$ and

$$\chi(x) \geq \langle z, x\rangle, \quad \forall\, x \in X.$$

Proof: Consider the following convex subsets A and B of $X \times \mathbb{R}$:

$$A = \{(x,t) \in X \times \mathbb{R} \,:\, \|x\| < -t\} \text{ and } B = \{(x,t) \in X \times \mathbb{R} \,:\, \chi(x) \leq t\}.$$

Notice that A is an open set, and due to the condition $\chi(x) \geq -\|x\|$, one has $A \cap B = \emptyset$. A well-known separation result (see [24, p. 5]) yields the existence of $\alpha, \beta \in \mathbb{R}$ and $w \in X^*$ such that $(w, \alpha) \neq (0,0)$,

$$\langle w, x\rangle - \alpha t \geq \beta\,, \quad \forall\ (x,t) \in \bar{A}$$

and

$$\langle w, x\rangle - \alpha t \leq \beta\,, \quad \forall\ (x,t) \in B.$$

We see that $\beta = 0$ because $(0,0) \in \bar{A} \cap B$. Set $t = -\|x\|$ in the first inequality above. It follows that $\langle w, x\rangle \geq -\alpha\|x\|$, $\forall\ x \in X$, which implies $\alpha > 0$ and $\|w\|_{X^*} \leq \alpha$. Set $z = \alpha^{-1}w$. Using $t = \chi(x)$, we deduce that $\langle z, x\rangle \leq \chi(x)$, $\forall\ x \in X$. As $\|w\|_{X^*} \leq \alpha$, we obtain $\|z\|_{X^*} \leq 1$. □

The following result establishes the equivalence between Definition 2.190 with $\Psi = 0$ and the (PS) condition in the sense of Chang [64].

Proposition 2.193. *A locally Lipschitz function $\Phi : X \to \mathbb{R}$ satisfies the* (PS) *condition in the sense of Definition 2.190 if and only if Φ verifies the Palais–Smale condition as defined in [64].*

Proof: Assume that the locally Lipschitz function $\Phi : X \to \mathbb{R}$ satisfies the (PS) condition formulated in Definition 2.190. Let a sequence $(u_n) \subset X$ with $\Phi(u_n)$ bounded and for which

$$\lambda(u_n) = \inf_{w \in \partial\Phi(u_n)} \|w\|_{X^*} \to 0 \ \text{ as } n \to \infty.$$

It is known from Proposition 2.171 (i) that an element $z_n \in \partial\Phi(u_n)$ can be found such that $\lambda(u_n) = \|z_n\|_{X^*}$. As

$$\Phi^o(u_n; v) \geq \langle z_n, v\rangle \geq -\|z_n\|_{X^*}\ \|v\|, \quad \forall\ v \in X,$$

the inequality in Definition 2.190 (with $\Psi = 0$) is verified with $\varepsilon_n = \|z_n\|$. It implies that (u_n) possesses a convergent subsequence, which ensures the Palais–Smale condition in the sense of [64].

Conversely, we suppose that Φ verifies the Palais–Smale condition in the sense of [64]. Let (u_n) be a sequence as in Definition 2.190. We can apply Lemma 2.192 for $\chi = \frac{1}{\varepsilon_n}\Phi^o(u_n; \cdot)$, which gives an element $w_n \in X^*$ with $\|w_n\|_{X^*} \leq 1$ and

$$\frac{1}{\varepsilon_n}\Phi^o(u_n; x) \geq \langle w_n, x\rangle,\ \forall\ x \in X.$$

It follows that $\varepsilon_n w_n \in \partial\Phi(u_n)$ and $\varepsilon_n w_n \to 0$ in X^* as $n \to \infty$. According to the Palais–Smale condition in [64], we have that (u_n) contains a convergent subsequence, so the (PS) condition in the sense of Definition 2.190 holds. □

2.5.4 Linking Theorem

The objective of this subsection is to provide sufficient conditions for the existence of critical points in the setting of functionals of type (H) in Sect. 2.5.3. We start with the following minimization result.

Theorem 2.194. *Assume that the function* $I = \Phi + \Psi : X \to \mathbb{R} \cup \{+\infty\}$ *satisfies hypothesis* (H), *is bounded from below, and verifies the* (PS) *condition. Then* $u \in X$ *exists such that* $I(u) = \inf_X I \in \mathbb{R}$ *and* u *is a critical point of* I *in the sense of Definition 2.182.*

Proof: Denote $m = \inf_X I \in \mathbb{R}$. We find a (minimizing) sequence $(u_n) \subset X$ such that

$$I(u_n) < m + \varepsilon_n^2,$$

for a sequence (ε_n) of positive numbers, with $\varepsilon_n \downarrow 0$. Applying Ekeland's variational principle (cf. [91]) to the function I, a sequence $(v_n) \subset X$ exists such that

$$I(v_n) < m + \varepsilon_n^2$$

and

$$I(v) \geq I(v_n) - \varepsilon_n \|v_n - v\|, \quad \forall\, v \in X,\ \forall\, n \in \mathbb{N}.$$

Setting $v = (1-t)v_n + tw$ in the above inequality, for arbitrary $0 < t < 1$ and $w \in X$, we obtain

$$\begin{aligned} &\Phi((1-t)v_n + tw) + \Psi((1-t)v_n + tw) \\ &\qquad \geq \Phi(v_n) + \Psi(v_n) - \varepsilon_n t \|w - v_n\|, \quad \forall\, w \in X,\ \forall\, t \in (0,1). \end{aligned}$$

The convexity of $\Psi : X \to \mathbb{R} \cup \{+\infty\}$ yields

$$\begin{aligned} &\Phi((1-t)v_n + tw) - t\Psi(v_n) + t\Psi(w) \\ &\qquad \geq \Phi(v_n) - \varepsilon_n t \|w - v_n\|, \quad \forall\, w \in X,\ \forall\, t \in (0,1). \end{aligned}$$

Dividing by t and letting $t \downarrow 0$, we deduce that for all $w \in X$, one has

$$\begin{aligned} &\Phi^o(v_n; w - v_n) + \Psi(w) - \Psi(v_n) \\ &\quad \geq \limsup_{t \downarrow 0} \frac{1}{t}(\Phi(v_n + t(w - v_n)) - \Phi(v_n)) + \Psi(w) - \Psi(v_n) \geq -\varepsilon_n \|w - v_n\|. \end{aligned}$$

On the other hand, we have $\Phi(v_n) + \Psi(v_n) \to m$ as $n \to \infty$. Then the (PS) condition (see Definition 2.190) implies that along a relabelled subsequence $v_n \to u$ in X, for some $u \in X$. The lower semicontinuity of I yields $I(u) \leq \liminf_{n\to\infty} I(v_n) \leq m$, so $I(u) = m$. Making use of Example 2.185, we derive that u is a critical point of I. □

We now focus on the existence of critical points for functionals of type (H) that are not obtained by minimization, thus, saddle-points. The subsequent minimax principle makes use of the notion of linking as given in [88].

Definition 2.195. *Let S be a nonempty closed subset of the Banach space X, and let Q be a compact topological submanifold of X with nonempty boundary ∂Q (in the sense of manifolds with boundary). We say that S and Q link if $S \cap \partial Q = \emptyset$ and $f(Q) \cap S \neq \emptyset$ whenever $f \in \Gamma$, where*

$$\Gamma := \{f \in C(Q, X) : \; f|_{\partial Q} = \mathrm{id}_{\partial Q}\}.$$

Example 2.196. Let $X = E \times \mathbb{R}$, with a Banach space E, and let $0 < \rho < r$. The sets $S = E \times \{\rho\}$ and $Q = \{(0, tr) \in E \times \mathbb{R} : t \in [0, 1]\}$ link.

The following result given in [171, Chap. 3] provides critical points of saddle-point type for nonsmooth functionals having the structure in (H).

Theorem 2.197. *Let the functional $I : X \to \mathbb{R} \cup \{+\infty\}$ satisfy assumptions* (H) *and* (PS)*. Let S and Q link in the sense of Definition 2.195. Assume further that*

$$\sup_Q I \in \mathbb{R}, \;\; b := \inf_S I \in \mathbb{R}, \;\; a := \sup_{\partial Q} I < b.$$

Then the number

$$c := \inf_{f \in \Gamma} \sup_{x \in Q} I(f(x)),$$

with Γ in Definition 2.195, is a critical value of I; that is, there is a critical point u of I in the sense of Definition 2.182 and $I(u) = c$. Moreover, $c \geq b$.

Proof: The inequality $c \geq b$ is a direct consequence of linking property in Definition 2.195. Arguing by contradiction we assume that c is not a critical value of I. Applying the deformation result in Theorem 3.1 in [171], with $\bar{\varepsilon} = c - a > 0$, we get an $\varepsilon \in (0, \bar{\varepsilon})$ as stated therein. Define Γ_1 as the set of all continuous mappings $\varphi : Q \to X$ such that

$$\varphi(\partial Q) \subset \left\{x \in X : \, I(x) \leq c - \frac{\varepsilon}{2}\right\}$$

and $\varphi|_{\partial Q}$, $\mathrm{id}_{\partial Q}$ are homotopic maps from ∂Q into $\left\{x \in X : I(x) \leq c - \frac{\varepsilon}{4}\right\}$. We have $\mathrm{id}_Q \in \Gamma_1$. Using the definitions of c and Γ_1, we obtain

$$c = \inf_{\varphi \in \Gamma_1} \sup_{x \in Q} I(\varphi(x)). \tag{2.42}$$

It is seen that Γ_1 is a closed subset of the Banach space $C(Q; X)$ with respect to the uniform norm $\|\varphi\| = \sup_{x \in Q} \|\varphi(x)\|$. Consider the lower semicontinuous functional $\Pi : \Gamma_1 \to \mathbb{R} \cup \{+\infty\}$ defined by

$$\Pi(\varphi) = \sup_{x \in Q} I(\varphi(x)), \quad \forall \varphi \in \Gamma_1.$$

Taking into account (2.42), Ekeland's variational principle [91] applied to the function Π on Γ_1 yields a $\varphi \in \Gamma_1$ satisfying $c \leq \Pi(\varphi) \leq c + \varepsilon$ and

$$\Pi(\psi) - \Pi(\varphi) \geq -\varepsilon\|\psi - \varphi\|, \quad \forall\, \psi \in \Gamma_1\,. \tag{2.43}$$

Theorem 3.1 in [171] provides a deformation $h_\varphi : W \times [0, \bar{s}] \to X$ corresponding to the compact set $A = \varphi(Q)$, where W is a closed neighborhood of A in X and $\bar{s}$ is a positive number, satisfying

$$\|v - h_\varphi(v, s)\| \leq s\,, \quad \forall\, v \in W,\ \forall\, s \in [0, \bar{s}], \tag{2.44}$$

$$\sup_{v\in A} I(h_\varphi(v, s)) - \sup_{v\in A} I(v) \leq -2\varepsilon s\,, \quad \forall\, s \in [0, \bar{s}]. \tag{2.45}$$

Let us show that for $\bar{s} > 0$ small enough, we have

$$h_\varphi(\varphi(\cdot), s) \in \Gamma_1\,, \quad \forall\, s \in [0, \bar{s}]. \tag{2.46}$$

To prove (2.46), it suffices to note that $h_\varphi(\varphi(\cdot), s)|_{\partial Q}$ and $\varphi|_{\partial Q}$ are homotopic maps from ∂Q into $\left\{x \in X : I(x) \leq c - \frac{\varepsilon}{2}\right\}$. Such a homotopy is $(x, t) \mapsto h_\varphi(\varphi(x), ts)$ as can be seen from (2.44) and (2.46). It follows from (2.45), (2.46), (2.43), and (2.44) that

$$\begin{aligned} -2\varepsilon s &\geq \Pi(h_\varphi(\varphi(\cdot), s)) - \Pi(\varphi) \\ &\geq -\varepsilon\|h_\varphi(\varphi(\cdot), s) - \varphi\| \geq -\varepsilon s\,, \quad \forall\ \ 0 \leq s \leq \bar{s}\,. \end{aligned}$$

This contradiction proves that our initial assumption that c is not a critical value of I is not possible, which completes the proof. □

Corollary 2.198. *Let E be a Banach space, let $\Phi : E \times \mathbb{R} \to \mathbb{R}$ be locally Lipschitz, and let $\Psi : E \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ be proper, convex, and lower semicontinuous. Suppose that the function*

$$F = \Phi + \Psi : E \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$$

satisfies the (PS) *condition and positive numbers ρ and r exist with $\rho < r$ such that $F(0, 0) \leq 0$, $F(0, r) \leq 0$, $0 < \inf_{v\in E} F(v, \rho) < +\infty$. Then*

$$c = \inf\{ \sup_{t\in[0,1]} F(g(t)) : \ g \in C([0, 1], E \times \mathbb{R}),\ g(0) = (0, 0),\ g(1) = (0, r)\}$$

is a critical value of F, and we have the estimate

$$\inf_{v\in E} F(v, \rho) \leq c \leq \sup_{t\in[0,1]} F(0, tr).$$

Proof: Apply Theorem 2.197 with $X = E \times \mathbb{R}$ and $I = F$ using the linking in Example 2.196. The last inequality above is obtained by taking the path $g(t) = (0, tr)$ for all $t \in [0, 1]$. □

Remark 2.199. In the general setting of nonsmooth functionals verifying hypothesis (H), Theorem 2.197 incorporates important minimax results in the critical point theory such as the mountain-pass theorem, saddle-point theorem, and generalized mountain-pass theorem.

Remark 2.200. A basic hypothesis in Theorem 2.197 is that $b > a$. Under the situation of linking in Theorem 2.197, we then have $c > a$. As from the expression of the minimax value c it is seen that always $c \geq a$, the situation that is not covered by Theorem 2.197 is the so-called liming case $c = a$. The specific situation $c = a$ is treated in [156].

http://www.springer.com/978-0-387-30653-7

Nonsmooth Variational Problems and Their Inequalities
Comparison Principles and Applications
Carl, S.; Le, V.K.; Motreanu, D.
2007, XII, 395 p. 30 illus., Hardcover
ISBN: 978-0-387-30653-7