Springer Monographs in Mathematics

Nikolaos S. Papageorgiou
Vicenţiu D. Rădulescu
Dušan D. Repovš

Nonlinear Analysis—Theory and Methods

Springer

Springer Monographs in Mathematics

This series publishes advanced monographs giving well-written presentations of the “state-of-the-art” in fields of mathematical research that have acquired the maturity needed for such a treatment. They are sufficiently self-contained to be accessible to more than just the intimate specialists of the subject, and sufficiently comprehensive to remain valuable references for many years. Besides the current state of knowledge in its field, an SMM volume should ideally describe its relevance to and interaction with neighbouring fields of mathematics, and give pointers to future directions of research.

More information about this series at http://www.springer.com/series/3733

Nikolaos S. Papageorgiou ·
Vicenţiu D. Rădulescu ·
Dušan D. Repovš

Nonlinear Analysis–Theory and Methods

Springer

Nikolaos S. Papageorgiou
Department of Mathematics
National Technical University
Athens, Greece

Institute of Mathematics, Physics
and Mechanics
Ljubljana, Slovenia

Dušan D. Repovš
Faculty of Education, Faculty
of Mathematics and Physics
University of Ljubljana
Ljubljana, Slovenia

Institute of Mathematics, Physics
and Mechanics
Ljubljana, Slovenia

Vicenţiu D. Rădulescu
Institute of Mathematics, Physics
and Mechanics
Ljubljana, Slovenia

Faculty of Applied Mathematics
AGH University of Science
and Technology
Kraków, Poland

ISSN 1439-7382 ISSN 2196-9922 (electronic)
Springer Monographs in Mathematics
ISBN 978-3-030-03429-0 ISBN 978-3-030-03430-6 (eBook)
https://doi.org/10.1007/978-3-030-03430-6

Library of Congress Control Number: 2018968386

Mathematics Subject Classification (2010): 35-02, 49-02, 58-02

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Introduction

Insight must precede application.
Max Planck (1858–1947),
Nobel Prize in Physics 1918

Answering the needs of specific applied problems, Nonlinear Analysis emerged as a separate field of research within Mathematical Analysis immediately after World War II, when Linear Functional Analysis (primarily Banach space theory) had reached a rather mature stage. Years of accumulated experience has convinced people that theory can no longer afford the luxury of dealing with linear, smooth and well-posed models. In many applications of interest, such requirements either exclude many important aspects of the problem or, even more dramatically, fail completely to provide a satisfactory model for the phenomena under investigation. Such considerations led to the development of Nonlinear Analysis, which today has developed significantly and is one of the most active areas of research. The advent of Nonlinear Analysis led to unifying theories describing different classical problems and permitted the investigation of a whole new range of applications. The theories, methods and techniques of Nonlinear Analysis proved to be indispensable tools in the analysis of various problems in many other fields. For this reason, Nonlinear Analysis eventually acquired an interdisciplinary character and it is a prerequisite for many nonmathematicians who want to conduct an in depth analysis of problems they face. This leads to an increasing demand for books that summarize the recent developments in various parts of Nonlinear Analysis.

Make no mistake, Nonlinear Analysis is a very broad subject and every such book focuses only on a part of it. Here the emphasis is on those aspects that are useful in the study of boundary value problems. In fact, Volume II will be devoted to the study of such problems. Given the orientation of this book project, it is natural to start in Chap. 1 with *Sobolev Spaces*, which are the main tools in the analysis of both stationary and nonstationary problems. Sobolev spaces play a central role in the modern theory of partial differential equations and they lead to a significant broadening of the notion of solution of a boundary value problem. They provide a natural analytical framework for the study of linear and nonlinear boundary value problems. We provide a concise but complete introduction to the subject, emphasizing those parts of the theory which are relevant to the study of boundary value problems. We deal with both functions of one and several variables. In the last section we also discuss capacities, which arise in the study of small sets in $\mathbb{R}^N$ and of the fine properties of Sobolev functions. We also present some related results which will be of interest to people dealing with boundary value problems.

In Chap. 2 we deal with *Compact Operators and Operators of Monotone Type*. Compact operators are in fact the starting point of Nonlinear Analysis, going back to the celebrated work of Leray and Schauder in the 1930s. Compactness was introduced as a first attempt to deal with infinite-dimensional nonlinear operator equations, since by its nature compactness (in all its forms) approximates infinite objects by finite objects. In parallel we develop the corresponding linear theory, leading to the spectral theorem for compact self-adjoint operators on a Hilbert space. This theorem is the basis of the spectral analysis of linear elliptic operators under different boundary conditions. Of course, compact operators have serious limitations, which researchers tried to overcome by introducing new classes of nonlinear maps. A broader framework for the analysis of infinite-dimensional problems is provided by monotone operators, which extend to an infinite-dimensional context the classical notion of an increasing real function. Monotone operators are rooted in variational problems. Of special interest are the so-called *maximal monotone operators*, which exhibit remarkable surjectivity properties. However, the development of a coherent theory of maximal monotone maps leads necessarily to multivalued maps (multifunctions). For this reason, in Sect. 2.5 we have a detour to *Set-Valued Analysis*. We point out that multivalued analysis provides basic tools in many applied areas such as optimization, optimal control, mathematical economics, game theory, etc. The "differential" theory of nonsmooth convex functions leads to a special class of multivalued maximal monotone operators (convex subdifferential). At the end of the chapter we also discuss useful generalizations of the notion of monotonicity.

In Chap. 3, we conduct a detailed study of the main degree theories. We start with Brouwer's theory (finite-dimensional spaces), following the analytical approach that goes back to the work of Nagumo in the 1950s. Brouwer's original approach to the definition of the degree was based on combinatorial and algebraic topology. Since most problems of interest are infinite-dimensional, it was necessary to extend Brouwer's theory to infinite-dimensional maps. This was done by Leray and Schauder in the 1930s, who used compact operators (namely operators of the form $I - K$ with I being the identity map and K a compact operator). The Leray–Schauder degree theory and its consequences are examined in Sect. 3.2. After that we examine degree theories for set-valued maps and for operators of monotone type. All these are recent theories and were introduced to deal with infinite-dimensional nonlinear problems for which the Leray–Schauder theory fails to address. At the end of the chapter, we also discuss some alternative generalizations of the Leray–Schauder theory using measures of noncompactness (condensing maps) and we examine the index of a ξ-point.

Chapter 4 deals with *Fixed Point Theory* and with some important *Variational Principles*. There is an informal classification of fixed-point theorems to “metric fixed points”, “topological fixed points” and “order fixed points”. Since the latter class involves some order structure in the underlying space (usually a Banach space), in Sect. 4.1 we start with a general discussion of cones and of the partial order they induce on the ambient space. Then we start discussing the three aforementioned classes of fixed points. Metric fixed-point theorems are always formulated in a metric space setting and the methods involved in their study exploit the metric structure and geometry of the spaces involved together with the metric properties of the maps. In contrast to the topological fixed point theory, the topological properties of the spaces and/or of the maps are involved. In particular, the notion of compactness is important in our considerations there. In “order fixed point theory”, the order on the space induced by a cone is the main ingredient and the basic hypotheses and conditions are based around this notion, as well as the notion of the “Leray–Schauder fixed point index”. We also discuss fixed points of multifunctions. In Sect. 4.6 we discuss some important abstract variational principles. Special emphasis is given on the celebrated “Ekeland variational principle” and its many interesting consequences. We conclude the chapter with a discussion of Young measures that arise in a large class of variational problems. When the direct method of the calculus of variations fails, the minimizing sequences (or appropriate subsequences of them) have a limit behavior (usually more and more oscillatory), which is captured by embedding the original functions in the space of Young measures (or parametric probabilities). This is the process of “relaxation”, familiar to people studying optimal control problems.

In Chap. 5 we study *Critical Point Theory*. When using variational methods, we are trying to find solutions of a given nonlinear equation, by looking for critical stationary points of a functional (energy or Euler functional) defined on the function space in which we want the solutions to be. If this functional is bounded from below, as above, we can look for local extrema and the direct method enters into play. If the functional is indefinite, we cannot expect to have local extrema and so other methods for locating critical points need to be found. These methods are based on minimax principles, which lead to critical points. These minimax methods are derived either using the deformation approach or the Ekeland variational principle. Here we follow the deformation approach which uses the change of the topological structure of the sublevel sets of the Euler functional along the flow produced by a kind of negative gradient vector field. We conduct a detailed study of critical point theory including an analysis of the structure of the critical set at the end of the chapter.

In Chap. 6, continuing the theme of locating and counting critical points of a given functional, we discuss *Morse Theory* and *Critical Groups*, which provide the tools to prove multiplicity theorems. Since these topics make use of tools from Algebraic Topology, in Sect. 6.1 we review the needed background from that field. Then we proceed with a self-contained presentation of the Morse theory related to the study of the existence and multiplicity of solutions for variational problems. Our presentation is rather complete, including continuity and homotopy invariance properties, which are valuable tools in the computation of critical groups.

Finally, we mention that each chapter has a final section called *Remarks*, which discusses the literature on the subject. We have tried to include a rather complete bibliography, although such a task seems impossible given the volume of the existing literature. We acknowledge the support of the Slovenian Research Agency program P1-0292 and grants J1-7025, J1-8131, N1-0064, and N1-0083 and ....

September 2018

Nikolaos S. Papageorgiou
Vicenţiu D. Rădulescu
Dušan D. Repovš

Keywords

Nonlinear elliptic equation • Boundary value problem • Critical point theory • Morse theory • Critical groups

Chapter 1
Sobolev Spaces

Pour atteindre les limites du possible, il faut rêver l'impossible.
René Thom (1923–2002), Fields Medal 1958

Sobolev spaces have been very important in the development of partial differential equations. They are based on the notion of "weak derivative", which defines partial derivatives for L^p-functions which are not differentiable in the classical sense. The weak derivative is based on the simple idea of integration by parts. In this way we transfer the burden of differentiation from a "bad" (nonsmooth) function, to a "good" (smooth) function. In this chapter, we do not conduct an exhaustive study of Sobolev spaces. This can be found in the specialized books included in the bibliography (see the Remarks at the end of the chapter). Instead, we focus on those items of the theory which are essential in the study of boundary value problems and which we will use in later chapters.

Throughout this chapter, Ω is a nonempty open set in $\mathbb{R}^N$ ($N \geqslant 1$). Additional conditions on Ω will be introduced as needed.

1.1 Definitions, Density, and Approximation Results

First we fix some standard notation and terminology. An element $\alpha \in \mathbb{N}^N$, $\alpha = (\alpha_k)_{k=1}^N$, is called a "multi-index". Given a multi-index α and an element $\hat{z} = (z_k)_{k=1}^N \in \mathbb{R}^N$, we introduce the following notation:

$$|\alpha| = \sum_{k=1}^{N} \alpha_k \text{ (the length of the multi-index)},$$

$$z^\alpha = z_1^{\alpha_1} \dots z_N^{\alpha_N}, \ \ D^\alpha = \frac{\partial^{|\alpha|}}{\partial z_1^{\alpha_1} \dots \partial z_N^{\alpha_N}}.$$

N. S. Papageorgiou et al., *Nonlinear Analysis–Theory and Methods*,
Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-030-03430-6_1

Mollifiers and regularizations by convolutions are two useful tools in the study of Sobolev spaces, since they allow us to approximate L^p-functions by smooth ones.

Definition 1.1.1 Let $\rho : \mathbb{R}^N \to [0, +\infty)$ be a bounded function such that

$$\operatorname{supp} \rho \subseteq \overline{B}_1(0) = \{z \in \mathbb{R}^N : \|z\|_{\mathbb{R}^N} \leqslant 1\} \text{ and } \int_{\mathbb{R}^N} \rho(z)dz = 1.$$

For $\varepsilon \in (0, 1)$ we set $\rho_\varepsilon(z) = \frac{1}{\varepsilon^N} \rho\left(\frac{z}{\varepsilon}\right)$ for all $z \in \mathbb{R}^N$. Evidently $\operatorname{supp} \rho_\varepsilon \subseteq \overline{B}_\varepsilon(0)$ and the functions $\{\rho_\varepsilon\}_{\varepsilon \in (0,1)}$ are called "mollifiers". Given $u \in L^1_{\text{loc}}(\mathbb{R}^N)$, we consider the convolution

$$u_\varepsilon(z) = (u * \rho_\varepsilon)(z) = \int_{\mathbb{R}^N} u(y)\rho_\varepsilon(z - y)dy \text{ for all } z \in \mathbb{R}^N.$$

The function $u_\varepsilon(\cdot)$ is a "mollification" or "regularization" of $u(\cdot)$.

Remark 1.1.2 The following function $\rho(\cdot)$ can be used to generate mollifiers:

$$\rho(z) = \begin{cases} c \exp\left(\dfrac{1}{\|z\|^2_{\mathbb{R}^N} - 1}\right) & \text{if } \|z\|_{\mathbb{R}^N} \leqslant 1 \\ 0 & \text{if } \|z\|_{\mathbb{R}^N} > 1. \end{cases}$$

Here $c > 0$ is such that $\int_{\mathbb{R}^N} \rho(z)dz = 1$. The corresponding family of mollifiers are called "standard mollifiers". Usually this is the family of mollifiers that we employ. When $u \in L^1_{\text{loc}}(\Omega)$, we may define

$$u_\varepsilon(z) = (u * \rho_\varepsilon)(z) = \int_\Omega u(y)\rho_\varepsilon(z - y)dy \text{ for all } z \in \Omega_\varepsilon,$$

where Ω_ε is the open set given by $\Omega_\varepsilon = \{z \in \Omega : d(z, \partial\Omega) > \varepsilon\}$. If $z \in \Omega$, then $u_\varepsilon(z)$ is well-defined for all $0 < \varepsilon < d(z, \partial\Omega)$ and so it makes sense to consider the limit $\lim_{\varepsilon \to 0^+} u_\varepsilon(z)$. In what follows we denote by $C_c(\Omega)$ the space of continuous functions with compact support.

Proposition 1.1.3 *(a) If $u \in C(\mathbb{R}^N)$, then $u_\varepsilon \to u$ as $\varepsilon \to 0^+$ uniformly on compact sets.*

(b) If $u \in L^p(\mathbb{R}^N)$, $1 \leqslant p < \infty$, then $u_\varepsilon \to u$ as $\varepsilon \to 0^+$ in $L^p(\mathbb{R}^N)$ and pointwise at every Lebesgue point of u.

Proof (a) Let $K \subseteq \mathbb{R}^N$ be a compact set. Then given $\eta > 0$, we can find $\delta = \delta(\eta, k) > 0$ such that

$$|u(z - y) - u(z)| < \eta \text{ for all } z \in K \text{ and all } \|y\|_{\mathbb{R}^N} < \delta. \tag{1.1}$$

We have

$$
\begin{aligned}
u_\varepsilon(z) - u(z) &= \int_{\mathbb{R}^N} [u(z-y) - u(z)]\rho_\varepsilon(y)dy \text{ (since } \int_{\mathbb{R}^N} \rho_\varepsilon(y)dy = 1) \\
&= \int_{B_\varepsilon(0)} [u(z-y) - u(z)]\rho_\varepsilon(y)dy \text{ (since } \operatorname{supp}\rho_\varepsilon \subseteq B_\varepsilon(0)).
\end{aligned}
$$

Then from (1.1) and for $\varepsilon \in (0, \delta)$, we have

$$
\begin{aligned}
&|u_\varepsilon(z) - u(z)| \leqslant \eta \int_{\mathbb{R}^N} \rho_\varepsilon(y)dy = \eta \text{ for all } z \in K \\
&\Rightarrow u_\varepsilon \to u \text{ as } \varepsilon \to 0^+ \text{ uniformly on compact sets in } \mathbb{R}^N.
\end{aligned}
$$

(b) Exploiting the density of $C_c(\mathbb{R}^N)$ in $L^p(\mathbb{R}^N)$, given $\eta > 0$, we can find $\hat{u} \in C_c(\mathbb{R}^N)$ such that $\|u - \hat{u}\|_p < \eta$. From part (a), we know that

$$
\hat{u}_\varepsilon \to \hat{u} \text{ as } \varepsilon \to 0^+ \text{ uniformly on compact subsets of } \mathbb{R}^N. \tag{1.2}
$$

We have

$$
\begin{aligned}
&\operatorname{supp}\hat{u}_\varepsilon \subseteq \operatorname{supp}\hat{u} + \overline{B}_\varepsilon(0) \subseteq \operatorname{supp}\hat{u} + \overline{B}_1(0) \text{ a compact subset of } \mathbb{R}^N \\
&\Rightarrow \|\hat{u}_\varepsilon - \hat{u}\|_p \to 0 \text{ as } \varepsilon \to 0^+ \text{ (see (1.3))}. \qquad (1.3)
\end{aligned}
$$

Then

$$
\begin{aligned}
\|u_\varepsilon - u\|_p &\leqslant \|u_\varepsilon - \hat{u}_\varepsilon\|_p + \|\hat{u}_\varepsilon - \hat{u}\|_p + \|\hat{u} - u\|_p \\
&\leqslant 2\|\hat{u} - u\|_p + \|\hat{u}_\varepsilon - \hat{u}\|_p \leqslant 3\eta \text{ for all } \varepsilon \in (0,1) \text{ small (see (1.3))}.
\end{aligned}
$$

Therefore we conclude that $u_\varepsilon \to u$ in $L^p(\mathbb{R}^N)$ as $\varepsilon \to 0^+$.

Finally, if $z \in \Omega$ is a Lebesgue point of u, we have

$$
\lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon^N} \int_{B_\varepsilon(z)} |u(y) - u(z)|dy = 0.
$$

But note that

$$
\begin{aligned}
|u_\varepsilon(z) - u(z)| &= \left| \int_\Omega u(y)\rho_\varepsilon(z-y)dy - u(z) \right| \\
&\leqslant \|\rho\|_\infty \frac{1}{\varepsilon^N} \int_{B_\varepsilon(z)} |u(y) - u(z)|dy \to 0 \text{ as } \varepsilon \to 0^+.
\end{aligned}
$$

So, we deduce the pointwise convergence on the set of Lebesgue points of u (recall that this set has full measure). $\square$

Remark 1.1.4 If $p = \infty$, then part (b) of the above proposition is no longer true since the uniform limit of continuous functions is continuous. Also, given $u \in C(\overline{\Omega})$, in general $\|u_\varepsilon - u\|_{C(\overline{\Omega})} \nrightarrow 0$ as $\varepsilon \to 0^+$. However, the following is true:

"If $u \in C(\overline{\Omega})$ and $\Omega_0 \subset\subset \Omega$ (that is, $\overline{\Omega}_0$ is compact and $\overline{\Omega}_0 \subseteq \Omega$), then $\|u_\varepsilon - u\|_{C(\overline{\Omega}_0)} \to 0$ as $\varepsilon \to 0^+$."

In the study of Sobolev spaces, the space $C_c^\infty(\Omega)$ of C^∞-functions with compact support (test functions) is important. From the theory of L^p-spaces we know that

$$C_c^\infty(\Omega) \text{ is dense in } L^p(\Omega) \text{ for all } 1 \leqslant p < \infty.$$

Proposition 1.1.5 *If* $u \in L^1_{\mathrm{loc}}(\Omega)$ *and* $\int_\Omega u(z)\vartheta(z)\,dz = 0$ *for all* $\vartheta \in C_c^\infty(\Omega)$, *then* $u(z) = 0$ *a.e. in* Ω.

Proof Let $g \in L^\infty(\mathbb{R}^N)$ be a function with compact support contained in Ω and let $g_n = g * \rho_n$ (here $\rho_n = \rho_{1/n}$). Then for $n \geqslant 1$ large we have $g_n \in C_c^\infty(\Omega)$. Hence by hypothesis

$$\int_\Omega u g_n\,dz = 0 \text{ for all } n \geqslant 1 \text{ large.}$$

From Proposition 1.1.3 (b), we know that $g_n \to g \in L^1(\mathbb{R}^N)$ and so, by passing to a suitable subsequence if necessary, we may assume that $g_n(z) \to g(z)$ for a.a. $z \in \mathbb{R}^N$. Also, we have $\|g_n\|_\infty \leqslant \|g\|_\infty$ for all $n \geqslant 1$. Therefore, applying the Lebesgue dominated convergence theorem, we have

$$\int_\Omega ug\,dz = 0. \tag{1.4}$$

Given K a compact set in Ω, we set $g(z) = \begin{cases} \text{sign}\,u & \text{if } z \in K \\ 0 & \text{if } z \in \mathbb{R}^N \setminus K. \end{cases}$ Then from (1.4) we have $\int_K |u|\,dz = 0$. Since K is arbitrary, we conclude that $u(z) = 0$ a.e. in Ω. $\square$

We now introduce the basic notion behind the Sobolev spaces, namely that of weak derivatives.

Definition 1.1.6 Let α be a multi-index and suppose that for $u, v \in L^1_{\mathrm{loc}}(\Omega)$ we have

$$\int_\Omega u(D^\alpha\vartheta)dz = (-1)^{|\alpha|}\int_\Omega v\vartheta dz \text{ for all } \vartheta \in C_c^\infty(\Omega).$$

Then v is called the "weak" or "distributional" partial derivative of u and is denoted by $D^\alpha u$. Clearly v is uniquely defined up to sets of measure zero.

Remark 1.1.7 From the above definition, we see that in order to define $D^\alpha u$, we do not need the existence of derivatives of smaller order (as is the case with classical derivatives). Moreover, for smooth functions, the weak and classical derivatives coincide. Finally, note that the weak derivative is a global notion and, unlike the classical derivative, it cannot be defined at a point.

Proposition 1.1.8 *If $u \in L^1_{\mathrm{loc}}(\Omega)$, then for all $\varepsilon \in (0,1)$, $u_\varepsilon \in C^\infty(\Omega_\varepsilon)$ (recall that $\Omega_\varepsilon = \{z \in \Omega : d(z, \partial\Omega) > \varepsilon\}$) and for every multi-index α, we have*

$$\frac{\partial^\alpha u_\varepsilon}{\partial z^\alpha}(z) = \left(u * \frac{\partial^\alpha \rho_\varepsilon}{\partial z^\alpha}\right)(z) = \int_{\mathbb{R}^N} u(y)\frac{\partial^\alpha \rho_\varepsilon}{\partial z^\alpha}(z-y)dy \text{ for all } z \in \Omega_\varepsilon.$$

Proof We fix $z \in \Omega_\varepsilon$ and $\lambda \in (0, d(z, \partial\Omega) - \varepsilon)$. Let $\{e_k\}_{k=1}^N$ be the canonical basis of $\mathbb{R}^N$. For every $h \in \mathbb{R}$ with $0 < |h| < \lambda$, we have

$$\begin{aligned}
&\frac{u_\varepsilon(z+he_k) - u_\varepsilon(z)}{h} - \int_\Omega u(y)\frac{\partial \rho_\varepsilon}{\partial z_k}(z-y)dy \\
&= \int_\Omega \left[\frac{\rho_\varepsilon(z+he_k-y) - \rho_\varepsilon(z-y)}{h} - \frac{\partial \rho_\varepsilon}{\partial z_k}(z-y)\right]u(y)dy \\
&= \int_\Omega \left(\frac{1}{h}\int_0^h \frac{\partial \rho_\varepsilon}{\partial z_k}(z-y+se_k)ds - \frac{\partial \rho_\varepsilon}{\partial z_k}(z-y)\right)u(y)dy \\
&= \frac{1}{h}\int_0^h \int_{B_{\varepsilon+\lambda}(z)} \left(\frac{\partial \rho_\varepsilon}{\partial z_k}(z-y+se_k) - \frac{\partial \rho_\varepsilon}{\partial z_k}(z-y)\right)u(y)dy\, ds \\
&\qquad\qquad \text{(by Fubini's theorem and since } \operatorname{supp}\rho_\varepsilon \subseteq \overline{B}_\varepsilon(0)).
\end{aligned}$$

Since $u_\varepsilon \in C_c^\infty(\mathbb{R}^N)$, its partial derivatives are uniformly continuous functions. So, given any $\eta > 0$, we can find $\delta = \delta(z, \lambda, \eta, \varepsilon) > 0$ such that

$$\left|\frac{\partial \rho_\varepsilon}{\partial z_k}(v) - \frac{\partial \rho_\varepsilon}{\partial z_k}(y)\right| \leqslant \frac{\eta}{1 + \|u\|_{L^1(B_{\varepsilon+\lambda}(z))}} \text{ for all } v, y \in B_{\varepsilon+\lambda}(z) \text{ with } \|v-y\| \leqslant \delta.$$

Then for $0 < |h| \leqslant \min\{\lambda, \delta\}$, we have

$$\begin{aligned}
&\left|\frac{u_\varepsilon(z+he_k) - u_\varepsilon(z)}{h} - \int_\Omega \frac{\partial \rho_\varepsilon}{\partial z_k}(z-y)u(y)dy\right| \leqslant \eta \\
&\Rightarrow \frac{\partial u_\varepsilon}{\partial z_k}(z) = \int_\Omega \frac{\partial \rho_\varepsilon}{\partial z_k}(z-y)u(y)dy.
\end{aligned}$$

In the above argument, we only used that $\rho_\varepsilon \in C_c(\mathbb{R}^N)$ and that $\operatorname{supp}\rho_\varepsilon \subseteq \overline{B}_\varepsilon(0)$. So, the argument above remains valid if we replace ρ_ε by $\frac{\partial \rho_\varepsilon}{\partial z_k}$. Therefore, by induction, for every multi-index α, we have

$$\frac{\partial^\alpha u_\varepsilon}{\partial z^\alpha}(z) = \left(u * \frac{\partial^\alpha \rho_\varepsilon}{\partial z^\alpha}\right)(z) = \int_\Omega u(y)\frac{\partial^\alpha \rho_\varepsilon}{\partial z^\alpha}(z-y)dy.$$

The proof is now complete. ∎

The next result essentially says that the operator D^α is closed.

Proposition 1.1.9 *Assume that* $\{u_n\}_{n\geqslant 1} \subseteq L^1_{\mathrm{loc}}(\Omega)$ *and for every* $\vartheta \in C_c^\infty(\Omega)$ *we have*

$$\int_\Omega u_n \vartheta dz \to \int_\Omega u\vartheta dz \text{ and } \int_\Omega (D^\alpha u_n)\vartheta dz \to \int_\Omega v\vartheta dz.$$

Then $v = D^\alpha u$.

Proof By Definition 1.1.6, we have

$$\begin{aligned}
&\int_\Omega u_n(D^\alpha \vartheta)dz = (-1)^{|\alpha|}\int_\Omega (D^\alpha u_n)\vartheta dz\\
\Rightarrow &\int_\Omega u(D^\alpha\vartheta)dz = (-1)^{|\alpha|}\int_\Omega v\vartheta dz \text{ for all } \vartheta \in C_c^\infty(\Omega)\\
\Rightarrow &\, D^\alpha u = v \text{ (see Definition 1.1.6)},
\end{aligned}$$

which concludes the proof. □

We are now ready to introduce the Sobolev spaces.

Definition 1.1.10 Let $m \geqslant 0$ be an integer and $p \in [1,\infty]$. The Sobolev space $W^{m,p}(\Omega)$ is defined by

$$W^{m,p}(\Omega) = \{u \in L^p(\Omega) : D^\alpha u \in L^p(\Omega) \text{ for every multi-index } \alpha \text{ with } |\alpha| \leqslant m\}.$$

The space $W^{m,p}(\Omega)$ is equipped with the norm

$$\begin{aligned}
\|u\|_{m,p} &= \left[\sum_{|\alpha|\leqslant m} \|D^\alpha u\|_p^p\right]^{1/p} \quad \text{when } 1 \leqslant p < \infty,\\
\|u\|_{m,\infty} &= \max_{|\alpha|\leqslant m} \|D^\alpha u\|_\infty \text{ when } p = \infty.
\end{aligned}$$

Remark 1.1.11 When $p = 2$, we write $H^m(\Omega)$ for the space $W^{m,2}(\Omega)$, to emphasize the Hilbert space structure of the space, which is induced by the inner product

$$(u, v)_m = \int_\Omega \sum_{|\alpha|\leqslant m} (D^\alpha u)(D^\alpha v)dz \text{ for all } u, v \in H^m(\Omega).$$

We also have local versions of the Sobolev spaces. Namely, $u \in W^{m,p}_{\mathrm{loc}}(\Omega)$ if and only if $u \in W^{m,p}(\Omega_0)$ for all $\Omega_0 \subset\subset \Omega$ (that is, Ω_0 has compact closure contained in Ω). Other equivalent norms that can be used on $W^{1,p}(\Omega)$ are the following

$$u \to \|u\|_p + \sum_{k=1}^{N} \left\| \frac{\partial u}{\partial z_k} \right\|_p$$
$$\text{and } u \to \|u\|_p + \|Du\|_{L^p(\Omega,\mathbb{R}^N)},$$

where for $u \in W^{1,p}(\Omega)$, we set $Du = \left(\frac{\partial u}{\partial z_1}, \dots, \frac{\partial u}{\partial z_N}\right)$ (the gradient of u).

Proposition 1.1.12 *(a) For $m \geqslant 0$ and $1 \leqslant p \leqslant \infty$, $W^{1,p}(\Omega)$ is a Banach space which is separable if $1 \leqslant p < \infty$ and reflexive and uniformly convex if $1 < p < \infty$.*
(b) For $m \geqslant 0$ and $p = 2$, $H^m(\Omega) = W^{m,2}(\Omega)$ is a separable Hilbert space.

Proof (a) Let $\{u_n\}_{n\geqslant 1} \subseteq W^{m,p}(\Omega)$ be a Cauchy sequence. For any multi-index α with $|\alpha| \leqslant m$ we have

$$\|D^\alpha u_n - D^\alpha u_k\|_p \leqslant \|u_n - u_k\|_{m,p} \text{ for all } n, k \geqslant 1$$
$$\Rightarrow \{D^\alpha u_n\}_{n\geqslant 1} \subseteq L^p(\Omega) \text{ is Cauchy.}$$

Therefore we have

$$u_n \to u \text{ and } D^\alpha u_n \to v \text{ in } L^p(\Omega) \text{ as } n \to \infty$$
$$\Rightarrow v = D^\alpha u \text{ (see Proposition 1.1.9).}$$

It follows that $u \in W^{m,p}(\Omega)$ and that $u_n \to u$ in $W^{m,p}(\Omega)$.

Let $\hat{n}$ be the number of multi-indices α satisfying $0 \leqslant |\alpha| \leqslant m$ ordered in some convenient way. Then let $T : W^{m,p}(\Omega) \to L^p(\Omega)^{\hat{n}}$ be the map defined by

$$T(u) = (D^\alpha u)_{|\alpha|\leqslant m}.$$

We have $\|T(u)\|_{L^p(\Omega)^{\hat{n}}} = \|u\|_{m,p}$ and so the Sobolev space $W^{m,p}(\Omega)$ is isometrically isomorphic to a closed subspace of $L^p(\Omega)^{\hat{n}}$. This implies that $W^{m,p}(\Omega)$ is separable for $1 \leqslant p < \infty$ and uniformly convex (hence reflexive, too) for $1 < p < \infty$.

(b) We set $(u, v)_m = \int_\Omega uv dz + \int_\Omega (Du, Dv)_{\mathbb{R}^N} dz$ for all $u, v \in H^m(\Omega)$. Then $(\cdot, \cdot)_m$ is an inner product on $H^m(\Omega)$ and $\|u\|_{m,2}^2 = (u, u)_m$ for all $u \in H^m(\Omega)$. Using (a) we conclude that $H^m(\Omega)$ is a separable Hilbert space. □

Remark 1.1.13 Let $\{u_n\}_{n\geqslant 1} \subseteq W^{1,p}(\Omega)$ and suppose that

$$u_n \to u \text{ in } L^p(\Omega) \text{ and } Du_n \to h \text{ in } L^p(\Omega, \mathbb{R}^N).$$

We have $u \in W^{1,p}(\Omega)$ and $u_n \to u$ in $W^{1,p}(\Omega)$. Furthermore, if $1 < p < \infty$, $u_n \to u$ in $L^p(\Omega)$ and $\{Du_n\}_{n\geqslant 1} \subset L^p(\Omega, \mathbb{R}^N)$ is bounded, then $u \in W^{1,p}(\Omega)$.

The following subspace of $W^{m,p}(\Omega)$ is important in the study of boundary value problems.

Definition 1.1.14 Set $W_0^{m,p}(\Omega) = \overline{C_c^\infty(\Omega)}^{\|\cdot\|_{m,p}}$ for every integer $m \geqslant 0$ and for $1 \leqslant p \leqslant \infty$.

Remark 1.1.15 As we will see later in this chapter, this space consists of those Sobolev functions that vanish in some sense on the boundary $\partial\Omega$. This will be made precise after the introduction of the notion of the trace of a Sobolev function. Note that since uniform convergence preserves continuity, the elements of $W^{m,\infty}(\Omega)$ are necessarily of class C^m. In particular then $W_0^{1,\infty}(\Omega)$ does not contain piecewise affine functions. For this reason some authors prefer to define $W_0^{1,\infty}(\Omega)$ as the closure of $C_c^\infty(\Omega)$ with respect to the weak topology of $W^{1,p}(\Omega)$. The two definitions are different. Finally, for $p = 2$, we write $H_0^m(\Omega)$ for the space $W_0^{m,2}(\Omega)$.

Example 1.1.16 Let $\Omega = (-1, 1)$ (that is, $N = 1$) and for $r \in (0, 1)$ we consider the function $u(t) = |t|^r$ for all $t \in (-1, 1)$. Then a simple computation yields that $u'(t) = r \operatorname{sign} t|t|^{r-1}$ (note that $u(\cdot)$, being locally Lipschitz, is differentiable a.e. in $\Omega = (-1, 1)$). Then

$$\int_{-1}^{1} (u')^2(t)dt = r^2 \int_{-1}^{1} t^{2r-2}dt.$$

The last integral is finite if $r > \frac{1}{2}$. Therefore $u(\cdot) = |\cdot|^r \in H^1(-1, 1)$ if and only if $r > \frac{1}{2}$.

Proposition 1.1.17 *If $u \in W_0^{m,p}(\Omega)$ and we set*

$$\hat{u}(z) = \begin{cases} u(z) & \text{if } z \in \Omega \\ 0 & \text{if } z \in \mathbb{R}^N \setminus \Omega, \end{cases}$$

then $u \in W_0^{m,p}(\Omega_1)$ for every open set $\Omega_1 \supseteq \Omega$; in particular $\hat{u} \in W_0^{m,p}(\mathbb{R}^N)$ and $u \to \hat{u}$ is a linear isometry.

Proof From Definition 1.1.14, we know that we can find $\{u_n\}_{n\geqslant 1} \subseteq C_c^\infty(\Omega)$ such that $u_n \to u$ in $W^{m,p}(\Omega)$. Let $\hat{u}_n(z) = \begin{cases} u_n(z) & \text{if } z \in \Omega \\ 0 & \text{if } z \in \mathbb{R}^N \setminus \Omega \end{cases}$. Then $\hat{u}_n \in C_c^\infty(\Omega_1)$ and $\hat{u}_n \to \hat{u}$ in $W^{m,p}(\Omega_1)$ (note that $\|\hat{u}_n - \hat{u}\|_{W^{m,p}(\Omega_1)} = \|u_n - u\|_{W^{m,p}(\Omega)}$). Hence $\hat{u} \in W^{m,p}(\Omega_1)$. Clearly, $u \to \hat{u}$ is a linear isometry. $\square$

Remark 1.1.18 In fact using the above proposition, we see at once that for every $u \in W_0^{m,p}(\Omega)$ we have $u_\varepsilon \to u$ in $W^{m,p}(\Omega)$ as $\varepsilon \to 0^+$. If $u \in W^{m,p}(\Omega)$ is arbitrary and $\hat{u}(\cdot)$ is defined as above, then in general $\hat{u}(\cdot)$ does not have weak derivatives. Therefore, in general we have $W^{m,p}(\Omega) \neq W_0^{m,p}(\Omega)$. However, for $\Omega = \mathbb{R}^N$ equality holds. To show this, first we prove the following generalization of the well-known Leibnitz differentiation rule for the product of two functions.

Lemma 1.1.19 *If $u \in W^{1,p}(\Omega)$ and $\varphi \in C_c^\infty(\Omega)$, then $\varphi u \in W^{1,p}(\Omega)$ and we have $\frac{\partial}{\partial z_k}(\varphi u) = \varphi \frac{\partial u}{\partial z_k} + u \frac{\partial \varphi}{\partial z_k}$ for all $k \in \{1, \dots, N\}$.*

Proof Let $\vartheta \in C_c^\infty(\Omega)$. From the classical differentiation rule we have

$$\frac{\partial}{\partial z_k}(\varphi\vartheta) = \varphi\frac{\partial\vartheta}{\partial z_k} + \vartheta\frac{\partial\varphi}{\partial z_k}.$$

Then

$$\begin{aligned}
\textstyle\int_\Omega \varphi u \frac{\partial\vartheta}{\partial z_k} dz &= \int_\Omega u\frac{\partial}{\partial z_k}(\varphi\vartheta)dz - \int_\Omega u\vartheta\frac{\partial\varphi}{\partial z_k}dz \\
&= -\int_\Omega \frac{\partial u}{\partial z_k}(\varphi\vartheta)dz - \int_\Omega u\vartheta\frac{\partial\varphi}{\partial z_k}dz \text{ (since } \varphi\vartheta \in C_c^\infty(\Omega)) \\
&= -\int_\Omega \left[\varphi\frac{\partial u}{\partial z_k} + u\frac{\partial\varphi}{\partial z_k}\right]\vartheta dz \text{ for all } \vartheta \in C_c^\infty(\Omega).
\end{aligned}$$

Note that $\varphi\frac{\partial u}{\partial z_k} + u\frac{\partial\varphi}{\partial z_k} \in L^p(\Omega)$. So, from Definition 1.1.10 we conclude that

$$\varphi u \in W^{1,p}(\Omega) \text{ and } \frac{\partial}{\partial z_k}(\varphi u) = \varphi\frac{\partial u}{\partial z_k} + u\frac{\partial\varphi}{\partial z_k} \text{ for all } k \in \{1, \ldots, N\}.$$

The proof is now complete. □

Now we can prove the stated result about the Sobolev spaces on $\mathbb{R}^N$.

In the sequel we will use the notion of the support of an L^p-function. Since L^p consists of equivalence classes of functions, the usual definition of the support of u (as the closure of the set $\{z \in \Omega : u(z) \neq 0\}$) is not adequate. We need a definition which is "intrinsic" in the sense that if $u_1 = u_2$ a.e. in Ω, then $\operatorname{supp} u_1 = \operatorname{supp} u_2$. The next definition achieves this.

Definition 1.1.20 Let $u : \mathbb{R}^N \to \mathbb{R}$ be a function and $\{V_i\}_{i\in I}$ be the family of all open subsets of $\mathbb{R}^N$ such that $u\big|_{V_i} = 0$ a.e. Let $V = \bigcup_{i\in I} V_i$. Then $u\big|_V = 0$ a.e. and we define the support of u (denoted by $\operatorname{supp} u$) to be the complement of the open set V.

Theorem 1.1.21 *For any $1 \leqslant p \leqslant \infty$ and any integer $m \geqslant 0$ we have*

$$W^{m,p}(\mathbb{R}^N) = \overline{C_c^\infty(\mathbb{R}^N)}^{\|\cdot\|_{m,p}} \text{ (that is, } W^{m,p}(\mathbb{R}^N) = W_0^{m,p}(\mathbb{R}^N)).$$

Proof For simplicity in the presentation, we assume that $m = 1$. Let $\varphi \in C_c^\infty(\mathbb{R}^N)$ such that $\varphi(0) = 1$ and let $\varphi_n(z) = \varphi(\frac{z}{n})$ for all $n \geqslant 1$. Then $\varphi_n \in C_c^\infty(\Omega)$ and for all $z \in \mathbb{R}^N$, $\lim_{n\to\infty}\varphi_n(z) = \varphi(0) = 1$.

Given $u \in W^{1,p}(\mathbb{R}^N)$, we set $u_n = \varphi_n u$ for all $n \geqslant 1$. We see that $\operatorname{supp} u_n \subseteq \operatorname{supp}\varphi_n \subseteq n \operatorname{supp}\varphi$ and so u_n has compact support. Also

$$\int_{\mathbb{R}^N} |u_n|^p dz \leqslant \|\varphi\|_\infty^p \int_{\mathbb{R}^N} |u|^p dz < \infty, \text{ that is, } u_n \in L^p(\Omega) \text{ for all } n \geqslant 1. \quad (1.5)$$

From Lemma 1.1.19 we have

$$\frac{\partial u_n}{\partial z_k} = \varphi_n \frac{\partial u}{\partial z_k} + u\frac{\partial \varphi_n}{\partial z_k} \text{ for all } k \in \{1, \dots, N\}, \text{ that is, } \frac{\partial u_n}{\partial z_k} \in L^p(\Omega). \quad (1.6)$$

From (1.5) and (1.6) it follows that $u_n \in W^{1,p}(\mathbb{R}^N)$.

Using the Lebesgue dominated convergence theorem, we have

$$\int_{\mathbb{R}^N} |u_n - u|^p dz = \int_{\mathbb{R}^N} |1 - \varphi_n|^p |u|^p dz \to 0 \text{ as } n \to \infty. \quad (1.7)$$

Moreover, again by the Lebesgue dominated convergence theorem, we have

$$\varphi_n \frac{\partial u}{\partial z_k} \to \frac{\partial u}{\partial z_k} \text{ in } L^p(\mathbb{R}^N). \quad (1.8)$$

In addition, we have

$$\begin{aligned} &\left| u(z)\frac{\partial \varphi_n}{\partial z_k}(z) \right| \leqslant \frac{1}{n} \left\| \frac{\partial \varphi}{\partial z_k} \right\|_\infty |u(z)| \text{ a.e. in } \Omega \\ \Rightarrow &\left\| u\frac{\partial \varphi_n}{\partial z_k} \right\|_p \leqslant \frac{1}{n} \left\| \frac{\partial \varphi}{\partial z_k} \right\|_\infty \|u\|_p \to 0 \text{ as } n \to \infty. \end{aligned} \quad (1.9)$$

From (1.8), (1.9) and Lemma 1.1.19 it follows that

$$\frac{\partial u_n}{\partial z_k} \to \frac{\partial u}{\partial z_k} \text{ in } L^p(\Omega) \text{ for all } k \in \{1, \dots, N\}. \quad (1.10)$$

From (1.7) and (1.10), we conclude

$$u_n \to u \text{ in } W^{1,p}(\mathbb{R}^N) \text{ and } \operatorname{supp} u_n \text{ is compact for all } n \geqslant 1. \quad (1.11)$$

In view of (1.11) it suffices to show that any $v \in W^{1,p}(\mathbb{R}^N)$ of compact support can be approximated in $W^{1,p}(\mathbb{R}^N)$ by a sequence in $C_c^\infty(\mathbb{R}^N)$.

To this end let $\{\rho_n\}_{n\geqslant 1}$ be a sequence of mollifiers and let $v_n = v * \rho_n$. Then by virtue of Proposition 1.1.8 we have $v_n \in C^\infty(\mathbb{R}^N)$ and $\operatorname{supp} v_n \subseteq \operatorname{supp} v + \overline{B}_{\frac{1}{n}}(0)$, hence $v_n \in C_c^\infty(\mathbb{R}^N)$.

Let $\eta_n(y) = -\rho_n(z - y)$. We have

$$\begin{aligned} \frac{\partial \eta_n}{\partial z_k}(y) &= \lim_{\lambda \to 0} \frac{1}{\lambda}[\eta_n(y + \lambda z_k) - \eta_n(y)] \\ &= \lim_{\lambda \to 0} \frac{1}{-\lambda}[\rho_n(z - y - \lambda z_k) - \rho_n(z - y)] \\ &= \frac{\partial \rho_n}{\partial z_k}(z - y). \end{aligned} \quad (1.12)$$

Then using Proposition 1.1.8 and (1.12), we can write

$$\begin{aligned}
\frac{\partial v_n}{\partial z_k}(z) &= \int_{\mathbb{R}^N} v(y)\frac{\partial \eta_n}{\partial z_k}(y)dy \\
&= -\int_{\mathbb{R}^N} \frac{\partial v}{\partial z_k}(y)\eta_n(y)dy \text{ (since } u \in W^{1,p}(\Omega)) \\
&= \int_{\mathbb{R}^N} \frac{\partial v}{\partial z_k}(y)\rho_n(z-y)dy \\
\Rightarrow \frac{\partial v_n}{\partial z_k} &= \left(\frac{\partial v}{\partial z_k}\right) * \rho_n \\
\Rightarrow \frac{\partial v_n}{\partial z_k} &\to \frac{\partial v}{\partial z_k} \text{ in } L^p(\Omega) \text{ as } n \to \infty \text{ (see Proposition 1.1.3 (b)).}
\end{aligned}$$

Since $v_n \to v$ in $L^p(\mathbb{R}^N)$ (again by Proposition 1.1.3 (b)), we conclude that $v_n \to v$ in $W^{1,p}(\mathbb{R}^N)$ and this proves the theorem (note that the convergence is pointwise on the set of Lebesgue points of u). □

Nevertheless, we can always approximate Sobolev functions in $W^{1,p}(\Omega)$ by smooth functions. This is the so-called "Meyers–Serrin theorem". To prove it, we will need the following simple lemma.

Lemma 1.1.22 *If $u \in W^{1,p}(\Omega)$ with $1 \leqslant p < \infty$, then*

$$\lim_{\varepsilon \to 0^+} \left[\|u - u_\varepsilon\|_{L^p(\Omega_\varepsilon)} + \|Du - Du_\varepsilon\|_{L^p(\Omega_\varepsilon, \mathbb{R}^N)}\right] = 0;$$

in particular if $\Omega_0 \subseteq \Omega$ and $d(\Omega_0, \partial\Omega) > 0$, then $u_\varepsilon \to u$ in $W^{1,p}(\Omega_0)$ as $\varepsilon \to 0^+$.

Proof From Proposition 1.1.8 we know that $u_\varepsilon \in C^\infty(\Omega_\varepsilon)$ and for $z \in \Omega_\varepsilon$ and $k \in \{1, \dots, N\}$ we have $\frac{\partial u_\varepsilon}{\partial z_k} = u * \frac{\partial \rho_\varepsilon}{\partial z_k}$. Moreover, in the proof of Theorem 1.1.21 we established that $u * \frac{\partial \rho_\varepsilon}{\partial z_k} = \frac{\partial u}{\partial z_k} * \rho_\varepsilon$. Then invoking Proposition 1.1.3 (b), we deduce the desired convergence. □

Theorem 1.1.23 (Meyers–Serrin) *If $1 \leqslant p < \infty$ and $m \geqslant 0$ is an integer, then $W^{m,p}(\Omega) \cap C^\infty(\Omega)$ is dense in $W^{m,p}(\Omega)$.*

Proof For every integer $k \geqslant 1$ we define $\Omega_k = \{z \in \Omega : \|z\|_{\mathbb{R}^N} \leqslant k \text{ and } d(z, \partial\Omega) > \frac{1}{k}\}$ and $\Omega_0 = \emptyset$. Let $U_k = \Omega_{k+1} \cap (\overline{\Omega}_{k-1})^c$. Evidently $\{U_k\}_{k\geqslant 1}$ is an open cover of Ω. Let $\{\psi_k\}_{k\geqslant 1}$ be a C^∞ partition of unity subordinate to the open cover $\{U_k\}_{k\geqslant 1}$. So, we have

$$\operatorname{supp} \psi_k \subseteq U_k, \;\; \psi_k \in C_c^\infty(U_k), \;\; \psi_k \geqslant 0 \text{ and } \sum_{k\geqslant 1} \psi_k = 1.$$

Let $u \in W^{m,p}(\Omega)$ and consider the truncated function $\psi_k u$. As in the proof of Theorem 1.1.21 we have $\psi_k u \in W_0^{1,p}(U_k)$. We set $(\psi_k u)(z) = 0$ for all $z \in \mathbb{R}^N \setminus U_k$ and we have an element of $W^{m,p}(\mathbb{R}^N)$ (see Proposition 1.1.17 and Theorem 1.1.21). Consider a sequence $\{\rho_n\}_{n\geqslant 1}$ of mollifiers. Then for $n = n(k) \geqslant 1$ large, we have

$$\|(\psi_k u) * \dot{\rho}_{n(k)} - \psi_k u\|_{m,p} < \frac{\varepsilon}{2^k} \text{ with } \varepsilon > 0 \tag{1.13}$$
$$\text{and } \operatorname{supp}[(\psi_k u) * \rho_{n(k)}] \subseteq U_k.$$

Let $v = \sum_{k \geqslant 1} (\psi_k u) * \rho_{n(k)}$ and note that for $z \in U_k$ we have

$$v(z) = \sum_{i=-1}^{1} \left[(\psi_{k+i} u) * \rho_{n(k+i)}\right](z).$$

So, $v \in C^\infty(\Omega)$ and we have

$$\begin{aligned}\|v - u\|_{m,p} &= \left\| \sum_{k \geqslant 1} \left((\psi_k u) - (\psi_k u) * \rho_{n(k)}\right) \right\|_{m,p} \\ &\leqslant \sum_{k \geqslant 1} \|(\psi_k u) - (\psi_k u) * \rho_{n(k)}\|_{m,p} \leqslant \varepsilon \text{ (see 1.13)}.\end{aligned}$$

This proves that $W^{m,p}(\Omega) \cap C^\infty(\Omega)$ is dense in $W^{m,p}(\Omega)$. $\square$

Proposition 1.1.24 *If $u \in W^{1,p}(\Omega)$ ($1 \leqslant p < \infty$) and has compact support, then $u \in W_0^{1,p}(\Omega)$.*

Proof Let $\varphi \in C_c^\infty(\Omega)$ such that $\varphi\big|_{\operatorname{supp} u} = 1$. By virtue of Theorem 1.1.23, we can find $\{\psi_k\}_{k \geqslant 1} \subseteq W^{1,p}(\Omega) \cap C^\infty(\Omega)$ such that $\psi_k \to u$ in $W^{1,p}(\Omega)$. Let $\eta_k = \varphi\psi_k$, $k \geqslant 1$. Then $\eta_k \in C_c^\infty(\Omega)$ and we have $\eta_k \to \varphi u = u$ in $W^{1,p}(\Omega)$. This proves that $u \in W_0^{1,p}(\Omega)$. $\square$

Another approximation by smooth functions of Sobolev functions is provided by the so-called "Friedrichs theorem". It is a partial extension of Theorem 1.1.21.

Theorem 1.1.25 (Friedrichs) *If $u \in W^{1,p}(\Omega)$ with $1 \leqslant p < \infty$, then there exists a sequence $\{u_n\}_{n \geqslant 1} \subseteq C_c^\infty(\mathbb{R}^N)$ such that*

$$u_n \to u \text{ in } L^p(\Omega) \text{ and } \frac{\partial u_n}{\partial z_k}\bigg|_{\Omega_0} \to \frac{\partial u}{\partial z_k}\bigg|_{\Omega_0} \text{ in } L^p(\Omega_0)$$

for every $k \in \{1, \ldots, N\}$ and every $\Omega_0 \subset\subset \Omega$ (that is, Ω_0 has compact closure which is contained in Ω).

Proof Let $\overline{u}$ denote the extension by zero outside of Ω of the function $u \in W^{1,p}(\Omega)$ and let $\{\rho_\varepsilon\}_{\varepsilon \in (0,1)}$ be a family of mollifiers. Then we have $\overline{u} * \rho_\varepsilon \to \overline{u}$ in $L^p(\mathbb{R}^N)$ (see Proposition 1.1.3 (b)). Hence $\overline{u} * \rho_\varepsilon \to u$ in $L^p(\Omega)$. Let α be any multi-index with $|\alpha| = 1$ and $\Omega_0 \subset\subset \Omega$. Then $d_0 = d(\Omega_0, \partial\Omega) > 0$ and we take $\varepsilon \in (0, d_0)$. Then for all $z \in \Omega_0$ we have $(\overline{u} * \rho_\varepsilon)(z) = (u * \rho_\varepsilon)(z)$ and so $D^\alpha(\overline{u} * \rho_\varepsilon) = (D^\alpha \overline{u}) * \rho_\varepsilon$ in Ω_0 (see the proof of Theorem 1.1.21), hence $D^\alpha(\overline{u} * \rho_\varepsilon) \in L^p(\Omega)$. We have $D^\alpha(\overline{u} * \rho_\varepsilon) \to D^\alpha u$ in $L^p(\Omega_0)$ and so we conclude that $\overline{u} * \rho_\varepsilon \to u$ in $W^{1,p}(\Omega_0)$.

Let $\varepsilon_n \downarrow 0$ and let $v_n = \overline{u} * \rho_{\varepsilon_n}$ for all $n \geqslant 1$. We have

$$v_n \to u \text{ in } L^p(\Omega) \text{ and } \left.\frac{\partial v_n}{\partial z_k}\right|_{\Omega_0} \to \left.\frac{\partial u}{\partial z_k}\right|_{\Omega_0} \text{ in } L^p(\Omega_0)\ (k \in \{1, \ldots, N\})$$

for every $\Omega_0 \subset\subset \Omega$. Let $\eta \in C_c^\infty(\mathbb{R}^N)$ such that $0 \leqslant \eta \leqslant 1$, $\eta\big|_{\overline{B_1(0)}} = 1$ and supp $\eta \subseteq \overline{B}_2(0)$ (such a function is usually called a "cut-off function"). Let $\{\eta_n\}_{n\geqslant 1}$ be defined by $\eta_n(z) = \eta(\frac{z}{n})$ for all $z \in \mathbb{R}^N$. Then $u_n = \eta_n v_n \in C_c^\infty(\mathbb{R}^N)$ for all $n \geqslant 1$ and $\{u_n\}_{n\geqslant 1}$ is the desired sequence. □

Remark 1.1.26 If $\Omega = \mathbb{R}^N$, then from the above proof it is clear that we can find $\{u_n\}_{n\geqslant 1} \subseteq C_c^\infty(\mathbb{R}^N)$ such that $u_n \to u$ in $W^{1,p}(\mathbb{R}^N)$, which is precisely the result of Theorem 1.1.21.

1.2 The One-Dimensional Case

In this section, $\Omega = I = (a, b)$ is an open interval, possibly unbounded.

We start with a simple property.

Lemma 1.2.1 *If $u \in L^1_{\text{loc}}(I)$ and $u' = 0$ in the distributional sense, then there exists a $c^* \in \mathbb{R}$ such that $u(t) = c^*$ for a.a. $t \in I$.*

Proof Since $u' = 0$ in the distributional sense, we have

$$\int_a^b u(t)\vartheta'(t)dt = 0 \text{ for all } \vartheta \in C_c^\infty(I). \tag{1.14}$$

Let $D = \{\vartheta' : \vartheta \in C_c^\infty(I)\}$. It is easy to see that $D = \{\varphi \in C_c^\infty(I) : \int_a^b \varphi(t)dt = 0\}$. Moreover, $\varphi \in D$ if and only if there exist $\vartheta \in C_c^\infty(I)$ and $\xi \in C_c^\infty(I)$ with $\int_a^b \xi(t)dt = 1$ such that

$$\varphi(t) = \vartheta(t) - \left(\int_a^b \vartheta(s)ds\right)\xi(t) \text{ for all } t \in I. \tag{1.15}$$

Rewriting (1.14) in terms of D, we have

$$\int_a^b u(t)\varphi(t)dt = 0 \text{ for all } \varphi \in D$$
$$\Rightarrow \int_a^b u(t)\vartheta(t)dt = \left(\int_a^b \vartheta(s)ds\right)\int_a^b u(t)\xi(t)dt \text{ (see (1.15))}. \tag{1.16}$$

Let $c^* = \int_a^b u(t)\xi(t)dt$. Then from (1.16) we have

$$\int_a^b [u(t) - c^*]\vartheta(t)dt = 0 \text{ for all } \vartheta \in C_c^\infty(I)$$
$$\Rightarrow u(t) = c^* \text{ for a.a. } t \in I \text{ (see Proposition 1.1.5)},$$

which completes the proof. □

This leads to the following important property of the one-dimensional Sobolev functions.

Theorem 1.2.2 *If $u \in W^{1,p}(I)$ with $1 \leqslant p \leqslant \infty$, then there exists a $\hat{u} \in C(\overline{I})$ such that $u(t) = \hat{u}(t)$ for a.a. $t \in I$ and*

$$\hat{u}(t) - \hat{u}(\tau) = \int_\tau^t u'(s)ds \ \textit{for all } t, \tau \in \overline{I} = [a, b].$$

Proof Let $u' \in L^p(I)$ be the weak (distributional) derivative of $u \in W^{1,p}(I)$. Let

$$\hat{u}(t) = \int_a^t u'(s)ds.$$

This function is continuous. In fact, when $p \in (1, +\infty]$, it is Hölder continuous. To see this, let $t, \tau \in I$. Then

$$\begin{aligned}
|\hat{u}(t) - \hat{u}(\tau)| = \left| \int_\tau^t u'(s)ds \right| &\leqslant \int_\tau^t |u'(s)|ds \\
&\leqslant |t - \tau|^{1/p'} \left(\int_\tau^t |u'(s)|^p ds \right)^{1/p} \\
&\quad \text{(by Hölder's inequality)} \\
&\leqslant \|u'\|_{L^p(I)} |t - \tau|^{1/p'}.
\end{aligned}$$

For every $\vartheta \in C_c^\infty(I)$ we have

$$\begin{aligned}
\int_a^b \hat{u}(s)\vartheta'(s)ds = \int_a^b \left(\int_a^s u'(r)dr \right) \vartheta'(s)ds &= \int_a^b \left(\int_a^b \chi_{[a,s]}(r)u'(r)dr \right) \vartheta'(s)ds \\
&= \int_a^b u'(r) \left(\int_a^b \chi_{[a,s]}(r)\vartheta'(s)ds \right) dr \\
&\quad \text{(by Fubini's theorem)} \\
&= \int_a^b u'(r) \left(\int_r^b \vartheta'(s)ds \right) dr \\
&= \int_a^b u'(r)\vartheta(r)dr \text{ (since } \vartheta(b) = 0).
\end{aligned}$$

It follows that $\hat{u}' = u'$, hence $(\hat{u} - u)' = 0$.

Invoking Lemma 1.2.1, we have that $\hat{u} - u = c^*$ for some $c* \in \mathbb{R}$. Therefore the function $\hat{u}(\cdot)$ has the desired properties. $\square$

Remark 1.2.3 From the above proof, we see that if $1 < p < \infty$, then $\hat{u} \in C^{0,\beta}(\overline{I})$ with $\beta = \frac{1}{p'}$ $\left(\frac{1}{p} + \frac{1}{p'} = 1\right)$, hence $|\hat{u}(t) - \hat{u}(\tau)| \leqslant c_0|t - \tau|^{\beta}$ for all $t, \tau \in \overline{I}$ and some $c_0 > 0$. In general the continuous representative $\hat{u}$ of $u \in W^{1,p}(I)$ is in fact absolutely continuous and it is unique up to an additive constant. In the sequel, when dealing with Sobolev functions of one variable, we always identify a function $u \in W^{1,p}(I)$ with its continuous representative. Keeping this in mind, note that if $u \in W^{1,p}(I)$ and $u' \in C(\overline{I})$, then $u \in C^1(\overline{I})$.

As a direct consequence of the fundamental theorem of the Lebesgue calculus, we have the converse of Theorem 1.2.2.

Proposition 1.2.4 *Assume that $u \in L^p(\Omega)$ with $1 \leqslant p \leqslant \infty$ and there exists an $h \in L^p(\Omega)$ such that*

$$u(t) - u(\tau) = \int_{\tau}^{t} h(s)ds \ for\ a.a.\ t, \tau \in I.$$

Then $u \in W^{1,p}(I)$ and $u' = h$ in the weak sense.

1.3 Duals of Sobolev Spaces

To better describe the duals of Sobolev spaces, we need to introduce the notion of distribution (in the sense of L. Schwartz). For this purpose we consider the usual space of test functions $C_c^{\infty}(\Omega)$. We introduce a notion of sequential convergence on $C_c^{\infty}(\Omega)$ and this will be the only topological notion on $C_c^{\infty}(\Omega)$ that we will use.

Definition 1.3.1 A sequence $\{u_n\}_{n\geqslant 1} \subseteq C_c^{\infty}(\Omega)$ is said to converge to $u \in C_c^{\infty}(\Omega)$ in the "$\mathscr{D}$-sense" if and only if the following conditions hold:

(a) There exists a compact set K contained in Ω such that supp $u_n \subseteq K$ for all $n \geqslant 1$ and supp $u \subseteq K$.
(b) For every multi-index α, $D^{\alpha}u_n \to D^{\alpha}u$ uniformly on K.

Remark 1.3.2 In fact, this is a topological notion. That is, there exists a locally convex topology $\hat{\tau}$ on $C_c^{\infty}(\Omega)$ for which the closed sets are the $\mathscr{D}$-sequentially closed sets. Then a linear functional L on $C_c^{\infty}(\Omega)$ is continuous if and only if it is $\mathscr{D}$-sequentially continuous, that is, if $u_n \to u$ in $C_c^{\infty}(\Omega)$ in the $\mathscr{D}$-sense, then $L(u_n) \to L(u)$. This topology is first countable and complete, but not metrizable. Usually in the literature, the space of test functions $C_c^{\infty}(\Omega)$ equipped with this topology is denoted by $\mathscr{D}(\Omega)$. This explains the use of the symbol $\mathscr{D}$ in the above definition.

Using this mode of convergence, we can introduce the notion of distribution.

Definition 1.3.3 A "distribution" on Ω is a linear functional on $C_c^\infty(\Omega)$ which is continuous in the following sense: if $\{u_n, u\}_{n\geqslant 1} \subseteq C_c^\infty(\Omega)$ and $u_n \to u$ in the $\mathscr{D}$-sense, then $L(u_n) \to L(u)$. Therefore the distributions are the elements of $\left(C_c^\infty(\Omega), \hat{\tau}\right)^*$.

Example 1.3.4 Every function $g \in L^1_{\text{loc}}(\Omega)$ defines a distribution $L_g : C_c^\infty(\Omega) \to \mathbb{R}^N$ by setting

$$L_g(\vartheta) = \int_\Omega g(z)\vartheta(z)dz \text{ for all } \vartheta \in C_c^\infty(\Omega).$$

Using the elementary tools of integration by parts (see also Definition 1.1.6), for a distribution $L : C_c^\infty(\Omega) \to \mathbb{R}^N$, we can define derivatives of all orders.

Definition 1.3.5 Let L be a distribution and α a multi-index. The αth-derivative of L is defined by

$$(D^\alpha L)(\vartheta) = \left(\frac{\partial^\alpha L}{\partial z^\alpha}\right)(\vartheta) = (-1)^{|\alpha|} L\left(\frac{\partial^\alpha \vartheta}{\partial z^\alpha}\right) \text{ for all } \vartheta \in C_c^\infty(\Omega).$$

It can be easily verified that $D^\alpha L = \frac{\partial^\alpha L}{\partial z^\alpha}$ is still a distribution.

Remark 1.3.6 In particular if $g \in L^1_{\text{loc}}(\Omega)$ and α is a multi-index, then $D^\alpha L_g$ (see Example 1.3.4) is the weak or distributional derivative of g introduced in Definition 1.1.6.

Now we turn our attention to the study of the dual of the Sobolev spaces $W^{1,p}(\Omega)$ and $W_0^{1,p}(\Omega)$. The idea is to view $W^{1,p}(\Omega)$ as a closed subspace of $L^p(\Omega, \mathbb{R}^{N+1})$. Thus, by the Hahn–Banach theorem, every element $L \in W^{1,p}(\Omega)^*$ can be extended to an element of $L^p(\Omega, \mathbb{R}^{N+1})^*$. Then the Riesz representation theorem for the dual of $L^p(\Omega, \mathbb{R}^{N+1})$ will give us a representation for the extension of L, hence for L too. We do the analysis for $W^{1,p}(\Omega)$ just for simplicity. Analogous results also hold for the Sobolev spaces $W^{m,p}(\Omega)$ with $m \geqslant 2$.

Theorem 1.3.7 *If $L \in W^{1,p}(\Omega)^*$ with $1 \leqslant p < \infty$, then there exist functions $\{f_k\}_{k=0}^N \subseteq L^{p'}(\Omega)$ $\left(\frac{1}{p} + \frac{1}{p'} = 1\right)$ such that*

$$L(u) = \int_\Omega \left[f_0(z)u(z) + \sum_{k=1}^N f_k(z)\frac{\partial u}{\partial z_k}(z) \right] dz \textit{ for all } u \in W^{1,p}(\Omega). \quad (1.17)$$

and

$$\|L\|_* = \left(\sum_{k=0}^N \|f_k\|_{p'}^{p'}\right)^{1/p'} \quad \left(\|\cdot\|_* \textit{ denotes the norm of } W^{1,p}(\Omega)^*\right).$$

Proof Let $A : W^{1,p}(\Omega) \to L^p(\Omega, \mathbb{R}^{N+1})$ be defined by $A(u) = (u, Du)$. Evidently A is continuous, injective and preserves the norm (that is, $\|A(u)\|_{L^p(\Omega,\mathbb{R}^{N+1})} = \|u\|_{1,p}$ for all $u \in W^{1,p}(\Omega)$). Then $V = A\left(W^{1,p}(\Omega)\right)$ is a closed subspace of

$L^p(\Omega, \mathbb{R}^{N+1})$. Given $L \in W^{1,p}(\Omega)^*$, let $L_0 : V \to \mathbb{R}$ be defined by $L_0\left(\{g_k\}_{k=0}^N\right) = L\left(A^{-1}(\{g_k\}_{k=0}^N)\right)$ for all $\{g_k\}_{k=0}^N \in L^p(\Omega, \mathbb{R}^{N+1})$. Evidently L_0 is linear, continuous and $\|L_0\|_{V^*} = \|L\|_*$. Invoking the Hahn–Banach theorem, we can find a continuous linear extension $\hat{L}_0 : L^p(\Omega, \mathbb{R}^{N+1}) \to \mathbb{R}$ of L_0 which preserves the norm, that is

$$\|\hat{L}_0\|_{L^p(\Omega,\mathbb{R}^{N+1})} = \|L_0\|_{V^*} = \|L\|_*.$$

Since V is not dense in $L^p(\Omega, \mathbb{R}^{N+1})$ (it is a closed subspace), this extension is not unique.

Invoking the Riesz representation theorem for $L^p(\Omega, \mathbb{R}^{N+1})^*$, we can find unique functions $f_0, \ldots, f_N \in L^{p'}(\Omega)$ such that

$$\hat{L}_0\left(\{g_k\}_{k=0}^N\right) = \int_\Omega \left[f_0 g_0 + \sum_{k=1}^N f_k g_k \right] dz \text{ for all } \{g_k\}_{k=0}^N \in L^p(\Omega, \mathbb{R}^{N+1})$$

$$\text{and } \|L\|_* = \|\hat{L}_0\|_{L^p(\Omega,\mathbb{R}^{N+1})^*} = \left(\sum_{k=0}^N \|f_k\|_{p'}^{p'} \right)^{1/p'}.$$

Then for all $u \in W^{1,p}(\Omega)$

$$L(u) = \int_\Omega \left[f_0 u + \sum_{k=1}^N f_k \frac{\partial u}{\partial z_k} \right] dz.$$

$\square$

Remark 1.3.8 Every element $L \in W^{1,p}(\Omega)^*$ $(1 \leqslant p < \infty)$ is an extension to the Sobolev space $W^{1,p}(\Omega)$ of a distribution $L_0 \in C_c^\infty(\Omega)^*$. To see this, suppose that L is given by (1.17) for some $\{g_k\}_{k=0}^N \in L^p(\Omega, \mathbb{R}^{N+1})$ and define the distributions $(L_0)_{g_k}$, L_0 by

$$(L_0)_{g_k}(\vartheta) = \int_\Omega g_k(z)\vartheta(z)dz \text{ and } L_0 = (L_0)_{g_0} - \sum_{k=1}^N \frac{\partial (L_0)_{g_k}}{\partial z_k}, \tag{1.18}$$

for all $\vartheta \in C_c^\infty(\Omega)$.

Evidently L is an extension of L_0. On the other hand, suppose that L_0 is a distribution having the form (1.18) for some $\{g_k\}_{k=0}^N \in L^p(\Omega, \mathbb{R}^{N+1})$. Then L_0 admits a possibly nonunique extension to $W^{1,p}(\Omega)$. However, it admits a unique extension to $W_0^{1,p}(\Omega)$. Indeed, if $u \in W_0^{1,p}(\Omega)$, then we can find $\{\vartheta_n\}_{n \geqslant 1} \subseteq C_c^\infty(\Omega)$ such that

$$\|\vartheta_n - u\|_{1,p} \to 0 \text{ as } n \to \infty \text{ (see Definition 1.1.14).}$$

Then we have

$$\begin{aligned}|L_0(\vartheta_n) - L_0(\vartheta_m)| &= \left|(L_0)_{g_0}(\vartheta_n - \vartheta_m) - \sum_{k=1}^{N} \frac{\partial (L_0)_{g_k}}{\partial z_k}(\vartheta_n - \vartheta_m)\right| \\ &\leqslant \|\vartheta_n - \vartheta_m\|_p \|g_0\|_{p'} + \sum_{k=1}^{N} \left\| \frac{\partial}{\partial z_k}(\vartheta_n - \vartheta_m)\right\|_p \|g_k\|_{p'} \\ &\leqslant c\|\vartheta_n - \vartheta_m\|_{1,p} \|\{g_k\}_{k=0}^N\|_{L^{p'}(\Omega,\mathbb{R}^{N+1})} \text{ for some } c > 0,\end{aligned}$$

hence $\{L_0(\vartheta_n)\}_{n\geqslant 1}$ is Cauchy.

Therefore we have $L_0(\vartheta_n) \to L(u)$, since the limit is easily seen to be independent of the particular approximating sequence from $C_c^\infty(\Omega)$. Evidently, L is linear and

$$\begin{aligned}&|L(u)| = \lim_{n\to\infty} |L_0(\vartheta_n)| \leqslant c\|\vartheta_n\|_{1,p} \|\{g_k\}_{k=0}^N\|_{L^{p'}(\Omega,\mathbb{R}^{N+1})} \\ &\qquad = c\|u\|_{1,p} \|\{g_k\}_{k=0}^N\|_{L^{p'}(\Omega,\mathbb{R}^{N+1})} \\ \Rightarrow \quad &L \in W_0^{1,p}(\Omega)^*.\end{aligned}$$

This uniqueness of the extension of the distribution L_0 defined in (1.18) leads to the following representation theorem for $W_0^{1,p}(\Omega)^*$.

Theorem 1.3.9 *The dual* $W_0^{1,p}(\Omega)^* = W^{-1,p'}(\Omega)$ $\left(\frac{1}{p} + \frac{1}{p'} = 1 \; with \; 1 \leqslant p < \infty\right)$ *can be identified with the subspace of distributions* L_0 *of the form*

$$L_0 = (L_0)_{g_0} - \sum_{k=1}^{N} \frac{\partial (L_0)_{g_k}}{\partial z_k} \tag{1.19}$$

for some $\{g_k\}_{k=0}^N \in L^{p'}(\Omega, \mathbb{R}^{N+1})$. *So, we have*

$$L_0(u) = \int_\Omega g_0(z)u(z)dz + \sum_{k=1}^{N} \int_\Omega g_k(z) \frac{\partial u}{\partial z_k} dz \; for \; all \; u \in W_0^{1,p}(\Omega).$$

Remark 1.3.10 Why the notation $W^{-1,p'}(\Omega)$ for the dual of $W_0^{1,p}(\Omega)$? Note that if $y \in W^{m,p'}(\Omega)$, then its first derivatives belong to $W^{m-1,p'}(\Omega)$. We would like this feature to be preserved for all integers. In particular, if $y \in L^{p'}(\Omega)$, then its first derivative must belong to $W^{-1,p'}(\Omega)$. Then from (1.19) we see why the notation $W^{-1,p'}(\Omega)$ is justified. For $p = \infty$, we use the Yosida–Hewitt theorem describing the dual of $L^\infty(\Omega)$ (see, for example, Denkowski et al. [143, p. 330]), in order to obtain a representation of the dual space $W_0^{1,\infty}(\Omega)^*$. So, $L \in W_0^{1,\infty}(\Omega)^*$ if and only if there exist unique, bounded, finitely additive signed measures $\mu_0, \mu_1 \ldots, \mu_N$ on Ω, all absolutely continuous with respect to the Lebesgue measure such that

$$L(u) = \int_\Omega u d\mu_0 + \sum_{k=1}^{N} \int_\Omega \frac{\partial u}{\partial z_k} d\mu_k \text{ for all } u \in W_0^{1,p}(\Omega).$$

1.4 Absolute Continuity on Lines, the Chain Rule and Consequences

In Theorem 1.2.2 we saw that the Sobolev functions of one variable admit a continuous (in fact, absolutely continuous) representative. When the dimension $N \geqslant 2$, the situation is more complex and in general the elements of $W^{1,p}(\Omega)$ do not have a continuous representative. In this subsection, we see what can be said in this context. So, we obtain an analog of Theorem 1.2.2.

In what follows, given $z_k' \in \mathbb{R}^{N-1}$ and a set $C \subseteq \mathbb{R}^N$, we write

$$C_{z_k'} = \{z_k \in \mathbb{R} : (z_k', z_k) \in C\}.$$

Also if $u : \Omega \to \mathbb{R}$ is Lebesgue integrable and $\Omega_{z_k'} = \emptyset$, then we set

$$\int_{\Omega_{z_k'}} u(z_k', z_k)dz_k = 0$$

while by Fubini's theorem, we have

$$\int_\Omega u(z)dz = \int_{\mathbb{R}^{N-1}} \int_{\Omega_{z_k'}} u(z_k', z_k)dz_k dz_k'. \tag{1.20}$$

Finally, for every $m \geqslant 2$ we denote by $\hat{\lambda}^m$ the Lebesgue measure on $\mathbb{R}^m$. If $m = 1$, then we simply write $\hat{\lambda}$.

Theorem 1.4.1 *A function $u \in L^p(\Omega)$ ($1 \leqslant p < \infty$) belongs to the Sobolev space $W^{1,p}(\Omega)$ if and only if it has a representative $\hat{u}$ which is absolutely continuous on $\hat{\lambda}^{N-1}$-a.e. line segment of Ω that is parallel to the coordinate axes and whose first-order (classical) partial derivatives belong to $L^p(\Omega)$. Moreover, the (classical) partial derivatives of $\hat{u}$ agree $\hat{\lambda}^N$-a.e. in Ω with the weak (distributional) derivatives of u.*

Proof First suppose that $u \in W^{1,p}(\Omega)$.

Let $\{\rho_\varepsilon\}_{\varepsilon\in(0,1)}$ be a family of mollifiers and let $u_\varepsilon = u * \rho_\varepsilon$ be defined on $\Omega_\varepsilon = \{z \in \Omega : d(z, \partial\Omega) > \varepsilon\}$. From Lemma 1.1.22 we have

$$\lim_{\varepsilon\to 0^+} \int_{\Omega_\varepsilon} \|Du_\varepsilon(z) - Du(z)\|^p_{\mathbb{R}^N} dz$$

$$\Rightarrow \lim_{\varepsilon\to 0^+} \int_{\mathbb{R}^{N-1}} \left(\int_{(\Omega_\varepsilon)_{z_k'}} \|Du_\varepsilon(z_k', z_k) - Du(z_k', z_k)\|^p_{\mathbb{R}^N} dz_k \right) dz_k' = 0 \text{ (see (1.20)),}$$

$$\text{for all } k \in \{1, \dots, N\}.$$

So, we can find a sequence $\varepsilon_n \to 0^+$ such that for all $k \in \{1, \dots, N\}$ and for $\hat{\lambda}^{N-1}$-a.a. $z_k' \in \mathbb{R}^{N-1}$, we have

$$\lim_{n\to\infty} \int_{(\Omega_{\varepsilon_n})_{z_k'}} \|Du_{\varepsilon_n}(z_k', z_k) - Du(z_k', z_k)\|_{\mathbb{R}^N}^p dz_k = 0. \tag{1.21}$$

Let $u_n = u_{\varepsilon_n}$ and consider the set

$$C = \{z \in \Omega : \lim_{n\to\infty} u_n(z) \text{ exists in } \mathbb{R}\}.$$

We claim that C contains all the Lebesgue points of u. Indeed, let $z \in \Omega$ be a Lebesgue point of u, that is,

$$\lim_{\varepsilon\to 0^+} \frac{1}{\varepsilon^N} \int_{B_\varepsilon(z)} |u(y) - u(z)| dy = 0. \tag{1.22}$$

We have

$$\begin{aligned} |u_n(z) - u(z)| = \left| \int_\Omega u(y)\rho_{\varepsilon_n}(z-y)dy \right| &= \frac{1}{\varepsilon_n^N} \left| \int_{B_{\varepsilon_n}(z)} [u(y) - u(z)]\rho\left(\frac{z-y}{\varepsilon_n}\right) dy \right| \\ &\leqslant \|\rho\|_\infty \frac{1}{\varepsilon_n^N} \int_{B_{\varepsilon_n}(z)} |u(y) - u(z)| dz \to 0 \\ &\qquad \text{(see (1.22))}. \end{aligned}$$

This proves that C contains the Lebesgue points of u. Therefore $\hat{\lambda}^N(\Omega \setminus C) = 0$. So, if we introduce the function $\hat{u} : \Omega \to \mathbb{R}$ defined by

$$\hat{u}(z) = \begin{cases} \lim\limits_{n\to\infty} u_n(z) & \text{if } z \in C \\ 0 & \text{if } z \in \Omega \setminus C, \end{cases}$$

then $\hat{u}$ is a representative of u. We need to show that $\hat{u}$ has the properties claimed by the theorem.

We first observe that Fubini's theorem implies that for all $k \in \{1, \dots, N\}$, we have

$$\int_{\mathbb{R}^{N-1}} \left(\int_{\Omega_{z_k'}} \|Du(z_k', z_k)\|_{\mathbb{R}^N} dz_k \right) dz_k' < \infty$$
$$\text{and } \int_{\mathbb{R}^{N-1}} \hat{\lambda}\left(\{z_k \in \Omega_{z_k'} : (z_k', z_k) \notin C\}\right) dz_k' = 0.$$

So, we can find $N_k \subseteq \mathbb{R}^{N-1}$, $\hat{\lambda}^{N-1}$-null such that for all $z_k' \in \mathbb{R}^{N-1} \setminus N_k$ for which $\Omega_{z_k'}$ is nonempty, we have

$$\int_{\Omega_{z'_k}} \|Du(z'_k, z_k)\|^p_{\mathbb{R}^N} dz_k < \infty. \tag{1.23}$$

Relation (1.23) holds for all $k \in \{1, \ldots, N\}$ and $(z'_k, z_k) \in C$ for $\hat{\lambda}$-a.a. $z_k \in \Omega_{z'_k}$.

We consider a rectangle $R = \prod_{k=1}^{N} [a_k, b_k] \subseteq \Omega$, with a_k, b_k rational numbers. Note that $d(R, \partial\Omega) > 0$ and so $R \subseteq \Omega_\varepsilon$ for $\varepsilon \in (0, 1)$ sufficiently small. Then by (1.21) for all $k \in \{1, \ldots, N\}$ and for all $z'_k \in \mathbb{R}'_k \setminus N_k$ where $R = R'_k \times [a_k, b_k]$, we have

$$\lim_{n\to\infty} \int_{a_k}^{b_k} \|Du_n(z'_k, z_k) - Du(z'_k, z_k)\|_{\mathbb{R}^N} dz_k = 0. \tag{1.24}$$

Let $y_n(t) = u_n(z'_k, t)$ with $t \in [a_k, b_k]$. Let $t_0 \in [a_k, b_k]$ such that $(z'_k, t_0) \in C$. Then $y_n(t_0) \to \hat{u}(z'_k, t_0)$. Since $y_n \in C^\infty([a_k, b_k])$, we have

$$y_n(t) = y_n(t_0) + \int_{t_0}^{t} y'_n(s)ds \text{ for all } t \in [a_k, b_k], \text{ all } n \geqslant 1.$$

From (1.24) it follows that for all $t \in [a_k, b_k]$ the following limit exists

$$\begin{aligned} \lim_{n\to\infty} y_n(t) &= \lim_{n\to\infty} \left[y_n(t_0) + \int_{t_0}^{t} y'_n(s)ds \right] \\ &= \hat{u}(z'_k, t_0) + \int_{t_0}^{t} \frac{\partial u}{\partial z_k}(z'_k, s)ds = y(t). \end{aligned}$$

It follows that

$$\{z'_k\} \times [a_k, b_k] \subseteq C \tag{1.25}$$

and

$$\hat{u}(z'_k, t) = y(t) = \hat{u}(z'_k, t_0) + \int_{t_0}^{t} \frac{\partial u}{\partial z_k}(z'_k, s)ds \text{ for all } t \in [a_k, b_k]. \tag{1.26}$$

From (1.26) and the fundamental theorem of Lebesgue calculus, we deduce that $\hat{u}(z'_k, \cdot)$ is absolutely continuous on $[a_k, b_k]$ and $\frac{\partial \hat{u}}{\partial z_k}(z'_k, t) = \frac{\partial u}{\partial z_k}(z'_k, t)$ for $\hat{\lambda}$-a.a. $t \in [a_k, b_k]$.

If $\widetilde{R} = \prod_{k=1}^{N} [\widetilde{a}_k, \widetilde{b}_k]$ is another rectangle contained in Ω and such that $[a_k, b_k] \cap [\widetilde{a}_k, \widetilde{b}_k] \neq \emptyset$, taking z'_k which is admissible for both R and $\widetilde{R}$ and $t_0 \in [a_k, b_k] \cap [\widetilde{a}_k, \widetilde{b}_k]$ it follows from (1.25) and (1.26) that y is absolutely continuous on $[a_k, b_k] \cup [\widetilde{a}_k, \widetilde{b}_k]$.

Now, note that Ω can be expressed as the countable union of such rectangles and a countable union of $\hat{\lambda}^{N-1}$-null sets is $\hat{\lambda}^{N-1}$-null. So, from (1.23), (1.26) and the Banach–Zaretsky theorem we can conclude that for $\hat{\lambda}^{N-1}$-a.a. $z'_k \in \mathbb{R}^{N-1} \setminus N_k$ for which $\Omega_{z'_k}$ is nonempty, the function y is absolutely continuous on any maximal interval of $\Omega_{z'_k}$.

Now assume that u admits a representative as postulated by the theorem.

Fix $k \in \{1, ..., N\}$ and let $z'_k \in \mathbb{R}^{N-1}$ be such that $\hat{u}(z'_k, \cdot)$ is absolutely continuous on the open set $\Omega_{z'_k}$. Then by integration by parts, for every $\vartheta \in C_c^\infty(\Omega)$ we have

$$\int_{\Omega_{z'_k}} \hat{u}(z'_k, t)\frac{\partial \vartheta}{\partial z_k}(z'_k, t)dt = -\int_{\Omega_{z'_k}} \frac{\partial \hat{u}}{\partial z_k}(z'_k, t)\vartheta(z'_k, t)dt.$$

This holds for $\hat{\lambda}^{N-1}$-a.a. $z'_k \in \mathbb{R}^{N-1}$ for which $\Omega_{z'_k} \neq \emptyset$. Integrating over $\mathbb{R}^{N-1}$ and using Fubini's theorem, we obtain

$$\begin{aligned}
&\int_\Omega \hat{u}(z)\frac{\partial \vartheta}{\partial z_k}(z)dz = -\int_\Omega \frac{\partial \hat{u}}{\partial z_k}(z)\vartheta(z)dz \\
&\Rightarrow \frac{\partial \hat{u}}{\partial z_k} \in L^p(\Omega) \text{ is the weak partial derivative of } \hat{u} = u \text{ a.e. in } \Omega \\
&\Rightarrow u \in W^{1,p}(\Omega).
\end{aligned}$$

Moreover, the classical and weak partial derivatives of u coincide $\hat{\lambda}^N$-a.e. in Ω. ☐

Using this theorem, we have a chain rule for Sobolev functions. Recall that the composition of a Lipschitz function $f : \mathbb{R} \to \mathbb{R}$ with an absolutely continuous function $u : [a, b] \to \mathbb{R}$, is absolutely continuous (see Natanson [315, p. 245]) and the chain rule holds (see Leoni [262, p. 104]). Then, invoking Theorem 1.4.1, we have the following chain rule for Sobolev functions.

Proposition 1.4.2 *Assume that $u \in W^{1,p}(\Omega)$ ($1 \leqslant p \leqslant \infty$), $f : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous, and in addition $f(0) = 0$ if $\hat{\lambda}^N(\Omega)$ is infinite. Then $f \circ u \in W^{1,p}(\Omega)$ and we have*

$$D(f \circ u)(z) = f^*(u(z))Du(z) \textit{ for a.a. } z \in \Omega,$$

where $f^* : \mathbb{R} \to \mathbb{R}$ is any Borel function such that $f^* = f'$ a.e. in $\mathbb{R}$.

Remark 1.4.3 The above chain rule remains true if $W^{1,p}(\Omega)$ is replaced by $W_0^{1,p}(\Omega)$.

Proposition 1.4.4 *If $u \in W^{1,p}(\Omega)$ (resp. $u \in W_0^{1,p}(\Omega)$) with $1 \leqslant p < \infty$, then $u^+, u^-, |u| \in W^{1,p}(\Omega)$ (resp. $u^+, u^-, |u| \in W_0^{1,p}(\Omega)$) and we have*

$$Du^+(z) = \begin{cases} 0 & \textit{for a.a. } z \in \{u \leqslant 0\} \\ Du(z) & \textit{for a.a. } z \in \{u > 0\} \end{cases}$$

$$Du^-(z) = \begin{cases} -Du(z) & \text{for a.a. } z \in \{u < 0\} \\ 0 & \text{for a.a. } z \in \{u \geqslant 0\} \end{cases}$$

$$D|u|(z) = \begin{cases} -Du(z) & \text{for a.a. } z \in \{u < 0\} \\ 0 & \text{for a.a. } z \in \{u = 0\} \\ Du(z) & \text{for a.a. } z \in \{u > 0\}. \end{cases}$$

Proof We apply Proposition 1.4.2 (the chain rule) with $f(t) = |t|$. Recall that

$$u^+ = \frac{|u| + u}{2} \text{ and } u^- = \frac{|u| - u}{2}.$$

□

Proposition 1.4.5 *Assume that $u, v \in W^{1,p}(\Omega)$ (resp. $u, v \in W_0^{1,p}(\Omega)$) with $1 \leqslant p < \infty$. Then $\hat{h} = \max\{u, v\}$ and $\tilde{h} = \min\{u, v\}$, both belong to $W^{1,p}(\Omega)$ (resp. to $W_0^{1,p}(\Omega)$) and we have*

$$D\hat{h}(z) = \begin{cases} Du(z) & \text{for a.a. } z \in \{u \geqslant v\} \\ Dv(z) & \text{for a.a. } z \in \{u \leqslant v\} \end{cases}$$

$$D\tilde{h}(z) = \begin{cases} Du(z) & \text{for a.a. } z \in \{u \leqslant v\} \\ Dv(z) & \text{for a.a. } z \in \{u \geqslant v\}. \end{cases}$$

Proof Note that $\hat{h} = \max\{u, v\} = u + (v - u)^+$ and $\tilde{h} = \min\{u, v\} = u - (u - v)^+$. Hence the result is a direct consequence of Proposition 1.4.4. □

Corollary 1.4.6 *If $u \in W_{\mathrm{loc}}^{1,p}(\Omega)$ $(1 \leqslant p < \infty)$ and $\vartheta \in \mathbb{R}$, then $Du(z) = 0$ for a.a. $z \in \{z \in \Omega : u(z) = \vartheta\}$.*

Remark 1.4.7 More generally, we can say that "if $u \in W^{1,p}(\Omega)$ $(1 \leqslant p < \infty)$ and $D \subseteq \mathbb{R}$ is Lebesgue-null, then $Du(z) = 0$ for a.a. $z \in u^{-1}(D)$". The result is known as "Stampacchia's theorem" (see Stampacchia [385]).

Proposition 1.4.8 *Assume that $\{u_n\}_{n \geqslant 1} \subseteq W^{1,p}(\Omega)$ (resp. $\{u_n\}_{n \geqslant 1} \subseteq W_0^{1,p}$) with $1 \leqslant p < \infty$ and*

$$\hat{u} = \sup_{n \geqslant 1} u_n \in L^p(\Omega), \; h = \sup_{n \geqslant 1} \|Du_n\|_{\mathbb{R}^N} \in L^p(\Omega).$$

Then $\hat{u} \in W^{1,p}(\Omega)$ (resp. $\hat{u} \in W_0^{1,p}(\Omega)$).

Proof Since $u_n = u_n^+ - u_n^-$, without any loss of generality we may assume that $u_n \geqslant 0$ for all $n \geqslant 1$. Let $v_m = \max_{1 \leqslant n \leqslant m} u_n$. Then from Proposition 1.4.5 we have that $v_m \in W^{1,p}(\Omega)$ (resp. $v_m \in W_0^{1,p}(\Omega)$) and for a.a. $z \in \Omega$, we have for all $m \geqslant 1$

$$0 \leqslant v_m(z) \leqslant \hat{u}(z) \text{ for all } m \geqslant 1, \;\; v_m(z) \uparrow \hat{u}(z) \text{ and } \|Dv_m(z)\|_{\mathbb{R}^N} \leqslant h(z). \tag{1.27}$$

From (1.27) and the dominated convergence theorem, we deduce that $v_m \to \hat{u}$ in $L^p(\Omega)$.

Also, from (1.27) and at least for a subsequence, we have

$$Dv_m \xrightarrow{w} g \text{ in } L^p(\Omega, \mathbb{R}^N).$$

Using the definition of weak derivative (see Definition 1.1.6), we obtain $g = D\hat{u}$. Therefore $\hat{u} \in W^{1,p}(\Omega)$ (resp. $\hat{u} \in W_0^{1,p}(\Omega)$). ☐

Proposition 1.4.9 *The maps $u \to |u|$ and $u \to u^+$ are continuous in $W^{1,p}(\Omega)$ ($1 < p < \infty$).*

Proof Since $u^\pm = \frac{1}{2}[|u| \pm u]$, it is enough to show that $u \to |u|$ is continuous. So, suppose $u_n \to u$ in $W^{1,p}(\Omega)$. Then $|u_n| \to |u|$ in $L^p(\Omega)$. Evidently $\|D|u_n|\|_p = \|Du_n\|_p$ for all $n \geqslant 1$ (see Proposition 1.4.4). So $\|D|u_n|\|_p \to \|D|u|\|_p$ and $\{D|u_n|\}_{n\geqslant 1} \subseteq L^p(\Omega, \mathbb{R}^N)$ is bounded. Hence we may assume that $D|u_n| \xrightarrow{w} g$ in $L^p(\Omega, \mathbb{R}^N)$ and using Definition 1.1.6, we see that $g = D|u|$. Therefore $D|u_n| \xrightarrow{w} D|u|$ in $L^p(\Omega, \mathbb{R}^N)$. Since $1 < p < \infty$, the space $L^p(\Omega, \mathbb{R}^N)$ has the Kadec–Klee property and so $Du_n \to Du$ in $L^p(\Omega, \mathbb{R}^N)$. We conclude that $u_n \to u$ in $W^{1,p}(\Omega)$. ☐

Corollary 1.4.10 *(a) If $u \in W^{1,p}(\Omega)$ ($1 < p < \infty$) and $u \geqslant 0$, then we can find $\{u_n\}_{n\geqslant 1} \subseteq W^{1,p}(\Omega) \cap C^\infty(\Omega)$, $u_n \geqslant 0$ such that $u_n \to u$ in $W^{1,p}(\Omega)$.*

(b) If $u \in W_0^{1,p}(\Omega)$ ($1 < p < \infty$) and $u \geqslant 0$, then we can find $\{u_n\}_{n\geqslant 1} \subseteq C_c^\infty(\Omega)$, $u_n \geqslant 0$ such that $u_n \to u$ in $W_0^{1,p}(\Omega)$.

Proof (a) From Theorem 1.1.23, we can find $\{v_n\}_{n\geqslant 1} \subseteq W^{1,p}(\Omega) \cap C^\infty(\Omega)$ such that $v_n \to u$ in $W^{1,p}(\Omega)$. Then from Proposition 1.4.9, we have $v_n^+ \to u$ in $W^{1,p}(\Omega)$. Let $\{\rho_{\varepsilon_n}\}_{n\geqslant 1}$ be a family of mollifiers and let $u_n(z) = \int_\Omega v_n(y)\rho_{\varepsilon_n}(z-y)dy$. Then $u_n \in C^\infty(\Omega)$, $u_n \geqslant 0$ and $u_n \to u$ in $W^{1,p}(\Omega)$.

(b) Similar as (a) using Definition 1.1.14. ☐

Corollary 1.4.11 *If $\{u_n, v_n\}_{n\geqslant 1} \subseteq W^{1,p}(\Omega)$ ($1 < p < \infty$) and $u_n \to u$, $v_n \to v$ in $W^{1,p}(\Omega)$, then $\max\{u_n, v_n\} \to \max\{u, v\}$ and $\min\{u_n, v_n\} \to \min\{u, v\}$ in $W^{1,p}(\Omega)$.*

Proof Recall that

$$\max\{u_n, v_n\} = u_n + (v_n - u_n)^+ \text{ and } \min\{u_n, v_n\} = u_n - (u_n - v_n)^+$$

and use Proposition 1.4.9. ☐

Proposition 1.4.12 *If $u \in W_0^{1,p}(\Omega)$ ($1 < p < \infty$) and $0 \leqslant v(z) \leqslant u(z)$ a.e. in Ω, then $v \in W_0^{1,p}(\Omega)$.*

Proof Let $\{u_n\}_{n\geqslant 1} \subseteq C_c^\infty(\Omega)$, $u_n \geqslant 0$ for all $n \geqslant 1$ such that $u_n \to u$ in $W^{1,p}(\Omega)$ (see Corollary 1.4.10 (b)). Let $v_n = \min\{v, u_n\}$, $n \geqslant 1$. Evidently v_n has compact support and so by virtue of Proposition 1.1.24, we have $v_n \in W_0^{1,p}(\Omega)$. Moreover, from Corollary 1.4.11 we have $v_n \to v$ in $W^{1,p}(\Omega)$, hence $v \in W_0^{1,p}(\Omega)$. □

The product rule holds for bounded Sobolev functions.

Proposition 1.4.13 *If $u, v \in W^{1,p}(\Omega) \cap L^\infty(\Omega)$ $(1 \leqslant p < \infty)$, then $uv \in W^{1,p}(\Omega)$ and $D(uv) = vDu + uDv$; the result remains true if $W^{1,p}(\Omega)$ is replaced by $W_0^{1,p}(\Omega)$.*

Proof From Theorem 1.1.25, we know that we can find sequences $\{u_n\}_{n\geqslant 1}$, $\{v_n\}_{n\geqslant 1} \subseteq C_c^\infty(\mathbb{R}^N)$ such that

$$\begin{aligned}&u_n \to u,\ v_n \to v \text{ in } L^p(\Omega) \text{ and for a.a. } z \in \Omega,\\ &Du_n \to Du,\ Dv_n \to Dv,\ \text{ in } L^p(\Omega_0, \mathbb{R}^N) \text{ for all } \Omega_0 \subset\subset \Omega.\end{aligned}$$

From the proof of Theorem 1.1.25, we have

$$\|u_n\|_{L^\infty(\mathbb{R}^N)} \leqslant \|u\|_{L^\infty(\Omega)} \text{ and } \|v_n\|_{L^\infty(\mathbb{R}^N)} \leqslant \|v\|_{L^\infty(\Omega)} \text{ for all } n \geqslant 1. \qquad (1.27')$$

Also, we have

$$\int_\Omega u_n v_n \frac{\partial \vartheta}{\partial z_k} dz = -\int_\Omega \left[v_n \frac{\partial u_n}{\partial z_k} + u_n \frac{\partial v_n}{\partial z_k} \right] \vartheta dz$$
$$\text{for all } k \in \{1, \ldots, N\},\ \text{ all } \vartheta \in C_c^\infty(\Omega).$$

Passing to the limit as $n \to \infty$ and using the Lebesgue dominated convergence theorem (see (1.27′)), we obtain

$$\int_\Omega uv \frac{\partial \vartheta}{\partial z_k} dz = -\int_\Omega \left[v \frac{\partial u}{\partial z_k} + u \frac{\partial v}{\partial z_k} \right] \vartheta dz \text{ for all } \vartheta \in C_c^\infty(\Omega),$$

hence $D(uv) = vDu + uDv$.

The assertion for Sobolev functions in $W_0^{1,p}(\Omega)$ follows as above by requiring that $\{u_n\}_{n\geqslant 1},\ \{v_n\}_{n\geqslant 1} \subseteq C_c^\infty(\Omega)$. □

Proposition 1.4.14 *Assume that $u \in W_0^{1,p}(\Omega)$ $(1 \leqslant p < \infty)$ and $|v(z)| \leqslant |u(z)|$ for a.a. $z \in \Omega \setminus K$ with K a compact subset of Ω. Then $v \in W_0^{1,p}(\Omega)$.*

Proof Let $\eta \in C_c^\infty(\Omega)$ such that $0 \leqslant \eta \leqslant 1$ and $\eta|_K = 1$. We set $\hat{u} = (1-\eta)|u| + \eta v^+$. From Proposition 1.4.4 we have that $(1-\eta)|u| \in W_0^{1,p}(\Omega)$. Also, since ηv^+ has compact support, Proposition 1.1.24 implies that $\eta v^+ \in W_0^{1,p}(\Omega)$. Therefore $\hat{u} \in W_0^{1,p}(\Omega)$. Also, $0 \leqslant v^+ \leqslant \hat{u}$ a.e. in Ω. Then Proposition 1.4.12 implies that $v^+ \in W_0^{1,p}(\Omega)$. Similarly we show that $v^- \in W_0^{1,p}(\Omega)$. Hence we conclude that $v = v^+ - v^- \in W_0^{1,p}(\Omega)$. □

Proposition 1.4.15 *If Ω is bounded and $u \in W^{1,p}(\Omega)$ $(1 \leqslant p < \infty)$ satisfies $\lim_{z \to y} u(z) = 0$ for all $y \in \partial\Omega$, then $u \in W_0^{1,p}(\Omega)$.*

Proof Since $u = u^+ - u^-$, without any loss of generality we may assume that $u \geqslant 0$. For $\epsilon > 0$ let $u_\epsilon = \max\{u + \epsilon, 0\}$. Then $u_\epsilon \in W^{1,p}(\Omega)$ and has compact support (due to the hypotheses). Then Proposition 1.1.24 implies that $u_\epsilon \in W_0^{1,p}(\Omega)$. Finally, note that $u_\epsilon \to u$ in $W^{1,p}(\Omega)$ as $\epsilon \to 0^+$ and so we conclude that $u \in W_0^{1,p}(\Omega)$. ☐

1.5 Trace Theory

The study of boundary value problems calls for a rigorous definition of the concept of $u|_{\partial\Omega}$ for a Sobolev function $u \in W^{1,p}(\Omega)$. In general, the fact that $u \in L^p(\Omega)$ is not sufficient information to make sense of $u|_{\partial\Omega}$, since the set $\partial\Omega$ is Lebesgue-null (provided it is a smooth enough topological manifold) and $L^p(\Omega)$-functions are defined modulo sets of measure zero. So, we need to exploit the additional information we have, namely that $Du \in L^p(\Omega, \mathbb{R}^N)$. If $N = 1$, there is no problem since the Sobolev functions have a $C(\overline{\Omega})$-representative (see Theorem 1.2.2). However, as we already remarked for $N \geqslant 2$, this is no longer true.

The study of the boundary behavior of Sobolev functions imposes some hypothesis on the geometry of the boundary $\partial\Omega$.

Definition 1.5.1 We say that $\partial\Omega$ is Lipschitz if for each $z \in \partial\Omega$ we can find $r > 0$ and a Lipschitz function $\xi : \mathbb{R}^{N-1} \to \mathbb{R}$ such that upon rotating and relabeling the coordinate axes if necessary, we have

$$\Omega \cap R_r(z) = \{z \in \Omega : \xi(z_1, ..., z_{N-1}) < z_N\} \cap R_r(z),$$

where $R_r(z) = \{y = (y_k)_{k=1}^N \in \Omega : |y_k - z_k| < r \text{ for all } k \in \{1, ..., N\}\}$.

Remark 1.5.2 According to this definition near every $z \in \partial\Omega$, the boundary $\partial\Omega$ is the graph of a Lipschitz function. Then by Rademacher's theorem (see, for example, Gasinski and Papageorgiou [182, p. 56]), we see that the outer unit normal $n(z)$ exists for $\mathscr{H}^{N-1}$-a.a. $z \in \partial\Omega$ (here $\mathscr{H}^{N-1}$ denotes the $(N-1)$-dimensional Hausdorff measure on $\partial\Omega$).

In the Meyers–Serrin theorem (see Theorem 1.1.23), we approximated Sobolev functions by other Sobolev functions which are smooth in Ω. Next, under additional conditions on Ω, we improve this to approximation by Sobolev functions which are smooth all the way up to the boundary. This global approximation will make possible the interpretation of the expression $u|_{\partial\Omega}$ for a Sobolev function $u \in W^{1,p}(\Omega)$.

Theorem 1.5.3 *Assume that Ω is bounded with Lipschitz boundary $\partial\Omega$ and $1 \leqslant p < \infty$. Then $\overline{C^\infty(\overline{\Omega})}^{\|\cdot\|_{1,p}} = W^{1,p}(\Omega)$.*

Proof Given $z \in \partial\Omega$, let $r > 0$ and $\xi : \mathbb{R}^{N-1} \to \mathbb{R}$ be the Lipschitz function as in Definition 1.5.1. We write $R = R_r(z)$ and $\hat{R} = R_{r/2}(z)$.

Let $u \in W^{1,p}(\Omega)$ and assume that u vanishes near $\Omega \cap \partial\hat{R}$. For $y \in \Omega \cap \hat{R}$, $\varepsilon \geqslant 0$ and $\beta = \mathrm{Lip}\,(\xi) + 2$ ($\mathrm{Lip}\,(\xi)$ is the Lipschitz constant of ξ), we define

$$y^\varepsilon = y + \varepsilon\beta e_N,$$

where $\{e_k\}_{k=1}^N$ is the canonical basis of $\mathbb{R}^N$. We have $B_\varepsilon(y^\varepsilon) \subseteq \Omega \cap R$, provided that $\varepsilon > 0$ is small enough.

We consider a family $\{\rho_\varepsilon\}_{\varepsilon\in(0,1)}$ of mollifiers and define

$$\begin{aligned} u_\varepsilon(y) &= \int_\Omega u(y^\varepsilon - z)\rho_\varepsilon(z)dz \\ &= \frac{1}{\varepsilon^N}\int_{B_\varepsilon(y^\varepsilon)} u(y)\rho\left(\frac{y-z}{\varepsilon} + \beta e_N\right)dz \text{ for all } y \in \Omega \cap \hat{R}. \end{aligned}$$

We have

$$u_\varepsilon \in C^\infty(\overline{\Omega \cap \hat{R}}) \text{ and } u_n \to u \text{ in } W^{1,p}(\Omega \cap \hat{R}) \text{ as } \varepsilon \to 0^+.$$

Since $u = 0$ near $\Omega \cap \partial\hat{R}$, for sufficiently small $\varepsilon > 0$ we also have $u_\varepsilon = 0$ near $\Omega \cap \partial\hat{R}$. Hence, we can extend u_ε to be 0 on $\Omega \cap \partial\hat{R}$.

The set $\partial\Omega$ is compact and so we can find a cover $\{\hat{R}_k = \hat{R}_{r_{k/2}}(z_k)\}_{k=1}^m$ of $\partial\Omega$. Then we can find smooth functions $\{\varphi_k\}_{k=0}^m$ such that

$$\begin{aligned} &0 \leqslant \varphi_k \leqslant 1,\ \operatorname{supp}\varphi_k \subseteq \hat{R}_k \text{ for all } k \in \{1, ..., m\}, \\ &0 \leqslant \varphi_0 \leqslant 1,\ \operatorname{supp}\varphi_0 \subseteq \Omega \text{ and } \sum_{k=1}^m \varphi_k = 1 \text{ in } \Omega. \end{aligned} \tag{1.28}$$

We set $u_k = \varphi_k u$ for all $k \in \{0, \dots, m\}$. We fix $\delta > 0$ and as before, using mollification, we produce $\{v_k = (u_k)_{\varepsilon_k}\}_{k=1}^m \subseteq C^\infty(\overline{\Omega})$ such that for all $k \in \{1, ..., m\}$

$$\operatorname{supp} v_k \subseteq \overline{\Omega} \cap R_k \text{ and } \|v_k - u_k\|_{W^{1,p}(\Omega\cap R_k)} < \frac{\delta}{2m}. \tag{1.29}$$

We also mollify u_0 and as in the proof of Theorem 1.1.23, we produce $v_0 \in C_c^\infty(\Omega)$ such that

$$\|v_0 - u_0\|_{1,p} < \frac{\delta}{2}.$$

Then we set

$$v = \sum_{k=0}^m v_k \in C^\infty(\overline{\Omega}). \tag{1.30}$$

We have

$$\|v - u\|_{1,p} \leqslant \|v_0 - u_0\|_{1,p} + \sum_{k=1}^{m} \|v_k - u_k\|_{W^{1,p}(\Omega \cap R_k)} < \delta \text{ (see (1.28), (1.29), (1.30)).}$$

This proves the density of $C^\infty(\overline{\Omega})$ in $W^{1,p}(\Omega)$. $\square$

This stronger density property of $W^{1,p}(\Omega)$ leads to the result that gives meaning to the expression $u|_{\partial\Omega}$ for $u \in W^{1,p}(\Omega)$.

Theorem 1.5.4 *Assume that Ω is bounded with Lipschitz boundary $\partial\Omega$ and $1 \leqslant p < \infty$. Then there exists a bounded linear operator $\gamma_0 : W^{1,p}(\Omega) \to L^p(\partial\Omega, \mathscr{H}^{N-1})$ such that $\gamma_0(u) = u|_{\partial\Omega}$ for all $u \in W^{1,p}(\Omega) \cap C(\overline{\Omega})$; the map γ_0 is called the trace operator of order zero.*

Proof First we consider $u \in C^1(\overline{\Omega}) \subseteq W^{1,p}(\Omega)$. Let $z \in \partial\Omega$ and let $r > 0$ and $\xi : \mathbb{R}^{N-1} \to \mathbb{R}$ be a Lipschitz function as in Definition 1.5.1.

Let $R = R_r(z)$ and assume momentarily that $u = 0$ in $\Omega \setminus R$. Note that

$$-(e_N, n)_{\mathbb{R}^N} \geqslant (1 + \mathrm{Lip}\,(\xi)^2)^{-\frac{1}{2}} > 0, \ \ \mathscr{H}^{N-1}\text{-a.e. on } \partial\Omega \cap R. \tag{1.31}$$

We fix $\varepsilon > 0$ and set

$$\mu_\varepsilon(t) = (t^2 + \varepsilon^2)^{\frac{1}{2}} - \varepsilon \text{ for all } t \in \mathbb{R}. \tag{1.32}$$

We have

$$\begin{aligned}
\int_{\partial\Omega} \mu_\varepsilon(u) d\mathscr{H}^{N-1} &= \int_{\partial\Omega \cap R} \mu_\varepsilon(u) d\mathscr{H}^{N-1} \\
&\leqslant c \int_{\partial\Omega \cap R} \mu_\varepsilon(u)(-(e_N, n)_{\mathbb{R}^N}) d\mathscr{H}^{N-1} \text{ with } c > 0 \text{ (see (1.31))} \\
&\leqslant -c \int_{\Omega \cap R} \frac{\partial}{\partial z_N}(\mu_\varepsilon(u)) dz \text{ (by the divergence theorem)} \\
&\leqslant c \int_{\Omega \cap R} |\mu_\varepsilon'(u)| \|Du\|_{\mathbb{R}^N} dz \\
&\leqslant c \int_{\Omega} \|Du\|_{\mathbb{R}^N} dz \text{ (since } |\mu_\varepsilon'(t)| \leqslant 1 \text{ for all } t \in \mathbb{R}, \text{ see (1.32)).}
\end{aligned}$$

Let $\varepsilon \to 0^+$ to obtain

$$\int_{\partial\Omega} |u| d\mathscr{H}^{N-1} \leqslant c \int_{\Omega} \|Du\|_{\mathbb{R}^N} dz. \tag{1.33}$$

We have proved (1.33) under the assumption that $u = 0$ on $\Omega \setminus R$. In the general case we cover ∂U by a finite number of cubes $\{R_k\}_{k=1}^m$ (recall that $\partial\Omega$ is compact).

Then we consider a partition of unity subordinate to this cover and as in the proof of Theorem 1.5.3, we show that

$$\int_{\partial\Omega} |u|^p d\mathscr{H}^{N-1} \leqslant c \int_{\Omega} (|u|^p + \|Du\|^p_{\mathbb{R}^N}) dz \text{ for all } u \in C^1(\overline{\Omega}). \tag{1.34}$$

So, if we set $\gamma_0(u) = u|_{\partial\Omega}$, by relation (1.34) we obtain that $\gamma_0 : C^1(\overline{\Omega}) \to L^p(\partial\Omega, \mathscr{H}^{N-1})$ is linear and continuous. The density of $C^1(\overline{\Omega})$ in $W^{1,p}(\Omega)$ (see Theorem 1.5.3) finally gives a linear continuous map $\gamma_0 : W^{1,p}(\Omega) \to \mathbb{R}$ such that $\gamma_0(u) = u|_{\partial\Omega}$ for all $u \in W^{1,p}(\Omega) \cap C(\overline{\Omega})$. □

As a direct consequence of Theorems 1.5.3 and 1.5.4, we have the following generalized integration by parts formula.

Proposition 1.5.5 *Assume that Ω is bounded with $\partial\Omega$ Lipschitz and $1 \leqslant p < \infty$. Then for all $\eta \in C^1(\mathbb{R}^N, \mathbb{R}^N)$ and all $u \in W^{1,p}(\Omega)$ we have*

$$\int_{\Omega} u(\operatorname{div} \eta) dz + \int_{\Omega} (Du, \eta)_{\mathbb{R}^N} dz = \int_{\partial\Omega} \gamma_0(u)(\eta, n)_{\mathbb{R}^N} d\mathscr{H}^{N-1}.$$

Proposition 1.5.6 *Assume that Ω is bounded with Lipschitz boundary $\partial\Omega$ and $1 \leqslant p < \infty$. Then $W_0^{1,p}(\Omega) = \ker \gamma_0$.*

Proof If $u \in C_c^\infty(\Omega)$, then $\gamma_0(u) = 0$. Since $W_0^{1,p}(\Omega)$ is the $\|\cdot\|_{1,p}$–closure of $C_c^\infty(\Omega)$ (see Definition 1.1.14) and γ_0 is continuous on $W^{1,p}(\Omega)$ (see Theorem 1.5.4), we infer that $\gamma_0(u) = 0$ for all $u \in W_0^{1,p}(\Omega)$. Therefore

$$W_0^{1,p}(\Omega) \subseteq \ker \gamma_0.$$

Next, let $u \in \ker \gamma_0$. Using partitions of unity and flattening out $\partial\Omega$, without any loss of generality, we may assume that $\Omega = \mathbb{R}^N_+ = \{z = (z', z_N) : z' \in \mathbb{R}^{N-1},\ z_N \geqslant 0\}$ and $u = 0$ for all $z \in \mathbb{R}^N_+$, with $\|z\|_{\mathbb{R}^N} \geqslant M > 0$. For $z \in \mathbb{R}^N$ we define

$$\hat{u}(z) = \begin{cases} 0 & \text{if } z_N \leqslant 0 \\ u(z) & \text{if } z_N > 0. \end{cases}$$

We have

$$\int_{\mathbb{R}^N} \hat{u} \frac{\partial \vartheta}{\partial z_k} dz = \int_{\mathbb{R}^N_+} u \frac{\partial \vartheta}{\partial z_k} dz = -\int_{\mathbb{R}^N_+} \vartheta \frac{\partial u}{\partial z_k} dz \text{ for all } \vartheta \in C_c^\infty(\mathbb{R}^N)$$
$$\text{for all } k \in \{1, ..., N\}$$

$$\Rightarrow \hat{u} \in W^{1,p}(\mathbb{R}^N) \text{ and } \frac{\partial \hat{u}}{\partial z_k}(z) = \begin{cases} 0 & \text{if } z_N \leqslant 0 \\ \dfrac{\partial u}{\partial z_k}(z) & \text{if } z_N > 0 \end{cases}$$
$$\text{for all } k \in \{1, \ldots, N\}.$$

For $t > 0$ and $z \in \mathbb{R}^N$, let

$$\hat{u}_t(z) = \hat{u}(z', z_N - t) \quad \text{(translation of } \hat{u} \text{ upwards)}.$$

We see that supp $\hat{u}_t$ is compact and contained in $\mathbb{R}^{N-1} \times (\frac{t}{2}, \infty)$. Given $\delta > 0$ we can find $t > 0$ small such that

$$\|u - \hat{u}_t\|_{W^{1,p}(\mathbb{R}^N_+)} \leqslant \delta. \tag{1.35}$$

Using regularization through mollification, we can find $0 < \varepsilon < \frac{t}{4}$ small such that

$$\|(\hat{u}_t)_\varepsilon - \hat{u}_\varepsilon\|_{W^{1,p}(\mathbb{R}^N_+)} \leqslant \delta. \tag{1.36}$$

If $\{\rho_\varepsilon\}_{\varepsilon\in(0,1)}$ is the family of mollifiers used above, then for all $z \in \mathbb{R}^N$ we have

$$(\hat{u}_t)_\varepsilon(z) = \int_{\mathbb{R}^N} \hat{u}_t(y)\rho_\varepsilon(z-y)dy = \int_{\mathbb{R}^{N-1}\times(\frac{t}{2},\infty)} \hat{u}_t(y)\rho_\varepsilon(z-y)dy.$$

Since supp $\rho_\varepsilon \subseteq \overline{B}_\varepsilon(0)$, it follows that if $0 < z_N < \varepsilon$, then $(\hat{u}_t)_\varepsilon(z) = 0$.

Therefore $\hat{u}_\varepsilon \in C_c^\infty(\mathbb{R}^N_+)$ and so we conclude that $u \in W_0^{1,p}(\mathbb{R}^N_+)$ (see (1.35) and (1.36) and recall that $\delta > 0$ is arbitrary). □

The trace theorem (see Theorem 1.5.4) leads to an extension of Green's theorem to Sobolev functions in $H^1(\Omega)$ with $\Omega \subseteq \mathbb{R}^N$ bounded with Lipschitz boundary $\partial\Omega$. In what follows, we denote by $n(z)$ the outward unit normal on $\partial\Omega$. We already remarked that the assumption on the boundary $\partial\Omega$ implies that $n(z)$ is defined uniquely for $\mathscr{H}^{N-1}$-a.a. $z \in \partial\Omega$ (Rademacher's theorem). We write generically $n(z) = (n_k(z))_{k=1}^N$.

Theorem 1.5.7 *Assume that Ω is bounded with Lipschitz boundary $\partial\Omega$. Then for all $u, v \in H^1(\Omega)$ and all $k \in \{1, ..., N\}$, we have*

$$\int_\Omega u\frac{\partial v}{\partial z_k}dz + \int_\Omega v\frac{\partial u}{\partial z_k}dz = \int_{\partial\Omega} \gamma_0(u)\gamma_0(v)n_k d\mathscr{H}^{N-1}.$$

Proof From Theorem 1.5.3, we know that there exist $\{u_n\}_{n\geqslant1},\ \{v_n\}_{n\geqslant1} \subseteq C^\infty(\overline{\Omega})$ such that

$$u_n \to u \text{ and } v_n \to v \text{ in } H^1(\Omega). \tag{1.37}$$

Then from the classical Green's theorem, we have

$$\int_\Omega u_n\frac{\partial v_n}{\partial z_k}dz + \int_\Omega v_n\frac{\partial u_n}{\partial z_k}dz = \int_{\partial\Omega} u_n v_n n_k d\mathscr{H}^{N-1} \text{ for all } n \geqslant 1.$$

Passing to the limit as $n \to \infty$ and using (1.37) and the continuity of the trace map, we obtain

$$\int_\Omega u \frac{\partial v}{\partial z_k} dz + \int_\Omega v \frac{\partial u}{\partial z_k} dz = \int_{\partial\Omega} \gamma_0(u)\gamma_0(v) n_k d\mathscr{H}^{N-1}. \tag{1.38}$$

$\square$

Setting $u \equiv 1$ and $v = v_k \in H^1(\Omega)$, from Theorem 1.5.7 we have

$$\int_\Omega \frac{\partial v_k}{\partial z_k} dz = \int_{\partial\Omega} v_k n_k d\mathscr{H}^{N-1} \text{ for all } k \in \{1, ..., N\}. \tag{1.39}$$

So, if we set $\bar{v} = (v_k)_{k=1}^N \in H^1(\Omega, \mathbb{R}^N)$, then from (1.39), we have

$$\int_\Omega \operatorname{div} \bar{v} dz = \int_{\partial\Omega} (\bar{v}, n)_{\mathbb{R}^N} d\mathscr{H}^{N-1},$$

which is the divergence theorem. If $u \in H^2(\Omega)$ and in (1.38) we replace u by $\frac{\partial u}{\partial z_k}$, then

$$\int_\Omega \frac{\partial u}{\partial z_k} \frac{\partial v}{\partial z_k} dz + \int_\Omega v \frac{\partial^2 u}{\partial z_k^2} dz = \int_{\partial\Omega} \frac{\partial u}{\partial z_k} n_k d\mathscr{H}^{N-1}.$$

If u is smooth, then $\sum_{k=1}^N \frac{\partial u}{\partial z_k} n_k = (Du, n)_{\mathbb{R}^N} = \frac{\partial u}{\partial n}$ (the directional derivative in the outward normal direction on $\partial\Omega$). Exploiting the continuity of the trace operator (see Theorem 1.5.4) and the density of $C^\infty(\overline{\Omega})$ in $H^1(\Omega)$ (see Theorem 1.5.3), we reach the following Green's identity.

Proposition 1.5.8 *Assume that Ω is bounded with Lipschitz boundary $\partial\Omega$, $u \in H^2(\Omega)$, and $v \in H^1(\Omega)$. Then*

$$\int_\Omega (Du, Dv)_{\mathbb{R}^N} dz + \int_\Omega (\Delta u) v dz = \int_{\partial\Omega} \left(\frac{\partial u}{\partial n}\right) v d\mathscr{H}^{N-1}.$$

Next, we extend this formula to all Sobolev spaces $W^{1,p}(\Omega)$ $(1 < p < \infty)$ and drop the extra regularity condition $u \in H^2(\Omega)$. To do this, we need to introduce Sobolev spaces of fractional order on manifolds.

Definition 1.5.9 Let M be a compact manifold in $\mathbb{R}^N$. For $s \in (0, 1)$, $p \in [1, +\infty)$ and $u \in C^\infty(M)$, we define

$$\|u\|_{W^{s,p}(M)} = \left[\int_M |u(z)|^p d\sigma + \int_{M \times M} \frac{|u(z) - u(y)|^p}{\|z - y\|_{\mathbb{R}^N}^{N-1+sp}} d\sigma d\sigma\right]^{1/p}$$

with σ being the surface measure on the manifold. This is a norm.

We define the Sobolev space of fractional order $W^{s,p}(M)$ to be the completion of $C^\infty(M)$ for this norm. For any $s > 0$, we write $s = k + \eta$ with k a nonnegative integer and $\eta \in (0, 1)$ (if s is not an integer). We define

$$W^{s,p}(M) = \{u \in W^{k,p}(M) : \ D^\alpha u \in W^{\eta,p}(M) \text{ for all } |\alpha| = k\}.$$

Remark 1.5.10 This definition makes sense for any bounded open set $\Omega \subseteq \mathbb{R}^N$. Also, if $s = 0$, then $W^{0,p}(M) = L^p(M)$.

Using this extended class of Sobolev spaces, we obtain the full version of the trace theorem (see Adams [2]).

Theorem 1.5.11 *Assume that Ω is bounded with Lipschitz boundary $\partial\Omega$, $m \geqslant 1$ is an integer, and $1 \leqslant p < \infty$. Then there exists a unique bounded linear operator $\gamma = (\gamma_k)_{k=0}^{m-1} : W^{m,p}(\Omega) \to L^p(\partial\Omega, \mathbb{R}^m)$ such that*

(a) if $u \in C^\infty(\overline{\Omega})$, then $\gamma_k(u) = \frac{\partial^k u}{\partial n^k}$ for all $k \in \{1, ..., m-1\}$;

(b) $\operatorname{range} \gamma = \prod_{k=0}^{m-1} W^{m-k-\frac{1}{p'},p}(\partial\Omega)$*;*

(c) $\ker \gamma = W_0^{m,p}(\Omega)$.

Remark 1.5.12 In particular, from this theorem we infer that the zero order trace operator γ_0 is not surjective. In fact, we have $\gamma_0(W^{1,p}(\Omega)) = W^{\frac{1}{p'},p}(\partial\Omega)$ (recall that $\frac{1}{p} + \frac{1}{p'} = 1$).

To have the desired generalization of Green's formula for $p \in (1, +\infty)$ and drop the requirement that $u \in H^2(\Omega)$ (see Proposition 1.5.8), we introduce the following space

$$V_p(\operatorname{div}, \Omega) = \{u \in L^p(\Omega, \mathbb{R}^N) : \ \operatorname{div} u \in L^p(\Omega)\},$$

where, as before, Ω is a bounded open set in $\mathbb{R}^N$ with Lipschitz boundary. The space $V_p(\operatorname{div}, \Omega)$ is endowed with the norm

$$\|u\|_{V_p} = \left[\|u\|^p_{L^p(\Omega,\mathbb{R}^N)} + \|\operatorname{div} u\|_p^p\right]^{\frac{1}{p}}.$$

Proposition 1.5.13 *The space $V_p(\operatorname{div}, \Omega)$ is separable, reflexive and $C^\infty(\overline{\Omega}, \mathbb{R}^N)$ is a dense subspace.*

Proof The separability and reflexivity of $V_p(\operatorname{div}, \Omega)$ are obvious. Moreover, from the classical Riesz representation theorem, we have that if $L \in V_p(\operatorname{div}, \Omega)^*$, then there exist $g \in L^{p'}(\Omega, \mathbb{R}^N)$ and $f \in L^{p'}(\Omega)$ such that

$$L(u) = \int_\Omega (g(z), u(z))_{\mathbb{R}^N} dz + \int_\Omega f(z) \operatorname{div} u(z) dz \text{ for all } u \in V_p(\operatorname{div}, \Omega). \quad (1.40)$$

To prove the density of $C^\infty(\overline{\Omega}, \mathbb{R}^N)$ in $V_p(\mathrm{div}, \Omega)$, we show that

$$\text{“}L|_{C^\infty(\overline{\Omega},\mathbb{R}^N)} = 0 \Rightarrow L \equiv 0.\text{”} \tag{1.41}$$

From (1.40) and (1.41), we have

$$\begin{aligned} &-\langle Df, y\rangle = \int_\Omega f \,\mathrm{div}\, y dz = -\int_\Omega g y dz \text{ for all } y \in C_c^\infty(\Omega, \mathbb{R}^N) \\ &\Rightarrow Df = g \text{ (see Definition 1.1.6)} \\ &\Rightarrow f \in W^{1,p}(\Omega). \end{aligned}$$

From (1.41) and Proposition 1.5.5, we obtain

$$\begin{aligned} 0 &= \int_\Omega f(\mathrm{div}\, u)dz + \int_\Omega (Df, u)_{\mathbb{R}^N} dz = \int_{\partial\Omega} \gamma_0(f)(u, n)_{\mathbb{R}^N} d\mathscr{H}^{N-1} \\ &\qquad \text{for all } u \in C^\infty(\overline{\Omega}, \mathbb{R}^N) \\ \Rightarrow 0 &= \int_{\partial\Omega} \gamma_0(f)\frac{\partial v}{\partial n} d\mathscr{H}^{N-1} \text{ for all } v \in C^\infty(\overline{\Omega}). \end{aligned} \tag{1.42}$$

But $\{\frac{\partial v}{\partial n} : v \in C^\infty(\overline{\Omega})\}$ is dense in $L^p(\partial\Omega)$. Then from (1.42) it follows that $\gamma_0(f) = 0$, hence $f \in W_0^{1,p}(\Omega)$. So, we can find $\{f_n\}_{n\geqslant 1} \subseteq C_c^\infty(\Omega)$ such that $f_n \to f$ in $W^{1,p}(\Omega)$. From (1.40) we have

$$\begin{aligned} L(u) &= \int_\Omega (Df, u)_{\mathbb{R}^N} dz + \int_\Omega f(\mathrm{div}\, u)dz \\ &= \lim_{n\to\infty} \left[\int_\Omega (Df_n, u)_{\mathbb{R}^N} dz + \int_\Omega f_n(\mathrm{div}\, u)dz \right] = 0 \text{ for all } u \in V_p(\mathrm{div}, \Omega) \end{aligned}$$

$$\Rightarrow L \equiv 0 \text{ and this proves the density of } C^\infty(\overline{\Omega}, \mathbb{R}^N) \text{ in } V_p(\mathrm{div}, \Omega).$$

The proof is now complete. □

Proposition 1.5.14 *Assume that Ω is bounded with Lipschitz boundary $\partial\Omega$ and $1 < p < \infty$. Then there exists a unique continuous map $\gamma_n : V_p(\mathrm{div}, \Omega) \to W^{-\frac{1}{p},p}(\partial\Omega)$ such that $\gamma_n(u) = (u, n)_{\mathbb{R}^N}$ for all $u \in C^\infty(\overline{\Omega}, \mathbb{R}^N)$. Moreover, we have $\int_\Omega (Dy, u)_{\mathbb{R}^N} dz + \int_\Omega y(\mathrm{div}\, u)dz = \langle \gamma_n(u), \gamma_0(y)\rangle_{\partial\Omega}$ for all $u \in V_p(\mathrm{div}, \Omega)$, all $y \in W^{1,p'}(\Omega)$ and with $\langle \cdot, \cdot \rangle_{\partial\Omega}$ being the duality brackets for the pair $(W^{-\frac{1}{p},p}(\partial\Omega) = W^{\frac{1}{p},p'}(\partial\Omega)^*, W^{\frac{1}{p},p'}(\partial\Omega))$.*

Proof From Proposition 1.5.5 we know that for all $u \in C^\infty(\overline{\Omega}, \mathbb{R}^N)$ and all $y \in W^{1,p'}(\Omega)$, we have

$$\int_\Omega (Dy, u)_{\mathbb{R}^N} dz + \int_\Omega y(div\, u)dz = \int_{\partial\Omega} \gamma_0(y)(u, n)_{\mathbb{R}^N} d\mathscr{H}^{N-1}. \tag{1.43}$$

Using Hölder's inequality, we have

$$\left| \int_{\partial\Omega} \gamma_0(y)(u,n)_{\mathbb{R}^N} d\mathscr{H}^{N-1} \right| \leqslant \|u\|_{V_p} \|\gamma_0(y)\|_{W^{\frac{1}{p},p'}(\partial\Omega)}.$$

Recalling that $\operatorname{im}\gamma_0 = W^{\frac{1}{p},p'}(\partial\Omega)$ (see Theorem 1.5.11), we deduce that for all $h \in W^{\frac{1}{p},p'}(\partial\Omega)$

$$\left| \int_{\partial\Omega} h(u,n)_{\mathbb{R}^N} d\mathscr{H}^{N-1} \right| \leqslant \|u\|_{V_p} \|h\|_{W^{\frac{1}{p},p'}}(\partial\Omega) \text{ (see (1.43))}. \tag{1.44}$$

By virtue of (1.44), given $u \in C^\infty(\overline{\Omega}, \mathbb{R}^N)$ we identify $\gamma_n(u) = (u,n)_{\mathbb{R}^N}$ with an element of $W^{-\frac{1}{p},p}(\partial\Omega) = W^{\frac{1}{p},p'}(\partial\Omega)^*$. We have

$$\begin{aligned} &\|\gamma_n(u)\|_{W^{-\frac{1}{p},p}(\partial\Omega)} \leqslant \|u\|_{V_p} \text{ for all } u \in C^\infty(\overline{\Omega}, \mathbb{R}^N) \\ \Rightarrow\ &\|\gamma_n(u)\|_{W^{-\frac{1}{p},p}(\partial\Omega)} \leqslant \|u\|_{V_p} \text{ for all } u \in V_p(\operatorname{div}, \Omega) \text{ (see Proposition 1.5.13).} \end{aligned}$$

The proof is now complete. □

Remark 1.5.15 In fact, using the theory of maximal monotone maps (see Chap. 2), one can show that $\operatorname{im}\gamma_n = W^{-\frac{1}{p},p}(\partial\Omega)$ and $\|\gamma_n\|_{\mathscr{L}} = 1$. For details, see Casas and Fernandez [105].

Consider a quasilinear differential operator A defined by

$$A(u) = -\sum_{i=1}^{N} D_k a_k(z,u,Du) + a_0(z,u,Du), \tag{1.45}$$

where $a_0,\ a_k : \Omega \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ ($k \in \{1, ..., N\}$) are Carathéodory functions (that is, for all $(x,y) \in \mathbb{R} \times \mathbb{R}^N$, $z \to a_0(z,x,y),\ a_k(z,x,y)$ are measurable and for a.a. $z \in \Omega$, $(x,y) \to a_0(z,x,y),\ a_k(z,x,y)$ is continuous) satisfying

$$|a_0(z,x,y)|,\ |a_k(z,x,y)| \leqslant c(|x|^{p-1} + \|y\|_{\mathbb{R}^N}^{p-1} + \xi(z))$$

for all a.a. $z \in \Omega$, all $(x,y) \in \mathbb{R} \times \mathbb{R}$, with $\xi \in L^{p'}(\Omega)$.

Corollary 1.5.16 *Assume that Ω is bounded with Lipschitz boundary $\partial\Omega$, $u \in W^{1,p}(\Omega)$ $(1 < p < \infty)$, and $A(u) \in L^{p'}(\Omega)$. Then there exists a unique element of $W^{-\frac{1}{p^*},p'}(\partial\Omega)$, which by extension is denoted by $\frac{\partial u}{\partial n_a} = \sum_{k=1}^{N} a_k(z,u,Du)n_k$, that satisfies the following nonlinear Green's identity*

$$\sum_{k=1}^{N} \int_{\Omega} a_k(z, u, Du) D_k v \, dz + \int_{\Omega} a_0(z, u, Du) Dv \, dz$$
$$= \int_{\Omega} A(u) v \, dz + \left\langle \frac{\partial u}{\partial n_a}, \ \gamma_0(v) \right\rangle_{\partial\Omega} \quad \textit{for all } v \in W^{1,p}(\Omega).$$

A particular case of interest is when $a_k(z, x, y) = ||y||^{p-2} y_k$ for all $k \in \{1, \dots, N\}$ and $a_0 = 0$. Then the corresponding differential operator is the p-Laplacian defined by

$$\Delta_p u = \mathrm{div}\,(||Du||^{p-2} Du) \text{ for all } u \in W^{1,p}(\Omega)\,.$$

Then Corollary 1.5.16 takes the following form.

Corollary 1.5.17 *Assume that Ω is bounded with Lipschitz boundary $\partial\Omega$, $u \in W^{1,p}(\Omega)$ $(1 < p < \infty)$ and $\Delta_p u \in L^{p'}(\Omega)$. Then there exists a unique element of $W^{-\frac{1}{p'},p'}(\partial\Omega)$, which by extension is denoted by*

$$\frac{\partial u}{\partial n_p} = \|Du\|_{\mathbb{R}^N}^{p-2} (Du, n)_{\mathbb{R}^N} = \|Du\|^{p-2} \frac{\partial u}{\partial n},$$

that satisfies the following nonlinear Green identity

$$\int_{\Omega} (\Delta_p u) v dz + \int_{\Omega} (Du, Dv)_{\mathbb{R}^N} dz = \left\langle \frac{\partial u}{\partial n_p}, \gamma_0(v) \right\rangle_{\partial\Omega} \quad \textit{for all} v \in W^{1,p}(\Omega).$$

1.6 The Extension Operator

In Proposition 1.1.17 we saw that starting with $u \in W_0^{1,p}(\Omega)$ and extending it by zero on $\mathbb{R}^N \setminus \Omega$, we still get a Sobolev function, that is, the extended function belongs in $W^{1,p}(\mathbb{R}^N) = W_0^{1,p}(\mathbb{R}^N)$ (see Theorem 1.1.21). We remarked that this is no longer true if $u \in W^{1,p}(\Omega)$. In fact, the true situation for $W^{1,p}(\Omega)$ is more delicate and in this case the geometry of the boundary plays an important role.

Theorem 1.6.1 *Assume that Ω is bounded with Lipschitz boundary $\partial\Omega$, $1 \leqslant p < \infty$, and $\Omega \subset\subset \Omega_1$. Then there exists a bounded linear operator $P : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^N)$ such that*

$$Pu|_{\Omega} = u \textit{ and } \mathrm{supp}\, Pu \subseteq \Omega_1,$$

$\|Pu\|_{L^p(\mathbb{R}^N)} \leqslant c\|u\|_p$ and $\|Pu\|_{W^{1,p}(\mathbb{R}^N)} \leqslant c\|u\|_{1,p}$ for all $u \in W^{1,p}(\Omega)$, with $c > 0$ depending only on Ω and p.

Proof We start by fixing our notation.

For $z = (z_k)_{k=1}^N \in \Omega$, we write $z = (z', z_N)$ with $z' = (z_k)_{k=1}^{N-1} \in \mathbb{R}^{N-1}$ and $z_N \in \mathbb{R}$. Also for $z \in \mathbb{R}^N$ and $r, \eta > 0$, we define the cylinder

$$C(z,r,\eta)=\{y\in\mathbb{R}^N:\|y'-z'\|_{\mathbb{R}^{N-1}}<r,\,|y_N-z_N|<\eta\}.$$

Exploiting the fact that $\partial\Omega$ is Lipschitz (see Definition 1.5.1), we can find $r,\eta>0$ and a Lipschitz function $\xi:\mathbb{R}^{N-1}\to\mathbb{R}$ such that

$$\left\{\begin{array}{l}\max\{|\xi(y')-z_N|:\|y'-z'\|_{\mathbb{R}^{N-1}}<r\}<\dfrac{\eta}{4},\\ \Omega\cap C(z,r,\eta)=\{y\in\Omega:\|y'-z'\|_{\mathbb{R}^{N-1}}<r,\ \xi(y')<y_N<z_N+\eta\},\\ C(z,r,\eta)\subseteq\Omega_1.\end{array}\right\}\tag{1.46}$$

Fix $z\in\partial\Omega$ and r,η,ξ as in (1.46). We set

$$C=C(z,r,\eta),\ \hat{C}=C(z,\frac{r}{2},\frac{\eta}{2}),\ \Omega_+=\Omega\cap\hat{C},\ \Omega_-=\hat{C}\setminus\overline{\Omega}.$$

We first suppose that $u\in C^1(\overline{\Omega})$ and assume that $\operatorname{supp}u\subseteq\overline{\Omega}\cap\hat{C}$. Let

$$\begin{aligned}&u^+(y)=u(y)\text{ for all }y\in\overline{\Omega}_+,\\&u^-(y)=u(y',2\xi(y')-y_N)\text{ for all }y\in\overline{\Omega}_-.\end{aligned}$$

Then $u^+=u^-=u$ on $\partial\Omega\cap\hat{C}$.

Claim 1. We have $\|u^-\|_{W^{1,p}(\Omega_-)}\leqslant c\|u\|_{1,p}$.
Let $\vartheta\in C_c^\infty(\Omega_-)$ and let $\{\xi_n\}_{n\geqslant1}$ be a sequence of C^∞-functions such that

$$\left\{\begin{array}{l}\xi_n\geqslant\xi,\ \xi_n\to\xi\text{ uniformly on }\mathbb{R}^{N-1},\\ D\xi_n(z)\to D\xi(z)\text{ a.e. in }\mathbb{R}^{N-1},\ \sup\limits_{n\geqslant1}\|D\xi_n\|_{L^\infty(\mathbb{R}^{N-1})}<\infty.\end{array}\right\}\tag{1.47}$$

For every $k\in\{1,...,N-1\}$, we have

$$\begin{aligned}\int_{\Omega_-}u^-\frac{\partial\vartheta}{\partial z_k}dz&=\int_{\Omega_-}u(z',2\xi(z')-z_N)\frac{\partial\vartheta}{\partial z_k}dz\\&=\lim_{n\to\infty}\int_{\Omega_-}u(z',2\xi_n(z')-z_N)\frac{\partial\vartheta}{\partial z_k}dz\\&=-\lim_{n\to\infty}\int_{\Omega_-}\Big[\frac{\partial u}{\partial z_k}(z',2\xi_n(z')-z_N)\\&\qquad\qquad+2\frac{\partial u}{\partial z_N}(z',2\xi_n(z')-z_N)\frac{\partial\xi_n}{\partial z_k}(z')\Big]\vartheta dz\\&=-\int_{\Omega_-}\Big[\frac{\partial u}{\partial z_k}(z',2\xi(z')-z_N)+2\frac{\partial u}{\partial z_N}(z',2\xi(z')-z_N)\frac{\partial\xi}{\partial z_k}(z')\Big]\vartheta dz\\&\qquad\qquad\text{(see (1.47)).}\end{aligned}$$

Similarly we show that

$$\int_{\Omega_-}u^-\frac{\partial\vartheta}{\partial z_N}dz=\int_{\Omega_-}\frac{\partial u}{\partial z_N}(z',2\xi(z')-z_N)\vartheta dz.$$

From (1.47) and using the change of variables formula, we have

$$\int_{\Omega_-} \|Du(z', 2\xi(z') - z_N)\|_{\mathbb{R}^N}^p dz \leqslant c_0 \int_{\Omega} \|Du\|_{\mathbb{R}^N}^p dz < \infty$$

with $c_0 > 0$ (depending on Ω and p). This proves Claim 1.

We define

$$Pu(z) = \bar{u}(z) = \begin{cases} u^+(z) & \text{if } z \in \overline{\Omega}_+ \\ u^-(z) & \text{if } z \in \overline{\Omega}_- \\ 0 & \text{if } z \in \mathbb{R}^N \setminus (\overline{\Omega}_+ \cup \overline{\Omega}_-). \end{cases}$$

Evidently $Pu \in C(\mathbb{R}^N)$.

Claim 2. We have $Pu \in W^{1,p}(\mathbb{R}^N)$, $\operatorname{supp} Pu \subseteq \hat{C} \subseteq \Omega_1$ and

$$\|Pu\|_{W^{1,p}(\mathbb{R}^N)} \leqslant c\|u\|_{1,p} .$$

Let $\vartheta \in C_c^1(\hat{C})$. For $k \in \{1, ..., N\}$ we have

$$\begin{aligned} \int_{\hat{C}} \bar{u}\frac{\partial\vartheta}{\partial z_k} dz &= \int_{\Omega_+} u^+ \frac{\partial\vartheta}{\partial z_k} dz + \int_{\Omega_-} u^- \frac{\partial\vartheta}{\partial z_k} dz \\ &= -\int_{\Omega_+} \frac{\partial u^+}{\partial z_k}\vartheta dz - \int_{\Omega_-} \frac{\partial u^-}{\partial z_k}\vartheta dz + \int_{\partial\Omega} [\gamma_0(u^+) - \gamma_0(u^-)]\vartheta n_k d\mathscr{H}^{N-1} \\ &\qquad \text{(see Theorem 1.5.7)} \\ &= -\int_{\Omega_+} \frac{\partial u^+}{\partial z_k}\vartheta\, dz - \int_{\Omega_-} \frac{\partial u^-}{\partial z_k}\vartheta dz \text{ (since } \gamma_0(u^+) = \gamma_0(u^-) = \gamma_0(u)). \end{aligned}$$

It follows that $\|Pu\|_{W^{1,p}(\mathbb{R}^N)} \leqslant c\|u\|_{1,p}$ (see Claim 1). This proves Claim 2.

So, we have proved the theorem for $u \in C^1(\overline{\Omega})$ with $\operatorname{supp} u \subseteq \overline{\Omega} \cap \hat{C}$. Now we remove the support restriction. The boundary $\partial\Omega$ is compact and so it can be covered by a finite number of cylinders $\{C_k = C(z_k, r_k, \eta_k)\}_{k=1}^m$. Let $\{\varphi_k\}_{k=0}^m$ be a partition of unity as in the proof of Theorem 1.5.3 (see (1.29)). From the previous step, we can define $P(\varphi_k u)$ for all $k \in \{1, ..., m\}$ and so finally we set $Pu = \varphi_0 u + \sum_{k=1}^m P(\varphi_k u)$.

Finally, if $u \in W^{1,p}(\Omega)$, then we use the approximation theorem 1.5.3 and the fact that for $C^\infty(\overline{\Omega})$ functions the result is true as we just proved. □

Using this theorem, we have an alternative simpler proof of Theorem 1.5.3.

Corollary 1.6.2 *Assume that Ω is bounded with Lipschitz boundary $\partial\Omega$ and $1 \leqslant p < \infty$. Then* $\overline{C^\infty(\overline{\Omega})}^{\|\cdot\|_{1,p}} = W^{1,p}(\Omega)$.

Proof Let $u \in W^{1,p}(\Omega)$. By Theorem 1.6.1, $Pu \in W^{1,p}(\mathbb{R}^N)$. From Theorem 1.1.21 we know that there exists a sequence $\{u_n\}_{n\geqslant 1} \subseteq C_c^\infty(\mathbb{R}^N)$ such that $u_n \to Pu$ in $W^{1,p}(\mathbb{R}^N)$. Then $u_n|_{\overline{\Omega}} \in C^\infty(\overline{\Omega})$ and $u_n|_\Omega \to Pu|_\Omega = u$ in $W^{1,p}(\Omega)$. □

In fact, there is an extension operator even if Ω is not bounded, provided we further strengthen the regularity of the boundary.

Definition 1.6.3 Let $\Omega \subseteq \mathbb{R}^N$ be an open set. We say that its boundary $\partial\Omega$ is "uniformly Lipschitz", if there exist $\varepsilon, L > 0, m_0 \in \mathbb{N}$, and a locally finite countable open cover $\{U_n\}_{n\geqslant 1}$ of $\partial\Omega$ such that

(a) if $z \in \partial\Omega$, then $B_\varepsilon(z) \subseteq U_n$ for some $n \geqslant 1$;
(b) no point $z \in \mathbb{R}^N$ is contained in more than m_0 of the sets $\{U_n\}_{n\geqslant 1}$;
(c) for each $n \geqslant 1$, there exist local coordinates $z = (z', z_N) \in \mathbb{R}^{N-1} \times \mathbb{R}$ and a Lipschitz function $\xi : \mathbb{R}^{N-1} \to \mathbb{R}$ (both depending on n) with $\operatorname{Lip}\xi \leqslant L$ such that
$$U_n \cap \Omega = U_n \cap \{(z', z_N) \in \mathbb{R}^{N-1} \times \mathbb{R} : \xi(z') < z_N\}.$$

Remark 1.6.4 If $\Omega \subseteq \mathbb{R}^N$ is open with $\partial\Omega$ bounded, then $\partial\Omega$ is uniformly Lipschitz if and only if it is Lipschitz (see Definition 1.5.1). If Ω is unbounded with uniformly Lipschitz boundary, then $\lambda^N(\Omega) = \infty$.

The proof of the next theorem can be found in Leoni [262, p. 356].

Theorem 1.6.5 *Assume that Ω has a uniformly Lipschitz boundary and $1 \leqslant p < \infty$. Then there exists a bounded linear operator $P : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^N)$ such that for all $u \in W^{1,p}(\Omega)$ we have $Pu|_\Omega = u$, $\|Pu\|_{L^p(\mathbb{R}^N)} \leqslant (1 + 2m_0)\|u\|_p$ and $\|Du\|_{L^p(\mathbb{R}^N,\mathbb{R}^N)} \leqslant c(1 + m_0(1 + L))(\frac{1}{\varepsilon}\|u\|_p + \|Du\|_p)$.*

1.7 The Rellich–Kondrachov Theorem

The Rellich–Kondrachov theorem is the main compactness theorem for Sobolev spaces. To prove it, we need to recall the Kolmogorov compactness theorem for $L^p(\mathbb{R}^N)$. In what follows, given $u \in L^p(\mathbb{R}^N)$ and $h \in \mathbb{R}^N$, we denote by $\tau_h(u)$ the translation by h of u, namely the function $\tau_h(u)(z) = u(z - h)$ for all $z \in \mathbb{R}^N$.

Theorem 1.7.1 (Kolmogorov) *A set $C \subset L^p(\mathbb{R}^N)$ ($1 \leqslant p < \infty$) is relatively compact if and only if the following conditions are satisfied:*

(a) *C is bounded in $L^p(\mathbb{R}^N)$;*
(b) $\displaystyle\lim_{\eta\to+\infty}\sup_{u\in C}\int_{\{\|z\|_{\mathbb{R}^N}>\eta\}}|u(z)|dz \to 0$;
(c) $\displaystyle\lim_{h\to 0}\sup_{u\in C}\|\tau_h(u) - u\|_{L^p(\mathbb{R}^N)} = 0$.

Proof "$\Downarrow$": Suppose that C is relatively compact in $L^p(\mathbb{R}^N)$.

Then C is bounded and so (a) is true. Also, C is totally bounded. So, given $\varepsilon > 0$ we can find $\{u_k\}_{k=1}^n \subseteq L^p(\mathbb{R}^N)$ such that for each $u \in L^p(\mathbb{R}^N)$ we can find $k \in \{1, ..., n\}$ such that $\|u - u_k\|_{L^p(\mathbb{R}^N)} \leqslant \varepsilon$. Let $\{s_k\}_{k=1}^n$ be simple functions such that $\|u_k - s_k\|_{L^p(\mathbb{R}^N)} \leqslant \varepsilon$ for all $k \in \{1, ..., n\}$. Simple functions have their values in a ball. So, for $\eta > 0$ large, we have for all $u \in L^p(\mathbb{R}^N)$, all $k \in \{1, ..., N\}$

$$\begin{aligned}&\left(\int_{\{\|z\|_{\mathbb{R}^N} > \eta\}} |u(z)|^p dz\right)^{\frac{1}{p}} \leqslant \left(\int_{\{\|z\|_{\mathbb{R}^N} > \eta\}} |u(z) - s_k(z)|^p dz\right)^{\frac{1}{p}}\\ \Rightarrow &\left(\int_{\{\|z\|_{\mathbb{R}^N} > \eta\}} |u(z)|^p dz\right)^{\frac{1}{p}} \leqslant \|u - s_k\|_{L^p(\mathbb{R}^N)} \leqslant 2\varepsilon \text{ for all } u \in C.\end{aligned}$$

This proves that (b) holds.

Next, note that (c) holds for simple functions. Therefore for any $u \in C$ we have

$$\begin{aligned}&\limsup_{h\to 0}\left(\int_{\mathbb{R}^N} |\tau_h(u)(z) - u(z)|^p dz\right)^{\frac{1}{p}}\\ &\leqslant \limsup_{h\to 0}\left(\int_{\mathbb{R}^N} |\tau_h(u)(z) - \tau_h(u_k)(z)|^p dz\right)^{\frac{1}{p}}\\ &+ \limsup_{h\to 0}\left(\int_{\mathbb{R}^N} |\tau_h(u_k)(z) - \tau_h(s_k)(z)|^p dz\right)^{\frac{1}{p}}\\ &+ \limsup_{h\to 0}\left(\int_{\mathbb{R}^N} |\tau_h(s_k)(z) - s_k(z)|^p dz\right)^{\frac{1}{p}} + \limsup_{h\to 0}\left(\int_{\mathbb{R}^N} |s_k(z) - u_k(z)|^p dz\right)^{\frac{1}{p}}\\ &+ \limsup_{h\to 0}\left(\int_{\mathbb{R}^N} |u_k(z) - u(z)|^p dz\right)^{\frac{1}{p}}\\ &\leqslant \varepsilon + \varepsilon + 0 + \varepsilon + \varepsilon = 4\varepsilon.\end{aligned}$$

This proves statement (c).

"$\Uparrow$": Now we assume that statements (a), (b), (c) hold.

By virtue of (b), we know that given $\varepsilon > 0$, we can find $\eta > 0$ large such that

$$\int_{\{\|z\|_{\mathbb{R}^N} > \eta\}} |u(z)|^p dz \leqslant \varepsilon \text{ for all } u \in C. \tag{1.48}$$

Let $\{\rho_n\}_{n\geqslant 1}$ be a sequence of mollifiers. From Proposition 1.1.3 we have for all $u \in L^p(\mathbb{R}^N)$, all $n \geqslant 1$

$$\begin{aligned}&\|u - u * \rho_n\|^p_{L^p(\mathbb{R}^N)} \leqslant \int_{\mathbb{R}^N} \|u - \tau_h(u)\|^p_{L^p(\mathbb{R}^N)} \rho_n(z) dz\\ &\Rightarrow \|u - u * \rho_n\|_{L^p(\mathbb{R}^N)} \leqslant \sup\left[\|u - \tau_h(u)\|_{L^p(\mathbb{R}^N)} : \|h\|_{\mathbb{R}^N} \leqslant \frac{1}{n}\right].\end{aligned}$$

By virtue of statement (c), we can find an integer $n_0 = n_0(\varepsilon) \geqslant 1$ such that

$$\|u - u * \rho_{n_0}\|_{L^p(\mathbb{R}^N)} \leqslant \varepsilon \text{ for all } u \in C. \tag{1.49}$$

Also for any $u \in L^p(\mathbb{R}^N)$, any $z, z' \in \mathbb{R}^N$ and any integer $n \geqslant 1$, we have

$$\begin{aligned}|(u * \rho_n)(z) - (u * \rho_n)(z')| &\leqslant \int_{\mathbb{R}^N} |u(z-y) - u(z'-y)|\rho_n(y)dy \\ &\leqslant \|\tau_{z-z'}(u) - u\|_{L^p(\mathbb{R}^N)} \|\rho_n\|_{L^{p'}(\mathbb{R}^N)}\end{aligned} \tag{1.50}$$

(due to the translation invariance of the Lebesgue measure).

In addition, we have

$$|(u * \rho_n)(z)| \leqslant \|u\|_{L^p(\mathbb{R}^N)} \|\rho_n\|_{L^{p'}(\mathbb{R}^N)} \text{ for all } z \in \Omega, \text{ all } n \geqslant 1. \tag{1.51}$$

We consider the family $D = \{u * \rho_{n_0} : \bar{B}_\eta(0) \to \mathbb{R} : u \in C\}$. Conditions (a) and (c) and (1.50), (1.51) imply that we can apply the Arzela–Ascoli theorem and infer that D is relatively compact in $C(\bar{B}_\eta(0))$. So, we can find $\{u_k\}_{k=1}^m \subseteq C$ such that

$$D \subseteq \bigcup_{k=1}^m B_{\varepsilon\eta^{-\frac{N}{p}}}(u_k * \rho_{n_0}).$$

So, given $u \in C$, we can find a $k \in \{1, ..., m\}$ such that

$$|(u * \rho_{n_0})(z) - (u_k * \rho_{n_0})(z)| \leqslant \varepsilon(\lambda^N(\bar{B}_\eta(0)))^{-\frac{1}{p}} \text{ for all } z \in \bar{B}_\eta(0). \tag{1.52}$$

We have

$$\begin{aligned}\|u - u_k\|_{L^p(\mathbb{R}^N)} \leqslant &\left(\int_{\{\|z\|_{\mathbb{R}^N} > \eta\}} |u|^p dz\right)^{\frac{1}{p}} + \left(\int_{\{\|z\|_{\mathbb{R}^N} > \eta\}} |u_k|^p dz\right)^{\frac{1}{p}} \\ &+ \|u - u * \rho_{n_0}\|_{L^p(\mathbb{R}^N)} + \|u_k - u_k * \rho_{n_0}\|_{L^p(\mathbb{R}^N)} \\ &+ \|u * \rho_{n_0} - u_k * \rho_{n_0}\|_{L^p(\bar{B}_\eta(0))}.\end{aligned} \tag{1.53}$$

Note that

$$\begin{aligned}&\|u * \rho_{n_0} - u_k * \rho_{n_0}\|_{L^p(\bar{B}_\eta(0))} \leqslant \\ &\left[\int_{\bar{B}_\eta(0)} |(u * \rho_{n_0})(z) - (u_k * \rho_{n_0})(z)|^p dz\right]^{\frac{1}{p}} \\ &\leqslant \varepsilon \text{ (see 1.52).}\end{aligned} \tag{1.54}$$

So, returning to (1.53) and using (1.48), (1.49) and (1.54), we obtain

$$\|u - u_k\|_{L^p(\mathbb{R}^N)} \leqslant 5\varepsilon$$

$\Rightarrow$ C is relatively compact in $L^p(\mathbb{R}^N)$ (being totally bounded). ☐

Remark 1.7.2 Condition (c) is an integral "equicontinuity" condition analogous to the equicontinuity assumption involved in the Arzela–Ascoli theorem.

The next estimate makes precise statement (c) in the above theorem and it will be useful in the proof of the Rellich–Kondrachov theorem for $W_0^{1,p}(\Omega)$ (see Theorem 1.7.4).

Proposition 1.7.3 *If $u \in W^{1,p}(\mathbb{R}^N)$ $(1 \leqslant p < \infty)$, then*

$$\|\tau_h(u) - u\|_{L^p(\mathbb{R}^N)} \leqslant \|Du\|_{L^p(\mathbb{R}^N,\mathbb{R}^N)} \|h\|_{\mathbb{R}^N} \text{ for all } h \in \mathbb{R}^N.$$

Proof Since $C_c^\infty(\mathbb{R}^N)$ is dense in $W^{1,p}(\mathbb{R}^N)$ (see Theorem 1.1.21), it suffices to prove the estimate for $u \in C_c^\infty(\mathbb{R}^N)$. We have

$$\tau_h(u)(z) - u(z) = u(z-h) - u(z) = -\int_0^1 (Du(z-th), h)_{\mathbb{R}^N} dt$$

$$\Rightarrow |\tau_h(u)(z) - u(z)|^p \leqslant \|h\|_{\mathbb{R}^N}^p \int_0^1 \|Du(z-th)\|_{\mathbb{R}^N}^p dt \text{ (by Jensen's inequality)}$$

$$\Rightarrow \|\tau_h(u) - u\|_{L^p(\mathbb{R}^N)}^p \leqslant \|Du\|_{L^p(\mathbb{R}^N,\mathbb{R}^N)}^p \|h\|_{\mathbb{R}^N}^p,$$

using Fubini's theorem and the translation invariance of the Lebesgue measure in $\mathbb{R}^N$. □

Now we can state the Rellich–Kondrachov theorem for the Sobolev space $W_0^{1,p}(\Omega)$. In that space the only requirement on Ω is that it is bounded. No regularity condition is assumed on the boundary $\partial\Omega$. The reason for this is that in this case the extension by zero on $\mathbb{R}^N \setminus \Omega$ works (see Proposition 1.1.17).

Theorem 1.7.4 (Rellich–Kondrachov) *Assume that $\Omega \subset \mathbb{R}^N$ is bounded. Then $W_0^{1,p}(\Omega)$ is embedded in $L^p(\Omega)$ compactly, that is, every bounded subset of $W_0^{1,p}(\Omega)$ is relatively compact in $L^p(\Omega)$.*

Proof Let $P_0 : W_0^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^N)$ be the extension by zero on $\mathbb{R}^N \setminus \Omega$ operator. From Proposition 1.1.17, P_0 is linear isometry. Also, the restriction operator $R_0 : L^p(\mathbb{R}^N) \to L^p(\Omega)$ defined by $R_0(u) = u|_\Omega$ is clearly linear continuous, with $\|R_0\|_{\mathscr{L}} \leqslant 1$. Finally, let $i_0 : W_0^{1,p}(\Omega) \to L^p(\Omega)$ and $i : W^{1,p}(\mathbb{R}^N) \to L^p(\mathbb{R}^N)$ be the embedding operators. We have

$$i_0 = R_0 \circ i \circ P_0.$$

Let B_1 be the open unit ball in the Sobolev space $W_0^{1,p}(\Omega)$. Evidently, we need to show that $i(P_0(B_1))$ is relatively compact. To this end, we use Theorem 1.7.1.

First note that the continuity of both i and P_0 implies that

$$(i \circ P_0)(B_1) \subseteq L^p(\mathbb{R}^N) \text{ is bounded.} \tag{1.54'}$$

Since by hypothesis Ω is bounded, we can find $\eta > 0$ big such that $\Omega \subseteq B_\eta(0)$. Then for all $u \in B_1$ we have $P_0(u) = 0$ on $\mathbb{R}^N \setminus B_\eta(0)$ and so

$$\int_{\{\|z\|_{\mathbb{R}^N} > \eta\}} |P_0(u)|^p dz = 0 \text{ for all } u \in B_1. \tag{1.55}$$

Finally, note that $R_0(\bar{B}_1)$ is contained in the unit ball of $W^{1,p}(\mathbb{R}^N)$. So, from Proposition 1.7.3 we see that we can find $c > 0$ such that

$$\|\tau_h(u) - u\|_{L^p(\mathbb{R}^N)} \leqslant c\|h\|_{\mathbb{R}^N} \text{ for all } u \in B_1. \tag{1.56}$$

Relations (1.54′), (1.55), (1.56) allow us to use Theorem 1.7.1 and so we conclude that $(i \circ P_0)(B_1) = i(P_0(B_1)) \subseteq L^p(\mathbb{R}^N)$ is relatively compact. This proves the compact embedding of $W_0^{1,p}(\Omega)$ into $L^p(\Omega)$. ☐

Another useful consequence of Theorem 1.7.1 is the following result.

Proposition 1.7.5 *Assume that $C \subseteq L^p(\mathbb{R}^N)$ $(1 \leqslant p < \infty)$ and*

(i) *C is $L^p(\mathbb{R}^N)$-bounded;*

(ii) $\displaystyle\lim_{\eta\to+\infty} \sup_{u\in C} \int_{\{\|z\|_{\mathbb{R}^N} > \eta\}} |u(z)|^p dz = 0.$

Then C is relatively compact in $L^p(\mathbb{R}^N)$.

Proof Hypotheses (i), (ii) and Proposition 1.7.3 make possible the use of Theorem 1.7.1 and so we conclude that C is relatively compact in $L^p(\mathbb{R}^N)$. ☐

We know that for the space $W^{1,p}(\Omega)$ the extension by zero on $\mathbb{R}^N \setminus \Omega$ does not work. So, we need to employ the extension operator (see Theorem 1.6.1). This requires that we impose some regularity assumption on the boundary of Ω.

Theorem 1.7.6 (Rellich–Kondrachov) *Assume that Ω is bounded with Lipschitz boundary and $1 \leqslant p < \infty$. Then $W^{1,p}(\Omega)$ is embedded compactly in $L^p(\Omega)$.*

Proof Let $\{u_n\}_{n\geqslant 1} \subseteq W^{1,p}(\Omega)$ be bounded. We need to show that it admits a subsequence which converges strongly in $L^p(\Omega)$. Choose $\eta > 0$ big such that $\Omega \subseteq B_\eta(0)$ and let $\varphi \in C_c^\infty(\mathbb{R}^N)$ be such that $\varphi|_{\bar{B}_\eta(0)} = 1$ and $\varphi|_{\mathbb{R}^N \setminus B_{2\eta}(0)} \equiv 0$ (cut-off function). Also, let $P : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^N)$ be the extension operator produced in Theorem 1.6.1. We consider the sequence $\{\varphi P(u_n)\}_{n\geqslant 1} \subseteq W^{1,p}(\mathbb{R}^N)$. This sequence is bounded and zero on $\mathbb{R}^N \setminus B_{2\eta}(0)$. So, we can apply Proposition 1.7.5 and infer that $\{\varphi P(u_n)\}_{n\geqslant 1} \subseteq L^p(\mathbb{R}^N)$ is relatively compact. By passing to a subsequence if necessary, we may assume that

$$\varphi P(u_n) \to u \text{ in } L^p(\mathbb{R}^N).$$

But note that $\varphi P(u_n)|_\Omega = u_n$. Therefore $u_n \to u$ in $L^p(\Omega)$ and this proves that $W^{1,p}(\Omega)$ is embedded compactly in $L^p(\Omega)$. ☐

Remark 1.7.7 Both Rellich–Kondrachov theorems (see Theorems 1.7.4 and 1.7.6) remain true if instead of Ω being bounded we assume that Ω has finite Lebesgue measure.

Corollary 1.7.8 *Assume that Ω is bounded with Lipschitz boundary $\partial\Omega$ and $1 \leqslant p < \infty$. Then $u_n \xrightarrow{w} u$ in $W^{1,p}(\Omega)$ if and only if $u_n \to u$ in $L^p(\Omega)$, $Du_n \xrightarrow{w} Du$ in $L^p(\Omega, \mathbb{R}^N)$.*

1.8 The Poincaré and Poincaré–Wirtinger Inequalities

The Poincaré inequality is a basic tool in the study of Dirichlet boundary value problems. We start with a formulation which can be proved easily using Theorem 1.7.6 (the Rellich–Kondrachov theorem for $W^{1,p}(\Omega)$).

Theorem 1.8.1 *Assume that Ω is bounded connected (that is, a bounded domain) with Lipschitz boundary $\partial\Omega$ and $V \subseteq W^{1,p}(\Omega)$ $(1 \leqslant p < \infty)$ is a closed linear subspace such that the only constant function belonging to V is the identically zero function. Then there exists a constant $c > 0$ such that*

$$||u||_p \leqslant c||Du||_{L^p(\Omega,\mathbb{R}^N)} \text{ for all } u \in V.$$

Proof We argue by contradiction. So, suppose the theorem is not true. Then we can find a sequence $\{u_n\}_{n\geqslant 1} \subseteq V$, $u_n \neq 0$ such that

$$||u_n||_p > n||Du_n||_{L^p(\Omega,\mathbb{R}^N)} \text{ for all } n \geqslant 1. \tag{1.57}$$

Let $y_n = \frac{u_n}{||u_n||_p}$, $n \geqslant 1$. Then $\{y_n\}_{n\geqslant 1} \subseteq V$ and $||y_n||_p = 1$ for all $n \geqslant 1$. From (1.57) we have

$$||Dy_n||_{L^p(\Omega,\mathbb{R}^N)} < \frac{1}{n} \text{ for all } n \geqslant 1. \tag{1.58}$$

So, $\{y_n\}_{n\geqslant 1} \subseteq V \subseteq W^{1,p}(\Omega)$ is bounded and by Theorem 1.7.6 we may assume that

$$y_n \xrightarrow{w} y \text{ in } W^{1,p}(\Omega),\ \ y_n \to y \text{ in } L^p(\Omega),\ \ y \in V. \tag{1.59}$$

From (1.58) we have that $Dy_n \to 0$ in $L^p(\Omega, \mathbb{R}^N)$, hence $Dy = 0$ (see (1.59)). Since Ω is connected, y is constant in Ω. Because $y \in V$, we must have $y = 0$ and so

$$y_n \to 0 \text{ in } L^p(\Omega),$$

which contradicts the fact that $||y_n||_p = 1$ for all $n \geqslant 1$. □

Remark 1.8.2 A possible choice is $V = W^{1,p}(\Omega)$. Also, let Ω be bounded, connected with Lipschitz boundary $\partial\Omega$ and let $\hat{\Gamma}$ be a subset of $\partial\Omega$ such that $\mathscr{H}^{N-1}(\hat{\Gamma}) > 0$. Let

$$V = \{u \in W^{1,p}(\Omega) :\ \gamma_0(u) = 0 \text{ on } \hat{\Gamma}\}.$$

Then the only constant function belonging to V is the identically zero function. Still a third possibility for the subspace V is

$$V = \left\{u \in W^{1,p}(\Omega) :\ \int_\Omega u\,dz = 0\right\}.$$

This choice leads to the so-called "Poincaré–Wirtinger inequality", which is a useful tool in the study of Neumann and periodic boundary value problems.

Theorem 1.8.3 *Assume that Ω is bounded connected with Lipschitz boundary and $1 \leqslant p < \infty$. Then there exists a constant $c > 0$ such that*

$$||u - \frac{1}{\lambda^N(\Omega)} \int_\Omega u(z)dz||_p \leqslant c||Du||_{L^p(\Omega,\mathbb{R}^N)} \text{ for all } u \in W^{1,p}(\Omega).$$

Proof Let $V = \{u \in W^{1,p}(\Omega) : \int_\Omega u(z)dz = 0\}$. Evidently V is a closed linear subspace of $W^{1,p}(\Omega)$ and $u \equiv 0$ is the only constant function belonging to V. Now note that for any $u \in W^{1,p}(\Omega)$, $u(\cdot) - \frac{1}{\lambda^N(\Omega)} \int_\Omega u\, dz \in V$ and so Theorem 1.8.1 implies that there exists a $c > 0$ such that

$$||u - \frac{1}{\lambda^N(\Omega)} \int_\Omega u\, dz||_p \leqslant c||Du||_{L^p(\Omega,\, \mathbb{R}^N)} \text{ for all } u \in W^{1,p}(\Omega).$$

□

In fact, Theorem 1.8.1 can be stated for more general open sets $\Omega \subseteq \mathbb{R}^N$. More precisely, we have the following general version of the Poincaré inequality.

Theorem 1.8.4 *Assume that Ω is bounded in one direction (that is, lies between two parallel hyperplanes) and $1 \leqslant p < \infty$. Then $||u||_p^p \leqslant \frac{d^p}{p}||Du||^p_{L^p(\Omega,\mathbb{R}^N)}$ for all $u \in W^{1,p}(\Omega)$, where d is the distance between the two hyperplanes.*

Proof By translating and rotating the coordinates if necessary, without any loss of generality, we may assume that Ω lies between the two parallel hyperplanes $z_N = 0$ and $z_N = d > 0$. Let $u \in C_c^\infty(\Omega)$. We have

$$\begin{aligned}|u(z', z_N)| = |u(z', z_N) - u(z', 0)| &= \left|\int_0^{z_N} \frac{\partial u}{\partial z_N}(z', s)ds\right| \\ &\leqslant z_N^{1/p'} \left(\int_0^d \left|\frac{\partial u}{\partial z_N}(z', s)\right|^p ds\right)^{1/p} \\ &\quad \text{(by Hölder's inequality)}.\end{aligned}$$

Using Fubini's theorem, we obtain

$$\begin{aligned}\int_{\mathbb{R}^{N-1}x[0,d]} |u(z', z_N)|^p dz &\leqslant \int_{\mathbb{R}^{N-1}} \int_0^d z_{z_N}^{\frac{p}{p'}} \int_0^d \left|\frac{\partial u}{\partial z_N}(z', s)\right|^p ds\, dz_N\, dz' \\ &= \left(\int_\Omega \left|\frac{\partial u}{\partial z_N}(z)\right|^p dz\right) \left(\int_0^d z_N^{p-1} dz_N\right) \\ &= \frac{d^p}{p} \int_\Omega \left|\frac{\partial u}{\partial z_N}(z)\right|^p dz.\end{aligned} \tag{1.60}$$

Recall that $\frac{\partial u}{\partial z_N} = (Du, n)_{\mathbb{R}^N}$, where n is the outward unit normal vector to the strip containing Ω. Therefore

$$\int_\Omega \left| \frac{\partial u}{\partial z_N}(z) \right|^p dz = \int_\Omega |(Du(z), n(z))_{\mathbb{R}^N}|^p dz$$
$$\leqslant \int_\Omega ||Du(z)||_{\mathbb{R}^N}^p \, dz = ||Du||_{L^p(\Omega,\mathbb{R}^N)}^p.$$

Returning to (1.60), we conclude that

$$||u||_p \leqslant \frac{d^p}{p} ||Du||_{L^p(\Omega,\mathbb{R}^N)} \text{ for all } u \in W_0^{1,p}(\Omega).$$

The proof is now complete. □

Remark 1.8.5 Therefore, if Ω has finite width, then $||Du||_{L^p(\Omega,\mathbb{R}^N)}$ is an equivalent norm on $W_0^{1,p}(\Omega)$.

1.9 The Sobolev Embedding Theorem

In Theorem 1.2.2 we saw that every $u \in W^{1,p}(a,b)$ $(1 \leqslant p \leqslant \infty)$ admits a continuous representative. This is no longer true in higher dimensions (that is, if $N \geqslant 2$). In this section we examine when we can claim that $W^{m,p}(\Omega) \hookrightarrow C(\bar{\Omega})$ and more generally we establish embeddings of various Sobolev spaces into others.

To simplify the presentation, we will conduct the analysis for the space $W^{1,p}(\Omega)$ and at the end we will formulate the general embedding theorem for the space $W^{m,p}(\Omega)$.

So, the fundamental question of this section is the following: "Given $u \in W^{1,p}(\Omega)$ can we conclude that u belongs to certain other spaces (besides $L^p(\Omega)$ of course)?". It turns out that the answer to this question depends on the relation between p and N. In fact, we consider three distinct cases

$$1 \leqslant p < N, \;\; p = N \text{ and } N < p \leqslant \infty,$$

which lead to different embeddings of $W^{1,p}(\Omega)$.

We want to find out for which exponents q we can assert that

$$||u||_{L^q(\mathbb{R}^N)} \leqslant c||Du||_{L^p(\mathbb{R}^N,\mathbb{R}^N)} \text{ for some } c > 0, \text{ all } u \in W^{1,p}(\mathbb{R}^N).$$

A simple homogeneity argument reveals that q must have a specific value depending on N and p. So, let $u \in W^{1,p}(\mathbb{R}^N)$, $u \neq 0$ and define

$$u_\lambda(z) = u(\lambda z) \text{ with } \lambda > 0.$$

We have
$$\left(\int_{\mathbb{R}^N} |u(\lambda z)|^q dz\right)^{1/q} \leqslant c \left(\int_{\mathbb{R}^N} \lambda^p ||Du(\lambda z)||^p_{\mathbb{R}^N} dz\right)^{1/p}.$$

We perform a change of variable $y = \lambda z$. Then
$$\frac{1}{\lambda^{N/q}} ||u||_{L^q(\mathbb{R}^N)} \leqslant c \, \frac{\lambda}{\lambda^{\frac{N}{p}}} ||Du||_{L^p(\mathbb{R}^N,\mathbb{R}^N)}$$
$$\Rightarrow \; ||u||_{L^q(\mathbb{R}^N)} \leqslant c \, \lambda^{\left(1-\frac{N}{p}+\frac{N}{q}\right)} ||Du||_{L^p(\mathbb{R}^N,\mathbb{R}^N)}. \tag{1.61}$$

Evidently, (1.61) makes sense only when
$$1 - \frac{N}{p} + \frac{N}{q} = 0 \;\Rightarrow\; q = p^* = \frac{Np}{N-p}.$$

Indeed, if $1 - \frac{N}{p} + \frac{N}{q} > 0$, then letting $\lambda \to 0$ in (1.61), we reach a contradiction. Similarly, if $1 - \frac{N}{p} + \frac{N}{q} < 0$, then letting $\lambda \to +\infty$ again we have a contradiction (since $u = 0$). In order for $q = p^* = \frac{Np}{N-p}$ to be positive, we need $p < N$.

Definition 1.9.1 For $p < N$, the number $p^* = \frac{Np}{N-p}$ is called the "Sobolev critical exponent".

So, we first deal with the case $1 \leqslant p < N$.

Note that, if $z = (z_k)_{k=1}^N \in \mathbb{R}^N$, then by z'_k we denote the $\mathbb{R}^{N-1}$-vector obtained by removing the kth component from z. By abuse of notation we write
$$z = (z'_k, z_k) \in \mathbb{R}^{N-1} \times \mathbb{R}.$$

If $k = N$, then we recover the notation used in previous sections
$$z = (z', z_N) \in \mathbb{R}^{N-1} \times \mathbb{R}.$$

Lemma 1.9.2 *Assume that* $N \geqslant 2$, $u_k \in L^{N-1}(\mathbb{R}^{N-1})$ *with* $k \in \{1, \ldots, N-1\}$ *and*
$$u(z) = u_1(z'_1)u_2(z'_2)\ldots u_N(z'_N) \text{ for all } z \in \mathbb{R}^N.$$

Then $u \in L^1(\mathbb{R}^N)$ *and* $||u||_{L^1(\mathbb{R}^N)} \leqslant \prod_{k=1}^{N} ||u_k||_{L^{N-1}(\mathbb{R}^{N-1})}$.

Proof The proof is by induction on N. First let $N = 2$. Then $u(z) = u_1(z_2)u_2(z_1)$ for all $z = (z_1, z_2) \in \mathbb{R}^2$. Using Fubini's theorem, we have
$$\int_{\mathbb{R}^2} |u(z)| dz = \int_{\mathbb{R}} |u_1(z_2)| dz_2 \int_{\mathbb{R}} |u_2(z_1)| dz_1,$$

and so we have verified the result.

Next we assume that the result is true for N (induction hypothesis) and we will prove that it also holds for $N+1$. So, for $u_k \in L^N(\mathbb{R}^N)$, $k \in \{1, \dots, N+1\}$, let

$$u(z) = u_1(z_1')u_2(z_2') \dots u_{N+1}(z_{N+1}') \text{ for all } z \in \mathbb{R}^{N+1}.$$

We fix $z_{N+1} \in \mathbb{R}$ and integrate with respect to $z_1, \dots, z_N$. Then via Hölder's inequality, we have

$$\int_{\mathbb{R}^N} |u(z)|\, dz_1 \dots dz_N \leqslant ||u_{N+1}||_{L^N(\mathbb{R}^N)} \times \left(\int_{\mathbb{R}^N} \prod_{k=1}^{N} |u_k(z_k')|^{\frac{N}{N-1}} dz_1 \dots dz_N \right)^{\frac{N-1}{N}}. \tag{1.62}$$

Expanding the notation introduced earlier, by z_k'' we denote the $\mathbb{R}^{N-1}$-vector obtained by removing the last component from z_k' and again by abuse of notation, we write $z_k' = (z_k'', z_{N+1}) \in \mathbb{R}^{N-1} \times \mathbb{R}$. Since z_{N+1} is fixed, by the induction hypothesis applied to the functions

$$\hat{u}_k(z_k'') = |u_k(z_k'', z_{N+1})|^{\frac{N}{N-1}} \text{ for all } z_k'' \in \mathbb{R}^{N-1}$$

with $k \in \{1, \dots, N\}$, we obtain

$$\begin{aligned} \int_{\mathbb{R}^N} \prod_{k=1}^{N} |u_k(z_k')|^{\frac{N}{N-1}} dz_1 \dots dz_N &\leqslant \prod_{k=1}^{N} ||\hat{u}_k||_{L^{N-1}(\mathbb{R}^{N-1})} \\ &= \prod_{k=1}^{N} \left(\int_{\mathbb{R}^{N-1}} |u_k(z_k'', z_{N+1})|^N dz_k'' \right)^{\frac{1}{N-1}}. \end{aligned}$$

Using this inequality, we obtain

$$\int_{\mathbb{R}^N} |u(z)| dz_1 \dots dz_N \leqslant ||u_{N+1}||_{L^N(\mathbb{R}^N)} \prod_{k=1}^{N} \left(\int_{\mathbb{R}^{N-1}} |u_k(z_k'', z_{N+1})|^N dz_k'' \right)^{\frac{1}{N}}.$$

Integrating both sides with respect to z_{N+1} and using Hölder's inequality (generalized version, see for example Denkowski et al. [143, p. 150]), we have

$$\int_{\mathbb{R}^{N+1}} |u(z)| dz = ||u||_{L^1(\mathbb{R}^N)} \prod_{k=1}^{N+1} ||u_k||_{L^N(\mathbb{R}^N)}.$$

This completes the induction and the proof of the lemma. $\square$

Now we are ready to prove the embedding theorem for the case $1 \leqslant p < N$. The result is known in the literature as the "Sobolev–Gagliardo–Nirenberg embedding theorem" or the "Sobolev–Gagliardo–Nirenberg inequality".

Theorem 1.9.3 *(Sobolev–Gagliardo–Nirenberg) Assume that $1 \leqslant p < N$ and $p^* = \frac{Np}{N-p}$ (see Definition 1.9.1). Then $W^{1,p}(\mathbb{R}^N)$ is continuously embedded into $L^{p^*}(\mathbb{R}^N)$; more precisely we can find a constant $c = c(p, N) > 0$ such that*

$$||u||_{L^{p^*}(\mathbb{R}^N)} \leqslant c||Du||_{L^p(\mathbb{R}^N,\mathbb{R}^N)} \text{ for all } u \in W^{1,p}(\mathbb{R}^N). \tag{1.63}$$

Proof We start with some simplifications.

First we show that it suffices to prove (1.63) for all $u \in C_c^\infty(\mathbb{R}^N)$. So, suppose for the moment that (1.63) holds for all $u \in C_c^\infty(\mathbb{R}^N)$. According to Theorem 1.1.21, given $u \in W^{1,p}(\mathbb{R}^N)$ we can find $\{u_n\}_{n\geqslant 1} \subseteq C_c^\infty(\mathbb{R}^N)$ such that

$$u_n \to u \text{ in } W^{1,p}(\mathbb{R}^N) \text{ and } u_n(z) \to u(z) \text{ a.e. in } \mathbb{R}^N. \tag{1.64}$$

By hypothesis we have

$$||u_n||_{L^{p^*}(\mathbb{R}^N)} \leqslant c||Du_n||_{L^p(\mathbb{R}^N,\mathbb{R}^N)} \text{ for all } n \geqslant 1.$$

Passing to the limit as $n \to \infty$ and using (1.64), we obtain

$$\begin{aligned} ||u||_{L^{p^*}(\mathbb{R}^N)} &\leqslant \liminf_{n\to\infty} ||u_n||_{L^{p^*}(\mathbb{R}^N)} \text{ (by Fatou's lemma)} \\ &\leqslant c \lim_{n\to\infty} ||Du_n||_{L^p(\mathbb{R}^N,\mathbb{R}^N)} = c\,||Du||_{L^p(\mathbb{R}^N,\mathbb{R}^N)}. \end{aligned}$$

So, we have established (1.63) also for all $u \in W^{1,p}(\mathbb{R}^N)$.

Now, the second simplification. We claim that it suffices to show (1.63) for $p = 1$, that is, it is enough to show that there exists a $c_1(N) > 0$ such that

$$||u||_{L^{1^*}(\mathbb{R}^N)} \leqslant c_1(N)||Du||_{L^1(\mathbb{R}^N,\mathbb{R}^N)} \text{ for all } u \in W^{1,1}(\mathbb{R}^N) \tag{1.65}$$

and from this we can have (1.63) for all $1 \leqslant p < N$.

So, suppose that (1.65) holds and let $u \in C_c^\infty(\mathbb{R}^N)$. Then $|u|^{\frac{p^*}{1^*}} \in W^{1,1}(\mathbb{R}^N)$ ($p > 1$). Indeed, note that $1^* < p^*$ (see Definition 1.9.1) and the function $|u(\cdot)|^{\frac{p^*}{1^*}}$ is continuously differentiable with compact support. Hence $|u|^{\frac{p^*}{1^*}} \in W^{1,1}(\mathbb{R}^N)$ and we have

$$D\left(|u|^{\frac{p^*}{1^*}}\right) = \frac{p^*}{1^*}\,|u|^{\frac{p^*}{1^*}-2}\,u\,Du. \tag{1.66}$$

In (1.65) we use $|u|^{\frac{p^*}{1^*}} \in W^{1,1}(\mathbb{R}^N)$. Then using (1.66), we have

$$
\begin{aligned}
\left(\int_{\mathbb{R}^N} |u|^{p^*} dz\right)^{1/1^*} &\leqslant c_1(N)\frac{p^*}{1^*}\int_{\mathbb{R}^N} |u|^{\frac{p^*}{1^*}-1}||Du||_{\mathbb{R}^N} dz \\
&\leqslant c_1(N)\frac{p^*}{1^*}\left(\int_{\mathbb{R}^N} |u|^{\left(\frac{p^*}{1^*}-1\right)p'} dz\right)^{1/p'} ||Du||_{L^p(\mathbb{R}^N,\mathbb{R}^N)} \\
&\qquad\qquad \text{(using Hölder's inequality).}
\end{aligned}
\tag{1.67}
$$

Note that $\frac{1}{1^*} - \frac{1}{p^*} = \frac{1}{p'}$, hence $p^* = \left(\frac{p^*}{1^*} - 1\right) p'$. So, (1.67) becomes

$$
||u||_{L^{p^*}(\mathbb{R}^N)} \leqslant c_1(N)\frac{p^*}{1^*}||Du||_{L^p(\mathbb{R}^N),\mathbb{R}^N},
$$

which is (1.63) for any $1 \leqslant p < N$ with $c(p, N) = c_1(N)\frac{p^*}{1^*}$.

These two simplifications reduce the problem to showing (1.65) for all $u \in C_c^\infty(\mathbb{R}^N)$.

For any $z = (z_k)_{k=1}^N \in \mathbb{R}^N$, we have

$$
|u(z)| = \left|\int_{-\infty}^{z_k} \frac{\partial u}{\partial z_k}(z_k', s)ds\right| \leqslant \int_{-\infty}^{z_k} \left|\frac{\partial u}{\partial z_k}(z_k', s)\right| ds. \tag{1.68}
$$

Similarly, we have

$$
|u(z)| \leqslant \int_{z_k}^{+\infty} \left|\frac{\partial u}{\partial z_k}(z_k', s)\right| ds. \tag{1.69}
$$

Adding (1.68) and (1.69), we obtain

$$
\begin{aligned}
&|u(z)| \leqslant \frac{1}{2}\int_{-\infty}^{\infty} \left|\frac{\partial u}{\partial z_k}(z_k', s)\right| ds \;=\; g_k(z_k') \text{ for all } k \in \{1, \dots, N\} \\
\Rightarrow \quad &|u(z)|^N \leqslant \frac{1}{2^N}\prod_{\mathrm{k}=1}^{N} g_k(z_k').
\end{aligned}
\tag{1.70}
$$

Since $1^* = \frac{N}{N-1}$, from (1.70) we have

$$
|u(z)|^{1^*} \leqslant \frac{1}{2^{\frac{N}{N-1}}}\prod_{\mathrm{k}=1}^{N} g_k(z_k')^{\frac{1}{N-1}}. \tag{1.71}
$$

Note that if $h_k(z_k') = g_k(z_k')^{\frac{1}{N-1}}$ for all $z_k' \in \mathbb{R}^{N-1}$, then $h_k \in L^{N-1}(\mathbb{R}^{N-1})$ and

$$
||h_k||_{L^{N-1}(\mathbb{R}^{N-1})} = \left\|\frac{\partial u}{\partial z_k}\right\|_{L^1(\mathbb{R}^N)}^{\frac{1}{N-1}}. \tag{1.72}
$$

We return to (1.71) and use Lemma 1.9.2 and (1.72). Then

$$\begin{aligned} ||u||_{L^{1^*}(\mathbb{R}^N)}^{1^*} &\leqslant \frac{1}{2^{\frac{N}{N-1}}} \left\| \prod_{k=1}^{N} h_k \right\|_{L^1(\mathbb{R}^N)} \\ &\leqslant \frac{1}{2^{\frac{N}{N-1}}} \prod_{k=1}^{N} ||h_k||_{L^{N-1}(\mathbb{R}^{N-1})} \\ &\leqslant \frac{1}{2^{\frac{N}{N-1}}} \prod_{k=1}^{N} \left\| \frac{\partial u}{\partial z_k} \right\|_{L^1(\mathbb{R}^N)}^{\frac{1}{N-1}} \\ \Rightarrow ||u||_{L^{1^*}(\mathbb{R}^N)} &\leqslant \frac{1}{2} \prod_{k=1}^{N} \left\| \frac{\partial u}{\partial z_k} \right\|_{L^1(\mathbb{R}^N)}^{\frac{1}{N}}, \text{ since } 1^* = \frac{N}{N-1}. \end{aligned} \tag{1.73}$$

By the geometric-arithmetic mean inequality, we have

$$\left(\prod_{k=1}^{N} \left\| \frac{\partial u}{\partial z_k} \right\|_{L^1(\mathbb{R}^N)} \right)^{1/N} \leqslant \frac{1}{N} \sum_{k=1}^{N} \left\| \frac{\partial u}{\partial z_k} \right\|_{L^1(\mathbb{R}^N)}.$$

So, relation (1.73) becomes

$$||u||_{L^{1^*}(\mathbb{R}^N)} \leqslant \frac{1}{2N} \sum_{k=1}^{N} \left\| \frac{\partial u}{\partial z_k} \right\|_{L^1(\mathbb{R}^N)} = \frac{1}{2N} ||Du||_{L^1(\mathbb{R}^N, \mathbb{R}^N)}$$

and so we have proved (1.65) for $u \in C_c^\infty(\mathbb{R}^N)$ and according to the previous discussion we have the theorem. ☐

Remark 1.9.4 The above proof suggests that $c = c(p, N) = \frac{1}{2N} \frac{p^*}{1^*} = \frac{p(N-1)}{2N(N-p)}$. However, this is not the best constant. The best constant is strictly less than the above quantity and is given by

$$c(p, N) = \frac{1}{\sqrt{\pi}} \frac{1}{N^{\frac{1}{p}}} \left(\frac{p-1}{N-p} \right) \left[\frac{\Gamma\left(1 + \frac{N}{2}\right) \Gamma(N)}{\Gamma\left(\frac{N}{p}\right) \Gamma\left(1 + N - \frac{N}{p}\right)} \right]^{\frac{1}{N}}$$

(see Aubin [23] and Talenti [399]). Moreover, equality is realized by the functions

$$u(z) = \left(a + b||z||_{\mathbb{R}^N}^{p'}\right)^{1-\frac{N}{p}} \text{ for all } z \in \mathbb{R}^N, \text{ with } a, b > 0.$$

Corollary 1.9.5 *Assume that* $1 \leqslant p < N$ *and* $p \leqslant q \leqslant p^*$. *Then* $W^{1,p}(\mathbb{R}^N)$ *is embedded continuously in* $L^q(\mathbb{R}^N)$.

Proof Evidently, the corollary is true for $q = p$ (see Definition 1.1.10) and for $q = p^*$ (see Theorem 1.9.3). So, we assume that $p < q < p^*$. Then we can find $t \in (0, 1)$ such that

$$\frac{1}{q} = \frac{t}{p} + \frac{1-t}{p^*}.$$

Then the interpolation inequality implies

$$\begin{aligned}
||u||_{L^q(\mathbb{R}^N)} &\leqslant ||u||^{\vartheta}_{L^p(\mathbb{R}^N)}\ ||u||^{1-\vartheta}_{L^{p^*}(\mathbb{R}^N)} \\
&\leqslant ||u||_{L^p(\mathbb{R}^N)} + ||u||_{L^{p^*}(\mathbb{R}^N)} \text{ (by Young's inequality)} \\
\Rightarrow ||u||_{L^q(\mathbb{R}^N)} &\leqslant ||u||_{L^p(\mathbb{R}^N)} + c\ ||Du||_{L^p(\mathbb{R}^N,\mathbb{R}^N)} \text{ (see Theorem 1.9.3)} \\
&\leqslant (1+c)\ ||u||_{W^{1,p}(\mathbb{R}^N)},
\end{aligned}$$

which shows that the Sobolev space $W^{1,p}(\mathbb{R}^N)$ is continuously embedded in $L^q(\mathbb{R}^N)$ for all $q \in [p, p^*]$. ☐

Theorem 1.9.3 leads to the following Poincaré-type inequality valid for any open set $\Omega \subseteq \mathbb{R}^N$ not necessarily bounded in any direction (compare with Theorems 1.8.1 and 1.8.4). The result is often referred as the "Poincaré–Sobolev inequality".

Proposition 1.9.6 *Assume that $\Omega \subseteq \mathbb{R}^N$ is an open set and $1 \leqslant p < N$. Then $||u||_{p^*} \leqslant c(p, N)||Du||_p$ for all $u \in W^{1,p}_0(\Omega)$.*

Proof Let $u \in W^{1,p}_0(\Omega)$ and let $\hat{u}$ be its extension by zero on $\mathbb{R}^N \backslash \Omega$. We know that $\hat{u} \in W^{1,p}(\Omega)$ (see Proposition 1.1.17). We apply Theorem 1.9.3 and have

$$\begin{aligned}
||\hat{u}||_{L^{p^*}(\mathbb{R}^N)} &\leqslant c(p, N)||D\hat{u}||_{L^p(\mathbb{R}^N,\mathbb{R}^N)} \\
\Rightarrow ||u||_{L^{p^*}(\Omega)} &\leqslant c(p, N)||Du||_{L^p(\Omega,\mathbb{R}^N)},
\end{aligned}$$

since $\hat{u} = 0,\ D\hat{u} = 0$ on $\mathbb{R}^N \backslash \Omega$. ☐

Remark 1.9.7 We stress that in the above inequality the constant $c(p, N)$ is independent of the set Ω, in contrast to the usual Poincaré inequality (see Theorem 1.8.4) where the constant depends on Ω. If in the above theorem we assume that Ω is bounded, then we can easily recover the usual Poincaré inequality. To see this note that

$$\begin{aligned}
&||u||_p \leqslant \lambda^N(\Omega)^{1-\frac{p}{p^*}}\ ||u||^{p/p^*}_{p^*} \text{ (by Hölder's inequality)} \\
\Rightarrow\ &||u||_p \leqslant \lambda^N(\Omega)^{\frac{1}{p}-\frac{1}{p^*}}\ c(p, N)||Du||_p \text{ (see Theorem 1.9.3)} \\
\Rightarrow\ &||u||_p \leqslant \lambda^N(\Omega)^{\frac{1}{N}}\ c(p, N)\ ||Du||_p \left(\text{since } \frac{1}{p} - \frac{1}{p^*} = \frac{1}{N}\right).
\end{aligned}$$

From this we see that the constant in Theorem 1.8.4 satisfies $c(\lambda\Omega) = \lambda\, c(\Omega)$ for all $\lambda > 0$.

Proposition 1.9.6 leads to the following embedding result, which can be proved as Corollary 1.9.5.

Proposition 1.9.8 *Assume that $\Omega \subseteq \mathbb{R}^N$ is an open set, $1 \leqslant p < N$ and $1 \leqslant q \leqslant p^*$. Then $W_0^{1,p}(\Omega)$ is embedded continuously in $L^q(\Omega)$.*

This result can also be stated with $W_0^{1,p}(\Omega)$ replaced by $W^{1,p}(\Omega)$ provided we strengthen the regularity of the boundary $\partial\Omega$ in order to exploit the extension operator produced in Theorem 1.6.5, since extension by zero does not work in this case (see the proof of Corollary 1.9.5).

Proposition 1.9.9 *Assume that Ω has uniformly Lipschitz boundary $\partial\Omega$ and $1 \leqslant p < N$. Then there exists a $c = c(p, N, \Omega) > 0$ such that*

$$||u||_q \leqslant c\, ||Du||_p \text{ for all } u \in W^{1,p}(\Omega) \text{ and all } p \leqslant q \leqslant p^*.$$

In particular, $W^{1,p}(\Omega)$ is continuously embedded in $L^q(\Omega)$ for all $q \in [p, p^*]$.

Next we consider the case $p > N$. We will show in this case that a Sobolev function $u \in W^{1,p}(\mathbb{R}^N)$ is in fact Hölder continuous, possibly after a modification on a Lebesgue-null set. The result is known in the literature as "Morrey's inequality" or "Morrey's embedding theorem".

Theorem 1.9.10 (Morrey) *Assume that $N < p \leqslant \infty$. Then there exists a constant $c(p, N) > 0$ such that for all $u \in W^{1,p}(\mathbb{R}^N)$ we have*

$$|u(y) - u(z)| \leqslant c(p, N)\, ||Du||_{L^p(\mathbb{R}^N)} ||y - z||^{\alpha}_{\mathbb{R}^N} \text{ for a.a. } y, z \in \mathbb{R}^N$$

with $\alpha = 1 - \frac{N}{p} > 0$; in particular $W^{1,p}(\mathbb{R}^N)$ is embedded continuously in $C^{0,\alpha}(\mathbb{R}^N)$ and

$$||u||_{C^{0,\alpha}(\mathbb{R}^N)} \leqslant \hat{c}||u||_{W^{1,p}(\mathbb{R}^N)}$$

for all $u \in W^{1,p}(\mathbb{R}^N)$ and some $\hat{c} = \hat{c}(p, N) > 0$.

Proof First we assume that $u \in W^{1,p}(\mathbb{R}^N) \cap C^\infty(\mathbb{R}^N)$.

We consider a ball $B_r(z) \subseteq \mathbb{R}^N$. Let $h \in \partial B_1(0)$ and $0 < t < r$. We have

$$\begin{aligned}
|u(z + th) - u(z)| &= \left|\int_0^t \frac{d}{ds}\, u(z + sh)\, ds\right| \\
&= \left|\int_0^t (Du(z + sh), h)_{\mathbb{R}^N}\, ds\right| \\
&\leqslant \int_0^t ||Du(z + sh)||_{\mathbb{R}^N}\, ds
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \int_{\partial B_1(0)} |u(z + th) - u(z)|\, d\mathscr{H}^{N-1} &\leqslant \int_0^t \int_{\partial B_1(0)} ||Du(z + sh)||_{\mathbb{R}^N} d\mathscr{H}^{N-1}\, ds \\
&= \int_0^s \int_{\partial B_1(0)} ||Du(z + sh)||_{\mathbb{R}^N} \frac{s^{N-1}}{s^{N-1}}\, d\mathscr{H}^{N-1} ds.
\end{aligned}$$

We set $y = z + sh$ and so $s = ||y - z||_{\mathbb{R}^N}$. Therefore

$$\int_{\partial B_1(0)} |u(z+th) - u(z)| \, dH^{H-1} \leqslant \int_{B_r(z)} \frac{||Du(y)||_{\mathbb{R}^N}}{||y-z||_{\mathbb{R}^N}^{N-1}} \, dy.$$

We multiply by t^{N-1} and integrate from 0 to r. Then

$$\begin{aligned} &\int_{B_r(0)} |u(y) - u(z)| \, dy \leqslant \frac{r^N}{N} \int_{B_r(z)} \frac{||Du(y)||_{\mathbb{R}^N}}{||y-z||_{\mathbb{R}^N}^{N-1}} \, dy \\ \Rightarrow \ &\frac{1}{\hat{\lambda}^N(B_r(z))} \int_{B_r(z)} |u(y) - u(z)| dy \leqslant \ c \int_{B_r(z)} \frac{||Du(y)||_{\mathbb{R}^N}}{||y-z||_{\mathbb{R}^N}^{N-1}} \, dy \end{aligned} \tag{1.74}$$

with $c > 0$ depending only on N, p (recall that $\hat{\lambda}^N(B_r(z)) = r^N \xi(N)$, where $\xi(N)$ is the volume of the unit ball given by $\xi(N) = \frac{\Gamma^{N/2}}{\Gamma(\frac{N}{2}+1)}$).

We fix $z \in \mathbb{R}^N$. We have

$$\begin{aligned} |u(z)| \ &\leqslant \frac{1}{\hat{\lambda}^N(B_1(z))} \int_{B_1(z)} |u(z) - u(y)| dy + \frac{1}{\hat{\lambda}^N(B_1(z))} \int_{B_1(z)} |u(y)| \, dy \\ &\leqslant c_1 \int_{B_1(z)} \frac{||Du(y)||_{\mathbb{R}^N}}{||y-z||_{\mathbb{R}^N}^{N-1}} \, dy + c_1 ||u||_{L^p(\mathbb{R}^N)} \\ &\qquad\qquad \text{for some } c_1 \geqslant c \text{ depending only on } N, p \text{ (see (1.74))} \\ &\leqslant c_1 ||Du||_p \left[\int_{B_1(z)} \frac{dy}{||y-z||^{(N-1)p'}} \right]^{1/p'} + c_1 ||u||_{L^p(\mathbb{R}^N)} \\ &\leqslant c_2 ||u||_{W^{1,p}(\mathbb{R}^N)} \text{ for some } c_2 \geqslant c_1 \text{ depending only on } N, p. \end{aligned} \tag{1.75}$$

To see this last inequality, note that since $N < p$, we have $(N-1)p' < N$ (recall that $p' = \frac{p}{p-1}$). Therefore

$$\int_{B_1(z)} \frac{dy}{||y-z||^{(N-1)p'}} \, dy < \infty.$$

Since $z \in \mathbb{R}^N$ is arbitrary, from (1.75) it follows that

$$||u||_{L^\infty(\mathbb{R}^N)} \leqslant c_2 ||u||_{W^{1,p}(\mathbb{R}^N)}. \tag{1.76}$$

Next, let $y, z \in \mathbb{R}^N$ and set $r = ||y - z||_{\mathbb{R}^N}, \ U = B_r(y) \cap B_r(z)$. Then

$$|u(y) - u(z)| \leqslant \frac{1}{\lambda^N(U)} \left(\int_U |u(y) - u(v)| dv + \int_U |u(v) - u(z)| dv \right). \tag{1.77}$$

We estimate the two integrals on the right-hand side

$$
\begin{aligned}
\frac{1}{\lambda^N(U)}\int_U |u(y)-u(v)|dv &\leqslant \frac{c_3\, r^N}{\lambda^N(B_r(y))}\int_{B_r(y)} |u(y)-u(v)|\,dv \\
&\qquad \text{for some } c_3>0 \text{ depending only on } N,\, p \\
&\leqslant c_4 ||Du||_{L^p(\mathbb{R}^N,\mathbb{R}^N)}\, r^{(N-(N-1)p')\frac{1}{p'}} \quad \text{(see (1.74))} \\
&= c_4\, r^{1-\frac{N}{p}} ||Du||_{L^p(\mathbb{R}^N,\mathbb{R}^N)} \qquad (1.78)\\
&\qquad \text{for some } c_4>0 \text{ depending only on } N,\, p.
\end{aligned}
$$

Similarly, we obtain

$$
\frac{1}{\lambda^N(U)}\int_U |u(y)-u(v)|\,dv \leqslant c_4\, r^{1-\frac{N}{p}} ||Du||_{L^p(\mathbb{R}^N,\mathbb{R}^N)}. \qquad (1.79)
$$

Returning to (1.77) and using (1.78) and (1.79), we have

$$
\begin{aligned}
|u(y)-u(z)| &\leqslant c_4\, r^{1-\frac{N}{p}} ||Du||_{L^p(\mathbb{R}^N,\mathbb{R}^N)} \qquad (1.80)\\
&= c_4 ||y-z||_{\mathbb{R}^N}^{1-\frac{N}{p}}\, ||Du||_{L^p(\mathbb{R}^N,\mathbb{R}^N)}.
\end{aligned}
$$

From (1.76) and (1.80) we conclude that

$$
||u||_{C^{0,\alpha}(\mathbb{R}^N)} \leqslant \hat{c}\, ||Du||_{L^p(\mathbb{R}^N,\mathbb{R}^N)}
$$

for all $u \in C^\infty(\mathbb{R}^N)$, some $\hat{c}>0$ depending on N, p and with $\alpha = 1-\frac{N}{p}>0$.

For general $u \in W^{1,p}(\mathbb{R}^N)$ we know that we can find $\{u_n\}_{n\geqslant 1} \subseteq C_c^\infty(\mathbb{R}^N)$ such that $u_n \to u$ in $W^{1,p}(\mathbb{R}^N)$ (see Theorem 1.1.21) and we have pointwise convergence on the set of Lebesgue points of u (see the proof of Theorem 1.1.21 and Proposition 1.1.3(b)). Given $y, z \in \mathbb{R}^N$ Lebesgue points of u, from the first part of the proof we have

$$
\begin{aligned}
&|u_n(y)-u(z)| \leqslant \hat{c}\, ||y-z||^{1-\frac{N}{p}} ||Du_n||_{L^p(\mathbb{R}^N,\mathbb{R}^N)} \\
&|u_n(z)| \leqslant c_2\, ||u_n||_{W^{1,p}(\mathbb{R}^N)}.
\end{aligned}
$$

Passing to the limit as $n\to\infty$ and recalling that the set of Lebesgue points of u has Lebesgue-null complement, we conclude the proof of the theorem. □

Remark 1.9.11 From the above proof it is clear that if $u \in W^{1,p}(\mathbb{R}^N)$ and $\bar{u} \in C^{0,\alpha}(\mathbb{R}^N)$ is its Hölder continuous representative, then $\bar{u}(z) \to 0$ as $||z||_{\mathbb{R}^N} \to \infty$.

Theorem 1.9.12 *Assume that Ω is bounded with Lipschitz boundary $\partial\Omega$ and $N < p < \infty$. Then $W^{1,p}(\Omega)$ is embedded continuously in $C^{0,\alpha}(\bar{\Omega})$ with $\alpha = 1-\frac{N}{p}>0$.*

Proof Theorem 1.6.1 implies that there exists a linear continuous map $P : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^N)$ such that $Pu|_\Omega = u$ and Pu has compact support. We can find $\{\hat{u}_n\}_{n\geqslant 1} \subseteq C_c^\infty(\mathbb{R}^N)$ such that

$$\hat{u}_n \to Pu \text{ in } W^{1,p}(\mathbb{R}^N) \text{ (see Theorem 1.1.21).} \tag{1.81}$$

From Theorem 1.9.10, we have

$$\begin{aligned} & ||\hat{u}_n - \hat{u}_k||_{C^{0,\alpha}(\mathbb{R}^N)} \leqslant \hat{c}\, ||\hat{u}_n - \hat{u}_k||_{W^{1,p}(\mathbb{R}^N)} \text{ for all } n, k \geqslant 1 \\ \Rightarrow\ & \hat{u}_k \to \hat{u} \text{ in } C^{0,\alpha}(\mathbb{R}^N) \\ \Rightarrow\ & \hat{u} = Pu \text{ (see (1.81)), that is } \hat{u} \text{ is a version of } Pu. \end{aligned}$$

Recall that

$$\begin{aligned} & ||\hat{u}_n||_{C^{0,\alpha}(\mathbb{R}^N)} \leqslant \hat{c}\, ||\hat{u}_n||_{W^{1,p}(\mathbb{R}^N)} \text{ for all } n \geqslant 1 \text{ (see Theorem 1.9.10)} \\ \Rightarrow\ & ||\hat{u}||_{C^{0,\alpha}(\mathbb{R}^N)} \leqslant \hat{c}\, ||Pu||_{W^{1,p}(\mathbb{R}^N)} \\ \Rightarrow\ & ||u||_{C^{0,\alpha}(\bar{\Omega})} \leqslant \hat{c}\, ||u||_{1,p} \text{ (recall that } P \text{ is continuous).} \end{aligned}$$

The proof is now complete. ☐

Finally, we treat the limit case $p = N$. Note that $p^* = \frac{Np}{N-p} \to +\infty$ as $p \to N^-$. So we may claim that in this limit case every $u \in W^{1,N}(\Omega)$ belongs to $L^\infty(\Omega)$. However, this is not true if $N \geqslant 2$. To see this, let $\Omega = B_1(0)$ and consider the function $u(z) = \ln\left(\ln\left(1 + \frac{1}{||z||_{\mathbb{R}^N}}\right)\right)$. Then $u \in W^{1,N}(\Omega)$, but $u \notin L^\infty(\Omega)$. For this case, we have the following result.

Theorem 1.9.13 *The space $W^{1,N}(\mathbb{R}^N)$ is continuously embedded in $L^q(\mathbb{R}^N)$ for all $q \in [N, +\infty)$.*

Proof Let $u \in W^{1,N}(\mathbb{R}^N)$ and define $y = |u|^t$ with $t > 1$ to be determined so that $y \in W^{1,1}(\mathbb{R}^N)$. Using Theorem 1.9.3 with $p = 1$ and Proposition 1.4.2, we have

$$\begin{aligned} \left(\int_{\mathbb{R}^N} |u|^{\frac{tN}{N-1}} dz\right)^{\frac{N-1}{N}} &= \left(\int_{\mathbb{R}^N} |y|^{\frac{N}{N-1}} dz\right)^{\frac{N-1}{N}} \\ &\leqslant \int_{\mathbb{R}^N} ||Dy||_{\mathbb{R}^N} dz \\ &\leqslant t \int_{\mathbb{R}^N} |u|^{t-1} ||Du||_{\mathbb{R}^N} dz \\ &\leqslant t \left(\int_{\mathbb{R}^N} |u|^{(t-1)N'} dz\right)^{\frac{1}{N'}} ||Du||_{L^N(\mathbb{R}^N,\mathbb{R}^N)} \\ &\qquad \text{(by Hölder's inequality).} \end{aligned}$$

Therefore

$$\left(\int_{\mathbb{R}^N} |u|^{\frac{tN}{N-1}} dz\right)^{\frac{N-1}{tN}} \leqslant c \left(\int_{\mathbb{R}^N} |u|^{(t-1)\frac{N}{N-1}} dz\right)^{\frac{N-1}{tN}} ||Du||^{\frac{1}{t}}_{L^N(\mathbb{R}^N,\mathbb{R}^N)}$$
$$\leqslant c \left[\left(\int_{\mathbb{R}^N} |u|^{(t-1)\frac{N}{N-1}} dz\right)^{\frac{N-1}{N}\ \frac{1}{t-1}} + ||Du||_{L^N(\mathbb{R}^N,\mathbb{R}^N)}\right]$$
$$\text{(by Young's inequality).} \tag{1.82}$$

Taking $t = N$, we have

$$\left(\int_{\mathbb{R}^N} |u|^{\frac{N^2}{N-1}} dz\right)^{\frac{N-1}{N^2}} \leqslant c \left[||u||_{L^N(\mathbb{R}^N)} + ||Du||_{L^N(\mathbb{R}^N,\mathbb{R}^N)}\right]$$
$$\Rightarrow u \in L^{\frac{N^2}{N-1}}(\mathbb{R}^N) \text{ with continuous embedding.}$$

As in the proof of Corollary 1.9.5, using the interpolation inequality, we obtain

$$||u||_{L^q(\mathbb{R}^N)} \leqslant c\ ||u||_{W^{1,p}(\mathbb{R}^N)} \text{ for all } q \in \left[N, \frac{N^2}{N-1}\right]. \tag{1.83}$$

Taking $t = N + 1 \leqslant \frac{N^2}{N-1}$ in (1.82) and using (1.83), we show

$$\left(\int_{\mathbb{R}^N} |u|^{\frac{N(N+1)}{N-1}} dz\right)^{\frac{N-1}{N(N+1)}} \leqslant c\ ||u||_{W^{1,p}(\mathbb{R}^N)}$$
$$\Rightarrow\ W^{1,p}(\mathbb{R}^N) \text{ is embedded continuously in } L^q(\mathbb{R}^N) \text{ for all } q \in \left[N, \frac{N(N+1)}{N-1}\right].$$

Continuing in this way, taking $t = N + 2,\ t = N + 3$, etc., we have

$$W^{1,p}(\mathbb{R}^N) \hookrightarrow L^q(\mathbb{R}^N) \text{ continuously for all } q \in [N, +\infty)\,.$$

The proof is now complete. □

As before (see for example Proposition 1.9.9 and Theorem 1.9.12), using the extension operator and Theorem 1.9.12, we obtain the following property.

Theorem 1.9.14 *Assume that Ω is bounded with Lipschitz boundary $\partial\Omega$. Then $W^{1,N}(\Omega)$ is embedded continuously in $L^q(\Omega)$ for all $q \in [1, +\infty)$.*

So, summarizing the situation for bounded sets, we can state the following theorem, known as the "Sobolev embedding theorem".

Theorem 1.9.15 *Assume that Ω is bounded with Lipschitz boundary $\partial\Omega$. Then the following properties are true:*

(a) *if* $1 \leqslant p < N$, *then* $W^{1,p}(\Omega) \hookrightarrow L^q(\Omega)$ *for all* $q \in [1, p^*]$ *and the embedding is compact for* $q \in [1, p^*)$;

(b) *if* $p = N$, *then* $W^{1,N}(\Omega) \hookrightarrow L^q(\Omega)$ *for all* $q \in [1, \infty)$ *and the embedding is compact;*

(c) *if* $N < p \leqslant \infty$, *then* $W^{1,N}(\Omega) \hookrightarrow C(\bar{\Omega})$ *and the embedding is compact.*

Proof (a) The continuous embedding of $W^{1,p}(\Omega)$ into $L^q(\Omega)$ for $q \in [1, p^*]$ follows from Proposition 1.9.9. We need to show that the embedding is compact if $q \in [1, p^*)$. By virtue of Theorem 1.7.6, we need to consider only the case $q \in (p, p^*)$.

Let $\{u_n\}_{n\geqslant 1} \subseteq W^{1,p}(\Omega)$ be bounded. We may assume that

$$u_n \xrightarrow{w} u \text{ in } W^{1,p}(\Omega). \tag{1.84}$$

Let $t \in (0, 1)$ such that $\frac{1}{q} = \frac{t}{p} + \frac{(1-t)}{p^*}$. From the interpolation inequality, we have

$$\begin{aligned} &||u_n - u||_q \leqslant ||u_n - u||_p^t \, ||u_n - u||_{p^*}^{1-t} \leqslant \\ &c\,||u_n - u||_p^t \text{ for some } c > 0, \text{ all } n \geqslant 1 \text{ (see (1.84))} \\ \Rightarrow\ &||u_n - u||_q \to 0 \text{ (see Theorem 1.7.6).} \end{aligned}$$

This proves the compactness of the embedding $W^{1,p}(\Omega) \hookrightarrow L^q(\Omega)$ for $q \in [1, p^*)$.

(b) Let $\{u_n\}_{n\geqslant 1} \subseteq W^{1,N}(\Omega)$ be bounded. We may assume that

$$u_n \xrightarrow{w} u \text{ in } W^{1,N}(\Omega). \tag{1.85}$$

Let $N < q < r < \infty$ and $t \in (0, 1)$ such that $\frac{1}{q} = \frac{t}{N} + \frac{1-t}{r}$. Using the interpolation inequality, we have

$$\begin{aligned} &||u_n - u||_q \leqslant ||u_n - u||_N^t \, ||u_n - u||_r^{1-t} \leqslant \\ &c\,||u_n - u||_p^t \text{ for some } c > 0, \text{ all } n \geqslant 1 \text{ (see (1.85))} \\ \Rightarrow\ &||u_n - u||_q \to 0 \text{ (see Theorem 1.7.6).} \end{aligned}$$

This proves the compactness of the embedding $W^{1,N}(\Omega) \hookrightarrow L^q(\Omega)$ for $q \in [1, +\infty)$.

(c) From Theorem 1.9.12 we have a continuous embedding of $W^{1,p}(\Omega)$ into the Hölder space $C^{0,\alpha}(\bar{\Omega})$ $\left(\alpha = 1 - \frac{N}{p}\right)$. By virtue of the Arzela–Ascoli theorem $C^{0,\alpha}(\bar{\Omega})$ is embedded compactly in $C(\bar{\Omega})$. So, we conclude the compact embedding of $W^{1,p}(\Omega)$ into $C(\bar{\Omega})$. $\square$

Remark 1.9.16 If Ω is not bounded, then the embedding $W^{1,p}(\Omega) \hookrightarrow L^p(\Omega)$ is not in general compact. In (a) the embedding of $W^{1,p}(\Omega)$ in $L^{p^*}(\Omega)$ is never compact.

Of course, we can also state the Sobolev embedding theorem for higher order Sobolev spaces (see Adams [2]).

Theorem 1.9.17 *Assume that Ω is bounded with Lipschitz boundary $\partial\Omega$, $1 \leqslant p, q < \infty$, and $m > k \geqslant 0$ integers. Then*

(a) $W^{m,p}(\Omega) \hookrightarrow W^{k,q}(\Omega)$ *continuously if* $d = \frac{1}{p} - \frac{m-k}{N} \leqslant \frac{1}{q}$ *and compactly if* $d < \frac{1}{q}$*;*
(b) $W^{m,p}(\Omega) \hookrightarrow C^{k,\alpha}(\bar{\Omega})$ *continuously if* $\alpha = m - k - \frac{N}{p}$ *and compactly if* $\alpha < m - k - \frac{N}{p}$*.*

Remark 1.9.18 Part (a) is still valid if $W^{1,p}(\Omega)$ is replaced by $W_0^{1,p}(\Omega)$. In fact, in this case Ω can have arbitrary boundary $\partial\Omega$.

1.10 Capacities. Miscellaneous Results

Capacity theory allows the study of small sets in $\mathbb{R}^N$. One can show that in $\mathbb{R}^N$ there are Lebesgue-null sets with capacity strictly bigger than zero. So, it makes sense to speak about the values of a function $u \in W^{1,p}(\Omega)$ on a set $D \subseteq \Omega$ with p-capacity bigger than zero. Such a set can be for example the boundary of an open set. Capacity theory is useful when $p \leqslant N$, because for $p > N$, we know that for Ω bounded with Lipschitz boundary $\partial\Omega$, we have $W^{1,p}(\Omega) \hookrightarrow C(\bar{\Omega})$ compactly and then the only set with p-capacity zero is the empty set. For this reason, in this section we assume that $1 < p < N$.

Definition 1.10.1 Let $\Omega \subseteq \mathbb{R}^N$ be an open set and $1 < p < N$.

(a) Given any open set $U \subseteq \Omega$, the "p-capacity of U with respect to Ω" is defined by

$$\mathrm{Cap}_p(U) = \inf \left\{ \int_\Omega ||Du||^p_{\mathbb{R}^N} dz : \ u \in W_0^{1,p}(\Omega), \ u(z) \geqslant 1 \text{ a.e. in } U \right\}.$$

(b) We can extend the notion of p-capacity to any subset $A \subseteq \Omega$ by setting

$$\mathrm{Cap}_p(A) = \inf \left[\mathrm{Cap}_p(U) : \ U \supseteq A, \ U \text{ open} \right].$$

The next proposition provides equivalent formulations of the p-capacity $\mathrm{Cap}_p(A)$ for an arbitrary set $A \subseteq \Omega$. We mention that $u \geqslant 1$ a.e. in a neighborhood of A means there exists an open set $U \supseteq A$, $U \subseteq \Omega$ such that $u \geqslant 1$ a.e. in U.

Proposition 1.10.2 *(a) For any $A \subseteq \Omega$, we have*

$$\mathrm{Cap}_p(A) = \inf \left\{ \int_\Omega ||Du||^p_{\mathbb{R}^N} dz : \ u \in W_0^{1,p}(\Omega), \ u \geqslant 1 \textit{ a.e. in a neighborhood of } A \right\}.$$

(b) We have

$$\operatorname{Cap}_p(A) = \inf\left\{\int_\Omega ||Du||^p_{\mathbb{R}^N} dz : \ u \in W_0^{1,p}(\Omega),\ u \geqslant 0,\ u = 1 \textit{ a.e. in a neighborhood of } A\right\}.$$

Proof (a) Let $u \in W_0^{1,p}(\Omega)$ be such that $u \geqslant 1$ a.e. in a neighborhood of A. So, by definition there exists an open set $U,\ A \subseteq U \subseteq \Omega$, such that $u \geqslant 1$ a.e. in U. From Definition 1.10.1 (a) we have

$$\operatorname{Cap}_p(U) \leqslant \int_\Omega ||Du||^p_{\mathbb{R}^N} dz \tag{1.86}$$

and from Definition 1.10.1 (b) we have

$$\operatorname{Cap}_p(A) \leqslant \operatorname{Cap}_p(U). \tag{1.87}$$

Then from (1.86) and (1.87) we have

$$\operatorname{Cap}_p(A) \leqslant \inf\left\{\int_\Omega ||Du||^p_{\mathbb{R}^N} dz : \ u \in W_0^{1,p}(\Omega),\ u \geqslant 1 \text{ a.e. in a neighborhood of } A\right\}.$$

We need to show that the opposite inequality is also true. We may assume that $\operatorname{Cap}_p(A) < +\infty$ or otherwise the reverse inequality is true. By virtue of Definition 1.10.1 (b) given $\epsilon > 0$ we can find U_ϵ open such that $A \subseteq U_\epsilon \subseteq \Omega$ and

$$\operatorname{Cap}_p(U_\epsilon) \leqslant \operatorname{Cap}_p(A) + \epsilon. \tag{1.88}$$

According to Definition 1.10.1 (a), we can find $u_\epsilon \in W_0^{1,p}(\Omega)$ such that $u_\epsilon \geqslant 1$ a.e. in U_ϵ and

$$\int_\Omega ||Du_\epsilon||^p_{\mathbb{R}^N} dz \leqslant \operatorname{Cap}_p(U_\epsilon) + \epsilon\,. \tag{1.89}$$

From (1.88) and (1.89) it follows that

$$\int_\Omega ||Du_\epsilon||^p_{\mathbb{R}^N} dz \leqslant \operatorname{Cap}_p(A) + 2\epsilon \tag{1.90}$$

with $u_\epsilon \in W_0^{1,p}(\Omega),\ u_\epsilon \geqslant 1$ a.e. in a neighborhood of A. Therefore

$$\inf\left\{\int_\Omega |||Du|^p_{\mathbb{R}^N} dz : \ u \in W_0^{1,p}(\Omega),\ u \geqslant 1 \text{ a.e. in a neighborhood of } A\right\} \leqslant$$
$$\operatorname{Cap}_p(A) + 2\epsilon \text{ (see (1.90)).}$$

We let $\epsilon \to 0^+$ to obtain the opposite inequality. So, in fact we have equality.

(b) Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(s) = \min\{1, s^+\}$ (recall that $s^+ = \max\{s, 0\}$). Evidently $f(\cdot)$ is nonexpansive, that is, $|f(s) - f(\tau)| \leqslant |s - \tau|$ for all $s, \tau \in \mathbb{R}$. Suppose that $u \in W_0^{1,p}(\Omega)$ satisfies $u \geqslant 1$ a.e. in a neighborhood of A. Then $f \circ u \in W_0^{1,p}(\Omega)$ (see Proposition 1.4.2 and Remark 1.4.3), $f \circ u \geqslant 0$ and $f \circ u = 1$ a.e. in a neighborhood of A. Moreover, Propositions 1.4.4 and 1.4.5 imply that

$$\int_\Omega ||D(f \circ u)||^p_{\mathbb{R}^N} dz \ \leqslant \ \int_\Omega ||Du||^p_{\mathbb{R}^N} dz. \tag{1.91}$$

It follows that

$$\begin{aligned}
&\inf \Big\{ \int_\Omega ||Du||^p_{\mathbb{R}^N} dz : \ u \in W_0^{1,p}(\Omega), \ u \geqslant 0, \ u = 1 \\
&\qquad\qquad\qquad\qquad \text{a.e. in a neighborhood of } \ A \} \\
&\leqslant \int_\Omega ||D(f \circ u)||^p_{\mathbb{R}^N} dz \ \leqslant \ \int_\Omega ||Du||^p_{\mathbb{R}^N} dz \text{ (see (1.91))} \\
\Rightarrow &\inf \Big\{ \int_\Omega ||Du||^p_{\mathbb{R}^N} dz : \ u \in W_0^{1,p}(\Omega), \ u \geqslant 0, \ u = 1 \\
&\qquad\qquad\qquad\qquad \text{a.e. in a neighborhood of } \ A \} \\
&\leqslant \inf \left\{ \int_\Omega ||Du||^p_{\mathbb{R}^N} dz : \ u \in W_0^{1,p}(\Omega), \ u \geqslant 1 \text{ a.e. in a neighborhood of } A \right\} \\
&= \mathrm{Cap}_p(A) \text{ (see } (a)).
\end{aligned}$$

The opposite inequality is clearly true. Therefore finally we have equality. ☐

Next we show that the notion of p-capacity is in fact a particular case of a Choquet capacity (see Choquet [121, 122]).

Proposition 1.10.3 *(a) If $k \subseteq \Omega$ is compact, then*

$$\mathrm{Cap}_p(K) = \inf \left\{ \int_\Omega ||Du||^p_{\mathbb{R}^N} dz : \ u \in C_c^\infty(\Omega), \ u(z) \geqslant 1 \textit{ for all } z \in K \right\}.$$

(b) For every $U \subseteq \Omega$ open we have

$$\mathrm{Cap}_p(U) = \sup \left\{ \mathrm{Cap}_p(K) : \ K \subseteq \Omega, \ K \textit{ is compact} \right\}.$$

Proof (a) For $\epsilon > 0$, we set $K_\epsilon = \{z \in \Omega : d(z, K) < \epsilon\}$. We can choose $\epsilon > 0$ small such that $K_\epsilon \subseteq \Omega$. Suppose $u \in W_0^{1,p}(\Omega)$ satisfies $u \geqslant 1$ a.e. in a neighborhood of K. Then we can find U open, $K \subseteq U \subseteq \Omega$, such that $u \geqslant 1$ a.e. in U. So, for $\epsilon > 0$ small we have $u \geqslant 1$ a.e. in K_ϵ. As in the proof of Theorem 1.1.12, through truncation of the domain and mollification, we can find $\{u_n\}_{n \geqslant 1} \subseteq C_c^\infty(\Omega)$ such that for all $n \geqslant 1$

$$u_n \geqslant 1 \text{ a.e. in a neighborhood of } K \text{ and } ||Du_n||_p \to ||Du||_p.$$

It follows that

$$\begin{aligned}
\mathrm{Cap}_p(K) &= \inf \left\{ \int_\Omega ||Du||^p_{\mathbb{R}^N} dz : \ u \in C_c^\infty(\Omega), \ u \geqslant 1 \text{ in a neighborhood of } K \right\} \\
&\geqslant \inf \left\{ \int_\Omega ||Du||^p_{\mathbb{R}^N} dz : \ u \in C_c^\infty(\Omega), \ u \geqslant 1 \text{ on } K \right\}.
\end{aligned}$$

On the other hand the opposite inequality holds since $\mathrm{Cap}_p(K) \leqslant \inf\limits_{\epsilon>0} \mathrm{Cap}_p(K_\epsilon)$. Therefore equality holds.

(b) Evidently, $\mathrm{Cap}_p(\cdot)$ is a monotone set function. Hence

$$\sup \left\{ \mathrm{Cap}_p(K) : \ K \subseteq U, \ K \text{ is compact} \right\} \leqslant \ \mathrm{Cap}_p(U).$$

To prove the opposite inequality, we assume that the left-hand side is finite. We can find $\{K_n\}_{n\geqslant 1}$ compact subsets of U such that $\cup_{n\geqslant 1} K_n = U$. For every $n \geqslant 1$, by virtue of part (a) we can find $u_n \in C_c^\infty(\Omega)$ such that

$$\int_\Omega ||Du_n||^p_{\mathbb{R}^N} dz \leqslant \mathrm{Cap}_p(K_n) + \frac{1}{n} \text{ and } u_n \geqslant 1 \text{ on } K_n.$$

It follows that $\{u_n\}_{n\geqslant 1} \subseteq W_0^{1,p}(\Omega)$ is bounded. So, we may assume that $u_n \xrightarrow{w} u$ in $W_0^{1,p}(\Omega)$ with $u \in W_0^{1,p}(\Omega)$. By Mazur's lemma we obtain $u \geqslant 1$ a.e. in U. We also have

$$\begin{aligned}
\mathrm{Cap}_p(U) &\leqslant \int_\Omega ||Du||^p_{\mathbb{R}^N} dz \\
&\leqslant \liminf_{n\to\infty} \int_\Omega ||Du_n||^p_{\mathbb{R}^N} dz \\
&\leqslant \liminf_{n\to\infty} \ \mathrm{Cap}_p(K_n) \\
&\leqslant \sup \left\{ \mathrm{Cap}_p(K) : \ K \text{ is compact}, \ K \subseteq U \right\}.
\end{aligned}$$

Therefore equality holds. □

Next we deal with restrictions of continuous functions on nodal domains.

Definition 1.10.4 Let $\Omega \subseteq \mathbb{R}^N$ be an open set and $u \in C(\Omega)$. Let $Z(u) = \{z \in \Omega; \ u(z) = 0\}$. Then a "nodal domain" of u is a connected component of $\Omega \backslash Z(u)$.

Proposition 1.10.5 *Assume that Ω is bounded, $X = W_0^{1,p}(\Omega)$ or $X = W^{1,p}(\Omega)$, $u \in X \cap C(\Omega)$, and U is a nodal domain of u. Then $\hat{u} = u\chi_U \in X$ and $D\hat{u} = (Du)\chi_U$ a.e. in U.*

Proof By definition, $\hat{u}$ is strictly positive or strictly negative in U, so we may assume that $u|_U > 0$. Then replacing u by u^+ (note that $u^+ \in X$, see Proposition 1.4.4), we may assume that $u \geqslant 0$.

First we assume that $X = W_0^{1,p}(\Omega)$. Then by virtue of Corollary 1.4.10, we can find $\{u_n\}_{n\geqslant 1} \subseteq C_c^\infty(\Omega)$, $u_n \geqslant 0$ such that $u_n \to u$ in $W_0^{1,p}(\Omega)$. Let $h_n = \min\{u, u_n\}$. Then $h_n \in W_0^{1,p}(\Omega) \cap C(\bar{\Omega})$ and has compact support. We claim that $h_n|_{\partial\Omega} = 0$. To see this, first consider $z \in \Omega \cap \partial U$. Then since U is a nodal domain, $u(z) = 0$. Next consider $z \in \partial\Omega \cap \partial U$. Then $h_n(z) = 0$ since u_n has compact support.

Let $\hat{h}_n = h_n\chi_U \in W^{1,p}(\Omega) \cap C(\bar{\Omega})$ and $\hat{h}_n|_{\partial\Omega} = 0$. Therefore $\hat{h}_n \in W_0^{1,p}(\Omega)$ for all $n \geqslant 1$ (see Theorem 1.5.4). Moreover, we have

$$\begin{aligned} & D\hat{u} = (Du)\chi_U, \ \ D\hat{h}_n = (Dh_n)\chi_U \text{ and } h_n \to u \text{ in } W^{1,p}(\Omega) \\ \Rightarrow\ & \hat{h}_n \to \hat{u} \text{ in } W^{1,p}(\Omega) \\ \Rightarrow\ & \hat{u} \in W_0^{1,p}(\Omega). \end{aligned}$$

Next assume that $X = W^{1,p}(\Omega)$. Let $z \in \Omega$, B a ball centered at z such that $\bar{B} \subseteq \Omega$ and $\varphi \in C_c^\infty(B)$ such that $\varphi \equiv 1$ in a neighborhood of z. We have $\varphi\hat{u} \in W^{1,p}(U \cap B) \cap C(U \cap B)$ and $\varphi\hat{u}|_{\partial(U\cap B)} = 0$. Therefore $\varphi\hat{u} \in W_0^{1,p}(U \cap B)$ (see Theorem 1.5.4). This proves that $\hat{u} \in W_{\mathrm{loc}}^{1,p}(\Omega)$. On the other hand we can easily see that $D\hat{u} = (Du)\chi_U$, hence $||D\hat{u}(z)||_{\mathbb{R}^N} \leqslant ||Du(z)||_{\mathbb{R}^N}$ for a.a. $z \in \Omega$. We conclude that $\hat{u} \in W^{1,p}(\Omega)$. $\square$

Proposition 1.10.6 *Assume that Ω is bounded with Lipschitz boundary $\partial\Omega$, $u \in W^{1,p}(\Omega) \cap C(\bar{\Omega})$, u has finitely many nodal domains, and $U \subseteq \Omega$ is such a nodal domain. Then $\hat{u} = u\chi_{\bar{U}} \in W^{1,p}(\Omega) \cap C(\bar{\Omega})$ and $\hat{u}|_{\partial\Omega}(z) = \begin{cases} u(z) & \text{if } z \in \partial\Omega \cap \partial U \\ 0 & \text{if } z \in \partial\Omega \setminus \partial U. \end{cases}$*

Proof From Proposition 1.10.5 we know that $\hat{u} \in W^{1,p}(\Omega) \cap C(\bar{\Omega})$.

Let $\Omega_+ = \{z \in \Omega :\ u(z) > 0\}$ and without any loss of generality assume that Ω_+ has two connected components U_1 and U_2.

Let $h_1 = u\chi_{\bar{U}_1}$ and $h_2 = u\chi_{\bar{U}_2}$. Let B be a ball containing Ω (recall that Ω is bounded) and extend u to a function in $C_c^1(B)$ (using a cut-off function). Then we can find two sequences $\{v_n\}_{n\geqslant 1}$, $\{y_n\}_{n\geqslant 1} \subseteq C^1(\bar{\Omega})$ such that

$$\operatorname{supp} v_n \subseteq \bar{U}_1,\ \operatorname{supp} y_n \subseteq \bar{U}_2 \quad \text{for all} n \geqslant 1$$

and

$$v_n \to h_1,\ y_n \to h_2 \quad \text{in} W^{1,p}(\Omega).$$

Then $\gamma_0(v_n) \to \gamma_0(h_1)$ and $\gamma_0(y_n) \to \gamma_0(h_2)$ in $L^p(\partial\Omega)$ (here γ_0 is the trace map, see Theorem 1.5.4). Also $v_n + u_n \to u^+$ in $W^{1,p}(\Omega)$, hence $\gamma_0(v_n + y_n) \to \gamma_0(u^+)$ in $L^p(\partial\Omega)$. We have

$$\gamma_0(u^+) = \begin{cases} u(z) & \text{if } z \in \partial\Omega \cap \partial U_1 \text{ or } z \in \partial\Omega \cap \partial U_2 \\ 0 & \text{otherwise} \end{cases} \quad \text{(see Theorem 1.5.4)} \tag{1.92}$$

Since $\operatorname{supp} v_n \subseteq \bar{U}_1$, $\operatorname{supp} y_n \subseteq \bar{U}_2$ and U_1, U_2 are disjoint, it follows that

$$h_k|_{\partial\Omega}(z) = \begin{cases} u(z) & \text{if } z \in \partial\Omega \cap \partial U_k \\ 0 & \text{otherwise} \end{cases} \quad \text{for } k = 1, 2.$$

The proof is now complete. □

Continuing with the presentation of some useful general results about Sobolev spaces, we give some equivalent norms for them.

The next proposition is a direct consequence of the Sobolev embedding theorem (see Theorem 1.9.15).

Proposition 1.10.7 *If $\Omega \subseteq \mathbb{R}^N$ $(N \geqslant 2)$ is bounded with Lipschitz boundary $\partial\Omega$, then $|u| = ||u||_q + ||Du||_p$ is an equivalent norm for $W^{1,p}(\Omega)$ provided*

$$\begin{aligned} &1 \leqslant q \leqslant p^* \text{ when } p \in [1, N)\,, \\ &1 \leqslant q < \infty \text{ when } p = N, \\ &1 \leqslant q \leqslant \infty \text{ when } p > N. \end{aligned}$$

In the one-dimensional case (that is, $N = 1$ and so $\Omega = I = (a, b)$, possibly unbounded) we know the following property (see Theorem 1.2.2 and Theorem 1.9.15).

Theorem 1.10.8 *(a) The Sobolev space $W^{1,p}(I)$ $(1 \leqslant p \leqslant \infty)$ is continuously embedded in $C(\bar{I})$.*

(b) If I is bounded, then $W^{1,p}(I)$ $(1 < p \leqslant \infty)$ is compactly embedded in $C(\bar{I})$ and $W^{1,1}(I)$ is compactly embedded in $L^q(I)$ for all $1 \leqslant q < \infty$.

Remark 1.10.9 The embedding $W^{1,1}(I)$ in $C(\bar{I})$ is always continuous (see Theorem 1.10.8 (a) and Theorem 1.2.2), but it is never compact (even if I is bounded). Nevertheless, we have the so-called "Helly's selection theorem" (see Denkowski et al. [143, p. 229], Kolmogorov and Fomin [245, p. 372] and Leoni [262, p. 59]), which says that:

"If $\{u_n\}_{n\geqslant 1} \subseteq W^{1,1}(I)$ is bounded, then we can extract a subsequence $\{u_{n_k}\}_{k\geqslant 1}$ of $\{u_n\}_{n\geqslant 1}$ which converges to a limit for every $t \in I$."

Also, if I is unbounded, then $W^{1,p}(I)$ $(1 < p \leqslant \infty)$ is embedded continuously but never compactly in $L^\infty(I)$. However, if $\{u_n\}_{n\geqslant 1} \subseteq W^{1,p}(I)$ $(1 < p \leqslant \infty)$ is bounded, then we can extract a subsequence $\{u_{n_k}\}_{k\geqslant 1}$ of $\{u_n\}_{n\geqslant 1}$ and a $u \in W^{1,p}(I)$ such that

$$u_{n_k} \to u \text{ in } L^\infty(T) \text{ for all } T \subseteq I \text{ bounded.}$$

Then we have the following equivalent norms for $W^{1,p}(I)$.

Proposition 1.10.10 *Assume that I is a bounded interval and $1 \leqslant p,\ q \leqslant \infty$. Then $|u| = ||u||_q + ||u'||_p$ is an equivalent norm for $W^{1,p}(I)$.*

To produce additional equivalent norms for the Sobolev spaces, we will need the following general theorem (see Mazja [294, p. 27]).

Theorem 1.10.11 *Assume that Ω is bounded with Lipschitz boundary, $1 \leqslant p < \infty$, $m \geqslant 1$ is an integer, and we set*

$$|u| = \left[\int_\Omega \sum_{||=\mathrm{m}} |D^\alpha u(z)|^p dz + \sum_{\mathrm{k}=1}^r f_k(u)^p \right]^{1/p} \quad \text{for all } u \in W^{1,p}(\Omega),$$

where $f_k : W^{m,p}(\Omega) \to \mathbb{R}$ are seminorms such that

$$0 \leqslant f_k(u) \leqslant c||u||_{m,p} \text{ for all } u \in W^{m,p}(\Omega), \text{ all } k \in \{1, \dots, r\}$$

(recall that $||u||_{m,p} = \left(\sum_{||\leqslant \mathrm{m}} ||D^\alpha u||_p^p \right)^{1/p}$ for all $u \in W^{m,p}(\Omega)$, see Definition 1.1.10). Assume that $f_k(P) = 0$ for all $k \in \{1, \dots, r\}$, where $P : \mathbb{R}^N \to \mathbb{R}$ is a polynomial of degree less or equal to $m - 1$, implies $P = 0$. Then $|\cdot|$ and $||\cdot||_{m,p}$ are equivalent norms for $W^{m,p}(\Omega)$.

Using this theorem, we can suggest some more equivalent norms for the Sobolev space $W^{1,p}(\Omega)$.

Proposition 1.10.12 *Assume that Ω is bounded with Lipschitz boundary $\partial\Omega$ and $1 \leqslant p < \infty$. Then the following norms are equivalent to the original norm $||\cdot||_{m,p}$ of $W^{m,p}(\Omega)$:*

$$\left[\sum_{\mathrm{k}=1}^N ||D_k u||_p^p + \left| \int_\Omega u \, dz \right|^p \right]^{1/p},$$

$$\left[\sum_{\mathrm{k}=1}^N ||D_k u||_p^p + \left| \int_{\partial\Omega} u \, d\mathscr{H}^{N-1} \right|^p \right]^{1/p},$$

$$\left[\sum_{\mathrm{k}=1}^N ||D_k u||_p^p + \int_{\partial\Omega} |u|^p \, d\mathscr{H}^{N-1} \right]^{1/p}.$$

If $N = 1$ and $\Omega = I = (a, b)$, $-\infty < a < b < \infty$, then $\int_{\partial I} u d\mathscr{H}^0 = -u(a) + u(b)$.

Proof This follows directly from Theorem 1.10.11. In this case P is a constant polynomial. □

Proposition 1.10.13 *Assume that $u \in W^{1,p}(\Omega)$ $1 < p \leqslant \infty$. Then there exists a $c > 0$ such that for all $\Omega_0 \subset\subset \Omega$ and all $h \in \mathbb{R}^N$ with $||h||_{\mathbb{R}^N} \leqslant d(\Omega_0, \partial\Omega)$ we have*

$$||\tau_h(u) - u||_{L^p(\Omega_0)} \leqslant c||h||_{\mathbb{R}^N},$$

where $\tau_h(u)(z) = u(z + h)$; in fact, c can be taken to be $||Du||_p$.

Proof Let $u \in C_c^\infty(\mathbb{R}^N)$ and $h \in \mathbb{R}^N$. We set

$$\vartheta(t) = u(z + th) \text{ for all } t \in \mathbb{R}\,.$$

Then we have

$$\vartheta'(t) = (Du(z + th), h)_{\mathbb{R}^N}.$$

Hence

$$\begin{aligned}
u(z+h) - u(z) &= \vartheta(1) - \vartheta(0) = \int_0^1 \vartheta'(t)dt = \int_0^1 (Du(z+th), h)_{\mathbb{R}^N} dt \\
\Rightarrow |\tau_h(u)(z) - u(z)|^p &\leqslant ||h||^p_{\mathbb{R}^N} \int_0^1 ||Du(z+th)||^p_{\mathbb{R}^N} dt \text{ (by Jensen's inequality)} \\
\Rightarrow \int_{\Omega_0} |\tau_h(u) - u|^p dz &\leqslant ||h||^p_{\mathbb{R}^N} \int_{\Omega_0} \int_0^1 ||Du(z+th)||^p_{\mathbb{R}^N} dt\, dz \\
&= ||h||^p_{\mathbb{R}^N} \int_0^1 \int_{\Omega_0} ||Du(z+th)||^p_{\mathbb{R}^N} dz\, dt \\
&\qquad \text{(by Fubini's theorem)} \\
&= ||h||^p_{\mathbb{R}^N} \int_0^1 \int_{\Omega_0 + th} ||Du(y)||^p_{\mathbb{R}^N} dy \\
&\qquad \text{(by a change of variables).}
\end{aligned}$$

If $||h||_{\mathbb{R}^N} < d(\Omega_0, \partial\Omega)$, then there exists an open set $\hat{\Omega}_0 \subset\subset \Omega$ such that $\hat{\Omega}_0 + th \subseteq \Omega_0$ for all $t \in [0, 1]$ and so

$$||\tau_h(u) - u||^p_{L^p(\Omega)} \leqslant ||h||^p_{\mathbb{R}^N}\, ||Du||^p_{L^p(\hat{\Omega}_0)}.$$

This proves the proposition for $u \in C_c^\infty(\mathbb{R}^N)$. For general $u \in W^{1,p}(\Omega)$ ($1 \leqslant p < \infty$), we use Theorem 1.1.25 (Friedrichs' theorem).

Finally, for $p = +\infty$, we use the result just proved for all $1 \leqslant p < \infty$ and let $p \to +\infty$. $\square$

Corollary 1.10.14 *Assume that* Ω *is connected and* $u \in W^{1,\infty}(\Omega)$. *Then* u *has a Lipschitz continuous representative* $\hat{u}$ *satisfying*

$$|\hat{u}(z) - \hat{u}(y)| \leqslant ||Du||_\infty d_\Omega(z, y),$$

where $d_\Omega(\cdot, \cdot)$ *denotes the "geodesic distance" from* z *to* y *in* Ω*; if* Ω *is convex, then* $d_\Omega(z, y) = ||z - y||_{\mathbb{R}^N}$.

Finally, we conclude this section by stating Hardy's inequality (see Brezis [74, p. 313]).

Theorem 1.10.15 *Assume that Ω is bounded with Lipschitz boundary and $1 < p < \infty$. Then $\left\| \frac{u}{d} \right\|_p \leqslant c||Du||_p$ for all $u \in W_0^{1,p}(\Omega)$ with $d(z) = d(z, \partial\Omega)$ for all $z \in \Omega$. Conversely, we have*

$$u \in W^{1,p}(\Omega) \text{ and } \frac{u}{d} \in L^p(\Omega) \Rightarrow u \in W_0^{1,p}(\Omega).$$

1.11 Remarks

1.1: The spaces $W^{m,p}(\Omega)$ were introduced by Sobolev [388] in the mid 1930s and they turned out to be very important in the development of partial differential equations. Similar spaces were also used by Morrey [305] and later by Dautray and Lions [137]. In the literature we can find several books devoted to the theory of Sobolev spaces. We mention the books of Adams [2], Adams and Fournier [3], Burenkov [96], Leoni [262] and Tartar [401]. Valuable information about Sobolev spaces can also by found in the books of Brezis [74], Evans [163], Evans and Gariepy [164] and Ziemer [429]. Initially there were two different approaches to the definition of the Sobolev spaces. The original one due to Sobolev based on weak derivatives (see Definition 1.1.10) and another one based on the completion of the space $\{u \in C^\infty(\Omega) : ||u||_{m,p} < \infty\}$. In fact, the latter spaces were denoted by $H^{m,p}(\Omega)$. The result of Meyers and Serrin [296] (see Theorem 1.1.23) clarified the situation and established that the two approaches are in fact equivalent.

1.2: For a more detailed presentation of the one-dimensional case, which is naturally related to the study of absolutely continuous functions (see Theorem 1.2.2 and Remark 1.2.3), we refer to Brezis [74, Chap. 8], Kannan and Krueger [228], and Natanson [315].

1.3: The important step of defining more general mathematical objects that permit us to define derivatives of any order for any locally integrable function, extending in this way the notion of weak derivative (see Definition 1.1.6), was performed by L. Schwartz, resulting in the theory of distributions (see Schwartz [377]). The analysis of Sobolev spaces and the study of their properties was facilitated significantly by the theory of distributions. This is evident in the description of $W_0^{1,p}(\Omega)^* = W^{-1,p'}(\Omega)$ (see Theorem 1.3.9) and why such a convenient representation is not available for $W^{1,p}(\Omega)^*$ (see Theorem 1.3.7 and Remark 1.3.8). In the theory of distributions, an important role is played by the "Radon measures". In fact, Schwartz considered distributions as the natural generalization of the previously developed Radon measures.

Definition 1.11.1 A Radon measure μ is a linear functional on $C_c(\Omega)$ such that for every compact $K \subseteq \Omega$, the restriction of μ on

$$C_K(\Omega) = \{u \in C_c(\Omega) : \operatorname{supp} u \subseteq K\}$$

is continuous, that is, there exists a constant $C_K > 0$ such that

$$|\mu(u)| \leqslant C_K ||u||_\infty \text{ for all } u \in C_K(\Omega).$$

Remark 1.11.2 To a Radon measure μ we associate its restriction to $C_c^\infty(\Omega)$. So, we can define the distribution

$$L_\mu(u) = \int_\Omega u(z)d\mu(z) \text{ for all } u \in C_c^\infty(\Omega).$$

Evidently, the density of $C_c^\infty(\Omega)$ in $C_c(\Omega)$ implies that μ is determined completely by the distribution L_μ. Therefore the space of Radon measure $M_r(\Omega)$ is embedded in the space of distributions $C_c^\infty(\Omega)^*$. Radon measures are important in variational analysis, in particular in connection with the theory of relaxation of integral functionals (see Buttazzo [97], Roubicek [365]).

1.4: Theorem 1.4.1 and the chain rule in Proposition 1.4.2 are due to Marcus and Mizel [286].

Using this chain rule, Marcus and Mizel [287] also proved the following result.

Proposition 1.11.3 *Assume that $\Omega \subseteq \mathbb{R}^N$ is open and $f : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous with $f(0) = 0$ when Ω is unbounded. Then for every $1 \leqslant p < \infty$, the Nemytskii (superposition) operator $N_f : W^{1,p}(\Omega) \to W^{1,p}(\Omega)$ defined by*

$$N_f(u) = f \circ u$$

is continuous.

1.5: Various other geometric conditions on the boundary $\partial\Omega$ more general than the Lipschitz condition (see Definition 1.5.1), which can be used for the development of the trace theory (as well as the embedding theorems), can be found in Adams [2, Chap. IV]. We mention that the first systematic study of traces was conducted by J.-L. Lions [274]. In fact it is Lions who first introduced the space $V_p(\mathrm{div}, \Omega)$ (see Lions [275, Sect. 2.4]). The generalized Green's identities in Proposition 1.5.14, Corollary 1.5.16 and Corollary 1.5.17 are due to Casas and Fernandez [105] and Kenmochi [231].

1.6: In the construction of the extension operator (see Theorem 1.6.1), we follow Evans and Gariepy [164] (see Sect. 4.4, Theorem 1). This method is based on reflections on smooth boundaries and has its origin in the work of Lichtenstein [270] and later of Hestenes [207] and Seeley [380]. There is an alternative approach due to Calderon [99] based on the Calderon–Zygmund theorem of singular integrals.

1.7: The Rellich–Kondrachov compact embedding result (see Theorem 1.7.4 and Theorem 1.7.6) originated in a lemma of Rellich [353] on the compactness in $L^2(\Omega)$ of a set bounded in the norm

$$\int_\Omega \left[u^2 + ||Du||^2_{\mathbb{R}^N}\right] dz.$$

It was proved specifically for Sobolev spaces $W^{1,p}(\Omega)$ $(1 \leqslant p < \infty)$ by Kondrachov [246]. As we will see several times, such compact embeddings are a basic tool in the spectral analysis of linear elliptic partial differential operators.

1.8: Apparently, the name of Poincaré is associated with this result because he used this kind of inequality in his work on tides. An extensive treatment of the Poincaré inequality can be found in the book of Ziemer [429].

Proposition 1.11.4 *(a) The Poincaré inequality does not hold on $W_0^{1,p}(\Omega)$, if Ω contains arbitrarily large balls, that is, if we can find $\{z_n\}_{n\geqslant 1} \subseteq \Omega$ and $\{r_n\}_{n\geqslant 1} \subseteq (0,\infty)$ such that $r_n \to \infty$ and $B_{r_n}(z_n) \subseteq \Omega$ for all $n \geqslant 1$.*

(b) If $V \subseteq W^{1,p}(\Omega)$ and $V \hookrightarrow L^p(\Omega)$ compactly, then the Poincaré inequality holds on V if and only if the constant function 1 does not belong to V.

Remark 1.11.5 If $\Omega_1 \subseteq \Omega_2$ and Poincaré's inequality holds for $W_0^{1,p}(\Omega_2)$ it also holds for $W_0^{1,p}(\Omega_1)$, since each function $u \in W_0^{1,p}(\Omega_1)$ can be extended by zero to a function in $W_0^{1,p}(\Omega_2)$.

In the next proposition, we provide an explicit value for the constant $c > 0$ in Theorem 1.8.3 (the Poincaré–Wirtinger inequality), see Gilbarg and Trudinger [187, p. 164] and Leoni [262, Sect. 12.2] and the papers of Acosta and Duran [1] (where $p = 1$), Bebendorf [40] and Chua and Wheeden [124]. In what follows for every $u \in L^1(\Omega)$ and every $D \subseteq \Omega$, we write

$$\bar{u}_D = \frac{1}{\lambda^N(D)} \int_D u(z)dz.$$

Also, for $z_0 \in \mathbb{R}^N$ and $r > 0$, we set

$$Q(z_0, r) = z_0 + \left(-\frac{r}{2}, \frac{r}{2}\right)^N.$$

Finally, we say that $D \subseteq \mathbb{R}^N$ is star-shaped with respect to $z_0 \in D$ if

$$ty + (1-t)z_0 \in D \text{ for all } t \in (0,1), \text{ all } y \in D.$$

Proposition 1.11.6 *(a) If $\Omega = R = \prod_{k=1}^{N}(0, b_k)$ (a rectangle) and $1 \leqslant p < \infty$ then $||u - \bar{u}_R||_p \leqslant N \ \max\{b_k\}_{k=1}^N \ ||Du||_p$ for all $u \in W^{1,p}(R)$.*

(b) If Ω is a bounded open convex set and $1 \leqslant p < \infty$, then there exists a constant $c = c(p, N) > 0$ such that

$$||u - \bar{u}_\Omega||_p \leqslant c \text{ (diam } \Omega) \, ||Du||_p \text{ for all } u \in W^{1,p}(\Omega).$$

(c) If Ω is an open star-shaped set with respect to $z_0 \in \Omega$ and

$$Q(z_0, 4r) \subseteq \Omega \subseteq B_R(z_0)$$

for some $r, R > 0$ *then there exists a constant* $c = c(N) > 0$ *such that*

$$||u - \bar{u}_\Omega||_p \leqslant cR^p \left(\frac{R}{r}\right)^{N-1} ||Du||_p \textit{ for all } u \in W^{1,p}(\Omega).$$

Remark 1.11.7 In (b) we can take $c = \frac{1}{4}$ when $p = 1$ and $c = \frac{1}{\pi^2}$ when $p = 2$. If $N = 1$, then these constants are optimal. Also in the scalar case $N = 1$, we have

$$||u||_\infty^2 \leqslant \frac{b}{12}||u'||_2^2 \text{ for all } u \in W^{1,2}(0,b) \text{ with } \int_0^b u(t)\,dt = 0$$

$$\text{and } ||u||_\infty^p \leqslant b^{p-1}||u'||_p^p \text{ for all } u \in W^{1,p}(0,b) \text{ with } \int_0^b u(t)\,dt = 0, \quad \text{all } 1 < p < \infty.$$

The constants are optimal.

For higher order Sobolev spaces, we have the following result.

Proposition 1.11.8 *Assume that* Ω *is bounded,* $1 \leqslant p < \infty$ *and* $0 \leqslant k \leqslant m$ *are integers. Then there exists a constant* $c = c(k, m, p, \operatorname{diam}\Omega)$ *such that*

$$||D^k u||_p \leqslant c\,||D^m u||_p \textit{ for all } u \in W_0^{m,p}(\Omega).$$

1.9: Theorem 1.9.3 is due to Sobolev [388] for $p > 1$ and due to Gagliardo [178] and Nirenberg [318] for $p = 1$. The best constant in this inequality (see Remark 1.9.4), was obtained independently by Aubin [23] and Talenti [399]. The case $N < p$ (see Theorem 1.9.10) is due to Morrey [305]. For the case $p = N$ (see Theorem 1.9.13), Trudinger [406] proved that

$$\int_\Omega \exp\left[|u|^{\frac{N}{N-1}}\right] dz < \infty \text{ for all } u \in W^{1,N}(\Omega).$$

1.10: A more detailed study of the p-capacity can be found in the books of Adams and Hedberg [4], Evans and Gariepy [164], Heinonen et al. [204] and Ziemer [429]. In these books the reader can find applications in the study of fine properties of functions. Friedrichs [173] proved that for Ω bounded, we have

$$||u||_2 \leqslant c\,\left[||Du||_2 + ||u||_{L^2(\partial\Omega)}\right] \text{ for some } c > 0, \text{ all } u \in H^1(\Omega)$$

from which the third equivalent norm in Proposition 1.10.12 follows at once for $p = 2$.

Chapter 2
Compact Operators and Operators of Monotone Type

What is now proved was once only imagined.
William Blake (1757–1827)

Compact maps constitute a class of maps to which we can extend many of the results which are valid for maps between finite-dimensional spaces. Compactness plays a central role in the infinite-dimensional extension of degree theory (Leray–Schauder degree, see Chap. 3) and in fixed point theory (see Chap. 4). The needs of problems in the calculus of variations and in nonlinear functional equations led to the class of operators of monotone type, which provides a broader framework than the class of compact maps.

2.1 Compact and Completely Continuous Maps

Definition 2.1.1 Let X, Y be Banach spaces, D a nonempty subset of X and $f : D \to Y$.

(a) We say that f is "compact" if it is continuous and maps bounded subsets of D into relatively compact subsets of Y. We denote the set of compact maps by $K(D, Y)$.
(b) We say that f is "completely continuous" if for every sequence $\{x_n\}_{n\geqslant 1} \subseteq D$ such that $x_n \xrightarrow{w} x$ in X with $x \in D$, we have $f(x_n) \to f(x)$ in Y. So, a completely continuous map is sequentially continuous from D with the relative weak topology into Y with the strong topology.

Remark 2.1.2 A word of caution. Although the above terminology seems to be the most popular among analysts, it is not universal. In the literature the names compact

N. S. Papageorgiou et al., *Nonlinear Analysis–Theory and Methods*,
Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-030-03430-6_2

and completely continuous are used to denote different things. So, the reader should check to see in which context each author uses these terms. In general, the two notions introduced in the above definition are distinct. Nevertheless, in some particular cases of interest, the two are related.

Proposition 2.1.3 *If X is reflexive, $D \subseteq X$ is nonempty, closed convex and $f : D \to Y$ is completely continuous, then $f : D \to Y$ is compact.*

Proof Evidently, $f(\cdot)$ is continuous. Let $B \subseteq D$ be bounded. We will show that $\overline{f(B)}$ is compact. So, let $\{y_n\}_{n\geqslant 1} \subseteq f(B)$. We will show that $\{y_n\}_{n\geqslant 1}$ admits a strongly convergent subsequence. We have $y_n = f(x_n)$ with $x_n \in B$ for all $n \geqslant 1$. Since X is reflexive and B is bounded, by passing to a subsequence if necessary, we may assume that $x_n \xrightarrow{w} x$ and $x \in D$ (recall that D is closed and convex, hence it is w-closed). So, the complete continuity of f implies $y_n = f(x_n) \to f(x) = y$ and proves the compactness of $f(\cdot)$. □

The set of linear elements of $K(X, Y)$ will be denoted by $\mathscr{L}_c(X, Y)$. So,

$$\mathscr{L}_c(X, Y) = \mathscr{L}(X, Y) \cap K(X, Y) .$$

As a direct consequence of Proposition 2.1.3, we have:

Proposition 2.1.4 *If X is reflexive, then $A \in \mathscr{L}_c(X, Y)$ if only if A is completely continuous.*

Clearly, $K(X, Y)$ is a vector space and it is closed under composition with a continuous, bounded map (recall that bounded means that it maps bounded sets to bounded sets).

Proposition 2.1.5 *Assume that $f_n, f : D \to Y$, f_n is compact and $f_n(x) \to f(x)$ in Y for all $x \in D$ and the convergence is uniform on bounded subsets of D. Then $f \in K(D, Y)$.*

Proof Let $B \subseteq D$ be nonempty and bounded. We need to show that $f(B)$ is totally bounded. By hypothesis, given $\epsilon > 0$, we can find $n_0 \geqslant 1$ such that

$$\|f(x) - f_{n_0}(x)\|_Y \leqslant \frac{\epsilon}{3} \text{ for all } x \in B. \tag{2.1}$$

Since f_{n_0} is compact, $f_{n_0}(B)$ is relative compact and so we can find $\{x_k\}_{k=1}^m \subseteq B$ such that

$$f_{n_0}(B) \subseteq \bigcup_{k=1}^{m} \overline{B}_{\frac{\epsilon}{3}}(f(x_k)) \ \left(\overline{B}_{\frac{\epsilon}{3}}(f(x_k)) = \{y \in Y : \|y - f(x_k)\|_Y \leqslant \frac{\epsilon}{3}\}\right). \tag{2.2}$$

Then for every $x \in B$, we can find $k \in \{1, 2, ...m\}$ such that

$$\begin{aligned}\|f(x) - f(x_k)\|_Y &\leqslant \|f(x) - f_{n_0}(x)\|_Y + \|f_{n_0}(x) - f_{n_0}(x_k)\|_Y \\ &\quad + \|f_{n_0}(x_k) - f(x_k)\|_Y \leqslant \varepsilon \text{ (see (2.1), (2.2))}\end{aligned}$$

$\Rightarrow f(B)$ is totally bounded, hence relatively compact in Y

$\Rightarrow f \in K(D, Y).$

The proof is now complete. □

By Definition 2.1.1, we see that if Y is finite-dimensional, then every continuous bounded map into Y is compact. This simple observation leads to the following definition.

Definition 2.1.6 Let X, Y be Banach spaces, $D \subseteq X$ a nonempty subset and $f : D \to Y$. We say that f is of "finite rank" if f is continuous, bounded and the range of f is a finite-dimensional subspace of Y. The set of finite rank maps from X into Y is denoted by $K_f(X, Y)$.

The next theorem, known in the literature as the "Schauder approximation theorem", explains why the family of compact maps is the right one to consider in order to extend results from the finite-dimensional theory.

Theorem 2.1.7 (Schauder) *If $D \subseteq X$ is bounded and $f : D \to Y$, then the following statements are equivalent:*

(a) $f \in K(D, Y)$;
(b) Given any $\epsilon > 0$, we can find $f_\epsilon \in K_f(D, Y)$ such that

$$\|f_\epsilon(x) - f(x)\|_Y \leqslant \varepsilon \textit{ for all } x \in D$$
$$\text{range } f_\varepsilon \subseteq \overline{\text{conv}}\, f(D)$$

Proof $(a) \Rightarrow (b)$ Since $f \in K(D, Y)$, $\overline{f(D)}$ is compact and so given $\epsilon > 0$, we can find $\{y_k\}_{k=1}^m \subseteq Y$ such that $\overline{f(D)} \subseteq \bigcup_{k=1}^m B_\epsilon(y_k)$ $((B_\epsilon(y_k) = \{y \in Y : \|y - y_k\|_Y < \epsilon\}))$.

Let $\{\varphi_k\}_{k=1}^m$ be a continuous partition of unity subordinate to this cover, that is, $\varphi_k \in C(Y)$, $\operatorname{supp} \varphi_k \subseteq B_\epsilon(y_k)$, $0 \leqslant \varphi_k \leqslant 1$ for all $k \in \{1, ..., m\}$ and $\sum_{k=1}^m \varphi_k(y) = 1$ for all $y \in f(D)$.

Let $f_\epsilon(x) = \sum_{k=1}^m \varphi_k(f(x))y_k$ for all $x \in D$. If $\varphi_k(f(x)) > 0$, then $\varphi_k(f(x))y_k \in B_\epsilon(y_\epsilon)$ and

$$\begin{aligned}\|f_\epsilon(x) - f(x)\|_Y &= \|\sum_{k=1}^m \varphi_k(f(x))y_k - f(x)\|_Y \\ &= \|\sum_{k=1}^m \varphi_k(f(x))(y_k - f(x))\|_Y \text{ (since } \sum_{k=1}^m \varphi_k(f(x)) = 1) \\ &\leqslant \epsilon \text{ (since } f(x) \in B_\epsilon(y_k)).\end{aligned}$$

Evidently, range $f_\epsilon \subseteq \overline{\text{conv}}\, f(D)$.

$(b) \Rightarrow (a)$ According to statement (b), we can find $\{f_n\}_{n\geqslant 1} \subseteq K_f(D, Y)$ such that

$$f_n \to f \text{ uniformly on } D.$$

Since $K_f(D, Y) \subseteq K(D, Y)$, we conclude from Proposition 2.1.5 that $f \in K(D, Y)$. $\square$

Proposition 2.1.8 *If $D \subseteq X$ is bounded and $f \in K(D, Y)$, then $f(x) = \sum_{n\geqslant 0} h_n(x)$ for all $x \in D$, with $h_n \in K_f(D, Y)$ for all $n \geqslant 0$ and $\|h_n(x)\|_Y \leqslant \frac{\epsilon}{2^n}$ for all $x \in D$ and all $n \geqslant 1$, and $\|h_0(x) - f(x)\|_Y \leqslant \frac{\epsilon}{4}$ for all $x \in D$.*

Proof According to Theorem 2.1.7, we can find $f_n \in K_f(D, Y)$ such that

$$\|f_n(x) - f(x)\|_Y \leqslant \frac{\epsilon}{2^{n+2}} \text{ for all } x \in D \text{ and all } n \geqslant 0.$$

Inductively, we define $\{h_n\}_{n\geqslant 0}$ by

$$h_0 = f_0,\ h_n = f_n - f_{n-1},\ n \geqslant 1.$$

Then $f_n = \sum_{k=1}^n h_k$ and $f_n = \sum_{k=1}^n h_k \to f$ on D. Also

$$\begin{aligned}
\|h_n(x)\|_Y &= \|f_n(x) - f_{n-1}(x)\|_Y \leqslant \|f_n(x) - f(x)\|_Y + \|f_{n-1}(x) - f(x)\|_Y \\
&\leqslant \frac{\epsilon}{2^{n+2}} + \frac{\epsilon}{2^{n+1}} < \frac{\epsilon}{2^n} \\
&\Rightarrow \sum_{n\geqslant 0} \|h_n(x)\|_Y \text{ is uniformly convergent on} D \\
&\Rightarrow \sum_{n\geqslant 0} h_n \to f \text{ uniformly on} D.
\end{aligned}$$

The proof is now complete. $\square$

Compact maps admit compact extensions. This can be established using the following extension result, known in the literature as the "Dugundji extension theorem".

Proposition 2.1.9 *If (X, d) is a metric space, Y is a normed space, $D \subseteq X$ is a nonempty closed and $f \in C(D, Y)$, then there exists an $\hat{f} \in C(X, Y)$ such that $\hat{f}\,|_D = f$ and $\hat{f}(X) \subseteq \operatorname{conv} f(D)$.*

Proof For each $x \in X \setminus D$ let B_x be an open ball with diam $B_x < d(B_x, D)$. Then $\{B_x\}_{x\in X\setminus D}$ is an open cover of $X \setminus D$. We can find a locally finite, continuous partition of unity $\{\varphi_i\}_{i\in J}$. Therefore

$$\varphi_i \in C(X \setminus D),\ \operatorname{supp} \varphi_i \subseteq B_{x(i)}, \sum_{i\in J} \varphi_i(x) = 1 \text{ for all } x \in X \setminus D \text{ and} \tag{2.3}$$

each $x \in X \setminus D$ has a neighborhood $V(x)$ such that all but finitely many φ_i's are zero on $V(x)$.

For each $u \in X \setminus D$, we choose $h_u \in D$ such that $d(h_u, B_u) < 2d(B_u, D)$ and define

$$\hat{f}(x) = \begin{cases} f(x) & \text{if } x \in D \\ \sum_{i \in J} \varphi_i(x) f(h_{u(i)}) & \text{if } x \in X \setminus D. \end{cases} \tag{2.4}$$

From (2.3) we see that $\hat{f}$ is well defined and $\hat{f}(x) \subseteq \operatorname{conv} f(D)$ for all $x \in X$.

Note that $\hat{f}(\cdot)$ is continuous in $\operatorname{int} D$ (may be empty) and in $X \setminus D$. It remains to show the continuity of $\hat{f}$ on ∂D. Let $x_0 \in \partial D$. Then $f(x_0) = \hat{f}(x_0)$ (see (2.4)). If $x \in X \setminus D$ and $\varphi_i(x) \neq 0$, then $x \in B_u$, where we set $u = u(i)$. Using the triangle inequality, we have

$$\begin{aligned} & d(h_u, x) \leqslant d(h_u, B_u) + \operatorname{diam} B_u \leqslant 3d(B_u, D) \leqslant 3d(x, x_0) \\ \Rightarrow\ & d(h_u, x_0) \leqslant d(h_u, x) + d(x, x_0) \leqslant 4d(x, x_0). \end{aligned}$$

Since $0 \leqslant \varphi_i(x) \leqslant$ and $\varphi_i(x_0) = 0$ for $d(h_{u(i)}, x_0) > 4d(x, x_0)$, we have

$$\begin{aligned} ||\hat{f}(x) - \hat{f}(x_0)||_Y &= ||\sum_{i \in J} \varphi_i(x)(f(h_{u(i)}) - f(x_0))||_Y \\ &\leqslant \sup\left\{||f(h_{u(i)}) - f(x_0)||_Y : i \in J,\ d(h_{u(i)}, x_0) \leqslant 4d(x, x_0)\right\}. \end{aligned}$$

If $\{x_n\}_{n \geqslant 1} \subseteq X \setminus D$ and $x_n \to x_0$ in (X, d), then the continuity of f implies

$$\begin{aligned} & \sup\{\|f(h_{u(i)}) - f(x_0)\|_Y : i \in J,\ d(h_{u(i)}, x_0) \leqslant 4d(x_n, x_0)\} \to 0 \\ \Rightarrow\ & \hat{f}(x_n) \to \hat{f}(x_0) \text{ in } Y. \end{aligned}$$

If $\{x_n\}_{n \geqslant 1} \subseteq D$ and $x_n \to x_0$, then $f(x_n) = \hat{f}(x_n)$ and so $\hat{f}(x_n) \to \hat{f}(x_0)$ in Y. This proves the continuity of $\hat{f}$. ☐

This leads to the following existence result for compact maps. Now we return to the setting of Definition 2.1.1 with X, Y being Banach spaces.

Proposition 2.1.10 *If $D \subseteq X$ is closed and bounded and $f \in K(D, Y)$, then there exists an $\hat{f} \in K(X, Y)$ such that $\hat{f}|_D = f$ and $\hat{f}(X) \subseteq \operatorname{conv} f(D)$.*

Proof By virtue of Proposition 2.1.9 we already have a continuous extension $\hat{f}$ of f. We know that $\hat{f}(X) \subseteq \operatorname{conv} f(D)$ and $f(D)$ is relatively compact (since $f \in K(D, Y)$) and hence so is $\operatorname{conv} f(D)$. So, we conclude that $\hat{f} \in K(X, Y)$. ☐

We also have an extension theorem without assuming that $D \subseteq X$ is bounded.

Proposition 2.1.11 *If $D \subseteq X$ is closed, $f \in K(D, Y)$ and $\delta > 0$, then there exists an $\hat{f}_\delta \in K(X, Y)$ such that $\hat{f}_\delta|_D = f$ and $d(\hat{f}_\delta(x), \overline{\operatorname{conv}}\, f(D)) \leqslant \delta$ for all $x \in X$.*

Proof By Proposition 2.1.8 we know that $f = \sum_{n\geqslant 0} h_n$ with $h_n \in K_f(D, Y)$ for all $n \geqslant 0$ and $\|h_n(x)\|_Y \leqslant \frac{\epsilon}{2^n}$ for all $n \geqslant 1$. By the Tietze extension theorem, we can find $\hat{h}_n \in C(X, Y_n)$ with $\dim Y_n < \infty$, $\hat{h}_n|_D = h_n$ for all $n \geqslant 0$ and $\|\hat{h}_n(x)\|_Y \leqslant \frac{\epsilon}{2^n}$ for all $x \in D$ and all $n \geqslant 0$, and $\|\hat{h}(x)\|_Y \leqslant \|f(x)\|_Y + \frac{\epsilon}{4}$ for all $x \in D$. So, for each $n \geqslant 0$ $\hat{h}_n$ is compact. We set

$$\hat{f}(x) = \sum_{n\geqslant 0} \hat{h}_n(x) \text{ for all } x \in X. \tag{2.5}$$

We have

$$d(\hat{h}_0(x), \overline{\text{conv}}\, f(D)) \leqslant \|\hat{h}_0(x) - f(x)\|_Y \leqslant \sum_{n\geqslant 0} \frac{\epsilon}{2^n} = \epsilon. \tag{2.6}$$

Also, for $x \in \text{conv}\, h_0(D)$, we have

$$x = \sum_{k=1}^{m} \vartheta_k h_0(x_k) \text{ with } \vartheta_k \in [0, 1], x_k \in D$$
$$\Rightarrow \|x - \sum_{k=1}^{m} \vartheta_k f(x_k)\|_Y \leqslant \sum_{k=1}^{m} \vartheta_k \|h_0(x_k) - f(x_k)\|_Y \leqslant \epsilon.$$

Choosing $\epsilon = \frac{\delta}{3}$, then $\hat{f}$ is the desired extension, since $\hat{f}$ is compact (see (2.5)) and

$$d(\hat{f}(x), \overline{\text{conv}}\, f(D)) \leqslant \|\hat{f}(x) - \hat{h}_0(x)\|_Y + d(\hat{h}_0(x), \overline{\text{conv}}\, f(D))$$
$$\leqslant \sum_{k\geqslant 1} \frac{\epsilon}{2^n} + \epsilon = 2\epsilon < \delta \text{ (see (2.5), (2.6))}.$$

The proof is now complete. ☐

Proposition 2.1.12 *If $D \subseteq X$ is open and $f \in K(D, Y)$ is Fréchet differentiable in D, then $f'(x) \in \mathscr{L}_c(X, Y)$ for all $x \in D$.*

Proof Let $x \in D$. We have

$$f(x + h) = f(x) + f'(x)h + w(x; h), \tag{2.7}$$

where $\frac{\|w(x;h)\|_Y}{\|h\|_X} \to 0$ as $\|h\|_X \to 0$. So, given $\epsilon > 0$, we can find $\delta = \delta(\epsilon, x) > 0$ such that

$$\|w(x; h)\|_Y \leqslant \epsilon\|h\|_X \text{ for all } \|h\|_X \leqslant \delta. \tag{2.8}$$

Then set $\overline{B}_1 = \{h \in X : \|h\|_X \leqslant 1\}$ and $\overline{B}_1(x) = x + \overline{B}_1$. We have

$$\begin{aligned} & \delta f'(x)\overline{B}_1 \subseteq -f(x) + f(\delta\overline{B}_1) + \delta\epsilon\overline{B}_1 \text{ (see (2.7), (2.8))} \\ \Rightarrow \quad & f'(x)\overline{B_1} \subseteq -\frac{1}{\delta} f(x) + \frac{1}{\delta} f(\delta\overline{B}_1) + \epsilon\overline{B}_1. \end{aligned} \tag{2.9}$$

Since $\frac{1}{\delta} f(\delta\overline{B}_1) - \frac{1}{\delta} f(x)$ is relatively compact, it follows by (2.9) that $f'(x)\overline{B}_1$ is relative compact and so $f'(x) \in \mathscr{L}_c(X, Y)$. □

2.2 Proper Maps and Gradient Maps

It is well known that the direct image of a compact set by a continuous map is compact. In this section we study continuous maps whose inverse image of a compact set is compact. Clearly not every continuous map has this property. The importance of this class of continuous maps comes from fact that if $f : X \to Y$ is such a map, then the solution set of the functional equation $f(x) = y$ is compact.

Definition 2.2.1 Let X, Y be metric spaces and $f \in C(X, Y)$. We say that f is "proper" if for any compact $K \subseteq Y$, $f^{-1}(K)$ is compact in X.

Remark 2.2.2 It is immediately clear from the definition that if X, Y are normed spaces and $L \in \mathscr{L}(X, Y)$ is proper, then L has closed range.

In the next proposition, we present alternative equivalent definitions of properness.

Proposition 2.2.3 *If X, Y are metric spaces and $f \in C(X, Y)$, then the following statements are equivalent:*

(a) f is proper;
(b) for every $y \in Y$, the set $f^{-1}(y) \subseteq X$ is compact and f is closed (that is, it maps closed sets to closed sets);
(c) every sequence $\{x_n\}_{n\geqslant 1} \subseteq X$ such that $f(x_n) \to y$ in Y has a subsequence $\{x_{n_k}\}_{k\geqslant 1} \subseteq X$ that converges to $x \in X$.

Proof $(a) \Rightarrow (b)$ Since $K = \{y\} \subseteq Y$ is compact, we have that $f^{-1}(K) = f^{-1}(y) \subseteq X$ is compact. We need to show that f is closed. So, let $C \subseteq X$ be closed and let $\{y_n\}_{n\geqslant 1} \subseteq f(C)$ such that $y_n \to y$ in Y with $y \in Y$. We have $y_n = f(x_n)$ with $x_n \in C$ for all $n \geqslant 1$ and $K = \{y_n, y\}_{n\geqslant 1} \subseteq X$ is compact. Then $f^{-1}(K)$ is compact in X and $\{x_n\}_{n\geqslant 1} \subseteq f^{-1}(K)$. So, by passing to a subsequence if necessary, we may assume that $x_n \to x$ in X with $x \in C$ (recall that C is closed in X). The continuity of f implies $y = f(x)$ and so $y \in f(C)$, which proves that f is closed.

$(b) \Rightarrow (c)$ Consider a sequence $\{x_n\}_{n\geqslant 1} \subseteq X$ such that $f(x_n) \to y$ in Y with $y \in Y$. We will show that the sequence $\{x_n\}_{n\geqslant 1}$ has a cluster point. So, let $D_m = \{x_n\}_{n\geqslant m}$. Since f is by hypothesis closed, we have $f(\overline{D_m}) = \overline{f(D_m)}$ for all $m \geqslant 1$. Also, because $f(x_n) \to y$, we have

$$\{y\} = \bigcap_{m\geqslant 1} \overline{f(D_m)} = \bigcap_{m\geqslant 1} f(\overline{D_m}).$$

If $x \in C = f^{-1}(y)$ then

$$\begin{aligned} & f(x) \in \bigcap_{m\geqslant 1} f(\overline{D_m}) \\ \Rightarrow \quad & E_m = C \cap \overline{D_m} \text{ is nonempty and closed for every } m \geqslant 1. \end{aligned}$$

But C is compact and the family of closed subsets $\{E_m\}_{m\geqslant 1}$ of C has the finite intersection property. Therefore $\cap_{m\geqslant 1} E_m = C \cap \left(\cap_{m\geqslant 1}\overline{D_m}\right) \neq \emptyset$ and so we conclude that $\{x_n\}_{n\geqslant 1}$ has a cluster point.

$(c) \Rightarrow (a)$ Let $K \subseteq Y$ be compact and let $\{x_n\}_{n\geqslant 1} \subseteq f^{-1}(K)$. It follows that $\{f(x_n)\}_{n\geqslant 1} \subseteq K$ and so we may assume that $f(x_n) \to y \in K$. Then by hypothesis, we can find a subsequence $\{x_{n_k}\}_{k\geqslant 1}$ of $\{x_n\}_{n\geqslant 1}$ such that $x_{n_k} \to x \in X$. Evidently, $f(x_{n_k}) \to f(x) = y \in K$ and so $x \in f^{-1}(K)$ which proves that $f^{-1}(K)$ is compact and so f is proper. □

If X, Y are finite-dimensional normed spaces, then properness is equivalent to coercivity.

Proposition 2.2.4 *If X, Y are finite-dimensional normed spaces and $f \in C(X, Y)$, then the following properties are equivalent:*

(a) f is proper;
(b) f is coercive (that is, $\|f(x)\|_Y \to +\infty$ when $\|x\|_X \to \infty$).

Proof $(a) \Rightarrow (b)$ Recall that in finite-dimensional normed spaces, bounded sets are relatively compact. Also, since $f \in C(X, Y)$, for every $B \subseteq Y$ we have $\overline{f^{-1}(B)} \subseteq f^{-1}(\overline{B})$. Therefore the properness of f implies that the inverse image of a bounded set in Y is a bounded set in X. This is a restatement of the property of coercivity.

$(b) \Rightarrow (a)$ Let $K \subseteq Y$ be compact. Then K is bounded and so the coercivity of f implies that $f^{-1}(K)$ is bounded, hence relatively compact in X. Note that $\overline{f^{-1}(K)} \subseteq f^{-1}(\overline{K}) = f^{-1}(K)$ (since $f \in C(X, Y)$) and so $f^{-1}(K)$ is closed, hence compact in X. This proves that f is proper. □

If the normed spaces X, Y are infinite-dimensional, then the above equivalence is no longer true. Nevertheless, in some particular cases we can still have some relation between coercivity and properness.

Proposition 2.2.5 *If X, Y are normed spaces, $f \in C(X, Y)$ is coercive and $f = g + h$ with g proper and h compact, then f is proper.*

Proof Let $\{x_n\}_{n\geqslant 1} \subseteq X$ such that $y_n = f(x_n) \to y \in Y$. The coercivity of f implies that $\{x_n\}_{n\geqslant 1} \subseteq X$ is bounded. Then the compactness of h implies that $\{h(x_n)\}_{n\geqslant 1}$ is relatively compact in Y. So, we may assume that $h(x_n) \to \overline{y} \in Y$. We have $g(x_n) = y_n - h(x_n) \to y - \hat{y}$ in Y. Since g is proper, from Proposition 2.2.3 we see that we may assume that $x_n \to x$ in X. The continuity of g implies $g(x_n) \to g(x)$ in Y. So $g(x) = y - \hat{y}$ and $f(x) = y$, which proves the properness of f (see Proposition 2.2.3). $\square$

Proposition 2.2.6 *If X is a reflexive Banach space, Y is a normed space, $f \in C(X, Y)$ and is coercive and $x_n \xrightarrow{w} x$ in X with $\{f(x_n)\}_{n\geqslant 1}$ strongly convergent in Y imply $x_n \to x$ in X, then f is proper.*

Proof Let $\{x_n\}_{n\geqslant 1}$ such that $f(x_n) \to y$ in Y. The coercivity of f implies that $\{x_n\}_{n\geqslant 1} \subseteq X$ is bounded. The reflexivity of X implies that we may assume that $x_n \xrightarrow{w} x$ in X. Then by hypothesis we have $x_n \to x$ in X. Therefore by virtue of Proposition 2.2.3, f is proper. $\square$

Proposition 2.2.7 *If X is a compact metric space, Y is a metric space and $f \in C(X, Y)$, then f is proper.*

Proof The continuity of f implies for all $y \in Y$, $f^{-1}(y)$ is closed in X, hence compact (recall that X is compact). Also, if $C \subseteq X$ is closed, it is compact. Hence $f(C)$ is compact in Y, in particular then closed too. So, we have proved that f is closed. By virtue of Proposition 2.2.3, f is proper. $\square$

Proposition 2.2.8 *If X, Y are metric spaces and $f \in C(X, Y)$ is injective, then the following statements are equivalent:*

(a) *f is proper;*
(b) *f is closed;*
(c) *f is bicontinuous from X onto $f(X)$.*

Proof $(a) \Rightarrow (b)$ See Proposition 2.2.3.

$(b) \Rightarrow (c)$ Since f is closed, $f(X)$ is closed in Y. Let $C \subseteq X$ be closed. Then $f(C) = (f^{-1})^{-1}(C)$ (since f is bijective from X onto $f(X)$) is closed in Y and so f^{-1} is continuous.

$(c) \Rightarrow (a)$ See Proposition 2.2.3. $\square$

The next proposition is an easy consequence of Definition 2.2.1 and Proposition 2.2.3.

Proposition 2.2.9 *If X,Y,Z are metric spaces and $f \in C(X, Y)$, $g \in C(Y, Z)$, then*

(a) *if f and g are proper, then $g \circ f$ is proper;*
(b) *if $g \circ f$ is proper, then f is proper;*
(c) *if $g \circ f$ is proper and f is surjective, then g is proper.*

Proper maps $f \in C(X, Y)$ exhibit a stability property of the solution set

$$S(f, y) = \{x \in X : f(x) = y\}$$

as f and y vary.

Proposition 2.2.10 *If X, Y are metric spaces and $f \in C(X, Y)$ is proper, then*

(a) for every $y \in Y$ and every $\epsilon > 0$, there exists a $\delta > 0$ such that

$$d_Y(f(x), y) \leqslant \delta \Rightarrow d_X(x, S(f, y)) \leqslant \epsilon;$$

(b) for every $\epsilon > 0$, there exists a $\delta > 0 >$ such that

$$g \in C(X, Y) \text{ and } d_Y(g(x), f(x)) \leqslant \delta \text{ for all } x \in X \Rightarrow d_X(S(g, y), S(f, y)) \leqslant \epsilon.$$

Proof (a) Arguing by contradiction, suppose that the statement is not true. Then we can find $\epsilon > 0$, $y \in Y$ and $\{x_n\}_{n \geqslant 1} \subseteq X$ such that

$$d_Y(f(x_n), y) \leqslant \frac{1}{n} \text{ and } d_X(x_n, S(f, y)) \geqslant \epsilon \text{ for all } n \geqslant 1. \tag{2.10}$$

Note that $f(x_n) \to y$ in Y. Then the properness of f and Proposition 2.2.3 imply that, passing to a suitable subsequence if necessary, we may assume that $x_n \to x$ in X with $x \in X$. The continuity of f implies $f(x) = y$, hence $x \in S(f, y)$. On the other hand, from (2.10), we have $d_X(x, S(f, y)) \geqslant \epsilon$, a contradiction.

(b) This is an immediate consequence of (2.10). ☐

Now we turn our attention to gradient maps.

Definition 2.2.11 Let X be a Banach space, X^* its topological dual and $D \subseteq X$ an open set. We say that $f \in C(D, X^*)$ is a "gradient map" if there is a function $F \in C^1(D)$ such that $F'(x) = f(x)$ for all $x \in D$. We say that F is the "potential" of f and we write $f(x) = \nabla F(x) = (\text{grad } F)(x)$.

The next example is important in the study of boundary value problems.

Example 2.2.12 Let $\Omega \subseteq \mathbb{R}^N$ be an open set and consider a Carathéodory function $g : \Omega \times \mathbb{R} \to \mathbb{R}$ (that is, for all $x \in \mathbb{R}$ the mapping $z \mapsto g(z, x)$ is measurable and for a.a. $z \in \Omega$ the mapping $x \mapsto g(z, x)$ is continuous). Suppose that

$$|g(z, x)| \leqslant \alpha(z) + c|x|^{p-1} \text{ for a.a. } z \in \Omega, \text{ all, } x \in \mathbb{R}, \tag{2.11}$$

where $\alpha \in L^{p'}(\Omega)$ $(\frac{1}{p} + \frac{1}{p'} = 1)$ and $c > 0$. Let $G(z, x) = \int_0^x g(z, x) ds$ and consider the functional $F : L^p(\Omega) \to \mathbb{R}$ defined by

$$F(u) = \int_\Omega G(z, u(z)) dz \text{ for all } u \in L^p(\Omega).$$

We will show that F is the potential of the Nemytskii operator $N_g : L^p(\Omega) \to L^{p'}(\Omega)$ defined by

$$N_g(u)(\cdot) = g(\cdot, u(\cdot)) \text{ for all } u \in L^p(\Omega).$$

So, we will show that $F \in C^1(L^p(\Omega))$ and $\nabla F(u) = N_g(u)$ for all $u \in L^p(\Omega)$. To this end, for every $h \in L^p(\Omega)$, we have

$$\begin{aligned}
&G(z,(u+h)(z)) - G(z,u(z)) = \int_0^1 \frac{d}{dt}G(z,(u+th)(z))dt = \\
&\int_0^1 g(z,(u+th)(z))h(z)dt \\
\Rightarrow\ & F(u+h) - F(u) = \int_\Omega \int_0^1 g(z,(u+th)(z))h(z)dtdz \\
\Rightarrow\ & F(u+h) - F(u) - \int_\Omega g(z,u(z))h(z)dz \\
&\qquad = \int_\Omega \int_0^1 [g(z,(u+th)(z)) - g(z,u(z))]h(z)dtdz.
\end{aligned} \tag{2.12}$$

We set

$$w(h) = F(u+h) - F(u) - \int_\Omega g(z,u(z))h(z)dz \text{ for all } h \in L^p(\Omega).$$

From (2.12) and using the Fubini theorem and Hölder's inequality, we have

$$\begin{aligned}
|w(h)| &\leqslant \int_\Omega \int_0^1 |g(z+(u+th)(z)) - g(z,u(z))||h(z)|dtdz \\
&\leqslant \int_0^1 \|N_g(u+th) - N_g(u)\|_{p'} dt\, \|h\|_p .
\end{aligned} \tag{2.13}$$

Exploiting the continuity of the Nemytskii operator and invoking the Lebesgue dominated convergence theorem (see (2.11)), we have

$$\begin{aligned}
&\int_0^1 \|N_g(u+th) - N_g(u)\|_{p'} dt \to 0 \text{ as } \|h\|_p \to 0 \\
\Rightarrow\ & \frac{|w(h)|}{\|h\|_p} \to 0 \text{ as } ||h||_p \to 0 \text{ (see (2.13))}.
\end{aligned}$$

Therefore we conclude that

$$F \in C^1(L^p(\Omega)) \text{ and } \nabla F(u) = N_g(u) \text{ for all } u \in L^p(\Omega).$$

Gradient operators are actually generalizations of self-adjoint operators.

Proposition 2.2.13 *If X is reflexive Banach space, X^* its topological dual, $\langle \cdot, \cdot \rangle$ the duality brackets for the pair (X^*, X), $D \subseteq X$ is open, convex and contains the origin and $f \in C^1(D, X^*)$, then the following statements are equivalent:*

(a) *f is a gradient map;*
(b) *$\int_\gamma f ds$ is independent of the path γ provided γ is simple and rectifiable in D;*
(c) *$\int_0^1 \langle f(tx), x\rangle dt - \int_0^1 \langle f(tu), u\rangle dt = \int_0^1 \langle f(c(t)), x-u\rangle dt$, where $c(t) = tx + (1-t)u$ for all $t \in [0,1]$ and with $x, u \in D$.*
(d) *$f'(x) \in \mathscr{L}(X, X^*)$ is self-adjoint for all $x \in D$.*

Proof $(a) \Rightarrow (b)$ Since f is a gradient map, we can find $F \in C^1(D)$ such that $f(x) = F'(x)$ for all $x \in D$. Let γ be a simple, rectifiable curve in D and suppose that γ has a parametric representation $\gamma = \{c(t) : 0 \leqslant t \leqslant 1\}$. We have

$$\begin{aligned}\textstyle\int_\gamma f ds &= \int_0^1 f(c(t))c'(t)\,dt = \int_0^1 F'(c(t))c'(t)dt\\ &= \int_0^1 \frac{d}{dt}F(c(t))dt\\ &= F(c(1)) - F(c(0)).\end{aligned}$$

$(b) \Rightarrow (c)$ Let γ_1 be the path in D obtained by joining linearly the origin with x and u. Also, let γ_2 be the path obtained by joining linearly directly x and u. The curve γ_1 contains two line segment joining $\{0, x\}$ and $\{0, u\}$ and so it is parametrized as follows:

$$\gamma_1 : \gamma_1 = {\gamma_1}^x \cup {\gamma_1}^u \text{ with } {\gamma_1}^x : c_1(t) = tx \text{ and } {\gamma_2}^u : c_2(t) = tu \text{ for all } t \in [0,1].$$

Hence, as above, we have

$$\begin{aligned}\int_{\gamma_1} f ds &= \int_{\gamma_1^*} f ds - \int_{\gamma_1^u} f ds = \int_0^1 \frac{d}{dt}F(tx)dt - \int_0^1 \frac{d}{dt}F(tu)dt\\ &= \int_0^1 \langle f(tx), x\rangle dt - \int_0^1 \langle f(tu), u\rangle dt.\end{aligned} \tag{2.14}$$

The parametric representation for γ_2 is:

$$\gamma_2 : c(t) = tx + (1-t)u \text{ for all } t \in [0,1].$$

Then

$$\int_{\gamma_2} f ds = \int_0^1 \frac{d}{dt}F(c(t))dt = \int_0^1 \langle f(c(t)), x-u\rangle dt. \tag{2.15}$$

From (2.14), (2.15) and (b), we get the desired equality.
$(c) \Rightarrow (d)$ Note that (c) implies

$$F(x+\lambda h) - F(x) = \lambda \int_0^1 \langle f(x+t\lambda h), h\rangle dt.$$

Letting $\lambda \to 0^+$ we see that the Gâteaux derivative of F is $F'(x)(h) = \langle f(x), h \rangle$. Since $f \in C^1(D, X^*)$, we see that $F \in C^2(D)$ and

$$\begin{aligned}
&F''(x)(h_1, h_2) = \langle f'(x)h_1, h_2 \rangle = \langle f'(x)h_2, h_1 \rangle \text{ for all } h_1, h_2 \in X \\
&\Rightarrow f'(x) \in \mathscr{L}(X, X^*) \text{ is self-adjoint.}
\end{aligned}$$

$(d) \Rightarrow (a)$ Let $F(x) = \int_0^1 \langle f(tx), x \rangle dt$. We have

$$F(x+h) - F(x) = \int_0^1 \langle f(tx+th), h \rangle dt + \int_0^1 \langle f(tx+th) - f(tx), x \rangle dt. \tag{2.16}$$

Then

$$\begin{aligned}
&\int_0^1 \langle f(tx+th) - f(tx), x \rangle dt \\
&= \int_0^1 \left[\int_0^t \frac{d}{ds} \langle f(tx+sh), x \rangle ds \right] dt \\
&= \int_0^1 \int_0^t \langle f'(tx+sh)x, h \rangle ds dt \text{ (see (d))} \\
&= \int_0^1 \int_s^1 \langle f'(tx+sh)x, h \rangle dt ds \text{ (Fubini's theorem)} \\
&= \int_0^1 \langle f(x+sh) - f(sx+sh), h \rangle ds \\
&\Rightarrow F(x+h) - F(x) = \int_0^1 \langle f(x+sh), h \rangle ds \text{ (see (2.16))} \\
&\Rightarrow \frac{1}{\|x\|_X} |F(x+h) - F(x) - \langle f(x), h \rangle| \to 0 \text{ as } \|h\|_X \to 0 \\
&\Rightarrow F'(x) = f(x) \text{ for all } x \in D.
\end{aligned}$$

The proof is now complete. □

2.3 Linear Compact Operators

In this section we focus on linear operators which are compact. So, let X and Y be Banach spaces. We consider the space $\mathscr{L}_c(X, Y)$, that is, the bounded linear operators which are compact. In the case of linear operators, the definition of compactness (see Definition 2.2.1) takes the following simple form.

Definition 2.3.1 Let X, Y be Banach spaces and $A : X \to Y$ a linear operator. We say that A is "compact" if it maps the closed unit ball $\overline{B_1} = \{x \in X : \|x\|_X \leqslant 1\}$ onto a relatively compact set. The space of such operators is denoted by $\mathscr{L}_c(X, Y)$ and if $X = Y$, then we write $\mathscr{L}_c(X)$.

Remark 2.3.2 Of course, not every linear operator is compact. For example, let X be an infinite-dimensional Banach space and let I be the identity operator on X. This operator is continuous but not compact.

Proposition 2.3.3 *If X, Y are Banach spaces, then $\mathscr{L}_c(X, Y)$ is a closed linear subspace of $\mathscr{L}(X, Y)$.*

Proof This is an immediate consequence of Proposition 2.1.5 □

The next proposition says that $\mathscr{L}_c(X)$ is closed two sided ideal of the Banach algebra $\mathscr{L}(X)$. Its proof is a straightforward consequence of Definition 2.1.1(a).

Proposition 2.3.4 *If X is a Banach space, $K \in \mathscr{L}_c(X)$ and $A \in \mathscr{L}(X)$, then $KA \in \mathscr{L}_c(X)$ and $AK \in \mathscr{L}_c(X)$.*

This result has an interesting consequence concerning the continuity of inverses.

Corollary 2.3.5 *If X is an infinite-dimensional Banach space and $K \in \mathscr{L}_c(X)$ such that K^{-1} exists, then K^{-1} is not continuous.*

Proof Suppose that $K^{-1} \in \mathscr{L}(X)$. Then by Proposition 2.3.3, we obtain that $I = K^{-1}K \in \mathscr{L}_c(X)$, which contradicts the infinite-dimensionality of X. □

Next, we show that compactness is a property passed to the adjoint operator.

Theorem 2.3.6 (Schauder) *If X, Y are Banach spaces and $K \in \mathscr{L}_c(X, Y)$, then $K^* \in \mathscr{L}_c(Y^*, X^*)$.*

Proof Let $\overline{B}_1^X = \{x \in X : \|x\|_X \leqslant 1\}$ (the closed unit ball in X). Since $K \in \mathscr{L}_c(X, Y)$, we have that $\overline{K(B_1^X)} \subseteq Y$ is compact.

Let $\{y_n^*\}_{n\geqslant 1} \subseteq \overline{B}_1^{Y^*} = \{y^* \in Y^* : ||y^*||_{Y^*} \leqslant 1\}$. Let $\hat{y}_n^* = y_n^* \mid_{K(\overline{B_1}^X)}, n \geqslant 1$. Then $\{\hat{y}^*\}_n \subseteq C(\overline{K(B_1{}^X)})$. For all $y_1, y_2 \in \overline{K(B_1{}^X)}$ we have

$$|\hat{y}_n^*(y_1) - \hat{y}_n^*(y_2)| = |\langle y_n^*, y_1 - y_2\rangle_{Y^*,Y}| \leqslant \|y_1 - y_2\|_Y$$

(recall that $\|y_n^*\|_{Y^*} \leqslant 1$ for all $n \geqslant 1$).

So the sequence $\{\hat{y}_n^*\}_{n\geqslant 1} \subseteq C(\overline{K(B_1^X)})$ is equicontinuous and equibounded. Invoking the Arzela–Ascoli theorem, we infer that $\{\hat{y}_n^*\}_{n\geqslant 1}$ is relatively compact in $C(K(\overline{B_1}^X))$. Hence we can find a subsequence $\{y_{n_k}^*\}_{k\geqslant 1}$ of $\{y_n^*\}_{n\geqslant 1}$ such that $\{y_{n_k}^*\}_{k\geqslant 1}$ is convergent in $C(K(\overline{B_1}^X))$. We have

$$\begin{aligned}
&\sup\{\langle y_{n_k}^* - y_{n_m}^*, K(x)\rangle : x \in \overline{B_1}^X\} = \|K^*(y_{n_k}^*) - K^*(y_{n_m}^*)\|_{X^*} \\
&\qquad\qquad\qquad\qquad\text{for all } k, m \geqslant 1 \\
&\Rightarrow\ \{K^*(y_{n_k}^*)\}_{k\geqslant 1} \text{ is Cauchy in } X^*,\ \text{hence convergent in } X^*.
\end{aligned}$$

This implies that $K^* \in \mathscr{L}_c(Y^*, X^*)$. □

In Theorem 2.1.7 we saw that every compact map can be approximated uniformly on bounded sets by finite rank maps. Suppose that the original compact map is linear (that is, it belongs to $\mathscr{L}_c(X, Y)$). It is natural to ask whether we can approximate it by linear finite rank operators. This is in fact a famous problem of functional analysis, which remained open for several decades.

Definition 2.3.7 A Banach space X is said to have the "approximation property" if for every compact set $C \subseteq X$ and every $\epsilon > 0$, there is an $A \in \mathscr{L}_f(X)$ such that

$$\|x - A(x)\| < \epsilon \text{ for every } x \in C.$$

Remark 2.3.8 Some authors employ a different definition of this property (see, for example, Megginson [295, p. 330]). So, they say that a Banach space X has the "approximation property" if for every Banach space Y, $\overline{\mathscr{L}_f(Y, X)}^{\|\cdot\|_{\mathscr{L}}} = \mathscr{L}_c(Y, X)$. It turns out that the two definitions are in fact equivalent (see, for example, Lindenstrauss and Tzafriri [272]).

The question of whether every Banach space X has the approximation property was first asked by Hildebrandt and was answered in the negative four decades later by Enflo [162], who found a separable reflexive Banach space which lacks the approximation property.

Nevertheless there are important classes of Banach spaces which exhibit the approximation property.

Proposition 2.3.9 *If H is a Hilbert space with inner product denoted by $(\cdot, \cdot)_H$, then the following statements are equivalent:*

(a) $A \in \mathscr{L}_c(H)$;
(b) there exists an $\epsilon > 0$ *and an orthonormal set* $\{e_\alpha\}_{\alpha \in J} \subseteq H$ *such that*

$$\{\alpha \in J : |(A(e_\alpha), e_\alpha)_H| \geqslant \epsilon\} \text{ is finite};$$

(c) there exists a sequence $\{L_n\}_{n \geqslant 1} \subseteq \mathscr{L}_f(H)$ *such that* $\|A - L_n\|_{\mathscr{L}} \to 0$ *as* $n \to \infty$.

Proof $(a) \Rightarrow (b)$ Suppose that the implication is not true. Then we can find an orthonormal sequence $\{e_n\}_{n \geqslant 1} \subseteq H$ and $\epsilon_0 > 0$ such that

$$|(A(e_n), e_n)_H| \geqslant \epsilon_0 \text{ for all } n \geqslant 1. \tag{2.17}$$

Since A is compact, we can find a sequence $\{e_{n_k}\}_{k \geqslant 1}$ of $\{e_n\}_{n \geqslant 1}$ such that $A(e_{n_k}) \to h$ in H. Without any loss of generality we may assume that

$$\|A(e_{n_k}) - x\| \leqslant \frac{\epsilon_0}{2} \text{ for all } k \geqslant 1. \tag{2.18}$$

Then we have

$$\begin{aligned}|(A(e_{n_k}), e_{n_k})_H - (x, e_{n_k})_H| &= |(A(e_{n_k}) - x, e_{n_k})_H| \\ &\leqslant \|A(e_{n_k}) - x\| \\ &\leqslant \frac{\epsilon_0}{2} \text{ for all } k \geqslant 1\end{aligned} \tag{2.19}$$

(see (2.18) and recall that $\|e_{n_k}\| = 1$ for all $k \geqslant 1$). Then from (2.17) and (2.19) it follows that

$$\frac{\epsilon_0}{2} \leqslant |(x, e_{n_k})_H| \quad \text{for all } k \geqslant 1,$$

which contradicts Bessel's inequality.

$(b) \Rightarrow (c)$ For $n \in \mathbb{N}$, let $\mathscr{S}$ be the family of all orthonormal sets $\{e_\alpha\}_{\alpha \in J}$ such that

$$|\langle A(e_\alpha), e_\alpha\rangle_H| \geqslant \frac{1}{n} \text{ for all } \alpha \in J.$$

By virtue of (b), each such orthonormal set is finite. We partially order $\mathscr{S}$ by inclusion and consider a chain C in $\mathscr{S}$. Then the union of the elements of C is still in $\mathscr{S}$ and is an upper bound for C. So, by the Kuratowski–Zorn lemma, $\mathscr{S}$ has a maximal element denoted by $\{e_\gamma\}_{\gamma \in J}$. Since it belongs in $\mathscr{S}$, it is finite. Let $V = \operatorname{span}\{e_\vartheta\}_{\vartheta \in J}$. Then by virtue of the maximality of $\{e_\gamma\}_{\gamma \in J}$ we have

$$|\langle A(x), x\rangle_H| < \frac{1}{n} \text{ for all } x \in V^\perp,\ \|x\| = 1. \tag{2.20}$$

Let $P_n \in \mathscr{L}(H)$ be the orthogonal projection of H onto V. Setting $x = (I - P_n)(u)$ $(u \in H)$ in (2.20), we obtain

$$\begin{aligned}&|\langle A(I - P_n)(u), (I - P_n)(u)\rangle_H| < \frac{1}{n} \\ &\Rightarrow |\langle (I - P_n)A(I - P_n)(u), u\rangle_H| < \frac{1}{n} \text{ for all } u \in H \text{ with } \|u\| \leqslant 1 \\ &\Rightarrow \|(I - P_n)A(I - P_n)\|_{\mathscr{L}} < \frac{2}{n}.\end{aligned}$$

Let $L_n = P_n A + A P_n - P_n A P_n \in \mathscr{L}_f(H)$. Then

$$\begin{aligned}&\|A - L_n\|_{\mathscr{L}} < \frac{2}{n} \\ &\Rightarrow L_n \to A \in \mathscr{L}(H).\end{aligned}$$

$(c) \Rightarrow (a)$ This implication follows from Proposition 2.1.5. ☐

More generally, we have the approximation property for Banach spaces with a Schauder basis.

Definition 2.3.10 Let X be an infinite-dimensional normed space. A sequence $\{e_n\}_{n\geqslant 1} \subseteq X$ is said to be a "Schauder basis of X" if for every $x \in X$, there exists a unique sequence of scalars $\{\lambda_n\}_{n\geqslant 1}$, called the coordinates of x, such that

$$x = \sum_{n\geqslant 1} \lambda_n e_n.$$

Remark 2.3.11 Evidently, the sequence $\{e_n\}_{n\geqslant 1}$ is a linearly independent set in X. A Banach space X with a Schauder basis is automatically separable. Indeed, note that the countable set of the finite rational combinations of the basic elements is dense in X. For finite-dimensional Banach spaces, the Schauder basis coincides with the algebraic (Hamel) basis. In his celebrated treatise [31] Banach asked whether every separable Banach space admits a Schauder basis. The counterexample of Enflo [162] provided a negative answer to this question.

Lemma 2.3.12 *If X is a Banach space with a Schauder basis $\{e_n\}_{n\geqslant 1}$, then there exists an $M \geqslant 1$ such that for any $x = \sum\limits_{\mathrm{n}\geqslant 1} \lambda_n e_n$ we have*

$$\|\sum_{n=1}^{m} \lambda_n e_n\| \leqslant M\|\sum_{\mathrm{n}\geqslant 1}\lambda_n e_n\| \textit{ for all } m \in \mathbb{N}.$$

Proof On X we introduce the following quantity:

$$|x| = \sup\left\{\|\sum_{n=1}^{m} \lambda_n e_n\| : m \in \mathbb{N}\right\}.$$

It is easy to see that $|\cdot|$ is a norm on X and $(X, |\cdot|)$ is a Banach space. Note that

$$\|x\| = \|\sum_{n\geqslant 1} \lambda_n e_n\| \leqslant \sup\,[\|\sum_{n=1}^{m} \lambda_n e_n\| : m \in \mathbb{N}] = |x|.$$

So, from Banach's theorem it follows that $\|\cdot\|$ and $|\cdot|$ are equivalent. Therefore, we can find $M \geqslant 1$ such that

$$\begin{aligned} &|x| \leqslant M\|x\| \text{ for all } x \in X \\ \Rightarrow\ &\|\sum_{n=1}^{m} \lambda_n e_n\| \leqslant M||\sum_{n\geqslant 1} \lambda_n e_n\| \text{ for all } m \in \mathbb{N}. \end{aligned}$$

The proof is now complete. □

In the next lemma we denote by $P_m \in \mathscr{L}(X)$ the projection on the vector space $X_m = \operatorname{span}\{e_n\}_{n=1}^{m}, m \in \mathbb{N}$. So, if $x = \sum\limits_{\mathrm{n}\geqslant 1} \lambda_n e_n$, then $P_m(x) = \sum\limits_{\mathrm{n}=1}^{m} \lambda_n e_n$ for all $m \geqslant 1$.

Lemma 2.3.13 *If X is a Banach space with a Schauder basis $\{e_n\}_{n\geqslant 1}$, then for every $C \subseteq X$ compact, we have* $\lim_{m\to\infty} \sup_{x\in C} \|x - P_m(x))\| = 0$.

Proof Since C is compact in X, it is totally bounded, So, given $\epsilon > 0$, we can find $\{x_n\}_{n=1}^k \subseteq C$ such that

$$C \subseteq \bigcup_{n=1}^{k} B_\epsilon(x_n) \text{ (recall that } B_\epsilon(x_n) = \{x \in X : \|x - x_n\| < \epsilon\}).$$

Thus, given $x \in C$, we can find $n_0 \in \{1, ..., k\}$ such that $\|x - x_{n_0}\| < \epsilon$. Also, there exists an $m_0 \in \mathbb{N}$ such that

$$\|x_{n_0} - P_m(x_{n_0})\| < \epsilon \text{ for all } m \geqslant m_0. \tag{2.21}$$

Then for all $m \geqslant m_0$, we have

$$\begin{aligned}
\|x - P_m(x)\| &\leqslant \|x - x_{n_0}\| + \|x_{n_0} - P_m(x_{n_0})\| + \|P_m(x_{n_0}) - P_m(x)\| \\
&\leqslant (1 + \|P_m\|_{\mathscr{L}})\|x - x_{n_0}\| + \|x_{n_0} - P_m(x_{n_0})\| \\
&\leqslant \epsilon(2 + \|P_m\|_{\mathscr{L}}) \text{ (see (2.21))}.
\end{aligned} \tag{2.22}$$

From Lemma 2.3.12, we have

$$\begin{aligned}
&\|P_m(x)\| = \|\sum_{n=1}^{m} \lambda_n e_n\| \leqslant M\|\sum_{n\geqslant 1} \lambda_n e_n\| = M\|x\| \text{ for all } x \in X \\
&\Rightarrow \|P_m\|_{\mathscr{L}} \leqslant M.
\end{aligned}$$

Therefore from (2.22) we see that

$$\|x - P_m(x)\| \leqslant \epsilon(2 + M) \text{ for all } x \in C \text{ and all } m \geqslant m_0 \tag{2.23}$$

$$\Rightarrow \lim_{n\to\infty} \sup_{x\in C} \|x - P_m(x)\| = 0.$$

The proof is now complete. □

Now we are ready to show that Banach spaces with a Schauder basis have the approximation property (see Definition 2.3.7) and so compact linear operators can be approximated by finite rank linear operators

Theorem 2.3.14 *If X is a Banach space with a Schauder basis, then* $\mathscr{L}_c(X) = \overline{\mathscr{L}_f(X)}^{\|\cdot\|_{\mathscr{L}}}$.

Proof Let $A \in \mathscr{L}_c(X)$. So, if $\overline{B_1} = \{x \in X : \|x\| \leqslant 1\}$, then $A(\overline{B_1})$ is compact in X. Hence by virtue of Lemma 2.3.13, given $\epsilon > 0$, we can find $m_0 \in \mathbb{N}$ such that

$$\|A(x) - P_m(A(x))\| \leqslant \epsilon(2 + M) \text{ for all } m \geqslant m_0 \ (see\ (2.23))$$
$$\Rightarrow \|A - P_m A\|_{\mathscr{L}} \leqslant \epsilon(2 + M) \text{ for all } m \geqslant m_0.$$

Evidently, $P_m A \in \mathscr{L}_f(X)$ and so we conclude that $\mathscr{L}_c(X) = \overline{\mathscr{L}_f(X)}^{\|\cdot\|_{\mathscr{L}}}$. □

2.4 Spectral Theory of Compact Linear Operators

Although in this book we consider linear spaces over the field of reals, in order to develop a complete spectral theory for linear operators, we need to consider linear spaces over $\mathbb{C}$. So, in what follows X is a Banach space over the field $\mathbb{C}$ of complex numbers.

Recall that $A \in \mathscr{L}(X)$ is regular if and only if A^{-1} exists and belongs in $\mathscr{L}(X)$. If X is finite-dimensional, then the situation is rather simple and A^{-1} exists on X if and only if A is injective (one-to-one). Moreover, the linearity of A and A^{-1} automatically implies their continuity. In infinite-dimensional spaces X, the situation is more involved. The noninvertibility of $A \in \mathscr{L}(X)$ can happen if one of following situations occurs:

- A is not injective;
- A is injective but $A(X)$ is not dense in X;
- A is injective, $A(X)$ is dense in X, but A^{-1} is not continuous on $A(X)$.

So, we are led to the following definition.

Definition 2.4.1 Let X be a complex Banach space and $A \in \mathscr{L}(X)$. The spectrum of A is the set

$$\tau(A) = \{\lambda \in \mathbb{C} : \ \lambda I - A \ \text{ is not invertible}\}.$$

The spectrum of A can be decomposed into three disjoint components.

The point spectrum of A, which is the set

$$\tau_\rho(A) = \{\lambda \in \mathbb{C} : \lambda I - A \text{ is not injective}\},$$

the residual spectrum of A, which is the set

$$\tau_r(A) = \{\lambda \in \mathbb{C} : \lambda I - A \text{ is injective but } (\lambda I - A)(X) \text{ is not dense in } X\},$$

and the continuous spectrum of A, which is the set

$$\tau_c(A) = \{\lambda \in D : \ \lambda J - A \text{ is injective } \ (\lambda I - A)(X) \text{ is dense in } X \text{ but}$$
$$(\lambda I - A)^{-1} \text{ is not continuous on } (\lambda I - A)(X)\}.$$

We have $\tau(A) = \tau_\rho(A) \cup \tau_r(A) \cup \tau_c(A)$ and the three component sets are disjoint.

The elements of $\tau_\rho(A)$ are called “eigenvalues” of A.

Remark 2.4.2 If X is finite-dimensional, then $\tau(A) = \tau_\rho(A)$. For an infinite-dimensional Banach space X, we recall that if $\lambda I - A$ is injective, $(\lambda I - A)(X)$ is dense in X and $(\lambda I - A)^{-1}$ is continuous on $(\lambda I - A)(X)$, then $\lambda I - A$ is surjective and $(\lambda I - A)^{-1} \in \mathscr{L}(X)$ (that is, $\lambda I - A$ is regular).

As a direct consequence of the above definition, we have:

Proposition 2.4.3 *If $A \in \mathscr{L}(X)$, then $\lambda \in \tau_\rho(A)$ if and only if the equation $A(x) = \lambda x$ has a nontrivial solution.*

Definition 2.4.4 For $\lambda \in \tau_\rho(A)$, the nontrivial elements of $N(\lambda I - A) = \ker(\lambda I - A)$ are called eigenvectors corresponding to the eigenvalue λ and $N(\lambda I - A)$ is the corresponding eigenspace. The dimension of $N(\lambda I - A)$ is the geometric multiplicity (or simply multiplicity) of the eigenvalue λ.

Example 2.4.5 Let $X = l^2$ and let $A \in \mathscr{L}(l^2)$ be the right shift operator defined by

$$A(s_1, s_2, ...) = (0, s_1, s_2, ...).$$

Then $0 \in \tau(A)$, $0 \notin \tau_\rho(A)$ and $0 \in \tau_r(A)$, since clearly $A(l^2)$ is not dense in l^2.

For the next observation concerning the spectrum of a bounded linear operator we will need the following well-known fact from operator theory (see, for example, Lang [258, p. 74]).

Proposition 2.4.6 *The set of regular (invertible) elements of $\mathscr{L}(X)$ is open and if $A \in \mathscr{L}(X)$ satisfies $\|I - A\|_{\mathscr{L}} < 1$, then A is regular.*

Using this fact, we can establish the topological properties of the set $\tau(A)$.

Proposition 2.4.7 *If $A \in \mathscr{L}(X)$, then $\tau(A)$ is compact and in fact $\tau(A) \subseteq \overline{B}_{\|A\|_{\mathscr{L}}} = \{\lambda \in \mathbb{C} : |\lambda| \leqslant \|A\|_{\mathscr{L}}\}$.*

Proof Let $\xi : \mathbb{C} \to \mathscr{L}(X)$ be the map defined by $\xi(\lambda) = \lambda I - A$. We have

$$\|\xi(\lambda) - \xi(\mu)\|_{\mathscr{L}} = |\lambda - \mu|, \text{ hence } \xi(\cdot) \text{ is continuous.}$$

Note that $\tau(A)$ is the inverse image of the set of singular (noninvertible) elements under ξ. So, by virtue of Proposition 2.4.6, $\tau(A)$ is closed.

Next, let $\lambda \in \mathbb{C}$ such that $|\lambda| > \|A\|_{\mathscr{L}}$. Then $\|\frac{1}{\lambda}A\|_{\mathscr{L}} < 1$ and so Proposition 2.4.6 implies that $\lambda I - A$ is regular. Hence we have $\tau(A) \subseteq \overline{B}_{\|A\|_{\mathscr{L}}}$. Therefore $\tau(A)$ is compact. $\square$

Proposition 2.4.8 *If H is a Hilbert space and $A \in \mathscr{L}(H)$ is self-adjoint, then $\tau_\rho(A) \subseteq \mathbb{R}$ and eigenvectors corresponding to different eigenvalues are orthogonal.*

Proof Since A is self-adjoint, $\langle A(x), x\rangle_H \in \mathbb{R}$ for all $x \in H$. Let $\lambda \in \tau_\rho(A)$ and $x \in H$ be a corresponding eigenvector. Then

$$\begin{aligned} \langle A(x), x\rangle_H &= \langle \lambda x, x\rangle_H = \lambda\langle x, x\rangle_H = \lambda\|x\|^2 \\ \Rightarrow\ \lambda &= \frac{\langle A(x), x\rangle}{\|x\|^2} \in \mathbb{R}. \end{aligned}$$

Also, let $\lambda, \mu \in \tau_\rho(A)$, $\lambda \neq \mu$ and let $x, u \in X$ be eigenvectors corresponding to λ, μ respectively. We have

$$\begin{aligned} \langle A(x), u\rangle_H &= \langle \lambda x, u\rangle_H = \lambda\langle x, u\rangle_H, \\ \langle A(x), u\rangle_H &= \langle x, A(u)\rangle_H = \langle x, \mu u\rangle_H = \mu\langle x, u\rangle_H \end{aligned}$$

(recall that A is self-adjoint and note that, as we just proved, $\lambda, \mu \in \mathbb{R}$).

So, we have

$$(\lambda - \mu)\langle x, u\rangle_H = 0, \text{ hence } \langle x, u\rangle_H = 0 \text{ (since } \lambda \neq \mu).$$

The proof is now complete. □

From Propositions 2.4.7 and 2.4.8, we obtain

Corollary 2.4.9 *If H is a Hilbert space then*

(a) for every $A \in \mathscr{L}(H)$ self-adjoint we have $\tau(A) \subseteq [-\|A\|_{\mathscr{L}}, \|A\|_{\mathscr{L}}]$;
(b) for every $A \in \mathscr{L}(H)$ self-adjoint and $A \geqslant 0$ (that is, $\langle A(x), x\rangle_H \geqslant 0$ for all $x \in H$), we have $\tau(A) \subseteq [0, \|A\|_{\mathscr{L}}]$.

Remark 2.4.10 For H a Hilbert space and $A \in \mathscr{L}(H)$ self-adjoint, we know that $\|A\|_{\mathscr{L}} = \sup\{|\langle A(x), x\rangle_H| : \|x\| \leqslant 1\}$ (see for example, Brezis [65]).

Proposition 2.4.11 *If H is a Hilbert space and $A \in \mathscr{L}(H)$ is self-adjoint, then $\lambda \in \tau(A)$ if and only if* $\inf\{\|(\lambda I - A)(x)\| : \|x\| = 1\} = 0$.

Proof $\underline{\Leftarrow}$ Suppose $\lambda \neq \tau(A)$. Then $(\lambda I - A)^{-1} \in \mathscr{L}(H)$ and for every $x \in H$ with $\|x\| = 1$, we have

$$\begin{aligned} &1 = \|x\| = \|(\lambda I - A)^{-1}(\lambda I - A)x\| \leqslant \|(\lambda I - A)^{-1}\|_{\mathscr{L}}\|(\lambda I - A)(x)\| \\ \Rightarrow\ &\frac{1}{\|(\lambda I - A)^{-1}\|_{\mathscr{L}}} \leqslant \|(\lambda I - A)(x)\| \text{ for all } x \in H \text{ with } \|x\| = 1. \end{aligned}$$

$\underline{\Rightarrow}$ Suppose that $\inf\{\|(\lambda I - A)(x)\| : \|x\| = 1\} = \eta > 0$. Then

$$\begin{aligned} &\|(\lambda I - A)(x)\| \geqslant \eta\|x\| \text{ for all } x \in H \qquad (2.24) \\ \Rightarrow\ &\lambda I - A \text{ is an isomorphism onto } (\lambda I - A)(H). \end{aligned}$$

If we can show that $\lambda I - A$ is dense in H, then we are done. Arguing by contradiction, suppose that $(\lambda I - A)(H)$ is not dense in H. Then we can find $\hat{u} \in H \setminus \{0\}$ such that

$$\begin{aligned}
&\langle (\lambda I - A)(x), \hat{u}\rangle_H = 0 \text{ for all } x \in H \\
\Rightarrow\ &\langle x, (\overline{\lambda} I - A)(\hat{u})\rangle_H = 0 \text{ for all } x \in H \text{ (sinceAis self-adjoint)} \\
\Rightarrow\ &(\overline{\lambda} I - A)(\hat{u}) = 0,\ \text{ hence } \overline{\lambda} \text{ is an eigenvalue of A.}
\end{aligned}$$

But from Proposition 2.4.8, we know that $\tau_\rho(A) \subseteq \mathbb{R}$ and so $\overline{\lambda} = \lambda$. Hence

$$(\lambda I - A)(\hat{u}) = 0,$$

which contradicts (2.24) ☐

For compact operators, the spectrum has a very simple structure. This is a consequence of the so-called "Fredholm alternative", which says that in a Banach space X for every $A \in \mathscr{L}_c(X)$ we have that $I - A$ is injective if and only if $I - A$ is surjective (notice the similarity with the finite-dimensional case). To prove this theorem, we need some auxiliary results.

Proposition 2.4.12 *If X is a Banach space, $A \in \mathscr{L}_c(X)$ and $I - A$ is injective, then $(I - A)^{-1}$ is continuous on $(I - A)(X)$.*

Proof We argue by contradiction. So, suppose that $(I - A)^{-1}$ is not continuous on $(I - A)(X)$. Then we can find $\{x_n\}_{n\geqslant 1} \subseteq X$ such that

$$\begin{aligned}
&\|(I - A)(x_n)\| < \frac{1}{n}\|x_n\| \text{ for all } n \geqslant 1 \\
\Rightarrow\ &\|(I - A)\left(\frac{x_n}{\|x_n\|}\right)\| < \frac{1}{n} \text{ for all } n \geqslant 1.
\end{aligned} \tag{2.25}$$

Since $A \in \mathscr{L}_c(X)$, it follows that $\{A(\frac{x_n}{\|x_n\|}\|\}_{n\geqslant 1}$ is relatively compact in X and so, we may assume that

$$\begin{aligned}
&A(x_n/\|x_n\|) \to u \text{ in } X \\
\Rightarrow\ &\frac{x_n}{\|x_n\|} \to u \text{ in } X \text{ (see (2.25)) and so } \|u\| = 1 \\
\Rightarrow\ &A(u) = u,\ \text{ which contradicts the injectivity of } I - A.
\end{aligned}$$

The proof is now complete. ☐

Proposition 2.4.13 *If X is a Banach space, $A \in \mathscr{L}_c(X)$ and $I - A$ is injective, then $(I - A)(X)$ is a closed subspace in X.*

Proof Proposition 2.4.12 implies that $(I - A)(X)$ is isomorphic to X, which is a Banach space. Hence $(I - A)(X)$ is complete in X, hence it is a closed subspace of X. ☐

Proposition 2.4.14 *If X is a Banach space and $A \in \mathscr{L}_c(X)$, then for each $n \geqslant 1$ we have $(I - A)^n = I - L_n$ with $L_n \in \mathscr{L}_c(X)$.*

Proof We have

$$(I - A)^n = \sum_{n=0}^{n} (-1)^k \binom{n}{k} A^k.$$

If we set

$$L_n = -\sum_{k=0}^{n} (-1)^k \binom{n}{k} A^k,$$

then $L_n \in \mathscr{L}_c(X)$ (see Proposition 2.3.4) and

$$(I - A)^n = I - L_n \text{ for all } n \geqslant 1,$$

which completes the proof. □

Now we are ready to state and prove the so-called "Fredholm alternative theorem".

Theorem 2.4.15 *If X is a Banach space and $A \in \mathscr{L}_c(X)$, then $I - A$ is injective if and only if $I - A$ is surjective*

Proof $\underline{\Rightarrow}$ Let $X_n = (I - A)^n(X)$ for all $n \geqslant 1$. Then

$$X \supseteq X_1 \supseteq \ldots \supseteq X_n \supseteq \ldots.$$

Suppose that $X_n \neq X_{n+1}$ for every $n \geqslant 1$. Propositions 2.4.13 and 2.4.14 imply that X_{n+1} is closed in X_n for every $n \geqslant 0$ with $X_0 = X$. Invoking the Riesz lemma (see, for example, Brezis [65, p. 160]), we can find $x_n \in X_n$, $\|x_n\| = 1$ and $d(x_n, X_{n+1}) \geqslant \frac{1}{2}$. Then for $n > m$ we have

$$\|A(x_m) - A(x_n)\| = \|x_m - (I - A)(x_m) - x_n + (I - A)(x_n)\| \geqslant \frac{1}{2}$$

since $(I - A)(x_m) + x_n - (I - A)(x_n) \in X_{m+1}$. This means that $\{A(x_n)\}_{n\geqslant 1}$ has no convergent subsequence. This contradicts the fact that A is compact (note that $\{x_n\}_{n\geqslant 1}$ is bounded). This implies that there exists an $m \in \mathbb{N}$ such that $X_m = X_{m+1}$ and so $X_n = X_{n+1}$ for all $n \geqslant m$. The injectivity of $I - A$ implies the injectivity of $(I - A)^n$ for all $n \geqslant 1$. Given $x \in X$, we have

$$\begin{aligned}
&(I - A)^m(x) = (I - A)^{2m}(u) \text{ for some } u \in X \\
&\Rightarrow (I - A)^m[x - (I - A)^m(u)] = 0 \\
&\Rightarrow x = (I - A)^m(u) \text{ (since } (I - A)^m \text{ is injective)} \\
&\Rightarrow X = X_m \text{ and so we conclude that } I - A \text{ is surjective.}
\end{aligned}$$

$\Leftarrow$ Let $V_n = N((I - A)^n) = \ker((I - A)^n)$ for all $n \geqslant 1$. We need to show that $V_1 = \{0\}$. Arguing by contradiction, suppose that this is not true. So, we can find $x_1 \in V_1, x_1 \neq 0$. Suppose we have produced $\{x_k\}_{k=1}^n \subseteq X$ such that

$$(I - A)(x_{k+1}) = x_k \text{ and } x_k \in V_k \setminus V_{k-1} \text{ for all } k \in \{1, .., n - 1\}$$

(with $V_0 = \{0\}$).

Then since $I - A$ is surjective, we can find $x_{n+1} \in X$ such that

$$\begin{aligned} & (I - A)(x_{n+1}) = x_n \\ \Rightarrow\ & (I - A)^n(x_{n+1}) = (I - A)^{n-1}(x_n) = \ldots = x_1 \neq 0 \\ & \qquad\qquad\qquad\qquad\qquad\qquad \text{(by the induction hypothesis)} \\ \Rightarrow\ & (I - A)^{n+1}(x_{n+1} = (I - A)(x_1) = 0. \end{aligned}$$

So, by induction we have produced a sequence $\{x_n\}_{n\geqslant 1}$ such that

$$(I - A)(x_{n+1}) = x_n \text{ and } x_n \in V_n \setminus V_{n-1} \text{ for all } n \geqslant 1.$$

But as in the first part of the proof, we can show that $V_m = V_{m+1}$ for some $m \geqslant 1$, a contradiction $\square$

Remark 2.4.16 Usually the Fredholm alternative is formulated as follows:
"For X a Banach space and $A \in \mathscr{L}_c(X)$, the equation $A(x) - \lambda x = h$ has a solution for every $h \in H$ if and only if the equation $A(x) = \lambda x$ has only the trivial solution $x = 0$."

The Fredholm alternative has the following implications on the spectrum of a compact linear operator.

Proposition 2.4.17 *If X is a Banach space, $A \in \mathscr{L}_c(X)$ and $\lambda \in \tau(A) \setminus \{0\}$, then $\lambda \in \tau_\rho(A)$.*

Proof Arguing by contradiction, suppose that $\lambda \notin \tau_\rho(A)$. Then $\ker(\lambda I - A) = \{0\}$ (see Proposition 2.4.3). Theorem 2.4.15 (the Fredholm alternative) implies that range $(\lambda I - A) = X$ and so $(\lambda I - A)$ is invertible, hence $\lambda \notin \tau(A)$, a contradiction. $\square$

Corollary 2.4.18 *If X is a Banach space and $A \in \mathscr{L}_c(X)$, then $\tau(A) = \tau_p(A) \cup \{0\}$.*

The next property will complete the picture for the spectrum of a compact operator. We will need the following lemma, which extends Proposition 2.4.6 to Banach spaces.

Lemma 2.4.19 *If X is a Banach space and $A \in \mathscr{L}(X)$, then eigenvectors corresponding to distinct eigenvalues are linearly independent.*

Proof We argue by induction. So, suppose that $\lambda_1, \ldots, \lambda_n$ are distinct eigenvalues and $e_1, ..., e_n$ are corresponding eigenvectors which we assume to be linearly independent. Let $\lambda_{n+1} \in \tau_p(A) \setminus \{\lambda_k\}_{k=1}^n$ and suppose that the eigenfunction e_{n+1} corresponding to λ_{n+1} is a linear combination of $\{e_k\}_{k=1}^n$. So, we have

$$e_{n+1} = \sum_{k=1}^{n} \xi_k e_k \text{ with } \xi_k \in \mathbb{R} \text{ for all } k \in \{1, \ldots, N\}.$$

Then

$$\begin{aligned} \sum_{k=1}^{n} \lambda_{n+1}\xi_k e_k &= \lambda_{n+1}e_{n+1} = A(e_{n+1}) = \sum_{k=1}^{n} \lambda_k \xi_k e_k \\ \Rightarrow \sum_{k=1}^{n} (\lambda_{n+1} - \lambda_k)\xi_k e_k &= 0. \end{aligned}$$

Since $\{e_k\}_{k=1}^n$ are linearly independent (induction hypothesis) and $\lambda_{n+1} \neq \lambda_k$ for all $k \in \{1, \ldots, n\}$, it follows that $\xi_k = 0$ for all $k \in \{1, \ldots, n\}$, hence $e_{n+1} = 0$, a contradiction. Therefore $e_{n+1} \notin \operatorname{span}\{e_k\}_{k=1}^n$ and so we are done. □

Theorem 2.4.20 *If X is a Banach space and $A \in \mathscr{L}_c(X)$, then for every $\epsilon > 0$, A has only finitely many eigenvalues with absolute value bigger than $\epsilon > 0$.*

Proof Arguing by contradiction, suppose that we can find $\{\lambda_n\}_{n\geqslant 1} \subseteq \tau_p(A)$ such that $|\lambda_n| \geqslant \epsilon$ for all $n \geqslant 1$. For every λ_n, let $e_n \in X$ be a corresponding eigenfunction. We let $X_n = \operatorname{span}\{e_k\}_{k=1}^n$. Then $A(X_n) = X_n$ and $X_n \neq X_{n+1}$ for all $n \geqslant 1$ (see Lemma 2.4.19).

Invoking the Riesz lemma, we can find $u_{n+1} \in X_{n+1}$, $\|u_{n+1}\| = 1$ such that $d(u_{n+1}, X_n) \geqslant \frac{1}{2}$ for all $n \geqslant 1$. Let $v_{n+1} = \frac{1}{\lambda_{n+1}} u_{n+1}$. Then $\|v_{n+1}\| = \frac{1}{|\lambda_{n+1}|} \leqslant \frac{1}{\epsilon}$ and we have

$$A(v_{n+1}) = \frac{1}{\lambda_{n+1}} A(u_{n+1}) \in X_{n+1}. \tag{2.26}$$

Also, since $u_{n+1} \in X_{n+1}$, we have $u_{n+1} = \sum_{k=1}^{n+1} \xi_k e_k$ with $\xi_k \in \mathbb{R}$. Then

$$u_{n+1} - A(v_{n+1}) = \sum_{k=1}^{n+1} \left(1 - \frac{\lambda_k}{\lambda_{n+1}}\right) \xi_k e_k = \sum_{k=1}^{n} \left(1 - \frac{\lambda_k}{\lambda_{n+1}}\right) \xi_k e_k \in X_n.$$

If $n > m$, then $A(v_{m+1}) \in X_{m+1} \subseteq X_n$ and $u_{n+1} - A(v_{n+1}) \in X_n$ (see (2.26)). Therefore

$$\begin{aligned}\|A(v_{n+1}) - A(v_{m+1})\| &\geqslant d(A(v_{n+1}), X_n) \\ &= d(A(v_{n+1}) + u_{n+1} - A(v_{n+1}), X_n) \\ &= d(u_{n+1}, X_n) \geqslant \frac{1}{2}\end{aligned}$$

$$\Rightarrow \ \{A(v_n)\}_{n \geqslant 1} \text{ has no convergent subsequence.}$$

But this contradicts the compactness of A. $\square$

So, summarizing the situation for the spectrum of a compact operator in an infinite-dimensional Banach space, we have:

Corollary 2.4.21 *If X is an infinite-dimensional Banach space and $A \in \mathscr{L}_c(x)$, then*

(a) $0 \in \tau(A)$;
(b) $\tau(A) \setminus \{0\} = \tau_p(A) \setminus \{0\}$;
(c) one of the following holds

- $\tau(A) = \{0\}$,
- $\tau(A) \setminus \{0\}$ *is a finite set,*
- $\tau(A) \setminus \{0\}$ *is a sequence converging to 0.*

Moreover, if $\lambda_k \in \mathbb{C} \setminus \{0\}$ $(k \geqslant 1)$ are the eigenvalues, then the corresponding eigenspaces are finite-dimensional (that is, $\dim N(\lambda_k I - A) = \dim \ker(\lambda_k I - A) < \infty$*).*

Proof The only new information here is the finite-dimensionality of the eigenspaces. To see this, let $E_k = N(\lambda_k I - A)$ and let $\overline{B_1}^{E_k} = \{x \in E_k : \|x\| \leqslant 1\}$, $\overline{B}_1 = \{x \in X : \|x\| \leqslant 1\}$. Then

$$\overline{B}_1^{E_k} \subseteq A(\bar{B}_1)$$

and the latter is compact since $A \in \mathscr{L}_c(X)$. Therefore the closed unit ball in E_k is compact, hence E_k is finite-dimensional. $\square$

Next we turn our attention to the spectral properties of compact self-adjoint operators on a Hilbert space. We have already made some preliminary observations concerning the spectrum of self-adjoint (not necessarily compact) operators on a Hilbert space (see Proposition 2.4.8, Corollary 2.4.9 and Proposition 2.4.11). Now we will see what happens if the property of compactness enters into the picture.

We start with a general result for self-adjoint operators (not necessarily compact), which refines Proposition 2.4.7.

Proposition 2.4.22 *If H is a Hilbert space, $A \in \mathscr{L}(H)$ is self-adjoint*

$$m_A = \inf\{\langle A(x), x\rangle_H : \|x\| = 1\} \text{ and } M_A = \sup\{\langle A(x), x\rangle_H : \|x\| = 1\},$$

then $\tau(A) \subseteq [m_A, M_A]$ and $m_A, M_A \in \tau(A)$.

Proof By considering $\widehat{A} = A + \mu I$ for $\mu > 0$ big enough if necessary (note that $\widehat{A}$ is still self-adjoint and $M_{\widehat{A}} = M_A + \mu$, $m_{\widehat{A}} = m_{\widehat{A}} + \mu$), without any loss of generality, we may assume that $0 \leqslant m_A \leqslant M_A$.

In this case, by virtue of Remark 2.4.10, we have $M_A = \|A\|_{\mathscr{L}}$. Then Corollary 2.4.9 implies that $\tau(A) \subseteq [-M_A, M_A]$. Let $\mathscr{E} > 0$ and consider $\lambda = m_A - \mathscr{E}$. We will show that $\lambda \notin \tau(A)$. To see this, let $x \in H$ with $\|x\| = 1$. We have

$$\langle (A - \lambda I)(x), x\rangle_H = \langle A(x), x\rangle_H - \lambda \|x\|^2 \geqslant m_A - \lambda = \mathscr{E}. \tag{2.27}$$

Also, we have

$$\begin{aligned}
&|\langle (A - \lambda I)(x), x\rangle_H| \leqslant \|(A - \lambda I)(x)\| \, \|x\| = \|(A - \lambda I)(x)\| \\
&\Rightarrow 0 < \mathscr{E} \leqslant \inf\{\|(A - \lambda I)(x)\| : \|x\| = 1\} \text{ see (2.27))} \\
&\Rightarrow \lambda \notin \tau(A) \text{ (see Proposition 2.4.11).}
\end{aligned}$$

So, we have proved that

$$\tau(A) \subseteq [m_A, M_A].$$

Next we show that $M_A \in \tau(A)$. To this end, let $\{x_n\}_{n\geqslant 1} \subseteq X$ with $\|x_n\| = 1$ for all $n \geqslant 1$ such that

$$\langle A(x_n), x_n\rangle_H \uparrow M = \|A\|_{\mathscr{L}} \text{ as } n \to \infty. \tag{2.28}$$

Then for all $n \geqslant 1$ we have

$$\begin{aligned}
0 \leqslant \|(M_A I - A)(x_n)\|^2 &= M_A^2 \|x_n\|^2 + \|A(x_n)\|^2 - 2M_A \langle A(x_n), x_n\rangle_H \\
&\leqslant 2M_A^2 - 2M_A \langle A(x_n), x_n\rangle_H \to 0 \text{ as} \\
&\quad n \to \infty \text{ (see (2.28))} \\
\Rightarrow\ & 0 = \inf\{\|(M_A I - A)(x_n)\| : \|x\| = 1\} \\
\Rightarrow\ & M_A \in \tau(A) \text{ (see Proposition 2.4.11).}
\end{aligned}$$

Finally, let $\widehat{A} = A - M_A I$. Then $m_{\widehat{A}} \leqslant M_{\widehat{A}} = 0$ and $|m_{\widehat{A}}| = \|\widehat{A}\|_{\mathscr{L}}$. Arguing as above, we show that $m_{\widehat{A}} \in \tau(\hat{A})$, hence $m_A \in \tau(A)$. ☐

Corollary 2.4.23 *If H is a Hilbert space and $A \in \mathscr{L}(H)$ is self-adjoint, then* $\|A\|_{\mathscr{L}} = \max\{|m_A|, |M_A|\} \in \tau(A)$.

Now we can focus on compact self-adjoint operators.

Proposition 2.4.24 *If H is a Hilbert space and $A \in \mathscr{L}_c(H)$ is self-adjoint, then* $\tau_p(A) \neq \emptyset$.

Proof If $A = 0$, then $\lambda = 0$ is an eigenvalue. So, suppose $A \neq 0$. From Corollary 2.4.23 we have $\|A\|_{\mathscr{L}} \in \tau(A)$ and so by Corollary 2.4.21 we obtain $0 \neq \|A\|_{\mathscr{L}} \in \tau_p(A)$. ☐

Proposition 2.4.25 *If H is an infinite-dimensional Hilbert space and $A \in \mathscr{L}_c(H)$, $A \neq 0$, is self-adjoint, then H has an orthonormal basis consisting of eigenvectors corresponding to the eigenvalues of A.*

Proof Let $\lambda \in \tau_p(A)$ and let B_λ be an orthonormal basis for the eigenspace $E_\lambda = N(\lambda I - A) = \ker(\lambda I - A)$. Proposition 2.4.8 implies that $B = \bigcup_{\breve{}\in \o_{\mathrm{p}}(\mathrm{A})} B_\lambda$ is an orthonormal set in H. We will show that $H = \overline{\operatorname{span}}\, B$. Suppose that $H \neq \overline{\operatorname{span}}\, B$ and let $V = [\overline{\operatorname{span}}\, B]^\perp$. Note that $\overline{\operatorname{span}}\, B$ is A-invariant, hence so is V. We have $\tau(A) = \tau(A|_{\overline{\operatorname{span}}\, B}) + \tau(A|_V)$. But $A|_V$ has an eigenvalue (see Proposition 2.4.24), hence an eigenvector v. Then v is also an eigenvector for A and so $v \in V \cap \overline{\operatorname{span}}\, B = \{0\}$, a contradiction. So, we conclude that $H = \overline{\operatorname{span}}\, B$. $\square$

Now we can state and prove the so-called "spectral theorem" for compact self-adjoint operators on a separable Hilbert space.

Theorem 2.4.26 *If H is an infinite-dimensional separable Hilbert space and $A \in \mathscr{L}_c(H)$ is self-adjoint, then there is an orthonormal basis $\{e_n\}_{n\geqslant 1}$ of H such that each e_n is an eigenvector corresponding to some eigenvalue $\lambda_n \in \mathbb{R}$ and for all $x \in H$, we have*

$$A(x) = \sum_{n\geqslant 1} \lambda_n \langle x, e_n\rangle_H e_n.$$

Also, for every $\lambda \notin \tau(A)$ and $x \in H$, we have

$$(\lambda I - A)^{-1}(x) = \sum_{n\geqslant 1} \frac{\langle x, e_n\rangle_H}{\lambda - \lambda_n} e_n\,.$$

Proof From Proposition 2.4.25 we know that we have a countable (since H is separable) orthonormal basis consisting of eigenvectors of A. For any $x \in H$, the series $\sum_{n\geqslant 1} \lambda_n \langle x, e_n\rangle_H e_n$ converges, since

$$\|\sum_{n=k}^{m} \lambda_n \langle x, e_n\rangle e_n\|^2 = \sum_{n=k}^{m} \lambda_n^2 |\langle x, e_n\rangle_H|^2 \leqslant \|A\|_{\mathscr{L}} \sum_{n=k}^{m} |\langle x, e_n\rangle_H|^2 \to 0 \text{ as } k, m \to \infty.$$

Also, for $x \in H$, $\|x\| \leqslant 1$ and $m \in \mathbb{N}$, we have

$$\begin{aligned}
\|\sum_{n=1}^{m} \lambda_n \langle x, e_n\rangle_H e_n\|^2 = \sum_{n=1}^{m} \lambda_n^2 |\langle x, e_n\rangle_H|^2 &\leqslant \|A\|_{\mathscr{L}}^2 \sum_{n=1}^{m} |\langle x, e_n\rangle_H|^2 \\
&\leqslant \|A\|_{\mathscr{L}}^2 \sum_{n\geqslant 1} |\langle x, e_n\rangle_H|^2 \\
&= \|A\|_{\mathscr{L}}^2 \|x\|^2.
\end{aligned}$$

So, if we define $L(x) = \sum_{n \geqslant 1} \lambda_n \langle x, e_n \rangle_H e_n$ for all $x \in H$, then $L \in \mathscr{L}(H)$. Since $A(e_n) = \lambda_n e_n$, we see that $A(e_n) = L(e_n)$ for all $n \geqslant 1$ and so $A = L$.

Next, let $\lambda \notin \tau(A)$. Recall that $\tau(A)$ is closed (Proposition 2.4.6). So, there exists an $\eta > 0$ such that $d(\lambda, \tau(A)) > \eta$. Hence $|\lambda - \lambda_n| > \eta$ for all $n \geqslant 1$. Then

$$\| \sum_{n=k}^{m} \frac{\langle x, e_n \rangle_H}{\lambda - \lambda_n} e_n \|^2 = \sum_{n=k}^{m} \frac{|\langle x, e_n \rangle_H|^2}{|\lambda - \lambda_n|^2} \leqslant \frac{1}{\eta^2} \sum_{n=k}^{m} |\langle x, e_n \rangle_H|^2 \to 0 \text{ as } k, m \to \infty.$$

So, we can define the linear operator

$$L(x) = \sum_{n=k}^{m} \frac{\langle x, e_n \rangle_H}{\lambda - \lambda_n} e_n \text{ for all } x \in H.$$

For $x \in H$ with $\|x\| \leqslant 1$, we have

$$\begin{aligned} \| \sum_{n \geqslant 1}^{m} \frac{\langle x, e_n \rangle_H}{\lambda - \lambda_n} e_n \|^2 &= \sum_{n=1}^{m} \frac{|\langle x, e_n \rangle_H|^2}{|\lambda - \lambda_n|^2} \leqslant \frac{1}{\eta^2} \sum_{n=1}^{m} |\langle x, e_n \rangle_H|^2 \\ &\leqslant \frac{1}{\eta^2} \sum_{n \geqslant 1} |\langle x, e_n \rangle_H|^2 = \frac{1}{\eta^2} \|x\|^2 \leqslant \frac{1}{\eta^2} \end{aligned}$$

$\Rightarrow L \in \mathscr{L}(H)$.

For $x = \sum_{n \geqslant 1} \langle x, e_n \rangle_H e_n$, we have $A(x) = \sum_{n \geqslant 1} \lambda_n \langle x, e_n \rangle_H e_n$ and

$$\begin{aligned} & (\lambda I - A)(x) = \sum_{n \geqslant 1} (\lambda - \lambda_n) \langle x, e_n \rangle_H e_n \\ \Rightarrow\ & (\lambda I - A)(L(x)) = \sum_{n \geqslant 1} (\lambda - \lambda_n) \langle \sum_{i \geqslant 1} \frac{\langle x, e_i \rangle_H}{\lambda - \lambda_i} e_i, e_n \rangle_H e_n \\ &= \sum_{n,i \geqslant 1} (\lambda - \lambda_n) \frac{\langle x, e_i \rangle_H}{\lambda - \lambda_i} \langle e_n, e_i \rangle_H e_n \\ &= \sum_{n \geqslant 1} \langle x, e_n \rangle_H e_n \text{ (since } \{e_n\}_{n \geqslant 1} \text{ is orthonormal)} \\ &= x \\ \Rightarrow\ & L = (\lambda I - A)^{-1}. \end{aligned}$$

The proof is now complete. $\square$

2.5 Multifunctions

In an attempt to produce a framework for the study of functional equations which is broader than the one provided by compact maps, in the next section, we will introduce nonlinear operators of monotone type. However, the development of a coherent theory of monotone maps leads necessarily to multivalued maps. Also, certain classes of problems with unilateral constraints (most notably variational inequalities) are related to nondifferentiable convex functionals, which in their analysis lead to multivalued operators. Finally, many parts of applied analysis such as the calculus of variations, optimal control, optimization, mathematical economics and game theory, employ as an indispensable tool multivalued maps. For all these reasons it is a good idea to get acquainted with some basic definitions and facts from multivalued analysis.

We start with some continuity concepts for multivalued maps (set valued maps). Let X be a Hausdorff topological space. We introduce the following hyperspaces:

$$P_f(X) = \{A \subseteq X : A \neq \emptyset \text{ and it is closed}\},$$
$$P_k(X) = \{A \subseteq X : A \neq \emptyset \text{ and it is compact}\}.$$

Evidently, $P_k(X) \subseteq P_f(X)$. Also, if X is a normed space, then we introduce the following additional hyperspaces:

$$P_{fc}(X) = \{A \in P_f(X) : A \text{ is convex}\},$$
$$P_{(w)kc}(X) = \{A \subseteq X : A \neq \emptyset,\ A \text{ is (weakly-) compact and convex}\},$$
$$P_{bf(c)}(X) = \{A \in P_f(X) : A \text{ is bounded (and convex)}\}.$$

In what follows if X is a Hausdorff topological space and $x \in X$, then by $\mathscr{N}(x)$ we denote the filter of neighborhoods of x. If (X, d) is a metric space, then for all $x \in X$ and all $r > 0$, we set

$$B_r(x) = \{u \in X : d(u, x) < r\} \text{ (the open } r\text{-ball centered at } x \in X),$$
$$\overline{B}_r(x) = \{u \in X : d(u, x) \leqslant r\} \text{ (the closed } r\text{-ball centered at } x \in X).$$

If X is a normed space and $x = 0$, then as before we write $B_r = B_r(0)$ and $\overline{B}_r = \overline{B}_r(0)$.

Definition 2.5.1 Let X, Y be two sets and $F : X \to 2^Y$ a multifunction.

(a) The "weak inverse image" of $A \subseteq X$ under F is the set

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

(b) The "strong inverse image" of $A \subseteq X$ under F is the set

$$F^+(A) = \{x \in X : F(x) \subseteq A\}.$$

Using these notions, we can introduce the following continuity concepts for a multifunction.

Definition 2.5.2 Let X, Y be Hausdorff topological spaces and $F : X \to 2^Y$ a multifunction.

(a) We say that F is upper semicontinuous (usc for short) at x_0 if for all $V \subseteq Y$ open such that $F(x_0) \subseteq V$, we can find $U \in \mathscr{N}(x_0)$ such that $F(U) \subseteq V$. We say that F is upper semicontinuous (usc) if it is usc at every $x_0 \in X$.
(b) We say that F is lower semicontinuous (lsc for short) at x_0 if for all $V \subseteq Y$ open such that $F(x_0) \cap V \neq \emptyset$, we can find $U \in \mathscr{N}(x_0)$ such that $F(x) \cap V \neq \emptyset$ for all $x \in V$. We say that F is lower semicontinuous (lsc) if it is lsc at every $x_0 \in X$.
(c) We say that $F(\cdot)$ is continuous (or Vietoris continuous) at x_0 if it is both usc and lsc at x_0. We say that F is continuous (or Vietoris continuous) if it is continuous at every $x_0 \in X$.

The next propositions are straightforward consequences of the above definitions and provide alternative characterizations of these notions.

Proposition 2.5.3 *If X, Y are Hausdorff topological spaces and $F : X \to 2^Y$ a multifunction, then the following statements are equivalent:*

(a) F is usc;
(b) for every $C \subseteq Y$ closed, the set $F^-(C)$ is closed in X;
(c) if $x \in X$, $\{x_\alpha\}_{\alpha \in J}$ is a net converging to x in X and $V \subseteq Y$ is an open set such that $F(x) \subseteq V$, then we can find $\alpha_0 \in J$ such that for all $\alpha \geqslant \alpha_0$, $F(x_\alpha) \subseteq V$.

Proposition 2.5.4 *If X, Y are Hausdorff topological spaces and $F : X \to 2^Y$ is a multifunction, then the following statements are equivalent:*

(a) F is lsc;
(b) for every $C \subseteq Y$ closed, the set $F^+(C)$ is closed in X;
(c) if $x \in X$, $\{x_\alpha\}_{\alpha \in J}$ is a net converging to x in X and $V \subseteq Y$ is an open set such that $F(x) \cap V \neq \emptyset$, then we can find $\alpha_0 \in J$ such that for all $\alpha \geqslant \alpha_0$ $F(x_\alpha) \cap V \neq \emptyset$;
(d) if $\{x_\alpha\}_{\alpha \in J}$ is a net in X converging to $x \in X$ and $y \in F(x)$, we can find $y_\alpha \in F(x_\alpha)$ $(\alpha \in J)$ such that $y_\alpha \to y$ in Y.

Proposition 2.5.5 *If X, Y are Hausdorff topological spaces and $F : X \to 2^Y$ a multifunction, then the following statements are equivalent:*

(a) F is continuous;
(b) for every $C \subseteq Y$ closed, the sets $F^-(C)$ and $F^+(C)$ are both closed;
(c) if $\{x_\alpha\}_{\alpha \in J}$ is a net in X converging to $x \in X$ and $V \subseteq Y$ is open such that $F(x) \subseteq V$ or $F(x) \cap V \neq \emptyset$, then we can find $\alpha_0 \in J$ such that for all $\alpha \geqslant \alpha_0$, $F(x_\alpha) \subseteq V$ or $F(x_\alpha) \cap V \neq \emptyset$.

Remark 2.5.6 In the above definitions and propositions, the arbitrary open set V can be replaced by a basic open set. The notions of upper and lower semicontinuity are distinct and both coincide with continuity when F is single-valued. If φ, ψ :

$\mathbb{R} \to \mathbb{R}$, $\varphi \leqslant \psi$ and $F(x) = [\varphi(x), \psi(x)]$ then F is usc (resp. lsc), if φ is lower semicontinuous and ψ is upper semicontinuous in the sense of single-valued maps (resp. φ is upper semicontinuous and ψ is lower semicontinuous).

Definition 2.5.7 Let X, Y be Hausdorff topological spaces and $F : X \to 2^Y$ a multifunction.

(a) The graph of F is the set $\operatorname{Gr} F = \{(x, y) \in X \times Y : y \in F(x)\}$;
(b) We say that F is "closed" if $\operatorname{Gr} F \subseteq X \times Y$ is closed.

Proposition 2.5.8 *If X, Y are Hausdorff topological spaces with Y regular and $F : X \to P_f(Y)$ is usc, then F is closed.*

Proof Let $\{(x_\alpha, y_\alpha)\}_{\alpha\in J} \subseteq \operatorname{Gr} F$ be a net and assume that $(x_\alpha, y_\alpha) \to (x, y)$ in $X \times Y$. Suppose that $y \notin F(x)$. Since Y is regular, we can find open sets $U \in \mathscr{N}(y)$ and $V \supseteq F(x)$ such that $U \cap V = \emptyset$. Because F is usc and $y_\alpha \to y$ in Y, we can find $\alpha_0 \in J$ such that for all $\alpha \geqslant \alpha_0$, $F(x_\alpha) \subseteq V$ and $y_\alpha \in U$. Therefore $y_\alpha \in U \cap V$ for all $\alpha \geqslant \alpha_0$, a contradiction. □

Remark 2.5.9 The converse is not true in general. To see this let $X = \mathbb{R}$, $Y = \mathbb{R}^2$ and consider the multifunction $F : \mathbb{R} \to P_f(\mathbb{R}^2)$ defined by $F(x) = \{(t, xt) : t \in \mathbb{R}\}$. Then it is easy to see that F is closed but not usc.

Directly from Definition 2.5.2(b), we see that:

Proposition 2.5.10 *If X, Y are Hausdorff topological spaces, $V \subseteq Y$ is open, $F : X \to 2^Y \setminus \{\emptyset\}$ is lsc and $F(x) \cap V \neq \emptyset$ for all $x \in X$, then*

(a) $x \mapsto F(x) \cap V$ *is lsc;*
(b) if $C \subseteq X$ is closed and $\widehat{F}(x) = \begin{cases} \overline{F(x) \cap V}, & \text{if } x \in C \\ \overline{F}(x), & \text{if } x \in X \setminus C, \end{cases}$ *then F is lsc.*

We can extend part (a) of this proposition as follows:

Proposition 2.5.11 *If X is Hausdorff topological space, Y is a topological vector space, $F : X \to 2^Y \setminus \{\emptyset\}$ is lsc, V is a neighborhood of the origin in Y and $u : X \to Y$ is a continuous map such that*

$$F(x) \cap [u(x) + V] \neq \emptyset \text{ for all } x \in X,$$

then the multifunction $G : X \to 2^Y \setminus \{\emptyset\}$ defined by

$$G(x) = F(x) \cap [u(x) + V]$$

is lsc.

Proof Let $\{x_\alpha\}_{\alpha\in J}$ be a net in X converging to $x \in X$ and let $y \in G(x)$. We have

$$y \in F(x) \text{ and } y \in u(x) + V.$$

Let W be a balanced neighborhood of the origin in Y such that $y + W + W \subseteq u(x) + V$. From the continuity of u, we have $u(x_\alpha) \to u(x)$ in Y. Then we can find $\alpha_0 \in J$ such that

$$u(x) - u(x_\alpha) \in W \text{ for all } \alpha \geqslant \alpha_0.$$

For every $w \in W$ and every $\alpha \geqslant \alpha_0$, we have

$$\begin{aligned} & y + (u(x) - u(x_\alpha)) + w \in u(x) + V \\ \Rightarrow\ & y + w \in u(x_\alpha) + V \text{ for all } \alpha \geqslant \alpha_0 \text{ and all } w \in W \\ \Rightarrow\ & y + W \subseteq u(x_\alpha) + V \text{ for all } \alpha \geqslant \alpha_0. \end{aligned} \tag{2.29}$$

Since $y \in F(x)$ and $F(\cdot)$ is lsc, Proposition 2.5.4, implies that we can find $y_\alpha \in F(x_\alpha)$ such that $y_\alpha \to y$ in Y. Then we can find $\alpha_1 \in J$ such that $\alpha_1 \geqslant \alpha_0$ and

$$\begin{aligned} & y_\alpha - y \in W \text{ for all } \alpha \geqslant \alpha_1 \\ \Rightarrow\ & y_\alpha \in y + W \subseteq u(x_\alpha) + V \text{ for all } \alpha \geqslant \alpha_1 \text{ (see (2.29))}. \end{aligned}$$

Therefore

$$y_\alpha \in F(x_\alpha) \cap [u(x_\alpha) + V] = G(x_\alpha) \text{ for all } \alpha \geqslant \alpha_1$$

and so by virtue of Proposition 2.5.4, we conclude that G is lsc. □

Definition 2.5.12 Let X, Y be Hausdorff topological spaces and $F : X \to 2^Y \setminus \{\emptyset\}$ a multifunction. A "continuous selection" of $F(\cdot)$ is a continuous map $f : X \to Y$ such that $f(x) \in F(x)$ for all $x \in X$.

A fundamental question in the topological theory of multifunctions is that of the existence of a continuous selection for a given multifunction. The next example suggests that upper semicontinuous multifunctions is not the right class to look for continuous selections.

Example 2.5.13 Let $F : \mathbb{R} \to P_f(\mathbb{R})$ be the multifunction defined by

$$F(x) = \begin{cases} -1 & \text{if } x < 0 \\ [-1, 1] & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases}$$

(this is the subdifferential of the convex function $x \mapsto |x|$ on $\mathbb{R}$, see Sect. 2.7). Evidently, $F(\cdot)$ is usc but it cannot have a continuous selection.

So, to produce continuous selections, we turn our attention to lsc multifunctions and we have the following result, known in the literature as the "Michael selection theorem".

Theorem 2.5.14 *If X is a paracompact space, Y is a Banach space and $F : X \to P_{f_c}(Y)$ is an lsc multifunction, then F admits a continuous selection.*

Proof In the first part of the proof, we produce an approximate selection.

Fix $\epsilon > 0$ and for every $y \in Y$, let $U_y = F^-(B_\epsilon(y))$. By virtue of Definition 2.5.2(b), U_y is open. Then $\{U_y\}_{y\in Y}$ is an open cover of X and since X is paracompact, we can find a locally finite refinement $\{U'_y\}_{y\in Y}$ of $\{U_y\}_{y\in Y}$ (we can always choose the refinement to be precise, that is, to be indexed by the same set, see Dugundji [150, p. 162]). Let $\{p_y\}_{y\in Y}$ be a continuous partition of unity subordinate to this cover. Let

$$\widehat{f}(x) = \sum_{y\in Y} p_y(x)y. \tag{2.30}$$

Clearly, $\widehat{f}$ is well-defined and continuous. Note that $p_y(x) > 0$ implies $x \in U'_y \subseteq U_y$ and so $y \in F(x) + \epsilon B_1$. Since F has convex values we have $\widehat{f}(x) \in F(x) + \epsilon B_1$ (see (2.30)). So, we have produced an approximate continuous selection for F.

Now, let $V_n = \frac{1}{2^n}B_1,\ n \geqslant 1$. Using induction we will produce a sequence of continuous functions $f_n : X \to Y,\ n \geqslant 1$, such that

$$f_n(x) \in f_{n-1}(x) + 2V_{n-1} \text{ for all } x \in X,\ \text{all } n \geqslant 2, \tag{2.31}$$

$$f_n(x) \in F(x) + V_n \text{ for all } n \geqslant 1. \tag{2.32}$$

From the first part of the proof, we know that there is a continuous map $f_1 : X \to Y$ satisfying (2.32). Suppose we were able to construct continuous maps $f_n : X \to Y,\ n = 1, \ldots, m$, satisfying (2.31) and (2.32). We set $G_m(x) = F(x) \cap [f_m(x) + V_m]$. The induction hypothesis implies that $G_m(x) \neq \emptyset$ for all $x \in X$, while by virtue of Proposition 2.5.11 the mapping $x \mapsto G_m(x)$ is lsc. So, we can apply the first part of the proof (with data the multifunction $G(\cdot)$ and $\epsilon = \frac{1}{2^{m+1}}$) and obtain a continuous map $f_{m+1} : X \to Y$ such that

$$f_{m+1}(x) \subseteq G_m(x) + V_{m+1} \subseteq F(x) + V_{m+1},$$
$$f_{m+1}(x) \subseteq f_m(x) + V_m + V_{m+1} \subseteq f_m(x) + 2V_m.$$

This completes the induction process.

So, we have a sequence of continuous maps $f_n : X \to Y, n \geqslant 1$, satisfying (2.31) and (2.32). From (2.31) we infer that $\{f_n(x)\}_{n\geqslant 1}$ is Cauchy, uniformly in $x \in X$. So, there exists a continuous map $f : X \to Y$ such that

$$f_n(x) \to f(x) \text{ in } Y \text{ as } n \to \infty.$$

From (2.32) in the limit as $n \to \infty$, we have $f(x) \in F(x)$ for all $x \in X$, that is, $f(\cdot)$ is a continuous selection of $F(\cdot)$. ∎

We can also produce a continuous selection passing from a prescribed point of $\mathrm{Gr}\, F$.

Corollary 2.5.15 *If X is a paracompact space, Y is a Banach space, $F : X \to P_{f_c}(Y)$ is lsc and $(x_0, y_0) \in \mathrm{Gr}\, F$, then there exists a continuous map $f : X \to Y$*

such that

$$f(x_0) = y_0 \textit{ and } f(x) \in F(x) \textit{ for all } x \in X.$$

Proof Let $G : X \to P_{f_c}(Y)$ be the multifunction defined by

$$G(x) = \begin{cases} F(x) \text{ if } x \neq x_0 \\ \{y_0\} \text{ if } x = x_0. \end{cases} \tag{2.33}$$

Evidently, G is lsc with values in $P_{f_c}(Y)$. So, we can apply Theorem 2.5.14 and produce a continuous map $f : X \to Y$ such that $f(x) \in G(x)$ for all $x \in X$. From (2.33) it follows that

$$f(x_0) = y_0 \text{ and } f(x) \in F(x) \text{ for all } x \in X.$$

The proof is now complete. □

When X is a metric space and Y is a separable Banach space, then Theorem 2.5.14 can be refined and have a whole sequence of continuous selections of $F(\cdot)$ which are dense in $F(x)$ for all $x \in X$.

Proposition 2.5.16 *If X is a metric space, Y is a separable Banach space and $F : X \to P_{f_c}(Y)$ is lsc then there exists a sequence of continuous selections $f_n : X \to Y$, $n \geqslant 1$ of $F(\cdot)$ such that*

$$F(x) = \overline{\{f_n(x)\}}_{n \geqslant 1} \textit{ for all } x \in X.$$

Proof Let $\{y_n\}_{n\geqslant 1}$ be dense in Y and let $V_m = \frac{1}{2^m}B_1$, $m \geqslant 1$. We set $U_{nm} = F^{-1}(B_{2^{-m}}(y_n))$. Since F is lsc, the set U_{nm} is open for all $n, m \geqslant 1$. In metric spaces open sets are F_δ-sets. So, we have $U_{nm} = \bigcup\limits_{k\geqslant 1} C_{nmk}$ with C_{nmk} closed for all $k \geqslant 1$. We define

$$F_{nmk}(x) = \begin{cases} \overline{F(x) \cap B_{2^{-m}}(y_n)} & \text{for } x \in C_{nmk} \\ F(x) & \text{for } x \in X \setminus C_{nmk}. \end{cases}$$

From Definition 2.5.1, we know that $F_{nmk}(\cdot)$ is lsc and has values in $P_{f_c}(Y)$. So, we can apply Theorem 2.5.14 and find a continuous selection $f_{nmk} : X \to Y$ of $F_{nmk}(\cdot)$. We claim that $\{f_{nmk}\}_{n,m,k\geqslant 1}$ is the desired dense sequence. Indeed, let $y \in F(x)$ and $m \geqslant 1$. We can find $y_n \in y + V_{m+2}$. Then $x \in U_{n(m+2)}$ and so $x \in C_{m(m+2)k}$ for some $k \geqslant 1$. But then $f_{n(m+2)k}(x) \in y_n + V_{m+2} \subseteq y_n + V_{m+1} \subseteq y + V_{m+2} + V_{m+1} \subseteq y + V_m$ and this completes the proof. □

In fact, in this setting we can relax the requirement that $F(\cdot)$ has values in $P_{f_c}(x)$. So, we have the following result due to Michael [298] (see Theorem 3.1''', p. 368) (see also Hu and Papageorgiou [218, p. 97]).

Theorem 2.5.17 *If X is a metric space, Y is a separable Banach space and $F : X \to 2^Y \setminus \{\emptyset\}$ is an lsc multifunction with convex values which either have nonempty interior or are finite-dimensional, then $F(\cdot)$ admits a continuous selection.*

We already saw that usc multifunctions in general do not have a continuous selection. Nevertheless, as Example 2.5.13 suggests, we can have a continuous approximate selection. More precisely, we have the following result.

Theorem 2.5.18 *If X is a metric space, Y is a Banach space and $F : X \to 2^Y \setminus \{\emptyset\}$ is a usc multifunction with convex values, then given $\epsilon > 0$ we can find a locally Lipschitz function $f_\epsilon : X \to Y$ such that*

$$f_\epsilon(X) \subseteq \operatorname{conv} F(X)$$

and $h^*(\mathrm{Gr}\, f_\epsilon, \mathrm{Gr}\, F) = \sup\{d((x, y), \mathrm{Gr}\, F) : y = f_\epsilon(x)\} < \epsilon$.

Proof Fix $\epsilon > 0$. Since F is usc, by virtue of Definition 2.5.2(a), for every $x \in X$ we can find $0 < \delta = \delta(\epsilon, x) < \dfrac{\epsilon}{2}$ such that if $x' \in B_{\delta(x)}(x)$, then $F(x') \subseteq F(x) + \dfrac{\epsilon}{2} B_1$. The family $\{B_{\frac{\delta(x)}{4}}(x)\}_{x\in X}$ is an open cover of X. So, we can find a precise locally finite refinement $\{U_x\}_{x\in X}$ of this open cover and a corresponding locally Lipschitz partition of unity $\{p_x\}_{x\in X}$ subordinate to this refinement. For each $x \in X$, let $(z_x, y_x) \in \mathrm{Gr}\, F \cap (U_x \times Y)$ and set

$$f_\epsilon(z) = \sum_{x\in X} p_x(z) y_x \text{ for all } z \in X.$$

Then $f_\epsilon : X \to Y$ is well-defined and locally Lipschitz. Moreover, we have $f_\epsilon(X) \subseteq \operatorname{conv} F(X)$.

Fix $z \in X$. We have $p_x(z) > 0$ for all $x \in J(z) \subseteq X$. For every $x \in J(z)$, let $v \in X$ such that $U_v \subseteq B_{\frac{\delta(v)}{4}}(v)$. Let $w \in J(v)$ and set $\delta_w = \max[\delta_v : v \in J(z)]$. Then $v \in B_{\frac{\delta_w}{2}}(w)$ and so $U_v \subseteq B_{\delta_w}(w)$. Hence for any $v \in J(z)$ we have $y_v \in F(U_v) \subseteq F(w) + \frac{\epsilon}{2} B_1$. But the last set is convex. So, we have $f_\epsilon(z) \in F(w) + \frac{\epsilon}{2} B_1$. Therefore we can find $y \in F(w)$ such that $\| f_\epsilon(z) - y\| < \epsilon$, which implies that $(z, f_\epsilon(z)) \in \mathrm{Gr}\, F + \epsilon B_1$ and so we conclude that $h^*(\mathrm{Gr}\, f_\epsilon, \mathrm{Gr}\, F) < \epsilon$. ☐

Another version of this approximate selection theorem is the following:

Theorem 2.5.19 *If X is a metric space, Y is a Banach space, $U \subseteq X$ is open, $K \subseteq U$ is compact and $F : \overline{U} \to 2^Y \setminus \{\emptyset\}$ is a usc multifunction with convex values, then for every $\epsilon > 0$, there is an open neighborhood V_ϵ of K and a locally Lipschitz function $f_\epsilon : V_\epsilon \to \operatorname{conv} F(K)$ with finite-dimensional range such that for every $x \in V_\epsilon$ we have $f_\epsilon(x) \in F(K \cap B_\epsilon(x)) + \epsilon B_1$.*

Proof We keep the notation from the proof of Theorem 2.5.18.

Since K is compact, from the cover $\{B_{\delta(x)}(x)\}_{x \in X}$ we can extract a finite subcover $\{B_{\delta(x_n)}(x_n)\}_{n=1}^m$. We can find a locally Lipschitz partition of unity $\{p_n\}_{n=1}^m$ subordinate to this finite cover and $y_n \in F(x_n)$ and then we define

$$f_\epsilon(x) = \sum_{n=1}^m p_n(x) y_n \text{ for all } x \in V_\epsilon = \bigcup_{n=1}^m B_{\delta(x_n)}(x_n).$$

Evidently, f_ϵ is locally Lipschitz and $f_\epsilon(x) \in \operatorname{conv} f(K) \cap Y_\epsilon$ with $Y_\epsilon = \operatorname{span}\{x_n\}_{n=1}^m$.

Reasoning as in the proof of Theorem 2.5.18, we show that

$$f_\epsilon(x) \in F(K \cap B_\epsilon(x)) + \epsilon B_1 \text{ for all } x \in V_\epsilon,$$

which completes the proof. □

Next, we turn our attention to the measurability of multifunctions. We start by introducing a class of functions of two variables, which appears in many different situations in nonlinear analysis. In what follows, for any Hausdorff topological space X, by $B(X)$ we denote its Borel σ-field.

Definition 2.5.20 Let (Ω, Σ) be a measurable space, and X, Y Hausdorff topological spaces. A function $f : \Omega \times X \to Y$ is said to be a "Carathéodory function" if

(i) for all $x \in X$, $\omega \mapsto f(\omega, x)$ is $(\Sigma, B(Y))$-measurable;
(ii) for all $\omega \in \Omega$, $x \mapsto f(\omega, x)$ is continuous.

Theorem 2.5.21 *If (Ω, Σ) is a measurable space, X is a separable metric space, Y a metric space, and $f : \Omega \times X \to Y$ is a Carathéodory function, then f is $(\Sigma \times B(X), B(Y))$-measurable.*

Proof Let $D \subseteq X$ be a countable dense subset and let $C \subseteq Y$ be closed. We introduce the open set $C_n = \{y \in Y : d_Y(y, C) < \frac{1}{n}\}$ with d_Y being the metric of Y. We have $f(\omega, x) \in C$ if and only if there exists a $u \in D$ such that $d_X(x, u) < \frac{1}{n}$ and $f(\omega, x) \in C_n$ (here d_X is the metric on X). So, it follows that

$$\begin{aligned} f^{-1}(C) = \bigcap_{n \geqslant 1} \bigcup_{u \in D} [\{\omega \in \Omega : f(\omega, u) \in C_n\} \times \{x \in X : d_X(x, u) < \frac{1}{n}\}] \\ \in \Sigma \times B(X) \end{aligned}$$

$$\Rightarrow f \text{ is } (\Sigma \times B(X), B(Y))\text{-measurable.}$$

The proof is now complete. □

Remark 2.5.22 The result fails if the requirement that $f(\omega, x)$ is Carathéodory, is replaced by the hypothesis that for all $x \in X$, $\omega \mapsto f(\omega, x)$ is $(\Sigma, B(Y))$-measurable and for all $\omega \in \Omega$, $x \mapsto f(\omega, x)$ is lower or upper semicontinuous.

Now we are ready to introduce the measurability notions for multifunctions.

Definition 2.5.23 Let (Ω, Σ) be a measurable space, X a Hausdorff topological space and $F : \Omega \to 2^X \setminus \{\emptyset\}$ a multifunction

(a) We say that $F(\cdot)$ is strongly measurable if for all $C \subseteq X$ closed

$$F^-(C) = \{\omega \in \Omega : F(\omega) \cap C \neq \emptyset\} \in \Sigma.$$

(b) We say that $F(\cdot)$ is measurable if for all $U \subseteq X$ open

$$F^-(U) = \{\omega \in \Omega : F(\omega) \cap U \neq \emptyset\} \in \Sigma.$$

(c) We say that $F(\cdot)$ is graph measurable if

$$\mathrm{Gr}\, F = \{(\omega, x) \in \Omega \times X : x \in F(\omega)\} \in \Sigma \times B(X).$$

Proposition 2.5.24 *If (Ω, Σ) is a measurable space, X a metric space and $F : \Omega \to 2^X \setminus \{\emptyset\}$ is strongly measurable, then $F(\cdot)$ is measurable.*

Proof In a metric space every open set is F_σ. So, if $U \subseteq X$ is open, then $U = \bigcup\limits_{n\geqslant 1} C_n$ with $C_n \subseteq X$ closed. Then

$$F^-(U) = F^-(\bigcup_{n\geqslant 1} C_n) = \bigcup_{n\geqslant 1} F^-(C_n) \in \Sigma$$
$$\Rightarrow \ F(\cdot) \text{ is measurable.}$$

The proof is now complete. $\square$

Proposition 2.5.25 *If (Ω, Σ) is a measurable space, X a separable metric space and $F : \Omega \to 2^X \setminus \{\emptyset\}$, then $F(\cdot)$ is measurable if and only if for all $x \in \mathbb{R}$, the mapping $\omega \mapsto d(x, F(\omega))$ is Σ-measurable (d being the metric on X).*

Proof $\underline{\Rightarrow}$ Suppose that $F(\cdot)$ is measurable. For every $\lambda > 0$ let $L_\lambda(x) = \{\omega \in \Omega : d(x, F(\omega)) < \lambda\}$. Clearly $L_\lambda(x) = F^-(B_\lambda(x)) \in \Sigma$, hence $\omega \mapsto d(x, F(\omega))$ is Σ-measurable.

$\underline{\Leftarrow}$ For every $x \in X$ and $\lambda > 0$, by hypothesis

$$F^-(B_\lambda(x)) = L_\lambda(x) = \{\omega \in \Omega : d(x, F(\omega)) < \lambda\} \in \Sigma.$$

Now let $U \subseteq X$ be open. Then $U = \bigcup_{n\geqslant 1} B_{\lambda_n}(x_n)$ (recall that X is separable). Hence

$$F^-(U) = F^-(\bigcup_{n\geqslant 1} B_{\lambda_n}(x_n)) = \bigcup_{n\geqslant 1} F^-(B_{\lambda_n}(x_n)) \in \Sigma$$
$$\Rightarrow \ F \text{ is measurable.}$$

The proof is now complete. □

Proposition 2.5.26 *If* (Ω, Σ) *is a measurable space, X is a separable metric space and* $F : \Omega \to P_f(X)$ *is a measurable multifunction, then* $F(\cdot)$ *is graph measurable.*

Proof Note that

$$\operatorname{Gr} F = \{(\omega, x) \in \Omega \times X : d(x, F(\omega)) = 0\}.$$

But by virtue of Proposition 2.5.25 $(\omega, x) \mapsto d(x, F(\omega))$ is a Carathéodory function. Hence invoking Theorem 2.5.21 we conclude that $\operatorname{Gr} F \in \Sigma \times B(X)$. So, $F(\cdot)$ is graph measurable. □

For compact-valued multifunctions, we can say more.

Proposition 2.5.27 *If* (Ω, Σ) *is a measurable space, X is a separable metric space and* $F : \Omega \to P_k(X)$ *is a multifunction then* $F(\cdot)$ *is strongly measurable if and only if* $F(\cdot)$ *is measurable.*

Proof $\underline{\Rightarrow}$ This is Proposition 2.5.24.
$\underline{\Leftarrow}$ Let $C \subseteq X$ be closed. Since $F(\cdot)$ is $P_k(X)$-valued, we see that

$$X \setminus F^-(C) = \{\omega \in \Omega : d(F(\omega), C) > 0\}. \tag{2.34}$$

If D is a countable dense subset of X, then

$$\begin{aligned} d(F(\omega), C) &= \inf\{d(x, F(\omega)) : x \in C\} \\ &= \inf\{d(x, F(\omega)) : x \in C \cap D\} \\ \Rightarrow \omega &\mapsto d(F(\omega), C) \text{ is } \Sigma\text{-measurable (see Proposition 2.5.25)} \\ \Rightarrow F^-(C) &\in \Sigma \text{ and so F is strongly measurable.} \end{aligned}$$

The proof is now complete. □

Proposition 2.5.28 *If* (Ω, Σ) *is a measurable space, X is a separable metric space and* $F : \Omega \to P_f(X)$ *is a measurable multifunction, then*

(a) for every $K \subseteq X$ *compact,* $F^-(K) \in \Sigma$,
(b) if X is σ*-compact, then* $F(\cdot)$ *is strongly measurable.*

Proof (a) Recall that a separable metric space is homeomorphic to a subset of the Hilbert cube $\mathscr{H} = [0, 1]^{\mathbb{N}}$ (this is established in the proof of the Urysohn metrization theorem). So, we can think of X as a dense subset of a compact metric space V. Let $G : \Omega \to P_k(V)$ be the multifunction defined by $G(\omega) = \overline{F(\omega)}^V$. Clearly, $G(\cdot)$ is measurable. For $K \subseteq X$ compact, we have

$$\begin{aligned} F^-(K) &= \{\omega \in \Omega : F(\omega) \cap K \neq \{\emptyset\}\} = \{\omega \in \Omega : G(\omega) \cap X \cap K \neq \{\emptyset\}\} \\ &= G^-(K) \in \Sigma \text{ (see Proposition 2.5.27).} \end{aligned}$$

(b) By hypothesis, $X = \bigcup_{n\geqslant 1} K_n$ with K_n compact. For every $C \subseteq X$ closed we have

$$\begin{aligned} F^-(C) &= F^-(C \cap (\bigcup_{n\geqslant 1} K_n)) = F^-(\bigcup_{n\geqslant 1} C \cap K_n) \\ &= \bigcup_{n\geqslant 1} F^-(C \cap K_n) \in \Sigma \text{ (see part (a))} \\ \Rightarrow \quad & \text{F is strongly measurable.} \end{aligned}$$

The proof is now complete. □

Remark 2.5.29 So, if X is σ-compact and $F(\cdot)$ is $P_k(X)$-valued, the notions of measurability and strong measurability are equivalent.

Next, we introduce two classes of spaces which are important in measure theory in general and in the measurability properties of multifunctions in particular.

Definition 2.5.30 (a) A Hausdorff topological space X is a Polish space if it is separable and there exists a metric on X for which the topology τ is complete.

(b) A Hausdorff topological space X is a Souslin space if there exists a Polish space Y and a continuous surjection from Y onto X.

Remark 2.5.31 In a Polish space the metric d is not a priori fixed. We only know that there exists one generating the topology of X which is complete. There are many Hausdorff topological spaces that are Polish, but have no complete metric which is particularly natural or simple. For example, an open set of a Polish space is itself Polish, but it is not immediately clear which is the complete metric topologizing the open set. An equivalent way to define a Souslin space is to say that it is a Hausdorff topological space for which there exists a stronger (finer) topology making the space homeomorphic to a quotient of a Polish space. A Souslin space is always separable but need not to be metrizable. Such a space is either an infinite-dimensional separable Banach space furnished with the weak topology or its dual furnished with the weak* topology. However, every locally compact Souslin space is Polish. The Souslin subsets of a Polish space are called "analytic sets".

Definition 2.5.32 Let (Ω, Σ) be a measurable space, X a Hausdorff topological space and $F : \Omega \to 2^X \setminus \{\emptyset\}$. A "measurable selection" of $F(\cdot)$ is a map $f : \Omega \to X$ which is $(\Sigma, B(X))$-measurable and $f(\omega) \in F(\omega)$ for all $\omega \in \Omega$.

The next result establishes the existence of such a selection. The result is known in the literature as the "Kuratowski–Ryll-Nardzewski selection theorem".

Theorem 2.5.33 *If (Ω, Σ) is a measurable space, X is a Polish space and $F : \Omega \to P_f(X)$ is a measurable multifunction, then $F(\cdot)$ admits a measurable selection.*

Proof As in the proof of the Michael selection theorem (see Proposition 2.5.24), first we produce an approximate measurable selection.

Without any loss of generality we may assume that $\operatorname{diam} X < 1$. Let $\{x_k\}_{k\geqslant 1}$ be dense in X. We construct a sequence of Σ-measurable maps $f_n : \Omega \to X$, $n \geqslant 0$, such that

(a) $d(f_n(\omega), F(\omega)) < \frac{1}{2^n}$ for all $n \geqslant 0$ and all $\omega \in \Omega$,
(b) $d(f_n(\omega), f_{n-1}(\omega)) < \frac{1}{2^{n-1}}$ for all $n \geqslant 1$ and all $\omega \in \Omega$.

Define $f_0(\omega) = x_1$ for all $\omega \in \Omega$. We have $d(f_0(\omega), F(\omega)) < 1$ for all $\omega \in \Omega$ (recall that diam $X < 1$). Suppose that $f_0, \dots, f_{n-1}$ have been constructed and satisfy (a), (b). We set

$$A_k^n = \{\omega \in \Omega : d(x_k, F(\omega)) < \frac{1}{2^n}\},$$
$$C_k^n = \{\omega \in \Omega : d(x_k, f_{n-1}(\omega)) < \frac{1}{2^{n-1}}\} \text{ and } D_k^n = A_k^n \cap C_k^n \text{ for all } k \geqslant 1.$$

We claim that $\Omega = \bigcup_{k \geqslant 1} D_k^n$. To this end, let $\omega \in \Omega$. From the induction hypothesis we know that we can find $z \in F(\omega)$ such that $d(f_{n-1}(\omega), z) < \frac{1}{2^{n-1}}$. Also, we can find x_k such that $d(x_k, z) < \frac{1}{2^n}$ and $d(x_k, z) + d(z, f_{n-1}(\omega)) < \frac{1}{2^{n-1}}$. So, by the triangle inequality $d(x_k, f_{n-1}(\omega)) < \frac{1}{2^{n-1}}$. Therefore $\omega \in D_k^n$ for some $k \geqslant 1$ and we have proved that $\Omega = \bigcup_{k \geqslant 1} D_k^n$. From Proposition 2.5.25, we have that $A_k^n \in \Sigma$. Also, the measurability of $f_{n-1}(\cdot)$ (by the induction hypothesis), implies that $C_k^n \in \Sigma$. Therefore $D_k^n \in \Sigma$. We define $f_n : \Omega \to X$ by setting

$$f_n(\omega) = x_k \text{ for all } \omega \in D_k^n \setminus \bigcup_{i=1}^{k-1} D_i^n.$$

Evidently, range $f_n = \{x_k\}_{k \geqslant 1}$ and f_n is Σ-measurable. Moreover, from (b) we have that $\{f_n(\omega)\}_{n \geqslant 0} \subseteq X$ is Cauchy uniformly in $\omega \in \Omega$. So, we can find $f : \Omega \to X$ such that $f_n(\omega) \to f(\omega)$ in X (recall that X is complete). Then f is Σ-measurable and from (a) we see that $f(\omega) \in F(\omega)$ for all $\omega \in \Omega$. □

Remark 2.5.34 A careful reading of the above proof reveals that we can drop the hypothesis that X is Polish and instead assume that X is a separable metrizable space and $F(\cdot)$ has complete values.

In fact, as we did with continuous selections (see Proposition 2.5.26), we can improve Theorem 2.5.33 and produce a whole sequence of measurable selections which is dense in $F(\omega)$ for all $\omega \in \Omega$.

Theorem 2.5.35 *If (Ω, Σ) is a measurable space, X is a Polish space and $F : \Omega \to P_f(X)$, then the following statements are equivalent:*

(a) F is measurable.
(b) There exists a sequence of Σ-measurable functions $f_n : \Omega \to X$ $n \geqslant 1$ such that

$$f_n(\omega) \in F(\omega) \textit{ for all } \omega \in \Omega,\ \textit{all } n \geqslant 1,$$
$$F(\omega) = \overline{\{f_n(\omega)\}}_{n \geqslant 1} \textit{ for all } \omega \in \Omega.$$

Proof $(a) \Rightarrow (b)$. Let $\{x_n\}_{n\geqslant 1}$ be dense in X. For every $k \geqslant 1$ we define

$$F_{nk}(\omega) = \begin{cases} F(\omega) \cap B_{2^{-k}}(x_n) & \text{if } \omega \in F^{-}(B_{2^{-k}}(x_n)) \\ F(\omega) & \text{if } \omega \in \Omega \setminus F^{-1}(B_{2^{-k}}(x_n)). \end{cases} \tag{2.35}$$

Note that the measurability of F implies $F^{-}(B_{2^{-k}}(x_n)) \in \Sigma$ and so $\omega \mapsto \overline{F_{nk}(\omega)}$ is measurable. Then Theorem 2.5.33 implies that we can find a Σ-measurable function $f_{nk} : \Omega \to X$ such that $f_{nk}(\omega) \in \overline{F_{nk}(\omega)}$ for all $\omega \in \Omega$.

We claim that $F(\omega) = \overline{\{f_{nk}(\omega)\}}_{n,k\geqslant 1}$ for all $\omega \in \Omega$. To this end, let $x \in F(\omega)$ and $\epsilon > 0$ be given. We choose $k \in \mathbb{N}$ such that $2^{-k} \leqslant \frac{\epsilon}{2}$ and $n \in \mathbb{N}$ such that $d(x, x_n) < 2^{-k}$. Therefore

$$\omega \in F^{-}(B_{2^{-k}}(x_n)) \text{ and } f_{nk}(\omega) \in \overline{B_{2^{-k}}(x_n)} \text{ (see (2.35)).}$$

So, finally we have

$$d(f_{nk}(\omega), x) \leqslant d(f_{nk}(\omega), x_n) + d(x_n, x) \leqslant \frac{1}{2^k} + \frac{\epsilon}{2} \leqslant \epsilon$$

and this proves the claim.

$(b) \Rightarrow (a)$. For every $x \in X$, we have $d(x, F(\omega)) = \inf_{n\geqslant 1} d(x, f_n(\omega))$ for all $\omega \in \Omega$. Hence the function $\omega \mapsto d(x, F(\omega))$ is Σ-measurable and this by Proposition 2.5.25 implies the measurability of $F(\cdot)$. $\square$

Definition 2.5.36 Let (Ω, Σ) be a measurable space. Given a measure μ on (Ω, Σ), by Σ_μ we denote the μ completion of Σ. Let $\widehat{\Sigma} = \bigcap_{\mu} \Sigma_\mu$ for all finite measures μ.

Remark 2.5.37 The σ-field $\widehat{\Sigma}$ is known as the "universal σ-field". In the definition of $\widehat{\Sigma}$, in the intersection we limit ourselves to finite measures μ, since if μ is σ-finite we can always find a finite measure with the same null sets. If (Ω, Σ, μ) is a σ-finite measure space, then $\widehat{\Sigma} = \Sigma_\mu$.

The next selection theorem is graph conditioned and is known in the literature as the "Yankov–von Neumann–Aumann selection theorem". For its proof we refer to Hu and Papageorgiou [218, p. 158].

Theorem 2.5.38 *If (Ω, Σ) is a measurable space, X is a Souslin space and $F : \Omega \to 2^X \setminus \{\emptyset\}$ is graph measurable, then there exists a $\widehat{\Sigma}$-measurable function $f : \Omega \to X$ such that $f(\omega) \in F(\omega)$ for all $\omega \in \Omega$.*

So, summarizing the situation for measurable multifunctions with closed values, we can state the following theorem.

Theorem 2.5.39 *Let (Ω, Σ) be a measurable space, X a separable metrizable space and $F : \Omega \to P_f(X)$. We consider the following statements:*

(1) F is strongly measurable;
(2) F is measurable;
(3) for every $x \in X$, $\omega \mapsto d(x, F(\omega))$ is Σ-measurable;
(4) there exists a sequence $f_n : \Omega \to X$ $(n \geqslant 1)$ of Σ-measurable selections such that $F(\omega) = \overline{\{f_n(\omega)\}_{n \geqslant 1}}$ for all $\omega \in \Omega$;
(5) F is graph measurable.

We have the following implications:

(a) $(1) \Rightarrow (2) \Leftrightarrow (3) \Rightarrow (5)$.
(b) If X is Polish, then $(2) \Leftrightarrow (3) \Leftrightarrow (4)$.
(c) If X is σ-compact, then $(1) \Leftrightarrow (2)$.
(d) If $\Sigma = \widehat{\Sigma}$ and X is Polish, then (1)–(5) *are equivalent.*

Remark 2.5.40 In case (d) (that is, $\Sigma = \widehat{\Sigma}$ and X Polish), statements (1)–(5) are also equivalent to $F^{-}(A) \in \Sigma$ for all $A \in B(X)$. Indeed, $F^{-}(A) = \{\omega \in \Omega : F(\omega) \cap A \neq \emptyset\} = \operatorname{proj}_{\Omega}[\operatorname{Gr} F \cap (\Omega \times A)] \in \widehat{\Sigma}$ by the projection theorem (see Hu and Papageorgiou [218, p. 149]).

2.6 Monotone Maps: Definition and Basic Results

A function $\varphi : \mathbb{R} \to \mathbb{R}$ is monotone nondecreasing when $t_1 \leqslant t_2$ implies $\varphi(t_1) \leqslant \varphi(t_2)$. Equivalently, we can rewrite this condition as follows

$$(\varphi(t_2) - \varphi(t_1))(t_2 - t_1) \geqslant 0 \text{ for all } t_1, t_2 \in \mathbb{R}.$$

The advantage of this second definition of monotonicity is that it does not employ the order structure on $\mathbb{R}$. Hence, it can be extended to the more general setting of a map from a Banach space X into its dual X^*, by replacing the product with the duality brackets for the pair (X^*, X). So, we make the following definitions.

Definition 2.6.1 Let X be a Banach space, X^* its topological dual, $\langle \cdot, \cdot \rangle$ the duality brackets for the pair (X^*, X) and $A : X \to 2^{X^*}$.

(a) We say that A is monotone if for all $x, u \in X$ and all $x^* \in A(x)$, $u^* \in A(u)$ we have
$$\langle u^* - x^*, u - x \rangle \geqslant 0.$$

(b) We say that A is strictly monotone if for all $x, u \in X$, $x \neq u$ and all $x^* \in A(x)$, $u^* \in A(u)$ we have
$$\langle u^* - x^*, u - x \rangle > 0.$$

(c) The set $\{x \in X : A(x) \neq \emptyset\}$ is called the domain of A and is denoted by $D(A)$; the set $\operatorname{Gr} A = \{(x, x^*) \in X \times X^* : x^* \in A(x)\}$ is called the graph of A; the inverse map $A^{-1} : X^* \to 2^X$ is defined by $A^{-1}(x^*) = \{x \in X : (x, x^*) \in \operatorname{Gr} A\}$.

(d) We say that A is maximal monotone if $\mathrm{Gr}\, A$ is not properly contained in the graph of another monotone map $\widehat{A} : X \to 2^{X^*}$.

Remark 2.6.2 From the definition of maximal monotonicity, we see that $A : X \to 2^{X^*}$ is maximal monotone if and only if the inequality $\langle u^* - x^*, u - x\rangle \geqslant 0$ for all $(x, x^*) \in \mathrm{Gr}\, A$ implies that $(u, u^*) \in \mathrm{Gr}\, A$.

From the remark the following simple observation follows at once.

Proposition 2.6.3 *If X is a reflexive Banach space and $A : X \to 2^{X^*}$, then A is maximal monotone if and only if A^{-1} is maximal monotone.*

Maximal monotone maps exhibit nice surjectivity properties, which of course are very important in the study of nonlinear boundary value problems. For this reason we will focus on them. Before continuing with our study of maximal monotone maps, let us give some typical examples of monotone maps.

Example 2.6.4 (a) Let H be a Hilbert space, $A \in \mathscr{L}(H)$ and $A \geqslant 0$ (that is, $(A(x), x)_H \geqslant 0$ for all $x \in H$). Then A is maximal monotone.

(b) Let H be a Hilbert space and $T : H \to H$ a nonexpansive map, that is,

$$\|T(x) - T(u)\| \leqslant \|x - u\| \text{ for all } x, u \in H.$$

Then $A = I - T$ is maximal monotone map.

(c) Let X be a Banach space and let $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ be a function which is convex lower semicontinuous, not identically $+\infty$. The subdifferential of φ is the set-valued map $\partial\varphi : X \to 2^{X^*}$ defined by

$$\partial\varphi(x) = \{x^* \in X^* : \langle x^*, u - x\rangle \leqslant \varphi(u) - \varphi(x) \text{ for all } u \in X\}.$$

This map is maximal monotone. The subdifferential of convex functions will be examined in more detail in Sect. 2.7.

(d) A monotone nondecreasing function $\varphi : \mathbb{R} \to \mathbb{R}$ is monotone, but to be maximal monotone, we need to have continuity of φ or otherwise pass to a set-valued map by filling in the gaps at the jump discontinuity points.

(e) Let $A : W^{1,p}(\Omega) \to W^{1,p}(\Omega)^*$ $(1 < p < \infty)$ be defined by

$$\langle A(u), y\rangle = \int_\Omega \|Du\|_{\mathbb{R}^N}^{p-2}(Du, Dy)_{\mathbb{R}^N} dz \text{ for all } u, y \in W^{1,p}(\Omega).$$

This map is maximal monotone and corresponds to the p-Laplace differential operator defined by

$$-\Delta_p u = -\mathrm{div}(\|Du\|_{\mathbb{R}^N}^{p-2} Du) \text{ for all } u \in W^{1,p}(\Omega).$$

The monotonicity of A is a consequence of the following elementary inequalities

$$(\|x\|_{\mathbb{R}^n}^{p-2}x - \|v\|_{\mathbb{R}^n}^{p-2}v, x - v)_{\mathbb{R}^m} \geqslant c_p \begin{cases} \dfrac{\|x-v\|_{\mathbb{R}^m}^2}{(1+\|x\|_{\mathbb{R}^m}+\|v\|_{\mathbb{R}^m})^{2-p}} & \text{if } 1 < p < 2 \\ \|x-v\|_{\mathbb{R}^m}^2 & \text{if } 2 \leqslant p \end{cases}$$

for all $x, v \in \mathbb{R}^m$ $(m \geqslant 1)$, with $c_p > 0$ a constant. The maximality of A is a consequence of Proposition 2.6.12 below.

Proposition 2.6.5 *If X is a Banach space and $A : X \to 2^{X^*}$ is a maximal monotone map, then for every $x \in X$, $A(x)$ is convex and w^*-closed in X^* and* Gr A *is closed in $X \times X^*_{w^*}$ and $X_w \times X^*$.*

Proof Let $u_0^*, u_1^* \in A(u)$ and set $u_t^* = (1-t)u_0^* + tu_1^*$ for all $t \in [0, 1]$. For every $(x, x^*) \in \operatorname{Gr} A$ we have

$$\begin{aligned} &\langle u_t^* - x^*, u - x\rangle = (1-t)\langle u_0^* - x^*, u - x\rangle + t\langle u_1^* - x^*, u - x\rangle \geqslant 0 \\ &\Rightarrow u_t^* \in A(u) \text{ (since} A \text{ is maximal monotone, see Remark 2.6.2)} \\ &\Rightarrow A(u) \text{ is convex.} \end{aligned}$$

Let $\{(x_\alpha, x_\alpha^*)\}_{\alpha\in J} \subseteq \operatorname{Gr} A$ be a net such that $x_\alpha \to x$ in X^* and $x_\alpha^* \overset{w^*}{\to} x^*$ in X^*. For every $(u, u^*) \in \operatorname{Gr} A$ we have

$$\begin{aligned} &\langle x_\alpha^* - u^*, x_\alpha - u\rangle \geqslant 0 \text{ for all } \alpha \in J \\ &\Rightarrow \langle x^* - u^*, x - u\rangle \geqslant 0 \text{ for all } (u, u^*) \in \operatorname{Gr} A \\ &\Rightarrow (x, x^*) \in \operatorname{Gr} A \text{ (due to the maximality of} A). \end{aligned}$$

This proves that Gr A is closed in $X \times X^*_{w^*}$ and so $A(x)$ is w^*-closed. Similarly we show the closedness of Gr A in $X_w \times X^*$. □

Definition 2.6.6 Let X be a Banach space and X^* its dual.

(a) A monotone map $A : X \to 2^{X^*}$ is locally bounded at $x \in D(A)$ if we can find $M > 0$ and $r > 0$ such that

$$\|x^*\|_* \leqslant M \text{ for all } x \in D(A) \cap B_r(x) \text{ and all } x^* \in A(x).$$

(b) A subset (not necessarily convex) $C \subseteq X$ which contains the origin is said to be absorbing if $X = \bigcup_{\lambda>0} \lambda C$. Equivalently, C is absorbing if for every $x \in X$ we can find $t > 0$ such that $tx \in C$. A point $x \in C$ is said to be an absorbing point of C if $C - x$ is an absorbing set. The set of absorbing points of C is called the core of C and denoted by core C.

Remark 2.6.7 Evidently every interior point of C is an absorbing point of C (that is, int $C \subseteq$ core C). However, C can have absorbing points even if int $C = \emptyset$. Consider the set $C = \partial B_1 \cup \{0\}$ (where $\partial B_1 = \{x \in X : \|x\| = 1\}$). Then C is absorbing, 0 is an absorbing point of C, but $\operatorname{int} C = \emptyset$.

The next proposition establishes a fundamental property of monotone maps.

Proposition 2.6.8 *If X is a Banach space, X^* its dual and $A : X \to 2^{X^*}$ is monotone, then A is locally bounded at every absorbing point of $D(A)$.*

Proof Let $x^* \in A(x)$ and consider instead of A the monotone map defined by $u \to A(x+u) - x^*$. So, without any loss of generality we may assume that $(0, 0) \in \operatorname{Gr} A$ and 0 is an absorbing point of $D(A)$. Let

$$\varphi(x) = \sup\{\langle u^*, x - u\rangle : u \in D(A), \|u\| \leqslant 1, u^* \in A(u)\}$$

and

$$C = \{x \in X : \varphi(x) \leqslant 1\}.$$

Note that since φ is the supremum of affine continuous functionals, it is convex and lower semicontinuous. So, C is closed, convex and it contains the origin. Since $(0, 0) \in \operatorname{Gr} A$ we see that $\varphi \geqslant 0$. Also, note that for all $(u, u^*) \in \operatorname{Gr} A$ from the monotonicity of A we have

$$\begin{aligned} &\langle u^* - 0, u - 0\rangle \geqslant 0 \\ \Rightarrow\ &\varphi(0) = 0. \end{aligned}$$

Let $D = C \cap (-C)$. This is a closed, symmetric set. We claim that A is absorbing. Recall that $D(A)$ is absorbing. So, if $x \in X$, we can find $t > 0$ such that $A(tx) \neq \emptyset$. Let $u^* \in A(tx)$. If $v \in D(A)$ and $v^* \in A(v)$, then

$$\begin{aligned} \langle v^*, tx - v\rangle &\leqslant \langle u^*, tx - v\rangle \\ \Rightarrow \varphi(tx) &\leqslant \sup\,[\langle u^*, tx - v\rangle : v \in D(A),\ \|v\| \leqslant 1] \\ &\leqslant \langle u^*, tx\rangle + \|u^*\|_*. \end{aligned}$$

Let $\lambda \in (0, 1)$ be such that $\lambda\varphi(tx) < 1$. The convexity of φ implies

$$\begin{aligned} &\varphi(\lambda tx) \leqslant \lambda\varphi(tx) + (1-\lambda)\varphi(0) = \lambda\varphi(tx) < 1 \\ \Rightarrow\ &\lambda tx \in C. \end{aligned}$$

Therefore C is absorbing as claimed and so it is a neighborhood of the origin (by the Baire category theorem). Hence there exists an $r > 0$ such that $\varphi(x) \leqslant 1$ for all $\|x\| \leqslant 2r$. Therefore

$$\|x\| \leqslant 2r \Rightarrow \langle v^*, x\rangle \leqslant \langle v^*, v\rangle \text{for all} v \in D(A) \text{ with } \|v\| \leqslant 1,\ \text{all } v^* \in A(v).$$

Then for $v \in D(A) \cap B_r(0)$ and $v^* \in A(v)$ we have

$$
\begin{aligned}
&2r\|v^*\|_* = \sup\left[\langle v^*, x\rangle : \|x\| \leqslant 2r\right] \leqslant \|v^*\|_*\|v\| + 1 \leqslant r\|v^*\|_* + 1 \\
\Rightarrow\ &\|v^*\|_* \leqslant \frac{1}{r}\,.
\end{aligned}
$$

The proof is now complete. □

Next, we present an important case when a monotone map is in fact maximal monotone.

Proposition 2.6.9 *If X is a Banach space, X^* its dual and $A : X \to 2^{X^*}$ is a monotone map with nonempty, convex and w^*-closed values and A is usc from line segments in X into $X^*_{w^*}$, then A is maximal monotone.*

Proof By virtue of Remark 2.6.2 it suffices to show that

$$\text{“if } \langle u^* - x^*, u - x\rangle \geqslant 0 \text{ for all } (x, x^*) \in \operatorname{Gr} A,\ \text{ then } (u, u^*) \in \operatorname{Gr} A''.$$

Arguing by contradiction, suppose that $u^* \notin A(u)$. Since $A(u)$ is w^*-closed and convex, by the strong separation theorem, we can find $h \in X \setminus \{0\}$ and $\varepsilon > 0$ such that

$$\langle u^*, h\rangle - \varepsilon \geqslant \sup\{\langle \hat{u}^*, h\rangle : \hat{u}^* \in A(u)\}. \tag{2.36}$$

For $t > 0$, let $x_t = u + th$ and let $x_t^* \in A(x_t)$. By virtue of Proposition 2.6.8, we see that for $t > 0$ small $\|A(x_t)\|_* \leqslant M$. So, by Alaoglu's theorem, we may assume that $x_t^* \xrightarrow{w^*} x^*$ in X^*. Then our hypothesis on A and Proposition 2.5.8 imply $x^* \in A(u)$. Since

$$
\begin{aligned}
&\langle x_t^* - u^*, h\rangle \geqslant 0 \text{ for all } t > 0 \\
\Rightarrow\ &\langle x^* - u^*, h\rangle \geqslant 0,
\end{aligned}
$$

which contradicts (2.36). □

When A is single-valued, the above continuity property has a special name.

Definition 2.6.10 An operator $A : X \to X^*$ is said to be hemicontinuous at $x \in D(A)$ if $D(A)$ is convex and for every $u \in D(A)$, the map $t \to A((1-t)x + tu)$ is continuous from $[0, 1]$ into $X^*_{w^*}$. We say that A is hemicontinuous if it is hemicontinuous at every $x \in D(A)$.

Remark 2.6.11 Evidently, the hemicontinuity of A at $x \in D(A)$ is equivalent to the continuity of the map $t \to A(x + tu)$ from $[0, t_0)$ into X_{w^*}.

Then immediately from Proposition 2.6.9, we have:

Proposition 2.6.12 *If X is a Banach space, X^* its dual and $A : X \to X^*$ is a monotone, hemicontinuous operator with $D(A) = X$, then A is maximal monotone.*

Another consequence of Proposition 2.6.8 is the following result.

Proposition 2.6.13 *If X is a Banach space, X^* its dual and $A : X \to 2^{X^*}$ is a maximal monotone map, then for all $x \in \operatorname{int} D(A)$, $A(x)$ is convex and w^* compact and A is usc from X into $X^*_{w^*}$.*

Proof Combining Proposition 2.6.5 and Proposition 2.6.8 it follows that for every $x \in \operatorname{int} D(A)$, $A(x)$ is convex and w^*-compact.

Next let $C \subseteq X^*$ be w^*-closed. According to Proposition 2.5.3 we need to show that $A^-(C) = \{x \in \operatorname{int} D(A) : A(x) \cap C \neq \emptyset\}$ is closed in X. So, let $\{x_\alpha\}_{\alpha\in J}$ be a net in $A^-(C)$ such that $x_\alpha \to x \in \operatorname{int} D(A)$. Then we can find $x^*_\alpha \in A(x_\alpha) \cap C$ for all $\alpha \in J$. On account of Proposition 2.6.8, we may assume that $\{x^*_\alpha\}_{\alpha\in J}$ is bounded. So, thanks to the Alaoglu theorem we may assume that $x^*_\alpha \xrightarrow{w^*} x^*$ in X^*. Evidently $x^* \in C$. Also $(x, x^*) \in \operatorname{Gr} A$ (see Proposition 2.6.5). Therefore $x^* \in A(x) \cap C$ and so $x \in A^-(C)$. This proves the desired upper semicontinuity of A. $\square$

2.7 The Subdifferential and Duality Maps

The subdifferential extends the classical concept of a derivative. Here we focus on the subdifferential of convex functions, since it leads to maximal monotone maps.

From the duality theory of convex analysis, we know that a convex set in a Banach space can be dually described using the notion of a supporting hyperplane. Just recall that a closed convex set is the intersection of all closed half-spaces which contain it. Suppose that the convex set is the epigraph of a convex function φ, that is, $C = \operatorname{epi}\varphi = \{(x, \lambda) \in X \times \mathbb{R} : \varphi(x) \leqslant \lambda\}$ with X being a Banach space. Then the supporting hyperplanes are described by continuous affine functionals minorizing φ. So, let $\xi : X \to \mathbb{R}$ be a continuous affine functional defined by $\xi(x) = \langle x^*, x\rangle + \eta$ with $x^* \in X^*$ and $\eta \in \mathbb{R}$. We require that ξ is an exact minorant of φ at $x \in X$, that is,

$$\begin{aligned} &\xi(u) \leqslant \varphi(u) \text{ for all } u \in X \text{ and } \xi(x) = \varphi(x) \\ \Rightarrow\ \ &\eta = \varphi(x) - \left\langle x^*, x\right\rangle, \text{ and so } \xi(u) = \varphi(x) + \left\langle x^*, u - x\right\rangle. \end{aligned}$$

Therefore we have

$$\left\langle x^*, u - v\right\rangle \leqslant \varphi(u) - \varphi(x) \text{ for all } u \in X.$$

This leads to the following notion already mentioned in Example 2.6.4 (c).

Definition 2.7.1 Let X be a Banach space and $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ be a convex function not identically $+\infty$. The subdifferential of φ is the multifunction $\partial\varphi : X \to 2^{X^*}$ defined by

$$\partial\varphi(x) = \{x^* \in X^* : \left\langle x^*, u - x\right\rangle \leqslant \varphi(u) - \varphi(x) \text{ for all } u \in X\}.$$

The elements of $\partial\varphi(x)$ are called subgradients of φ at x.

Remark 2.7.2 In the terminology of Convex Analysis, $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$-valued functions which are not identically $+\infty$ are called "proper". However, in this book we have reserved this name for a different class of maps (see Definition 2.2.1). In the sequel when considering $\overline{\mathbb{R}}$-valued functions we will always assume that they are not identically $+\infty$. From Definition 2.7.1 it is clear that $\partial\varphi(x)$ is always a w^*-closed, convex set in X^*. It may be empty. The domain of $\partial\varphi$, denoted by $D(\partial\varphi)$, is the set $D(\partial\varphi) = \{x \in X : \partial\varphi(x) \neq \emptyset\}$. Evidently, $D(\partial\varphi) \subseteq \operatorname{dom}\varphi = \{x \in X : \varphi(x) < +\infty\}$ (the effective domain of φ) and this inclusion can be strict. If $x \in D(\partial\varphi)$ then we say that φ is subdifferentiable at x. All this terminology reflects the fact that if $\varphi \in C^1(X)$, then $\langle\varphi'(x), u - x\rangle \leqslant \varphi(u) - \varphi(x)$for all$u \in X$ and this inequality characterizes $\varphi'(x)$. It is easily seen that $\partial\varphi : X \to 2^{X^*}$ is monotone. In fact, it is maximal monotone, but the proof of this basic result is postponed until the next section.

Since the notion of subdifferential is linked with the duality theory of convex functions, to better understand it, we will need some basic definitions and facts from this theory.

Definition 2.7.3 Let X be a Banach space and $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ (recall that we always assume $\varphi \not\equiv +\infty$). The conjugate of φ is the function $\varphi^* : X^* \to \overline{\mathbb{R}}$ defined by

$$\varphi^*(x^*) = \sup\{\big\langle x^*, x\big\rangle - \varphi(x) : x \in X\}.$$

Remark 2.7.4 Let $x_0 \in \operatorname{dom}\varphi$. Then $\varphi^*(x^*) \geqslant \langle x^*, x_0\rangle - \varphi(x_0)$. So, φ^* admits a continuous affine minorant and therefore we see that φ^* cannot take the value $-\infty$ (that is, φ^* is $\overline{\mathbb{R}}$-valued). If $u^* \in \operatorname{dom}\varphi^*$, then we can find $\eta \in \mathbb{R}$ such that

$$\begin{aligned} &\big\langle u^*, u\big\rangle - \varphi(u) \leqslant \eta \text{ for all } u \in X \\ \Rightarrow\ &\big\langle u^*, u\big\rangle - \eta \leqslant \varphi(u) \text{ for all } u \in X. \end{aligned}$$

Recall that if φ is convex and lower semicontinuous, then φ admits such a continuous affine minorant. Therefore for φ convex and lower semicontinuous, φ^* is not identically $+\infty$ and being the supremum of affine continuous functionals, it is convex and lower semicontinuous.

This remark suggests that we focus on convex and lower semicontinuous functions. A basic result for this class of functions is the following (see, for example, Ioffe and Tichomirov [221, p. 177]).

Proposition 2.7.5 *If X is a Banach space and $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is convex and lower semicontinuous, then φ is the supremum of all continuous affine functionals minorizing φ.*

Since $\varphi^* : X^* \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is convex and lower semicontinuous, we can compute its conjugate $\varphi^{**} : X^{**} \to \overline{\mathbb{R}}$, which is convex and lower semicontinuous. Exploiting the canonical embedding of X into X^{**}, we can restrict φ^{**} to X to obtain

$$\varphi^{**}(u) = \sup\{\langle u^*, u\rangle - \varphi^*(u^*) : u^* \in X^*\} \text{ for all } u \in X.$$

Then we can reformulate Proposition 2.7.5 as follows:

Proposition 2.7.6 *If X is a Banach space and $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is convex and lower semicontinuous, then $\varphi = \varphi^{**}$ on X.*

Example 2.7.7 Let X be a normed space and $C \subseteq X$ a nonempty, closed convex set. The indicator function of C is the function defined by

$$\delta_C(x) = \begin{cases} 0 & \text{if } x \in C \\ +\infty & \text{if } x \in X \setminus C. \end{cases}$$

This function is convex and lower semicontinuous. We have

$$(\delta_C)^*(x^*) = \sup\{\langle x^*, x\rangle - \delta_C(x) : x \in X\} = \sup\{\langle x^*, x\rangle : x \in C\} = \sigma_C(x^*).$$

So, the conjugate of δ_C is the support function of C. From Proposition 2.7.6 we have

$$\begin{aligned} &\delta_C = (\delta_C)^{**} \text{ on } X \\ \Rightarrow\ &\delta_C(u) = \sup\{\langle u^*, u\rangle - \sigma_C(u^*) : u^* \in X^*\}. \end{aligned}$$

Note that $u \in C$ if and only if $\delta_C(u) = 0$. Hence

$$C = \{x \in X : \langle x^*, x\rangle \leqslant \sigma_C(x^*) \text{ for all } x^* \in X^*\}.$$

Now we return to the study of the subdifferential.

Proposition 2.7.8 *If X is a normed space and $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is convex and lower semicontinuous, then the following conditions are equivalent:*

(a) $x^* \in \partial\varphi(x)$;
(b) $\varphi(x) + \varphi^*(x^*) = \langle x^*, x\rangle$.

Proof $(a) \Rightarrow (b)$ From Definition 2.7.3 we see that we have

$$\varphi^*(x^*) \geqslant \langle x^*, x\rangle - \varphi(x). \tag{2.37}$$

Since $x^* \in \partial\varphi(x)$, we have $\langle x^*, u - x\rangle \leqslant \varphi(u) - \varphi(x)$ for all $u \in X$. Then

$$\begin{aligned} &\langle x^*, u\rangle - \varphi(u) \leqslant \langle x^*, x\rangle - \varphi(x) \text{ for all } u \in X \\ \Rightarrow\ &\varphi^*(x^*) \leqslant \langle x^*, x\rangle - \varphi(x) \\ \Rightarrow\ &\varphi(x) + \varphi^*(x^*) = \langle x^*, x\rangle \,(\text{see } (2.37)). \end{aligned}$$

$(b) \Rightarrow (a)$ The continuous affine function

$$u \to \xi(u) = \langle x^*, u \rangle + \varphi(x) - \langle x^*, x \rangle$$

minorizes φ (note that the constant term is by hypothesis equal to $\varphi^*(x^*)$) and it is exact at x. Therefore $x^* \in \partial\varphi(x)$. $\square$

Proposition 2.7.9 *If X is a Banach space and $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is convex and lower semicontinuous, then $\langle x, x^* \rangle \in \mathrm{Gr}\,\partial\varphi$ if and only if $(x^*, x) \in \mathrm{Gr}\,\partial\varphi^*$.*

Proof We have

$$\begin{aligned}
(x, x^*) \in \mathrm{Gr}\,\partial\varphi \quad &\Leftrightarrow\; x^* \in \partial\varphi(x) \\
&\Leftrightarrow\; \varphi(x) + \varphi^*(x^*) = \langle x^*, x \rangle \;\; \text{(see Proposition 2.7.8)} \\
&\Leftrightarrow\; \varphi^{**}(x) + \varphi^*(x^*) = \langle x^*, x \rangle \;\; \text{(see Proposition 2.7.6)} \\
&\Leftrightarrow\; x \in \partial\varphi^*(x^*) \;\; \text{(again Prop. 2.7.8)}.
\end{aligned}$$

The proof is now complete. $\square$

Remark 2.7.10 Using the notation introduced in Definition 2.6.1, we have $(\partial\varphi)^{-1} = \partial\varphi^*$.

The next proposition provides a simple criterion for subdifferentiability of convex functions.

Proposition 2.7.11 *If X is a Banach space, $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is convex, $x \in \mathrm{dom}\,\varphi$ and φ is continuous at x, then $\partial\varphi(u) \neq \emptyset$ for all $u \in \mathrm{int\,dom}\,\varphi$ and in particular $\partial\,\varphi(x) \neq \emptyset$.*

Proof Since φ is continuous at x, we have that $\mathrm{int\,dom}\,\varphi \neq \emptyset$ and $\varphi\,|_{\mathrm{int\,dom}\,\varphi}$ is continuous (in fact locally Lipschitz, see for example Gasinski and Papageorgiou [182, p. 494]). So, it suffices to show that $\partial\varphi(x) \neq \emptyset$.

The continuity of φ at x implies that $\mathrm{int\,epi}\,\varphi \neq \emptyset$ (recall that $\mathrm{epi}\,\varphi = \{(x, \lambda) \in X \times \mathbb{R} : \varphi(x) \leqslant \lambda\}$) (the epigraph of φ). Since $(x, \varphi(x)) \in \partial(\mathrm{epi}\,\varphi)$, by the weak separation theorem we can find $x^* \in X^*$ and $\eta, \mu \in \mathbb{R}$ not all zero such that

$$\langle x^*, u \rangle + \mu\lambda \geqslant \eta = \langle x^*, x \rangle + \mu\varphi(x) \text{for all } (u, \lambda) \in \mathrm{epi}\,\varphi.$$

If $\mu = 0$, then

$$\begin{aligned}
&\langle x^*, u - x \rangle \geqslant 0 \;\text{ for all } u \in \mathrm{dom}\,\varphi \\
\Rightarrow\;& x^* = 0 \text{ (dom } \varphi \text{ contains a neighborhood of } x) \\
\Rightarrow\;& x^* = 0 \text{ and } \mu = \eta = 0, \text{ a contradiction.}
\end{aligned}$$

Therefore $\mu > 0$ and then dividing by $\mu > 0$, we have

$$\frac{\eta}{\mu} - \left\langle \frac{1}{\mu} x^*, u \right\rangle \leqslant \varphi(u) \text{ for all } u \in \operatorname{dom} \varphi, \tag{2.38}$$

$$\frac{\eta}{\mu} - \left\langle \frac{1}{\mu} x^*, x \right\rangle = \varphi(x). \tag{2.39}$$

From (2.38) and (2.39) it follows that

$$\left\langle -\frac{1}{\mu} x^*, u - x \right\rangle \leqslant \varphi(u) - \varphi(x) \text{ for all } u \in X$$
$$\Rightarrow -\frac{1}{\mu} x^* \in \partial\varphi(x), \text{ hence } \partial\varphi(x) \neq \emptyset.$$

The proof is now complete. □

Corollary 2.7.12 *If X is a Banach space and $\varphi : X \to \mathbb{R}$ is convex and lower semi-continuous, then* $\operatorname{int} \operatorname{dom} \varphi \subseteq D(\partial\varphi) \subseteq \operatorname{dom} \varphi$.

Proof Just recall that $\varphi|_{\operatorname{int} \operatorname{dom} \varphi}$ is continuous and use Proposition 2.7.11. □

As we already mentioned the subdifferential is a generalization of the classical derivative. In the next results, we make this more precise.

Definition 2.7.13 Let X be a Banach space and $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$. We say that

$$\varphi'(x; h) = \lim_{\lambda \to 0^+} \frac{\varphi(x + \lambda h) - \varphi(x)}{\lambda},$$

if it exists, is the directional derivative of φ at x in the direction h.

Remark 2.7.14 If φ is convex and $x \in \operatorname{dom} \varphi$, then the quotient $[\varphi(x + \lambda h) - \varphi(x)]/\lambda$ is nondecreasing as a function of λ and so the limit exists and we have

$$\varphi'(x; h) = \lim_{\lambda \to 0^+} \frac{\varphi(x + \lambda h) - \varphi(x)}{\lambda} = \inf_{\lambda > 0} \frac{\varphi(x + \lambda h) - \varphi(x)}{\lambda}. \tag{2.40}$$

Moreover, for all $x \in \operatorname{dom} \varphi$, the mapping $h \mapsto \varphi'(x; h)$ is sublinear (that is, positively homogeneous and subadditive). If φ is Gâteaux differentiable, then $\varphi'(x; \cdot) \in X^*$.

Proposition 2.7.15 *If X is a Banach space and $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is convex and continuous at $x \in X$, then*

$$\varphi'(x; h) = \sigma_{\partial\varphi(x)}(h) = \sup\{\langle x^*, h \rangle : x^* \in \partial\varphi(x)\} \textit{ for all } h \in X.$$

Proof Let $\hat{\sigma}_x(h) = \varphi'(x; h)$ and $\vartheta_\lambda(h) = \frac{\varphi(x+\lambda h)-\varphi(x)}{\lambda}$ for all $h \in X$ and all $\lambda > 0$. We have

$$\hat{\sigma}_x = \inf_{\lambda > 0} \vartheta_\lambda \text{ (see (2.40))}.$$

Directly from Definition 2.7.3 we obtain

$$(\hat{\sigma}_x)^* = \sup_{\lambda>0} \vartheta_\lambda^*. \tag{2.41}$$

For every $h^* \in X^*$ we have

$$\begin{aligned}\vartheta_\lambda^*(h^*) &= \sup\left\{\langle h^*, h\rangle - \frac{1}{\lambda}(\varphi(x+\lambda h) - \varphi(x)) : h \in X\right\}\\ &= \sup \frac{1}{\lambda}\left\{\langle h^*, u\rangle - \varphi(u) + \varphi(x) - \langle h^*, x\rangle : u \in X\right\}\\ &= \frac{1}{\lambda}\left\{\varphi^*(h^*) + \varphi(x) - \langle h^*, x\rangle\right\}.\end{aligned} \tag{2.42}$$

According to Proposition 2.7.8, we have

$$\partial\varphi(x) = \{h^* \in X^* : \varphi^*(h^*) + \varphi(x) = \langle h^*, x\rangle\}.$$

Therefore from (2.41) and (2.42) it follows that

$$(\hat{\sigma}_x)^*(h^*) = \begin{cases} 0 & \text{if } h^* \in \partial\varphi(x)\\ +\infty & \text{if } h^* \notin \partial\varphi(x)\end{cases}$$

$$\Rightarrow (\hat{\sigma}_x)^* = \delta_{\partial\varphi(x)}. \tag{2.43}$$

Note that

$$\langle u^*, h\rangle \leqslant \varphi'(x; h) \leqslant \varphi(x+h) - \varphi(x) \text{ for all } u^* \in \partial\varphi(x) \text{ and all } h \in X.$$

The continuity of $\varphi(\cdot)$ at x implies that $\varphi'(x; \cdot)$ is finite and continuous on X. Therefore, we have

$$\begin{aligned}&(\hat{\sigma}_x)^{**} = \hat{\sigma}_x \text{ (see Proposition 2.7.6)}\\ \Rightarrow\ &\hat{\sigma}_x(h) = \sigma_{\partial\varphi(x)}(h) \text{ for all } h \in X \text{ (see Example 2.7.7).}\end{aligned}$$

The proof is now complete. $\square$

Recall that a convex and lower semicontinuous function $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is continuous in $\operatorname{int}\operatorname{dom}\varphi$. In fact an argument similar to that in the proof of Proposition 2.6.8 reveals that φ is continuous on $\operatorname{core}\operatorname{dom}\varphi$ (see Definition 2.6.6).

This observation leads to the following result, which is another illustration of the power of the convexity condition, which although a purely algebraic condition, leads to remarkable topological conclusions.

Proposition 2.7.16 *If X is a Banach space and $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is convex, lower semicontinuous and also Gâteaux differentiable at x, then φ is continuous at x.*

Remark 2.7.17 For general functions $\varphi : X \to \mathbb{R}$ it is well-known that Gâteaux differentiability does not imply continuity. To see this, let us consider the function

$$\psi(x,u) = \begin{cases} \dfrac{x^6}{x^8 + (u - x^2)^2} & \text{if } (x,u) \neq (0,0) \\ 0 & \text{if } (x,u) = (0,0). \end{cases}$$

The Gâteaux derivative of ψ at $(0, 0)$ exists and it is zero. However, $\psi(x, x^2) = x^{-2}$ and so ψ is not continuous (and a fortiori not differentiable) at the origin.

Proposition 2.7.18 *If X is a Banach space and $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is convex and Gâteaux differentiable at x, then $\varphi'(x) \in \partial\varphi(x)$.*

Proof We have that $\langle \varphi'(x), h\rangle = \lim\limits_{\lambda \to 0^+} \frac{\varphi(x+\lambda h)-\varphi(x)}{\lambda}$. Then the convexity of φ implies that for all $h \in \partial\mathscr{B}_1 = \{x \in X : \|x\| = 1\}$ and all $t > 0$, we have

$$\begin{aligned} & \langle \varphi'(x), th\rangle \leqslant \varphi(x + th) - \varphi(x) \\ \Rightarrow\ & \langle \varphi'(x), u\rangle \leqslant \varphi(x + u) - \varphi(x) \ \text{ for all } u \in X \\ \Rightarrow\ & \varphi'(x) \in \partial\varphi(x) \ \text{ (see Definition 2.7.1).} \end{aligned}$$

The proof is now complete. □

Proposition 2.7.19 *If X is a Banach space and $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is a convex function continuous at x, then φ is Gâteaux differentiable at x if and only if $\partial\varphi(x) \subseteq X^*$ is a singleton.*

Proof $\underline{\Rightarrow}$ If φ is Gâteaux differentiable at x, then $\varphi'(x) \in \partial\varphi(x)$ (see Proposition 2.7.18). Suppose that $u^* \in \partial\varphi(x)$ and $u^* \neq \varphi'(x)$. Then, for some $h \in \partial\mathscr{B}_1$, we have

$$\begin{aligned} & \langle \varphi'(x), h\rangle < \left\langle u^*, h\right\rangle \leqslant \frac{1}{\lambda}[\varphi(x + \lambda h) - \varphi(x)] \ \text{for all } \lambda > 0 \\ \Rightarrow\ & \langle \varphi'(x), h\rangle < \varphi'(x; h), \ \text{ a contradiction.} \end{aligned}$$

$\underline{\Leftarrow}$ Suppose that $\partial\varphi(x) = \{x^*\}$. Then from Proposition 2.7.15 we have

$$\varphi'(x; h) = \left\langle x^*, h\right\rangle \text{for all } h \in X,$$

hence φ is Gâteaux differentiable at $x \in X$ and $\varphi'(x) = x^*$. □

In general, if $\varphi, \psi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ are convex and lower semicontinuous, then $\partial\varphi(x) + \partial\psi(x) \subseteq \partial(\varphi + \psi)(x)$ for all $x \in X$. The inclusion may be strict. However, we have a simple situation where equality holds.

Proposition 2.7.20 *If X is a Banach space, $\varphi, \psi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ are convex and lower semicontinuous and there exists an $\overline{x} \in dom\, \varphi \cap dom\, \psi$ such that one of them is continuous, then $\partial(\varphi + \psi)(x) = \partial\varphi(x) + \partial\psi(x)$ for all $x \in X$.*

Proof As we already mentioned, it is an easy consequence of Definition 2.7.1 that $\partial\varphi(x) + \partial\psi(x) \subseteq \partial(\varphi + \psi)(x)$ for all $x \in X$. Let $x^* \in \partial(\varphi + \psi)(x)$ and assume that φ is continuous at $\overline{x}$. We have

$$\left\langle x^*, u - x \right\rangle \leqslant \tau(u) - \tau(x) \text{ for all } u \in X, \quad \text{where } \tau = \varphi + \psi.$$

We introduce the sets $C_1 = \{(u, \lambda) \in X \times \mathbb{R} : \varphi(u) - \langle x^*, u - x\rangle - \varphi(x) \leqslant \lambda\}$ and $C_2 = \{(u, \lambda) \in X \times \mathbb{R} : \lambda \leqslant \psi(x) - \psi(u)\}$. Both sets are convex, the continuity of φ at $\overline{x}$ implies $\operatorname{int} C_1 \neq \emptyset$ and $\operatorname{int} C_1 \cap C_2 = \emptyset$. So, by the weak separation theorem we can find $u^* \in X^* \setminus \{0\}$ and $\eta \in \mathbb{R}$ such that

$$\psi(x) - \psi(u) \leqslant \left\langle u^*, u \right\rangle + \eta \leqslant \varphi(u) - \left\langle x^*, u - x \right\rangle - \varphi(x) \text{ for all } u \in X.$$

Choosing $u = x$, we obtain $\eta = -\langle u^*, x\rangle$ and so

$$\begin{aligned} &\left\langle -u^*, u - x \right\rangle \leqslant \psi(u) - \psi(x) \text{ for all } u \in X, \\ &\left\langle u^* + x^*, u - x \right\rangle \leqslant \varphi(u) - \varphi(x) \text{ for all } u \in X \\ \Rightarrow\ & -u^* \in \partial\psi(x) \text{ and } u^* + x^* \in \partial\varphi(x) \\ \Rightarrow\ & x^* = x_1^* + x_2^* \text{ with } x_1^* = u^* + x^* \in \partial\varphi(x)\,,\ x_2^* = -u^* \in \partial\psi(x). \end{aligned}$$

We have proved the desired equality. □

In this last part of this section, we focus on the duality map. Duality maps have become an important tool in nonlinear analysis, in particular in connection with monotone operators, and the geometry of Banach spaces.

Definition 2.7.21 Let X be a Banach space and let $\varphi : X \to \mathbb{R}$ be the convex function defined by $\varphi(x) = \frac{1}{2}\|x\|^2$. The map $x \to J(x) = \partial\varphi(x)$ is the duality map for the Banach space.

Remark 2.7.22 The continuity of φ implies that $\operatorname{dom} J = X$. Also, $J(\cdot)$ is monotone and, in fact, maximal monotone as we will prove in the next section.

Proposition 2.7.23 *We have*

$$J(x) = \{x^* \in X^* : \left\langle x^*, x \right\rangle = \|x\|^2 \quad \textit{and} \quad \|x^*\|_* = \|x\|\}.$$

Proof Let $\tau(x) = \|x\|$ for all $x \in X$. Then $\varphi'(x; h) = \|x\|\tau(x; h)$ for all $x, h \in X$. So, if $x = 0$, then

$$\begin{aligned} &\varphi'(x; h) = 0 \text{ for all } h \in X \\ \Rightarrow\ &\partial\varphi(0) = 0 \text{ (see Proposition 2.7.15).} \end{aligned}$$

So, suppose that $x \neq 0$. Then $x^* \in \partial\varphi(x)$ if and only if $\langle x^*, h\rangle \leqslant \varphi'(x; h)$ for all $h \in X$. Hence $x^* \in \partial\varphi(x)$ if and only if $\left\langle \frac{x^*}{\|x\|}, h\right\rangle \leqslant \tau'(x; h)$ for all $h \in X$. The latter inequality is equivalent to saying that $\frac{x^*}{\|x\|} = u^* \in \partial\tau(x)$. Therefore $x^* \in \partial\varphi(x)$ if and only if $\langle u^*, y - x\rangle \leqslant \tau(y) - \tau(x)$ for all $y \in X$, with $u^* = \frac{x^*}{\|x\|}$.

In the last inequality, we choose $y = x + v$ with $\|v\| \leqslant 1$. Then

$$\begin{aligned}&\left\langle u^*, v\right\rangle \leqslant \tau(x+v) - \tau(x) \leqslant \|v\| \text{ (by the triangle inequality)}\\ \Rightarrow\ &\|u^*\|_* \leqslant 1.\end{aligned}$$

Also, if we take $y = 0$, then $\|x\| \leqslant \langle u^*, x\rangle \leqslant \|u^*\|_*\|x\|$, hence $1 \leqslant \|u^*\|_*$. So, we conclude that $\|u^*\|_* = 1$ and $\langle u^*, x\rangle = \|x\|$. Therefore $\langle x^*, x\rangle = \|x\|^2$ and $\|u^*\|_* = \|x\|$.

Conversely, let $\|u^*\|_* = 1$ and $\langle x^*, x\rangle = \|x\|$. Then for all $y \in X$ we have

$$\begin{aligned}&\left\langle u^*, y - x\right\rangle \leqslant \|y\| - \|x\| = \tau(y) - \tau(x)\\ \Rightarrow\ &u^* \in \partial\tau(x) = \frac{1}{\|x\|}\,\partial\varphi(x).\end{aligned}$$

The proof is now complete. □

Remark 2.7.24 In general, the duality map is multivalued. Also, the above proposition shows that the duality map depends on the norm considered on X or on X^*. So, if we consider an equivalent norm on either of the spaces, the duality map changes.

Definition 2.7.25 (a) A Banach space X is said to be strictly convex (rotund) if for all $x, u \in X$, $x \neq u$, $\|x\| = 1 = \|u\|$ we have

$$\|\lambda x + (1-\lambda)u\| < 1 \quad \text{for all} \quad \lambda \in (0, 1).$$

(b) A Banach space X is said to be smooth if for every $x \in X$ with $\|x\| = 1$ there exists a unique $x^* \in X^*$ such that $\|x^*\| = 1$ and $\langle x^*, x\rangle = 1$.

Remark 2.7.26 Evidently, X is strictly convex if the boundary of the unit ball has no flat parts. On the other hand, X is smooth if $J(x)$ is single-valued for all $x \neq 0$ and this is equivalent to saying that the norm of X is Gâteaux differentiable at every $x \neq 0$. We know that if X^* is strictly convex (resp. smooth), then X is smooth (resp. strictly convex), see Megginson [295, p. 481].

Proposition 2.7.27 *If X is a reflexive Banach space with strictly convex dual X^*, then the duality map $J : X \to X^*$ is single-valued, maximal monotone bounded (that is, maps bounded sets to bounded sets) and coercive (that is, $\|J(x)\|_* \to +\infty$ as $\|x\| \to +\infty$).*

Proof Let $x_1^*, x_2^* \in J(x)$. Then from Proposition 2.7.23 we have

$$\left\langle x_k^*, x\right\rangle = \|x\|^2 = \|x_k^*\|_*^2 \text{ for } k = 1, 2.$$

Hence

$$
\begin{aligned}
& 2\|x_1^*\|_*\|x\| \leqslant \|x_1^*\|_*^2 + \|x_2^*\|_*^2 = \langle x_1^* + x_2^*, x \rangle \leqslant \|x_1^* + x_2^*\|_*\|x\| \\
\Rightarrow \ & \|x_1^*\|_* = \|x_2^*\|_* \leqslant \frac{1}{2}\|x_1^* + x_2^*\|_* .
\end{aligned}
$$

The strict convexity of X^* implies that $x_1^* = x_2^*$.

Suppose that $x_n \to x$ in X. Then $\{J(x_n)\}_{n\geqslant 1} \subseteq X^*$ is bounded. Since X is reflexive from the Eberlein–Smulian theorem, by passing to a subsequence if necessary, we may assume that $J(x_n) \xrightarrow{w} u^*$ in X^*. Then

$$
\langle u^*, x \rangle = \lim_{n\to\infty} \langle J(x_n), x_n \rangle = \lim_{n\to\infty} \|x_n\|^2 = \|x\|^2. \tag{2.44}
$$

Also, for every $h \in X$ we have

$$
\langle u^*, h \rangle = \lim_{n\to\infty} \langle J(x_n), h \rangle = \lim_{n\to\infty} \|x_n\|\|h\| = \|x\|\|h\|. \tag{2.45}
$$

From (2.44), (2.45) and Proposition 2.7.23, we infer that $u^* = J(x)$. By Urysohn's criterion, for the original sequence we have $J(x_n) \xrightarrow{w} J(x)$ in X^*. Since $J(\cdot)$ is monotone (see Remark 2.7.2), from Proposition 2.6.12 we infer that $J(\cdot)$ is maximal monotone.

Finally, since $\|J(x)\|_* = \|x\|$ for all $x \in X$, we conclude that $J(\cdot)$ is bounded and coercive. □

Remark 2.7.28 Here we proved the maximal monotonicity using Corollary 2.7.12. In fact, the maximal monotonicity is a consequence of the maximality of the subdifferential map. However this important fact will be proved in the next section. Moreover, the results of the next section will imply $J(\cdot)$ is also surjective.

Definition 2.7.29 (a) A Banach space X is said to be uniformly convex if for all sequences $\{x_n\}_{n\geqslant 1}, \{u_n\}_{n\geqslant 1} \subseteq X$ such that $\|x_n\| = \|u_n\| = 1$ and $\|x_n + u_n\| \to 2$, we have $\|x_n - u_n\| \to 0$ as $n \to \infty$.

(b) A Banach space X is said to be locally uniformly convex if for every $\|x\| = 1$ and every sequence $\{x_n\}_{n\geqslant 1} \subseteq X$ such that $\|x_n\| = 1$ for all $n \geqslant 1$ and $\|x_n + x\| \to 2$ as $n \to \infty$ we have $\|x_n - x\| \to 0$ as $n \to \infty$.

Remark 2.7.30 By the Milman–Pettis theorem (see Megginson [295, p. 452]) every uniformly convex Banach space is reflexive. Using the parallelogram identity, we see that every Hilbert space is uniformly convex. Locally uniformly convex Banach spaces exhibit the so-called Kadec–Klee property. Namely, if $x_n \xrightarrow{w} x$ in X and $\|x_n\| \to \|x\|$, then $x_n \to x$ in X. Of course, a uniformly convex Banach space is locally uniformly convex. The converse is not true in general. To see this let $\{p_n\}_{n\geqslant 1} \subseteq (1, +\infty)$, $\|\cdot\|_{p_n}$ be the usual norm of l^{p_n} and let

$$X = \left\{ (x_n)_{n \geqslant 1} : x_n \in l^{p_n} \text{ and } \|(x_n)\| = (\sum_{n \geqslant 1} \|(x_n)\|_{p_n}^2)^{1/2} < \infty \right\}.$$

This Banach space is locally uniformly convex, but it is uniformly convex if and only if $p_n \in [a, b]$ for all $n \geqslant 1$ with $1 < a \leqslant b < \infty$ (see Day [138, p. 146]). In general, we have the following implications between Banach spaces:

$$\begin{array}{ccc} X = \text{Uniformly convex} & \longrightarrow & X = \text{locally Uniformly Convex} \\ \downarrow & & \downarrow \\ X = \text{Reflexive} & & X = \text{Strictly convex} \\ \updownarrow & & \\ X^* = \text{Reflexive} & & \end{array}$$

Proposition 2.7.31 *If X is a reflexive Banach space with locally uniformly convex dual X^*, then the duality map $J : X \to X^*$ is continuous.*

Proof Let $x_n \to x$ in X. Then $\|J(x_n)\|_* \to \|J(x)\|_*$. Also, from proof of Proposition 2.7.27, we know that J is sequentially continuous from X into X_w^*. Hence $J(x_n) \xrightarrow{w} J(x)$ in X^*. Then from the Kadec–Klee property (see Remark 2.7.30) we have that $J(x_n) \to J(x)$ in X^* and this proves the continuity of $J(\cdot)$. $\square$

Proposition 2.7.32 *If X is a reflexive Banach space with locally uniformly convex dual X^*, then the norm functional $\tau(u) = \|u\|$ for all $u \in X$ is Fréchet differentiable on $X \setminus \{0\}$ and $\tau'(u) = \frac{J(u)}{\|u\|}$ for all $u \in X \setminus \{0\}$.*

Proof Recall that $J(x) = \varphi'(x) = \|x\|\tau(x)$ for all $x \in X$ (see the proof of Proposition 2.7.23). From Proposition 2.7.31 we know that J is continuous, hence $x \to \tau'(x)$ is continuous on $X \setminus \{0\}$ and Gâteaux differentiable. Therefore we conclude that $\tau(\cdot)$ is Fréchet differentiable on $X \setminus \{0\}$. Moreover we have $\tau'(x) = \frac{J(x)}{\|x\|}$ for all $x \in X \setminus \{0\}$. $\square$

Proposition 2.7.33 *If X is a reflexive Banach space and both X and its dual X^* are locally uniformly convex, then the duality map $J : X \to X^*$ is a homeomorphism.*

Proof From Proposition 2.7.31 we know that J is continuous. Also, from Remark 2.7.28 we know that J is surjective. Identifying $X = X^{**}$ (recall that X is reflexive) we can consider the duality map $\hat{J} : X^* \to X$, which is continuous (see Proposition 2.7.31).

We have

$$\langle x^*, \hat{J}(x^*) \rangle = \|x^*\|_*^2 \text{ and } \|x^*\|_* = \|\hat{J}(x^*)\|.$$

Then from 2.7.23 and 2.7.27, we infer that

$$\begin{aligned} & J(\hat{J}(x^*)) = x^* \text{ and } \hat{J}(J(x)) = x \text{ for all } x \in X \\ \Rightarrow\ & J^{-1} = \hat{J}, \text{ which is continuous by virtue of Proposition 2.7.31} \\ \Rightarrow\ & J \text{ is a homeomorphism.} \end{aligned}$$

This completes the proof. $\square$

Remark 2.7.34 As a byproduct of the above proof, we have that in the setting of Proposition 2.7.33, $J^{-1} : X^* \to X$ is the duality map of X^*.

Proposition 2.7.35 *If X is a reflexive Banach space with X^* uniformly convex, then the duality map $J : X \to X^*$ is uniformly continuous on bounded sets of X.*

Proof We first show that J is uniformly continuous on $\partial\mathscr{B}_1 = \{x \in X : \|x\| = 1\}$. Arguing by contradiction, suppose that $J\,|_{\partial\mathscr{B}_1}$ is not uniformly continuous. Then we can find $\varepsilon > 0$ and $\{x_n\}_{n\geqslant 1}, \{u_n\}_{n\geqslant 1} \subseteq \partial\mathscr{B}_1$ such that

$$\|x_n - u_n\| \to 0 \text{ as } n \to \infty \text{ and } \|J(x_n) - J(u_n)\|_* \geqslant \varepsilon \text{ for all } n \geqslant 1. \tag{2.46}$$

For all $x, u \in X$ we have

$$\begin{aligned} \|J(x) + J(u)\|_* \, \|x\| &\geqslant \langle J(x) + J(u), x\rangle \\ &= \langle J(x), x\rangle + \langle J(u), u\rangle + \langle J(u), x - u\rangle \\ &= \|x\|^2 + \|u\|^2 - \|u\|\|x - u\|. \end{aligned} \tag{2.47}$$

In (2.47) let $x = x_n$ and $u = u_n$, $n \geqslant 1$. Then

$$\frac{1}{2}\|J(x_n) + J(u_n)\|_* \geqslant 1 - \frac{1}{2}\|x_n - u_n\|\,. \tag{2.48}$$

Note that $\|J(x_n)\|_* = 1 = \|J(u_n)\|_*$ for all $n \geqslant 1$. Then by virtue of (2.47) we see that (2.48) contradicts the uniform convexity of X^*. This proves the uniform continuity of $J\,|_{\partial\mathscr{B}_1}$.

Next, note that $J(\lambda x) = \lambda J(x)$ for all $\lambda > 0$ and all $x \in X$. Then for all $x, u \in X \setminus \{0\}$ we have

$$\begin{aligned} \|J(x) - J(u)\|_* &= \Big\| \|x\| J\Big(\frac{x}{\|x\|}\Big) - \|u\| J\Big(\frac{u}{\|u\|}\Big)\Big\|_* \\ &\leqslant \|x\|\, \Big\|J\Big(\frac{x}{\|x\|}\Big) - J\Big(\frac{u}{\|u\|}\Big)\Big\|_* + \big|\|x\| - \|u\|\big|\, \Big\|J\Big(\frac{u}{\|u\|}\Big)\Big\|. \end{aligned} \tag{2.49}$$

Then from the first part of the proof and (2.49), we see that $J(\cdot)$ is uniformly continuous on bounded sets not containing the origin. Since $J(0) = 0$ and $J(\cdot)$ is continuous at the origin, we conclude the uniform continuity of $J(\cdot)$ on any bounded subset of X. □

We conclude with a renorming theorem due to Troyanski [405], which as we will see in the next sections is a valuable tool in the study of maximal monotone operators.

Theorem 2.7.36 (Troyanski) *If X is a reflexive Banach space, then there exist equivalent norms on X and X^* such that both spaces (which remain dual to each other) are locally uniformly convex.*

2.8 Surjectivity and Characterizations of Maximal Monotonicity

The power of maximal monotone operators comes from their surjectivity properties. Surjectivity results correspond to existence results for nonlinear boundary value problems. In this section, we present the main surjectivity results involving maximal monotone maps. The approach is to start with the finite-dimensional case, establish the result there and then use Galerkin approximations to pass to the infinite-dimensional case. For this reason, we start with a finite-dimensional result. First we formally define a notion already encountered in Sect. 2.2 in the context of single-valued proper maps.

Definition 2.8.1 Let X be a Banach space and $A : X \to 2^{X^*}$ a map.

(a) We say that A is coercive if $D(A)$ is bounded or $D(A)$ is unbounded and $\inf\{\, \|x^*\|_* : x^* \in A(x)\} \to +\infty$ as $\|x\| \to \infty$, $x \in D(A)$.
(b) We say that A is strongly coercive if $D(A)$ is bounded or $D(A)$ is unbounded and
$$\frac{\inf\{\langle x^*, x\rangle : x^* \in A(x)\}}{\|x\|} \to +\infty \text{ as } \|x\| \to \infty,\ x \in D(A).$$

Remark 2.8.2 Evidently, strong coercivity implies coercivity. The duality map $J : X \to 2^{X^*}$ is strongly coercive (see Proposition 2.7.23). Also, note that coercivity implies that A^{-1} is locally bounded (see Definition 2.6.6 (a)).

Proposition 2.8.3 *If X is a finite-dimensional Banach space, $C \subseteq X$ is nonempty, closed convex, $A : X \to 2^{X^*}$ is a monotone map with $D(A) \subseteq C$ and $F : C \to X^*$ is a monotone, continuous and strongly coercive map, then we can find $x_0 \in C$ such that*
$$\left\langle x^* + F(x_0), x - x_0\right\rangle \geqslant 0 \ \text{for all } (x, x^*) \in \operatorname{Gr} A.$$

Proof We may assume that $(0, 0) \in \operatorname{Gr} A$. Indeed, if this is not the case, we fix $(\hat{x}, \hat{x}^*) \in \operatorname{Gr} A$ and replace the maps A and F by
$$\hat{A}(x) = A(x + \hat{x}) - \hat{x}^* \text{ and } \hat{F}(x) = F(x + \hat{x}) - \hat{x}^*.$$

Evidently these translations do not alter the properties of the original maps and so $\hat{A}$ remains monotone, while $\hat{F}$ remains monotone, continuous and strongly coercive. Moreover, we see that $(0, 0) \in \operatorname{Gr}\hat{A}$.

Initially, we assume that $D(A)$ is bounded. Then $K = \overline{\operatorname{conv}}\, D(A)$. So, K is compact and convex. Arguing by contradiction, suppose that the proposition is not true. So for every $u \in K$, we can find $(x, x^*) \in \operatorname{Gr} A$ such that
$$\left\langle x^* + F(u), x - u\right\rangle < 0.$$

It follows that $K = \bigcup_{(x,x^*)\in \mathrm{Gr}\, A} \{u \in K : \langle x^* + F(u), x - u\rangle < 0\}$. In this union every set is open. So, by virtue of the compactness of K, we can find $\{(x_k, x_k^*)\}_{k=1}^n \subseteq \mathrm{Gr}\, A$ such that

$$K = \bigcup_{k=1}^{n} \{u \in K : \left\langle x_k^* + F(u), x_k - u\right\rangle < 0\}.$$

Let $\{\varphi_k\}_{k=1}^n$ be a continuous partition of unity subordinate to this cover and consider the map $\eta : K \to K$ defined by

$$\eta(u) = \sum_{k=1}^{n} \varphi_k(u) x_k.$$

This map is continuous and so we can apply Brouwer's fixed point theorem. So, we can find $x_0 \in K$ such that $\eta(x_0) = x_0$. For every $u \in K$, we have

$$\begin{aligned} g(u) &= \left\langle \sum_{k=1}^{n} \varphi_k(u) x_k^* + F(u), \eta(u) - u \right\rangle \\ &= \left\langle \sum_{k=1}^{n} \varphi_k(u) x_k^* + F(u), \sum_{m=1}^{n} \varphi_m(u)(x_m - u) \right\rangle \text{ (recall that } \sum_{m=1}^{n} \varphi_m(u) = 1) \\ &= \sum_{k,m=1}^{n} \varphi_k(u)\varphi_m(u) \left\langle x_k^* + F(u), x_m - u\right\rangle. \end{aligned}$$

If $k = m$ and $\varphi_k(u)^2 \neq 0$, then $u \in K$ and so $\left\langle x_k^* + F(u), x_k - u\right\rangle < 0$.

If $k \neq m$ and $\varphi_k(u)\varphi_m(u) \neq 0$, then $u \in \{u \in K : \left\langle x_k^* + F(u), x_k - u\right\rangle < 0\} \cap \{u \in K : \left\langle x_m^* + F(u), x_m - u\right\rangle < 0\}$. Exploiting the monotonicity of A we have

$$\begin{aligned} &\left\langle x_k^* + F(u), x_m - u\right\rangle + \left\langle x_m^* + F(u), x_k - u\right\rangle = \\ &\left\langle x_k^* + F(u), x_k - u\right\rangle + \left\langle x_m^* + F(u), x_m - u\right\rangle + \left\langle x_k^* - x_m^*, x_m - x_k\right\rangle < 0. \end{aligned}$$

So, we see that

$$g(u) < 0 \text{ for all } u \in K.$$

But note that for $x_0 \in K$ we have $\eta(x_0) = x_0$ and so $g(x_0) = 0$, a contradiction. Therefore, the proposition is true when $D(A)$ is bounded.

Now we drop the boundedness hypothesis on $D(A)$. From the previous case, we know that we can find $x_n \in C$ such that

$$\left\langle x^* + F(x_n), x - x_n\right\rangle \geqslant 0 \text{ for all } (x, x^*) \in \mathrm{Gr}\, A \,|_{\overline{\mathscr{B}}_n} \quad (\overline{\mathscr{B}}_n = \{x \in X : \|x\| \leqslant n\}).$$

Recall that $(0, 0) \in \mathrm{Gr}\, A$. So we have

$$\langle F(x_n), x_n \rangle \leqslant 0 \text{ for all } n \geqslant 1.$$

The strong coercivity of F implies that $\{x_n\}_{n\geqslant 1} \subseteq X$ is bounded. So, we may assume that $x_n \to x_0$ in X as $n \to \infty$. Evidently, $x_0 \in C$ and

$$\left\langle x^* + F(x_0), x - x_0 \right\rangle \geqslant 0 \text{ for all } (x, x^*) \in \operatorname{Gr} A.$$

The proof is now complete. □

Using Galerkin approximations, we extend the result to infinite-dimensional Banach spaces and then use the extension to establish surjectivity results for maximal monotone maps.

Proposition 2.8.4 *If X is a reflexive Banach space, $C \subseteq X$ is nonempty, closed convex, $A : X \to 2^{X^*}$ is a monotone map with $D(A) \subseteq C$ and $F : C \to X^*$ is monotone, hemicontinuous, bounded and strongly coercive, then there exists an $x_0 \in C$ such that*

$$\left\langle x^* + F(x_0), x - x_0 \right\rangle \geqslant 0 \textit{ for all } (x, x^*) \in \operatorname{Gr} A.$$

Proof We will use finite-dimensional approximations (Galerkin approximations) in order to exploit Proposition 2.8.3. So, let $\{X_\alpha\}_{\alpha\in J}$ be a directed family of finite-dimensional Banach spaces such that $X = \bigcup\limits_{\alpha\in J} X_\alpha$. Let $p_\alpha \in \mathscr{L}(X, X_\alpha)$ be the corresponding projection operator. Then $p_\alpha^* \in \mathscr{L}(X_\alpha^*, X^*)$ is the corresponding embedding map. We introduce the Galerkin approximations of A, F and C, namely we define

$$A_\alpha = p_\alpha^* \circ A \circ p_\alpha \ , \ F_\alpha = p_\alpha^* \circ F \circ p_\alpha \text{ and } C_\alpha = C \cap X_\alpha \ , \ \alpha \in J.$$

For each $\alpha \in J$, we can apply Proposition 2.8.3 on the triple $(A_\alpha, F_\alpha, C_\alpha)$ and find $x_\alpha \in C$ such that

$$\begin{aligned} &\left\langle \hat{x}^* + F_\alpha(x_\alpha), \hat{x} - x_\alpha \right\rangle_{X_\alpha} \geqslant 0 \text{ for all } (\hat{x}, \hat{x}^*) \in \operatorname{Gr} A_\alpha \\ \Rightarrow \ &\left\langle x^* + F(x_\alpha), x - x_\alpha \right\rangle \geqslant 0 \text{ for all } (x, x^*) \in \operatorname{Gr}(A \circ p_\alpha). \end{aligned} \tag{2.50}$$

Since $(0, 0) \in \operatorname{Gr} A$, we have $\langle F(x_\alpha), x_\alpha \rangle \leqslant 0$ for all $\alpha \in J$ and this by virtue of the strong coercivity and boundedness of F implies that we can find $M > 0$ such that

$$\|x_\alpha\| \leqslant M \text{ and } \|F(x_\alpha)\|_* \leqslant M \text{ for all } \alpha \in J.$$

The reflexivity of X and the Eberlein–Smulian theorem imply that we can find a sequence $\{x_{\alpha n} = x_n\}_{n\geqslant 1} \subseteq C$ such that

$$x_n \xrightarrow{w} x_0 \text{ in } X \text{ and } F(x_n) \xrightarrow{w} x_0^* \text{ in } X^*, \text{ with } (x_0, x_0^*) \in C \times X^*.$$

From (2.50) we have

$$\limsup_{n\to\infty} \langle F(x_n), x_n\rangle \leqslant \left\langle x^*, x - x_0\right\rangle + \left\langle x_0^*, x\right\rangle \text{for all } (x, x^*) \in \operatorname{Gr} A. \tag{2.51}$$

Using Zorn's lemma, without any loss of generality we may assume that A is maximal monotone on $D(A)$. We claim that we can find $(\tilde{x}, \tilde{x}^*) \in \operatorname{Gr} A$ such that

$$\langle \tilde{x}^*, \tilde{x} - x_0\rangle + \left\langle x_0^*, \tilde{x}\right\rangle \leqslant \langle x_0^*, x_0\rangle\,. \tag{2.52}$$

Suppose that the claim is not true. Then

$$\begin{aligned} &\left\langle x^* + x_0^*, x - x_0\right\rangle > 0 \text{ for all } (x, x^*) \in \operatorname{Gr} A \qquad (2.53)\\ \Rightarrow\ &(x_0, -x_0^*) \in \operatorname{Gr} A \ \ \text{(due to maximality of } A \text{ on } D(A)). \end{aligned}$$

So, if in (2.53) we use $(x_0, -x_0^*) \in \operatorname{Gr} A$, we have a contradiction. Therefore (2.52) holds for some $(\tilde{x}, \tilde{x}^*) \in \operatorname{Gr} A$. Using (2.52) in (2.51), we have

$$\begin{aligned} &\limsup_{n\to\infty} \langle F(x_n), x_n\rangle \leqslant \left\langle x_0^*, x_0\right\rangle\\ \Rightarrow\ &\limsup_{n\to\infty} \langle F(x_n), x_n - x_0\rangle \leqslant 0. \end{aligned} \tag{2.54}$$

Let $x \in D(A)$ and let $x_t = tx_o + (1-t)x$ with $t \in [0, 1]$. Then $x_t \in C$ and the monotonicity of F implies that

$$\begin{aligned} &\langle F(x_n) - F(x_t), x_n - x_t\rangle \geqslant 0 \text{ for all } n \geqslant 1, \text{ all } t \in [0,1]\\ \Rightarrow\ &t\,\langle F(x_n), x_n - x_0\rangle + (1-t)\,\langle F(x_n), x_n - x\rangle\\ &\geqslant t\,\langle F(x_t), x_n - x_0\rangle + (1-t)\,\langle F(x_t), x_n - x\rangle\\ \Rightarrow\ &\liminf_{n\to\infty} \langle F(x_n), x_n - x\rangle \geqslant \langle F(x_t), x_0 - x\rangle \quad \text{for all } t \in [0,1] \ \text{ see (2.54).} \end{aligned}$$

Because of the hemicontinuity of $F(\cdot)$, we have

$$\begin{aligned} &\liminf_{n\to\infty} \langle F(x_n), x_n - x\rangle \geqslant \langle F(x_0), x_0 - x\rangle\\ \Rightarrow\ &\liminf_{n\to\infty} \langle F(x_n), x_n - x\rangle \geqslant \langle F(x_0), x_0 - x\rangle + \left\langle x_0^*, x\right\rangle \ \text{ for all } x \in D(A). \end{aligned} \tag{2.55}$$

From (2.51) and (2.55) it follows that

$$\left\langle x^* + F(x_0), x - x_0\right\rangle \geqslant 0 \ \text{ for all } (x, x^*) \in \operatorname{Gr} A.$$

The proof is now complete. □

Now we are ready for the first surjectivity result.

Theorem 2.8.5 *If X is a reflexive Banach space, $C \subseteq X$ is nonempty, closed convex, $A : X \to 2^{X^*}$ is monotone map with $D(A) \subseteq C$ and $F : C \to X^*$ is monotone,* *hemicontinuous, bounded and strongly coercive, then $A + F$ is surjective (that is, $R(A + F) = X^*$).*

Proof Let $x_0^* \in X^*$ and consider the map

$$\hat{A}(x) = A(x) - x_0^* \ \text{ for all } x \in D(A).$$

Evidently, $\hat{A}$ is still monotone with $D(\hat{A}) \subseteq C$. We apply Proposition 2.8.4 on the triple $(\hat{A}, F, C)$ and obtain $x_0 \in C$ such that

$$\begin{aligned} &\langle \hat{x}^* + F(x_0), x - x_0 \rangle \geqslant 0 \ \text{ for all } (x, \hat{x}^*) \in \operatorname{Gr} \hat{A} \\ \Rightarrow\ &\big\langle x^* - (x_0^* - F(x_0)), x - x_0 \big\rangle \ \text{ for all } (x, x^*) \in \operatorname{Gr} A. \end{aligned}$$

The maximal monotonicity of A implies

$$\begin{aligned} &(x_0, x_0^* - F(x_0)) \in \operatorname{Gr} A \\ \Rightarrow\ & x_0^* \in A(x_0) + F(x_0). \end{aligned}$$

Since $x_0^* \in X^*$ is arbitrary, we conclude that $A + F$ is surjective. □

Now we present a necessary and sufficient condition for the surjectivity of maximal monotone maps.

Theorem 2.8.6 *If X is a reflexive Banach space and $A : X \to 2^{X^*}$ is a maximal monotone map, then A is surjective if and only if A^{-1} is locally bounded.*

Proof $\underline{\Rightarrow}$ Recall that A^{-1} is maximal monotone (see Proposition 2.6.3) and so Proposition 2.6.8 implies that A^{-1} is locally bounded.

$\underline{\Leftarrow}$ We will show that $R(A) \subseteq X^*$ is both closed and open, hence $R(A) = X^*$.

First we show the closedness of $R(A)$. So, let $\{x_n^*\}_{n\geqslant 1} \subseteq R(A)$ such that $x_n^* \to x^*$ in X^*. We have $x_n^* \in A(x_n)$ $n \geqslant 1$. Since by hypothesis A^{-1} is locally bounded, it follows that $\{x_n\}_{n\geqslant 1} \subseteq X$ is bounded. Hence, due to the reflexivity of X, we may assume that $x_n \xrightarrow{w} x$ in X. We have $(x_n, x_n^*) \in \operatorname{Gr} A$ for all $n \geqslant 1$. Then Proposition 2.6.5 implies $(x, x^*) \in \operatorname{Gr} A$ and so we have proved that $R(A)$ is closed.

Next we show that $R(A)$ is open. Invoking Theorem 2.7.36, we may assume that both X and X^* are locally uniformly convex. Of course, this does not affect the maximal monotonicity of A.

Let $x^* \in A(x)$. Since maximal monotonicity is invariant under translation without any loss of generality, we may assume that $x = 0$. Since by hypothesis A^{-1} is locally bounded, we can find $r > 0$ such that $A^{-1}|_{B_r^*(x^*)}$ is bounded (here, $B_r^*(x^*) = \{u^* \in X^* : ||u^* - x^*||_* < r\}$). We will show that $B_{r/2}^*(x^*) \subseteq R(A)$.

To this end, let $u^* \in B_{r/2}^*(x^*)$. Let $\lambda > 0$ and let $F = \lambda J$, with J being the duality map. By virtue of Proposition 2.7.33 we can apply Theorem 2.8.5 for the triple $(A,\ F = \lambda J,\ X)$ and infer that there exists $(u_\lambda, u_\lambda^*) \in \operatorname{Gr} A$ such that

$$u_\lambda^* + \lambda J(u_\lambda) = u^*. \tag{2.56}$$

The monotonicity of A implies

$$
\begin{aligned}
&\langle u^* - \lambda J(u_\lambda) - x^*,\ u_\lambda \rangle \geqslant 0 \text{ (recall that } x = 0) \\
\Rightarrow\ & \lambda||u_\lambda|| \leqslant ||u^* - x^*||_* < \frac{r}{2} \\
\Rightarrow\ & ||u^* - u_\lambda^*||_* = \lambda||J(u_\lambda)||_* = \lambda||u_\lambda|| < \frac{r}{2} \qquad (2.57) \\
\Rightarrow\ & ||x^* - u_\lambda^*||_* < r \text{ for all } \lambda > 0. \qquad (2.58)
\end{aligned}
$$

Since $A^{-1}|_{B_r^*(x^*)}$ is bounded and $u_\lambda \in A^{-1}(u_\lambda^*)$, from (2.58) it follows that $\{u_\lambda\}_{\lambda>0} \subseteq X$ is bounded. Then (2.57) implies that $u_\lambda^* \to u^*$ in X^* as $\lambda \to 0^+$. We already know that $R(A)$ is closed. So, $u^* \in R(A)$ and we have proved that $B_{r/2}^*(x^*) \subseteq R(A)$. This proves that $R(A)$ is open, hence $R(A) = X^*$. □

Corollary 2.8.7 *If X is a reflexive Banach space and $A : X \to 2^{X^*}$ is maximal monotone and coercive, then A is surjective (that is, $R(A) = X^*$).*

Proof The coercivity of A implies that A^{-1} is locally bounded. Hence we can apply Theorem 2.8.6 and conclude that A is surjective. □

Corollary 2.8.8 *If X is a reflexive Banach space and $A : X \to X^*$ is monotone, hemicontinuous coercive with $D(A) = X$, then A is surjective.*

Proof Proposition 2.6.12 implies A is maximal monotone. So, we can apply Corollary 2.8.7 and conclude that A is surjective. □

Theorem 2.8.5 leads to a convenient characterization of maximal monotonicity.

Theorem 2.8.9 *If X is a reflexive Banach space such that both X and its dual X^* are strictly convex and $A : X \to 2^{X^*}$ is a monotone map, then A is maximal monotone if and only if for every $\lambda > 0$ (equivalently for some $\lambda > 0$) $A + \lambda J$ is surjective (J is the duality map).*

Proof $\underline{\Downarrow}$ From Proposition 2.7.27 and its proof, we know that the duality map $J : X \to X^*$ is single-valued, monotone and sequentially continuous from X into $X^*_{w^*}$ (this type of continuity is known as demicontinuity). In particular then J is hemicontinuous and of course bounded and strongly coercive. So, we can use Theorem 2.8.5 and infer that $R(A + \lambda J) = X^*$ for all $\lambda > 0$.

$\underline{\Uparrow}$ Suppose that for some $\lambda > 0$, we have $R(A + \lambda J) = X^*$. Without any loss of generality we may assume that $\lambda = 1$. Suppose that for some $(x, x^*) \in X \times X^*$, we have

$$\langle x^* - u^*,\ x - u \rangle \geqslant 0 \text{ for all } (u, u^*) \in \operatorname{Gr} A. \qquad (2.59)$$

We can find $(v, v^*) \in \operatorname{Gr} A$ such that

$$v^* + J(v) = x^* + J(x) \text{ (recall that } R(A + J) = X^*). \qquad (2.60)$$

Then in (2.59) we choose $u = v$ and $u^* = v^*$. We obtain

$$\begin{aligned} 0 &\leqslant \left\langle v^* + J(v) - J(x) - v^*,\ x - v \right\rangle \\ &= \langle J(v) - J(x),\ x - v\rangle \leqslant 0. \end{aligned}$$

Since X, X^* are strictly convex, J is strictly monotone and so we conclude that $v = x$. Hence $v^* = x^*$ (see (2.60)) and so $(x, x^*) \in \mathrm{Gr}\, A$ which proves the maximality of A. $\square$

Using this characterization, we can show the maximal monotonicity of the subdifferential map.

Theorem 2.8.10 *If X is a reflexive Banach space and $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is convex and lower semicontinuous, then $\partial\varphi : X \to 2^{X^*}$ is maximal monotone.*

Proof On account of Theorem 2.7.36, we may assume that both X and X^* are locally uniformly convex. According to Theorem 2.8.9, it suffices to show that $R(\partial\varphi + J) = X^*$. To this end, let $x^* \in X$ and consider the convex and lower semicontinuous function $\psi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ defined by

$$\psi(x) = \varphi(x) + \eta(x) - \left\langle x^*, x \right\rangle, \text{ where } \eta(x) = \frac{1}{2}||x||^2 \text{ for all } x \in X.$$

Recall that φ is bounded below by an affine continuous function $x \to \langle u^*, x\rangle - c$ for all $x \in X$. Then

$$\begin{aligned} &\psi(x) \geqslant \eta(x) + \left\langle u^* - x^*, x \right\rangle - c \geqslant \frac{1}{2}||x||^2 - ||u^* - x^*||_*||x|| - c \\ \Rightarrow\ &\psi \text{ is coercive, that is, } \psi(x) \to +\infty \text{ as } ||x|| \to \infty. \end{aligned}$$

Invoking the Weierstrass theorem, we can find $\hat{x} \in X$ such that

$$\psi(\hat{x}) = \inf\{\psi(x) : x \in X\}. \tag{2.61}$$

From (2.61) and Definition 2.7.1 it follows that

$$0 \in \partial\psi(\hat{x}). \tag{2.62}$$

We have

$$\partial\psi(\hat{x}) = \partial\varphi(\hat{x}) + \partial\eta(\hat{x}) - x^* = \partial\varphi(\hat{x}) + J(\hat{x}) - x^* \tag{2.63}$$

(see Definition 2.7.21).

Therefore, from (2.62) and (2.63) we infer that

$$x^* \in \partial\varphi(\hat{x}) + J(\hat{x}).$$

But $x^* \in X^*$ is arbitrary. Hence

$$\begin{aligned} & R(\partial\varphi + J) = X^* \\ \Rightarrow\ & \partial\varphi \text{ is maximal monotone (see Theorem 2.8.9).} \end{aligned}$$

The proof is now complete. □

Remark 2.8.11 In fact, the result is true for any Banach space, not necessarily reflexive. For a proof of this more general result, we refer to Rockafellar [359].

Actually the subdifferential has more structure which distinguishes it from general maximal monotone operators. To see this, we introduce the following notion.

Definition 2.8.12 Let X be a Banach space and X^* its dual.

(a) A multivalued map $A : X \to 2^{X^*}$ is said to be n-cyclically monotone, ($n \geqslant 2$) provided

$$\sum_{k=1}^{n} \langle x_k^*, x_k - x_{k-1} \rangle \geqslant 0$$

whenever $x_0, x_1, \ldots, x_n \in X,\ x_n = x_0,\ x_k^* \in A(x_k)$ for all $k \in \{1, \ldots, n\}$.

(b) A multivalued map $A : X \to 2^{X^*}$ is said to be cyclically monotone if it is n-cyclically monotone for every $n \geqslant 2$.

(c) A multivalued map $A : X \to 2^{X^*}$ is maximal cyclically monotone, if $A = S$ whenever $S : X \to 2^{X^*}$ is cyclically monotone and $\mathrm{Gr}\, A \subseteq \mathrm{Gr}\, S$.

Remark 2.8.13 Clearly, a 2-cyclically monotone map is monotone. Also, a maximal monotone map which is cyclically monotone is necessarily maximal cyclically monotone. The linear map $A : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $A(x, u) = (u, -x)$ is positive, hence monotone, but it is not 3-monotone. To see this consider the points $(1, 1)$, $(0, 1)$ and $(1, 0)$.

It turns out that the subdifferentials are the only maximal cyclically monotone maps.

Theorem 2.8.14 *If X is a Banach space and $A : X \to 2^{X^*}$, then the following statements are equivalent*

(a) $A = \partial\varphi$ *with* $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ *convex and lower semicontinuous;*

(b) A *is maximal cyclically monotone.*

Proof (a) $\Rightarrow$ (b) From Definition 2.7.1 we have

$$\varphi(x_{k-1}) - \varphi(x_k) \geqslant \langle x_k^*, x_{k-1} - x_k \rangle$$

for $x_k^* \in \partial\varphi(x_k)$, $k = 1, \ldots, n$, and $x_n = x_0$. Adding, we obtain

$$0 \leqslant \sum_{\text{k}=1}^{n} \left\langle x_k^*, x_k - x_{k-1} \right\rangle$$

$\Rightarrow \partial\varphi$ is maximal cyclically monotone (see Theorem 2.8.10 and Remark 2.8.13).

(b) $\Rightarrow$ (a) Fix $x_0 \in D(A)$ and $x_0^* \in A(x_0)$. Then for $u \in X$, we define

$$\psi(u) = \sup\left\{\left\langle x_n^*, u - x_n\right\rangle + \left\langle x_{n-1}^*, x_n - x_{n-1}\right\rangle + \ldots + \left\langle x_0^*, x_1 - x_0\right\rangle\right\},$$

where the supremum is taken over all finite sets of elements $x_k \in D(A)$ and $x_k^* \in A(x_k)$, $k = 1, \ldots, n$ and $n \geqslant 1$. Since ψ is the pointwise supremum of a family of continuous affine functions, it follows that $\psi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is convex and lower semicontinuous. Also, the cyclical monotonicity of A implies that $\psi(x_0) = 0$ and so we see that ψ is not identically $+\infty$. To reach (a), it suffices to show that $\operatorname{Gr} A \subseteq \operatorname{Gr} \partial\psi$. To this end, let $(x, x^*) \in \operatorname{Gr} A$. We will have that $(x, x^*) \in \operatorname{Gr} \partial\psi$, provided we show that

$$\left\langle x^*, u - x\right\rangle \leqslant \psi(u) - \lambda \text{ for all } u \in X \text{ and } \lambda < \psi(x). \tag{2.64}$$

From the definition of ψ, we know that there exist $x_k \in D(A)$ and $x_k^* \in A(x_k)$, $k = 1, \ldots, n$, such that

$$\lambda < \left\langle x_n^*, x - x_n\right\rangle + \left\langle x_{n-1}^*, x_n - x_{n-1}\right\rangle + \ldots + \left\langle x_0^*, x_1 - x_0\right\rangle. \tag{2.65}$$

Let $x_{n+1} = x$ and $x_{n+1}^* = x^*$. Then by definition, for any $u \in X$ we have

$$\begin{aligned}\psi(u) &\geqslant \left\langle x_{n+1}^*, u - x_{n+1}\right\rangle + \left\langle x_n^*, x_{n+1} - x_n\right\rangle + \ldots + \left\langle x_0^*, x_1 - x_0\right\rangle \\ &> \left\langle x_{n+1}^*, u - x_{n+1}\right\rangle + \lambda \text{ (see (2.65))}.\end{aligned}$$

From this inequality, relation (2.64) follows and so we have proved the theorem. □

Remark 2.8.15 In (a) the function φ is unique up to an additive constant.

The sum of two maximal monotone maps need not be maximal monotone. So, we need sufficient conditions for the maximality of the sum of maximal monotone maps. The main result in this direction is the following theorem. Its proof can be found in Barbu [32, p. 46] and Zeidler [427, p. 888].

Theorem 2.8.16 *If X is a reflexive Banach space and $A, F : X \to 2^{X^*}$ are maximal monotone maps such that $D(A) \cap \operatorname{int} D(F) \neq \emptyset$, then $A + F : X \to 2^{X^*}$ is maximal monotone.*

Remark 2.8.17 Proposition 2.7.20 is a particular case of this theorem, if we observe that $\operatorname{int} D(\partial\varphi) = \operatorname{int} \operatorname{dom} \varphi$.

2.9 Regularizations and Linear Monotone Operators

When $X = H$ is a Hilbert space, then there are two useful single-valued and Lipschitz continuous maps associated with a maximal monotone operator. These maps are the resolvent and the Yosida approximation. In this section we study these maps and at the end we characterize linear monotone operators.

Definition 2.9.1 Let H be a Hilbert space and $A : H \to 2^H$ a multivalued map.

(a) For every $\lambda > 0$, $J_\lambda^A = (I + \lambda A)^{-1}$ is the resolvent of A.
(b) For every $\lambda > 0$, $A_\lambda = \frac{1}{\lambda}\left(I - J_\lambda^A\right)$ is the Yosida approximation of A.

The next proposition shows that these operators are useful for monotone maps.

Proposition 2.9.2 *If H is a Hilbert space and $A : H \to 2^H$ is monotone, then A is monotone if and only if for all $(x, x^*) \in \mathrm{Gr}\, A,\ (u, u^*) \in \mathrm{Gr}\, A$ and all $\lambda > 0$ we have $||x - u|| \leqslant ||(x - u) + \lambda(x^* - u^*)||$.*

Proof $\underline{\Rightarrow}$ We have

$$\begin{aligned}||(x-u)+\lambda(x^*-u^*)||^2 &= ((x-u)+\lambda(x^*-u^*),(x-u)+\lambda(x^*-u^*))_H\\ &= ||x-u||^2 + 2\lambda(x^*-u^*, x-u)_H + \lambda^2||x^*-u^*||^2\\ &\geqslant ||x-u||^2.\end{aligned}$$

$\underline{\Leftarrow}$ We have $2\lambda(x^* - u^*, x - u)_H + \lambda^2||x^* - u^*||^2 \geqslant 0$, hence

$$2(x^* - u^*, x - u)_H + \lambda||x^* - u^*|| \geqslant 0.$$

Let $\lambda \to 0$ to get $(x^* - u^*, x - u)_H \geqslant 0$ for all $(x, x^*) \in \mathrm{Gr}\, A,\ (u, u^*) \in \mathrm{Gr}\, A$, hence A is monotone. $\square$

From this proposition, we infer the following result.

Corollary 2.9.3 *If H is a Hilbert space and $A : H \to 2^H$, then the following statements are equivalent:*

(a) A is monotone;
(b) J_λ^A is nonexpansive for every $\lambda > 0$ (that is, $\left\| J_\lambda^A(x) - J_\lambda^A(u)\right\| \leqslant ||x - u||$ for all $x, u \in R(I + \lambda A)$).

Often it is said that if H is a Hilbert space, then the duality map J is the identity. Strictly speaking this is not correct. The Riesz–Fréchet theorem asserts that H and its dual H^* are isometrically isomorphic (in the case of complex Hilbert spaces the canonical isometry is conjugate-linear). So, we may identify H with its dual H^*. We often do this, but not always. In the next remark, we outline a situation arising in the study of boundary value problems where this identification cannot be done.

Remark 2.9.4 Let H be a Hilbert space and $X \subseteq H$ a linear subspace which is embedded continuously and densely into H, that is, there exists a $c > 0$ such that

$$||u||_H \leqslant c||u||_X \text{ for all } u \in X.$$

Then there is a canonical map $j : H^* \to X^*$ defined by

$$\left\langle j(h^*), x\right\rangle_X = \langle h^*, x\rangle_H \text{ for all } h^* \in H^*, \text{ all } x \in X$$

(so we simply consider the restriction to X of an element in H^*). We have

(a) $||j(h^*)||_{X^*} \leqslant c||h^*||_{H^*}$ for all $h^* \in H^*$;
(b) j is injective;
(c) if X is reflexive, then $R(j)$ is dense in X^*.

Identifying H with its dual H^* via the Riesz–Fréchet representation theorem, we have

$$X \hookrightarrow H = H^* \hookrightarrow X^*$$

with all injections being continuous and dense (under the condition that X is reflexive). Such a triple of spaces (X, H, X^*) is often called an "evolution triple", because of their importance in the theory of evolution equations. Note that $\langle\cdot,\cdot\rangle_X|_{H\times X} = (\cdot,\cdot)_H$. So, we see that when X is a Hilbert space too, since we identify H with its dual, we cannot do the same thing for X. A typical example is the triple $(H_0^1(\Omega), L^2(\Omega), H^{-1}(\Omega))$. The Hilbert space H for which we assume $H = H^*$ is called a "pivot Hilbert space". In this section we always assume that the Hilbert space H is pivot, that is, $H = H^*$. Hence $J = I$ (the duality map is the identity).

With this remark in mind, the next proposition is a consequence of Theorem 2.8.9 and Corollary 2.9.3.

Proposition 2.9.5 *If H is a Hilbert space and $A : H \to 2^H$, then the following statements are equivalent:*

(a) A is maximal monotone;
(b) A is monotone and $D\left(J_\lambda^A\right) = R(I + \lambda A) = H$ for all $\lambda > 0$.

Proposition 2.9.6 *If H is a Hilbert space and $A : H \to 2^H$ is maximal monotone, then for all $\lambda, \mu > 0$ we have the so-called resolvent identity*

$$J_\lambda^A = J_\mu^A \circ \left[\frac{\mu}{\lambda}I + \left(1 - \frac{\mu}{\lambda}\right) J_\lambda^A\right].$$

Proof Note that $u = J_\lambda^A(x)$ is equivalent to saying that $\frac{1}{\lambda}\left[x - J_\lambda^A(x)\right] \in A\left(J_\lambda^A(x)\right)$ for all $x \in H$ and all $\lambda > 0$. We rewrite the last inclusion as follows:

$$\frac{1}{\mu}\left[\frac{\mu}{\lambda}x + \left(1 - \frac{\mu}{\lambda}\right) J_\lambda^A(x) - J_\lambda^A(x)\right] \in A\left(J_\lambda^A(x)\right) \text{ for all } \mu > 0$$

and this is equivalent to

$$\frac{\mu}{\lambda}x + \left(1 - \frac{\mu}{\lambda}\right) J_\lambda^A(x) \in (I + \mu A)\left(J_\lambda^A(x)\right),$$

from which we deduce the resolvent identity. $\square$

Proposition 2.9.7 *If H is a Hilbert space and $A : H \to 2^H$ is a maximal monotone map, then $\overline{D(A)}$ is convex and $\lim\limits_{\lambda\to 0^+} J_\lambda^A(x) = \mathrm{proj}_{\overline{D(A)}}(x)$ (here $\mathrm{proj}_{\overline{D(A)}}$ denotes the metric projection on $\overline{D(A)}$).*

Proof Let $C = \overline{\mathrm{conv}}\, D(A)$ and $x \in H$. Let $u_\lambda = J_\lambda^A(x)$. Then $\frac{1}{\lambda}(x - u_\lambda) \in A(u_\lambda)$. From the monotonicity of A, we have

$$\begin{aligned} &\left(\frac{1}{\lambda}(x - u_\lambda) - v^*, u_\lambda - v\right)_H \geqslant 0 \text{ for all } (v, v^*) \in \mathrm{Gr}\, A \\ \Rightarrow\ & ||u_\lambda||^2 \leqslant (x, u_\lambda - v)_H + (u_\lambda, v)_H - \lambda(v^*, u_\lambda - x)_H \text{ for all } \lambda > 0 \quad (2.66) \\ \Rightarrow\ & \{u_\lambda\}_{\lambda\in(0,1]} \subseteq H \text{ is bounded.} \end{aligned}$$

So, we can find a subsequence $\{u_{\lambda_n} = u_n\}_{n\geqslant 1} \subseteq H$ such that $u_n \xrightarrow{w} \hat{u}$ in H, $\lambda_n \to 0$. From (2.66) we have

$$||\hat{u}||^2 \leqslant (x, \hat{u} - v)_H + (\hat{u}, v)_H \text{ for all } v \in D(A).$$

We have

$$\begin{aligned} &\hat{u} \in C \text{ and } (x - \hat{u}, v - \hat{u})_H \leqslant 0 \text{ for all } v \in C \\ \Rightarrow\ & \hat{u} = \mathrm{proj}_C(x). \end{aligned}$$

So, the uniqueness of the weak limit implies that $u_\lambda \xrightarrow{w} \hat{u}$ in H as $\lambda \to 0^+$. Then

$$\limsup_{\lambda\to 0^+} ||u_\lambda||^2 \leqslant (x, \hat{u} - v)_H + (\hat{u}, v)_H \text{ for all } v \in C \text{ (see (2.66))}.$$

Let $v = \hat{u}$. Then

$$\limsup_{\lambda\to 0} ||u_\lambda|| \leqslant ||\hat{u}||. \tag{2.67}$$

On the other hand, since $u_\lambda \xrightarrow{w} \hat{u}$ in H, we have

$$||\hat{u}|| \leqslant \liminf_{\lambda\to 0^+} ||u_\lambda||. \tag{2.68}$$

From (2.67) and (2.68) it follows that $||u_\lambda|| \to ||\hat{u}||$ as $\lambda \to 0^+$. The Kadec–Klee property of Hilbert spaces (see Remark 2.7.30) implies that $u_\lambda \to \hat{u} = \mathrm{proj}_C(x)$ in H. Finally, note that if $x \in C$, then $u_\lambda \in D(A)$ and so it follows that $\overline{D(A)} = C$. $\square$

From Proposition 2.6.5 we know that if $A : H \to 2^H$ is maximal monotone, then for every $x \in D(A)$ we have that $A(x)$ is nonempty, convex and w-closed (hence closed too). So, the next definition makes sense.

Definition 2.9.8 Let H be a Hilbert space and $A : H \to 2^H$ a maximal monotone map. The minimal section operator $A^0 : H \to H$ is defined by

$$A^0(x) = \mathrm{proj}_{A(x)}(0) = \{x^* \in H : x^* \in A(x),\ x^* \text{ has minimal norm}\}.$$

Proposition 2.9.9 *If H is a Hilbert space and $A : H \to 2^H$ is a maximal monotone map, then*

(a) for every $\lambda > 0$, A_λ is single-valued maximal monotone and Lipschitz continuous with constant $\frac{1}{\lambda}$;
(b) $(A_\lambda)_\mu = A_{\lambda+\mu}$ for all $\lambda, \mu > 0$;
(c) for every $x \in D(A)$, $||A_\lambda(x)||$ increases up to $||A^0(x)||$ as $\lambda \to 0^+$, $A_\lambda(x) \to A^0(x)$ as $\lambda \to 0^+$ and

$$||A_\lambda(x) - A^0(x)||^2 \leqslant ||A^0(x)||^2 - ||A_\lambda(x)||^2 \text{ for all } \lambda > 0;$$

(d) for every $x \notin D(A)$, $||A_\lambda(x)||$ is increasing and unbounded as $\lambda \to 0^+$.

Proof (a) Recall that $\frac{1}{\lambda}\left[x - J_\lambda^A(x)\right] \in A\left(J_\lambda^A(x)\right)$ for all $x \in H$ and all $\lambda > 0$. From this and Definition 2.9.1, we have

$$\begin{aligned}&(A_\lambda(u) - A_\lambda(x), u - x)_H = \\ &\left(A_\lambda(u) - A_\lambda(x),\ \lambda\left(A_\lambda(u) - A_\lambda(x)\right) + \left(J_\lambda^A(u) - J_\lambda^A(x)\right)\right)_H \geqslant \\ &\lambda||A_\lambda(u) - A_\lambda(x)||^2 \text{ (since } A \text{ is monotone).}\end{aligned}$$

Therefore A_λ is monotone and $||A_\lambda(u) - A_\lambda(x)|| \leqslant \frac{1}{\lambda}||u - x||$. So, Proposition 2.6.12 implies that A_λ is maximal monotone.

(b) This follows from the characterization of $u = A_\lambda(x)$ by $u \in A(x - \lambda u)$.

(c) Let $x \in D(A)$. Then

$$\begin{aligned}0 &\leqslant \left(A^0(x) - A_\lambda(x),\ x - J_\lambda(x)\right)_H = \lambda\left(A^0(x) - A_\lambda(x), A_\lambda(x)\right)_H \\ &\Rightarrow ||A_\lambda(x)||^2 \leqslant \left(A^0(x), A_\lambda(x)\right)_H \\ &\Rightarrow ||A_\lambda(x)|| \leqslant ||A^0(x)||.\end{aligned} \tag{2.69}$$

Applying this inequality to A_μ together with (b), we have

$$\begin{aligned}
&||A_{\lambda+\mu}(u)||^2 \leqslant \left(A_\lambda(x), A_{\lambda+\mu}(x)\right)_H \text{ for } \mu > 0 \\
\Rightarrow\ &||A_{\lambda+\mu}(x)|| \leqslant ||A_\lambda(x)|| \\
\Rightarrow\ &\{||A_\lambda(x)||\}_{\lambda>0} \text{ increases as } \lambda > 0 \text{ decreases.}
\end{aligned}$$

Also, we have

$$||A_\lambda(x) - A_{\lambda+\mu}(x)||^2 \leqslant ||A_\lambda(x)||^2 - ||A_{\lambda+\mu}(x)||^2$$
$$\text{for all } \lambda, \mu > 0 \text{ and all } x \in H. \quad (2.70)$$

If $\{||A_\lambda(x)||\}_{\lambda>0}$ is bounded, then from (2.70) it follows that

$$\lim_{\lambda\to 0^+} A_\lambda(x) = u \text{ in } H.$$

Since $\lambda A_\lambda(x) = \left(I - J_\lambda^A\right)(x)$, it follows that $J_\lambda^A(x) \to x$ in H. Next, since $\left(J_\lambda^A(x), A_\lambda(x)\right) \in \operatorname{Gr} A$, it follows that $(x, u) \in \operatorname{Gr} A$. Also

$$\begin{aligned}
&||u|| = \lim_{\lambda\to 0^+} ||A_\lambda(x)|| \leqslant ||A^0(x)|| \text{ (see (2.69))} \\
\Rightarrow\ &u = A^0(x).
\end{aligned}$$

(d) Follows from the previous argument. □

Summarizing the main properties of the two operators introduced in Definition 2.9.1, we can state the following theorem.

Theorem 2.9.10 *If H is a Hilbert space, $A : H \to 2^H$ is a maximal monotone map and $\lambda > 0$, then*

(a) J_λ^A *is defined on all of H and it is nonexpansive;*
(b) $\overline{D(A)}$ *is convex and* $\lim_{\lambda\to 0^+} J_\lambda(x) = \operatorname{proj}_{\overline{D(A)}}(x)$ *for all $x \in H$;*
(c) A_λ *is defined on all of H, it is monotone and Lipschitz continuous with constant $\frac{1}{\lambda}$ (hence maximal monotone, too) and*

$$A_\lambda(x) \in A\left(J_\lambda^A(x)\right) \textit{ for all } x \in H;$$

(d) *for all $x \in D(A)$, we have*

$$\begin{aligned}
&||A_\lambda(x)|| \leqslant ||A^0(x)||, \\
&\{||A_\lambda(x)||\}_{\lambda>0} \textit{ is increasing as } \lambda \textit{ decreases}, \\
&A_\lambda(x) \to A^0(x) \textit{ in } H \textit{ as } \lambda \to 0^+.
\end{aligned}$$

Remark 2.9.11 Analogous regularizations can be defined in the more general setting of a reflexive Banach space X such that both X and X^* are locally uniformly convex.

So, we consider a maximal monotone map $A : X \to 2^{X^*}$. Then by virtue of Theorem 2.8.9 for every $\lambda > 0$ and every $x \in X$, the operator inclusion

$$0 \in J(u - x) + \lambda A(u) \text{ (J being the duality map of X)}$$

has a solution. The hypotheses on X and X^* imply that this solution is unique. We denote it by $J_\lambda^A(x)$. Also, we set $A_\lambda(x) = \frac{1}{\lambda} J\left(x - J_\lambda^A(x)\right)$. These are the resolvent and Yosida approximation operators. However, their properties are not as good as in the Hilbert space case.

In the particular case where $A = \partial\varphi$ with $\varphi : H \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ convex and lower semicontinuous, the resolvent and Yosida approximation of A are related to some useful regularizations of φ.

Proposition 2.9.12 *If H is a Hilbert space, $\varphi : H \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is convex and lower semicontinuous, $A = \partial\varphi$ and for every $\lambda > 0$ we define*

$$\varphi_\lambda(x) = \inf\left\{\varphi(u) + \frac{1}{2\lambda}||u - x||^2 : u \in H\right\},$$

then

(a) $\varphi_\lambda(x) = \varphi(J_\lambda(x)) + \frac{\lambda}{2}||A_\lambda(x)||^2$ *for all $x \in H$ and all $\lambda > 0$;*
(b) *φ_λ is convex and Fréchet differentiable and $\varphi'_\lambda = \partial\varphi_\lambda = (\partial\varphi)_\lambda = A_\lambda$;*
(c) *for every $x \in H$, $\varphi_\lambda(x)$ increases as λ decreases and converges to $\varphi(x)$ as $\lambda \to 0^+$;*
(d) *if $\lambda_n \to 0^+$, $x_n \to x$ in H and $\partial\varphi_{\lambda_n}(x_n) \xrightarrow{w} x^*$ in H, then $(x, x^*) \in \operatorname{Gr}\partial\varphi$;*
(e) *if $\lambda_n \to 0^+$, $x_n \to x$ in H and $\{\partial\varphi_{\lambda_n}(x_n)\}_{n\geqslant 1} \subseteq H$ is bounded, then $\varphi_{\lambda_n}(x_n) \to \varphi(x)$ as $n \to \infty$;*
(f) $\varphi_\lambda(x) \leqslant \varphi_\lambda(u) + ||\partial\varphi_\lambda(x)||\,||x - u||$ *for all $x, u \in H$ and all $\lambda > 0$;*
(g) *we have for all $x, u, v \in H$ and all $\lambda > 0$*

$$|\varphi_\lambda(x) - \varphi_\lambda(u)| \leqslant \left(2||\partial\varphi_\lambda(v)|| + \frac{||x - v||}{\lambda} + \frac{||u - v||}{\lambda}\right)||x - u||.$$

Proof (*a*) We claim that $\hat{u} \in H$ realizes the infimum involved in the definition of φ_λ if and only if $\frac{1}{\lambda}(x - \hat{u}) \in \partial\varphi(\hat{u})$. To see this, note that if $\frac{1}{\lambda}(x - \hat{u}) \in \partial\varphi(\hat{u})$, then $\hat{u} \in \operatorname{dom}\varphi$ and

$$\varphi(u) - \varphi(\hat{u}) \geqslant \frac{1}{\lambda}(x - \hat{u}, u - \hat{u})_H \geqslant \frac{1}{2}\lambda\left[||\hat{u} - x||^2 - ||u - x||^2\right] \text{ for all } u \in H$$

$$\Rightarrow \varphi(\hat{u}) + \frac{1}{2\lambda}||\hat{u} - x||^2 \leqslant \varphi(u) + \frac{1}{2\lambda}||u - x||^2 \text{ for all } u \in H.$$

Conversely, let $w = (1 - t)\hat{u} + tv$ with $v \in H$ and $t \in (0, 1)$. We have

$$
\begin{aligned}
&t[\varphi(v) - \varphi(\hat{u})] \geqslant \varphi(w) - \varphi(\hat{u}) \text{ (due to the convexity of } \varphi) \geqslant \\
&\frac{1}{2\lambda}\left[||\hat{u} - x||^2 - ||(1-t)\hat{u} + tv - x||^2\right] \\
\Rightarrow\ &\varphi(v) - \varphi(\hat{u}) \geqslant \frac{1}{\lambda}(x - \hat{u}, v - \hat{u})_H \text{ for all } v \in H \\
\Rightarrow\ &\frac{1}{\lambda}(x - \hat{u}) \in \partial\varphi(\hat{u}).
\end{aligned}
$$

From this claim it follows that the infimum is realized at $J_\lambda^A(x)$ and so

$$
\varphi_\lambda(x) = \varphi\left(J_\lambda^A(x)\right) + \frac{1}{2\lambda}||x - J_\lambda^A(x)||^2 = \varphi\left(J_\lambda^A(x)\right) + \frac{\lambda}{2}||A_\lambda(x)||^2.
$$

(b) Let $x, u \in H$. Recall that $A_\lambda(x) \in \partial\varphi(J_\lambda(x))$. So, we have

$$
\begin{aligned}
&(A_\lambda(x), J_\lambda(u) - J_\lambda(x))_H \leqslant \varphi(J_\lambda(u)) - \varphi(J_\lambda(x)) \\
\Rightarrow\ &\frac{\lambda}{2}\left[||A_\lambda(u)||^2 - ||A_\lambda(x)||^2 + \frac{2}{\lambda}(A_\lambda(x), J_\lambda(u) - J_\lambda(x))_H\right] \\
&\leqslant \varphi_\lambda(u) - \varphi_\lambda(x) \\
\Rightarrow\ &(A_\lambda(x), u - x)_H \leqslant \varphi_\lambda(u) - \varphi_\lambda(x) \text{ (since } J_\lambda = I - \lambda A_\lambda).
\end{aligned} \tag{2.71}
$$

Interchanging the roles of $x, u \in H$ in the above argument, we also have

$$
\begin{aligned}
&\varphi_\lambda(x) - \varphi_\lambda(u) \geqslant (A_\lambda(u), x - u)_H = (A_\lambda(x), x - u)_H \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad + (A_\lambda(u) - A_\lambda(x), x - u)_H \\
\Rightarrow\ &0 \leqslant \varphi_\lambda(u) - \varphi_\lambda(x) - (A_\lambda(x), u - x) \leqslant \frac{1}{\lambda}||x - u||^2 \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(see (71))and Theorem 2.9.10(c))} \\
\Rightarrow\ &\varphi_\lambda'(x) = A_\lambda(x).
\end{aligned}
$$

This proves that φ_λ is continuously Fréchet differentiable and convex.
(c) Clearly for all $0 < \lambda < \mu$, we have

$$
\varphi_\mu(x) \leqslant \varphi_\lambda(x) \leqslant \varphi(x) \text{ for all } x \in H \tag{2.72}
$$

and from (a) we also have

$$
\varphi(J_\mu(x)) \leqslant \varphi_\mu(x). \tag{2.73}
$$

From Proposition 2.9.7 it follows that for each $x \in \overline{D(A)}$ we have $J_\lambda(x) \to x$ in H as $\lambda \to 0^+$. Hence

$$\varphi(x) \leqslant \liminf_{\lambda\to 0^+} \varphi(J_\lambda(x)) \leqslant \liminf_{\lambda\to 0^+} \varphi_\lambda(x) \leqslant \limsup_{\lambda\to 0^+} \varphi_\lambda(x) \leqslant \varphi(x) \text{ (see (2.72), (2.73))}$$
$$\Rightarrow \varphi_\lambda(x) \uparrow \varphi(x) \text{ as } \lambda \to 0^+ \text{ for all } x \in \overline{D(A)}.$$

If $x \notin \overline{D(A)}$, then $||x - J_\lambda(x)|| \to ||x - \mathrm{proj}_{\overline{\mathrm{D(A)}}}(x)|| > 0$ as $\lambda \to 0^+$. We have

$$\lambda||A_\lambda(x)||^2 = ||A_\lambda(x)||\,||x - J_\lambda(x)|| \to +\infty \text{ as } \lambda \to 0^+$$
$$\text{(since } ||A_\lambda(x)|| \to \infty \text{ and } ||x - J_\lambda(x)|| \geqslant d(x, \overline{D(A)})).$$

Therefore

$$\varphi_\lambda(x) \to +\infty = \varphi(x) \text{ as } \lambda \to 0^+.$$

(d) Recall that $\left(J_{\lambda_n}(x_n), \partial\varphi_{\lambda_n}(x_n)\right) \in \mathrm{Gr}\,\partial\varphi$ and $||J_{\lambda_n}(x_n) - J_{\lambda_n}(x)|| \leqslant ||x_n - x||$ for all $n \geqslant 1$. Therefore

$$J_{\lambda_n}(x_n) \longrightarrow \mathrm{proj}_{\overline{D(A)}}(x) \text{ as } n \to \infty \text{ (see Theorem 2.9.10}(b)).$$

Note that

$$(\partial\varphi_\lambda(x_n), u - x_n)_H \leqslant \varphi_{\lambda_n}(u) - \varphi_{\lambda_n}(x_n) \leqslant \varphi(u) - \varphi_{\lambda_n}(x_n)$$
$$\text{for all } u \in H,\ \text{all } n \geqslant 1$$
$$\Rightarrow \varphi\left(J_\lambda^A(x_n)\right) \leqslant \varphi_{\lambda_n}(x_n) \leqslant M \text{ for some } M > 0,\ \text{all } n \geqslant 1 \text{ (chose } u \in \mathrm{dom}\,\varphi)$$
$$\Rightarrow x \in \mathrm{dom}\,\varphi.$$

But from Corollary 2.7.12, we have $\overline{D(A)} = \overline{\mathrm{dom}\,\varphi}$ and so $x \in \overline{D(A)}$. Hence

$$J_{\lambda_n}(x_n) \to x \text{ in } H$$
$$\Rightarrow (x, x^*) \in \mathrm{Gr}\,\partial\varphi \text{ (see Proposition (2.6.5))}.$$

(e) From (a) we have

$$\varphi_{\lambda_n}(x_n) = \varphi(J_{\lambda_n}(x_n)) + \frac{\lambda_n}{2}||A_{\lambda_n}(x_n)||^2.$$

By hypothesis, $\left\{A_{\lambda_n}(x_n) = \partial\varphi_{\lambda_n}(x_n)\right\}_{n\geqslant 1} \subseteq H$ is bounded. So

$$\frac{\lambda_n}{2}||A_{\lambda_n}(x_n)||^2 \to 0 \text{ as } n \to \infty.$$

Therefore

$$\liminf_{n\to\infty} \varphi_{\lambda_n}(x_n) = \liminf_{n\to\infty} \varphi(J_{\lambda_n}(x_n)) \geqslant \varphi\left(\mathrm{proj}_{\overline{D(A)}}(x)\right). \tag{2.74}$$

If $x \notin \overline{D(A)}$, then

$$\lambda_n ||A_{\lambda_n}(x)||^2 \to +\infty \text{ (see the proof of } (c))$$

and $\lambda_n ||A_{\lambda_n}(x_n)|| \leqslant ||x - x_n|| + \lambda_n ||A_{\lambda_n}(x_n)|| \to 0$ as $n \to \infty$, a contradiction. This proves that $x \in \overline{D(A)}$ and so

$$\liminf_{n\to\infty} \varphi_{\lambda_n}(x_n) \geqslant \varphi(x) \quad \text{(see ((2.74)))}. \tag{2.75}$$

On the other hand from the convexity of φ_{λ_n} we have

$$\begin{aligned} &\varphi_{\lambda_n}(x) - (\partial\varphi_{\lambda_n}(x_n), x - x_n)_H \geqslant \varphi_{\lambda_n}(x_n) \text{ for all } n \geqslant 1 \\ \Rightarrow\ &\varphi(x) \geqslant \limsup_{n\to\infty} \varphi_{\lambda_n}(x_n). \end{aligned} \tag{2.76}$$

From (2.75) and (2.76) we conclude that

$$\varphi_{\lambda_n}(x_n) \to \varphi(x).$$

(f) This follows directly from the definition of the subdifferential (see Definition 2.7.1).

(g) We have

$$||\partial\varphi_\lambda(x)|| \leqslant ||\partial\varphi_\lambda(x) - \partial\varphi_\lambda(v)|| + ||\partial\varphi_\lambda(v)|| \leqslant \frac{1}{\lambda}||x - v|| + ||\partial\varphi_\lambda(v)|| \tag{2.77}$$
$$\text{(see Theorem 2.9.10 } (c)).$$

So, from (f) it follows that

$$\begin{aligned} \varphi_\lambda(x) - \varphi_\lambda(u) &\leqslant ||\partial\varphi_\lambda(x)||\ ||x - u|| \\ &\leqslant \left(\frac{1}{\lambda}||x - v|| + ||\partial\varphi_\lambda(v)||\right) ||x - u|| \text{ (see (2.77))}. \end{aligned} \tag{2.78}$$

Similarly, we have

$$\varphi_\lambda(u) - \varphi_\lambda(x) \leqslant \left(\frac{1}{\lambda}||u - v|| + ||\partial\varphi_\lambda(v)||\right) ||x - u||. \tag{2.79}$$

From (2.78) and (2.79) we conclude that

$$|\varphi_\lambda(x) - \varphi_\lambda(u)| \leqslant \left(2||\partial\varphi_\lambda(v)|| + \frac{1}{\lambda}||x - v|| + \frac{1}{\lambda}||u - v||\right) ||x - u||$$
$$\text{for all } x, u, v \in H, \text{ all } \lambda > 0.$$

The proof is now complete. $\square$

Remark 2.9.13 The regularization φ_λ of φ, is usually called the "Moreau–Yosida regularization" of φ. The minimization involved in the definition of φ_λ in the language of convex analysis is called "infimal convolution". Finally, in part (*e*) if we drop the requirement that $\{\partial\varphi_\lambda(x_n)\}_{n\geqslant 1} \subseteq H$ is bounded, then we have $\varphi(x) \leqslant \liminf_{n\to\infty} \varphi_{\lambda_n}(x_n)$.

An interesting by-product of this theorem is the following result:

Corollary 2.9.14 *If H is a Hilbert space and $\varphi : H \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is convex and lower semicontinuous, then $D(\partial\varphi) \subseteq \operatorname{dom}\varphi \subseteq \overline{\operatorname{dom}\varphi} = \overline{D(\partial\varphi)}$.*

Example 2.9.15 Suppose H is a Hilbert space and $C \subseteq H$ a nonempty closed, convex set. Let $\delta_C(x) = \begin{cases} 0 & \text{if } x \in C \\ +\infty & \text{if } x \notin C \end{cases}$ (the indicator function of C, see Example 2.7.7). We know that $\delta_C : H \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is convex and lower semicontinuous. Also

$$\partial\delta_C(x) = \{x^* \in H : (x^*, u - x)_H \leqslant 0 \text{ for all } u \in C\}.$$

So, $\partial\delta_C(x)$ is a closed, convex cone in H, known as the normal cone to C at x and denoted by $N_C(x)$. If $x \notin C$, then $N_C(x) = \emptyset$ and $x \in \operatorname{int} C$, then $N_C(x) = \{0\}$. The normal cone is the polar of the tangent cone $T_C(x) = \overline{\bigcup_{\lambda>0} \lambda(C - x)}$. If $A = \partial\delta_C$, then

$$J_\lambda^A = (I + \lambda\partial\delta_C)^{-1} = \operatorname{proj}_C.$$

Indeed, $u = (I + \lambda\partial\delta_C)^{-1}(x) \Leftrightarrow x - u \in \lambda\partial\delta_C(u) \Leftrightarrow (x - u, v - u)_H \leqslant 0$ for all $v \in C \Leftrightarrow u = \operatorname{proj}_C(x)$.

Moreover, we have

$$A_\lambda(x) = \frac{1}{\lambda}[x - \operatorname{proj}_C(x)] \text{ and } (\delta_C)_\lambda(x) = \frac{1}{2\lambda}||x - \operatorname{proj}_C(x)||^2.$$

Finally, let us look at linear monotone operators.

Proposition 2.9.16 *If H is a Hilbert space and $A : H \to H$ linear, unbounded, maximal monotone, then A is cyclically monotone if and only if $A = A^*$ (that is, A is self-adjoint); moreover, we have $A = \partial\varphi$ with*

$$\varphi(x) = \begin{cases} \frac{1}{2}||A^{1/2}x||^2 & \textit{if } x \in D(A^{1/2}) \\ +\infty & \textit{otherwise.} \end{cases}$$

Proof First recall that a linear maximal monotone operator has a unique linear maximal monotone square root $A^{1/2}$ such that $(A^{1/2})^2 = A$ and if A is self-adjoint, then so is $A^{1/2}$ (see Kato [229, p. 281]).

$\underline{\Rightarrow}$ Since by hypothesis A is cyclically monotone, by virtue of Theorem 2.8.14, we can find $\varphi : H \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ a convex and lower semicontinuous function such that $A = \partial\varphi$. Since $A(0) = 0$, without any loss of generality we may assume that $\varphi(0) = 0$. From Proposition 2.9.12 we know that $A_\lambda = (\partial\varphi)_\lambda = \partial\varphi_\lambda$ for all $\lambda > 0$. Then by the chain rule, we have

$$\begin{aligned}
&\frac{d}{dt}\varphi_\lambda(tu) = (A_\lambda(tu), u)_H = t(A_\lambda(u), u)_H \text{ for all } u \in H\\
&\Rightarrow \varphi_\lambda(u) = \int_0^1 (A_\lambda(u), u)_H t\, dt = \frac{1}{2}(A_\lambda(u), u)_H \text{ (since } \varphi_\lambda(0) = 0)\\
&\Rightarrow \partial\varphi_\lambda = A_\lambda = \frac{1}{2}(A_\lambda + A_\lambda^*).
\end{aligned}$$

So, we have $A_\lambda = A_\lambda^*$ for all $\lambda > 0$, hence $A = A^*$.
$\underline{\Leftarrow}$ Note that for all $u \in D(A)$ and all $v \in D(A^{1/2})$ we have

$$\begin{aligned}
&\frac{1}{2}||A^{1/2}v||^2 + \frac{1}{2}||A^{1/2}u||^2 = (A(u), v)_H = (A^{1/2}(u), A^{1/2}(v))_H\\
&\Rightarrow A(u) \subseteq \partial\varphi(u), \text{ hence } A = \partial\varphi \text{ (recall that } A \text{ is maximal monotone).}
\end{aligned}$$

□

We conclude with the following simple maximality criterion for linear monotone operators. The result can be found in Kato [229, p. 279].

Proposition 2.9.17 *If H is a Hilbert space and $A : H \to H$ is linear monotone, then A is maximal monotone if and only if A is closed and A^* is monotone.*

Remark 2.9.18 A linear maximal monotone operator is necessarily densely defined (see Kato [229, p. 279]).

2.10 Operators of Monotone Type

In this section we introduce some generalizations of maximal monotonicity which are important in the study of nonlinear boundary value problems and in degree theory.

Definition 2.10.1 Let X be a reflexive Banach space and $A : X \to 2^{X^*}$.

(a) The multivalued map A is said to be pseudomonotone if the following hold:

- a_1 For every $x \in X$, $A(x)$ is nonempty, convex and w-compact;
- a_2 A is usc from every finite-dimensional subspace V of X into X_w^* (that is, X^* furnished with the weak topology);
- a_3 If $x_n \xrightarrow{w} x$ in X and $x_n^* \in A(x_n)$ satisfy $\limsup_{n\to\infty} \langle x_n^*, x_n - x\rangle \leqslant 0$, then for all $u \in X$, we can find $x^*(u) \in A(x)$ such that

$$\langle x^*(u), x - u\rangle \leqslant \liminf_{n\to\infty} \langle x_n^*, x_n - u\rangle .$$

(b) The multivalued map A is said to be generalized pseudomonotone if the following is true:
"for every sequence $(x_n, x_n^*) \in \mathrm{Gr}\, A$, $n \geqslant 1$, such that $x_n \xrightarrow{w} x$ in X, $x_n^* \xrightarrow{w} x^*$ in X^* and

$$\limsup_{n\to\infty} \langle x_n^*, x_n - x\rangle \leqslant 0,$$

we have $x^* \in A(x)$ and $\langle x_n^*, x_n\rangle \to \langle x^*, x\rangle$."

Proposition 2.10.2 *If X is a reflexive Banach space and $A : X \to 2^{X^*}$ is pseudomonotone, then A is generalized pseudomonotone.*

Proof Let $\{(x_n, x_n^*)\}_{n\geqslant 1} \subseteq \mathrm{Gr}\, A$ and assume that

$$x_n \xrightarrow{w} x \text{ in } X,\ \ x_n^* \xrightarrow{w} x^* \text{ in } X^* \text{ and } \limsup_{n\to\infty} \langle x_n^*, x_n - x\rangle \leqslant 0. \tag{2.80}$$

Since A is pseudomonotone, from $[a_3]$ of Definition 2.10.1(a), for every $u \in X$, we can find $x^*(u) \in A(x)$ such that

$$\langle x^*(u), x - u\rangle \leqslant \liminf_{n\to\infty} \langle x_n^*, x_n - u\rangle . \tag{2.81}$$

Note that $\{\langle x_n^*, x_n\rangle\}_{n\geqslant 1} \subseteq \mathbb{R}$ is bounded (see (2.80)). So, we may assume that

$$\begin{aligned} &\langle x_n^*, x_n\rangle \longrightarrow \vartheta \text{ in } \mathbb{R} \\ &\Rightarrow \limsup_{n\to\infty} \langle x_n^*, x_n - x\rangle = \vartheta - \langle x^*, x\rangle \leqslant 0 \text{ (see (2.80))}. \end{aligned} \tag{2.82}$$

Also, we have

$$\begin{aligned} &\vartheta - \langle x^*, u\rangle \geqslant \liminf_{n\to\infty} \langle x_n^*, x_n - u\rangle \geqslant \langle x^*(u), x - u\rangle \ \text{(see (2.81))} \\ \Rightarrow\ &\langle x^*, x - u\rangle \geqslant \langle x^*(u), x - u\rangle \text{ for all } u \in X \text{ (see (2.82))}. \end{aligned} \tag{2.83}$$

We show that $x^* \in A(x)$. Arguing by contradiction, suppose that $x^* \notin A(x)$. Then by the strong separation theorem, we can find $h \in X\setminus\{0\}$ such that

$$\langle x^*, h\rangle < m \leqslant \inf\{\langle u^*, h\rangle : u^* \in A(x)\}. \tag{2.84}$$

In (2.83) we choose $u = x - h$ and have

$$\langle x^*(u), h\rangle \leqslant \langle x^*, h\rangle \text{ with } x^*(u) \in A(x),$$

which contradicts (2.84). This proves that $x^* \in A(x)$.

Next, we show that $\left\langle x_n^*, x_n \right\rangle \to \langle x^*, x\rangle$. By virtue of the pseudomonotonicity, by choosing $u = x$, we have

$$\begin{aligned} 0 &\leqslant \liminf_{n\to\infty} \left\langle x_n^*, x_n - x \right\rangle \\ \Rightarrow \left\langle x^*, x \right\rangle &\leqslant \liminf_{n\to\infty} \left\langle x_n^*, x_n \right\rangle . \end{aligned} \tag{2.85}$$

On the other hand, from (2.80) we have

$$\limsup_{n\to\infty} \left\langle x_n^*, x_n \right\rangle \leqslant \left\langle x^*, x \right\rangle . \tag{2.86}$$

From (2.85) and (2.86) we conclude that

$$\left\langle x_n^*, x_n \right\rangle \to \left\langle x^*, x \right\rangle .$$

So, we have that A is generalized pseudomonotone. $\square$

By strengthening a little the conditions on A, we have a converse of this proposition.

Proposition 2.10.3 *If X is a reflexive Banach space, $A : X \to 2^{X^*}$ is locally bounded and generalized pseudomonotone and for all $x \in X$, $A(x)$ is nonempty, closed and convex, then A is pseudomonotone.*

Proof Since A is locally bounded, we see that the values of A are nonempty, convex and w-compact. So, condition $[a_1]$ in Definition 2.10.1 (a) is satisfied.

Let $V \subseteq X$ be a finite-dimensional subspace and consider $A|_V : V \to 2^{X^*}$. Let $C \subseteq X^*$ be nonempty, w-closed and let $x_n \in (A|_V)^-(C) = \{x \in V : A(x) \cap C \neq \emptyset\}$, $n \geqslant 1$, such that $x_n \to x$ in V. We can find $x_n^* \in A|_V(x_n) \cap C$. Since A is locally bounded, $\{x_n^*\}_{n\geqslant 1} \subseteq C$ is bounded. So, the reflexivity of X and the Eberlein–Smulian theorem imply that by passing to a subsequence if necessary, we may assume that $x_n^* \xrightarrow{w} x^*$ in X^*. Evidently $x^* \in C$ and the generalized pseudomonotonicity of A implies $x^* \in A(x)$. Hence $x \in (A|_V)^- (C)$ and so we conclude that $A|_V : V \to 2^{X^*}$ is usc into X_w^* (see Proposition 2.5.3). So, condition $[a_2]$ in Definition 2.10.1 (a) is satisfied.

Next, suppose that $x_n \xrightarrow{w} x$ in X, $x_n^* \in A(x_n)$ and $\limsup_{n\to\infty} \left\langle x_n^*, x_n - x \right\rangle \leqslant 0$. The local boundedness of A implies that $\{x_n^*\}_{n\geqslant 1} \subseteq X^*$ is bounded and so we may assume that $x_n^* \xrightarrow{w} x^*$ in X^*. The generalized pseudomonotonicity of A implies

$$x^* \in A(x) \text{ and } \left\langle x_n^*, x_n \right\rangle \to \left\langle x^*, x \right\rangle . \tag{2.87}$$

Arguing by contradiction, suppose that condition $[a_3]$ in Definition 2.10.1 (a) is not true. Then we can find $u \in X$ such that

$$\liminf_{n\to\infty} \langle x_n^*, x_n - u \rangle < \inf \left\{ \langle v^*, x - u \rangle : v^* \in A(x) \right\}$$
$$\Rightarrow \langle x^*, x - u \rangle < \inf[\langle v^*, x - u \rangle : v^* \in A(x)] \text{ (see (2.87)), a contradiction.}$$

So $[a_3]$ holds and we have proved that A is pseudomonotone. ☐

Next, we show that the concepts introduced in Definition 2.10.1 extend the notion of maximal monotonicity, at least for maps A with $D(A) = X$.

Proposition 2.10.4 *If X is a reflexive Banach space and $A : X \to 2^{X^*}$ is maximal monotone, then A is generalized pseudomonotone.*

Proof Let $\{(x_n, x_n^*)\}_{n\geqslant 1} \subseteq \mathrm{Gr}\, A$ such that

$$x_n \xrightarrow{w} x \text{ in } X,\ x_n^* \xrightarrow{w} x^* \text{ in } X^* \text{ and } \limsup_{n\to\infty} \langle x_n^*, x_n - x \rangle \leqslant 0. \tag{2.88}$$

For every $(u, u^*) \in \mathrm{Gr}\, A$, we have

$$\begin{aligned}
\langle x_n^*, x_n \rangle &= \langle x_n^* - u^*, x_n - u \rangle + \langle x_n^*, u \rangle + \langle u^*, x_n \rangle - \langle u^*, u \rangle \\
&\geqslant \langle x_n^*, u \rangle + \langle u^*, x_n \rangle - \langle u^*, u \rangle \ \text{ (since } A \text{ is monotone)} \\
\Rightarrow \langle x^*, x \rangle &\geqslant \langle x^*, u \rangle + \langle u^*, x \rangle - \langle u^*, u \rangle \ \text{ (see (2.88))} \\
\Rightarrow \langle x^* - u^*, x - u \rangle &\geqslant 0 \\
\Rightarrow (x, x^*) &\in \mathrm{Gr}\, A \ \text{ (since } A \text{ is maximal monotone).}
\end{aligned}$$

Also, for every $n \geqslant 1$, we have

$$\begin{aligned}
& \langle x_n^* - x^*, x_n - x \rangle \geqslant 0 \\
\Rightarrow & \lim_{n\to\infty} \langle x_n^* - x^*, x_n - x \rangle = 0 \text{ (see (2.88))} \\
\Rightarrow & \langle x_n^*, x_n \rangle \to \langle x^*, x \rangle .
\end{aligned}$$

Therefore A is generalized pseudomonotone. ☐

Proposition 2.10.5 *If X is a reflexive Banach space and $A : X \to 2^{X^*}$ is a maximal monotone map with $D(A) = X$, then A is pseudomonotone.*

Proof From Proposition 2.10.4 we know that A is generalized pseudomonotone. From Proposition 2.6.8 we know that A is locally bounded, while from Proposition 2.6.13 we have that the values of A are nonempty, w-compact and convex. So, we can apply Proposition 2.10.3 and conclude that A is pseudomonotone. ☐

Pseudomonotonicity is preserved under addition.

Proposition 2.10.6 *If X is a reflexive Banach space and $A, F : X \to 2^{X^*}$ are pseudomonotone maps, then $x \to (A + F)(x) = A(x) + F(x)$ is pseudomonotone, too.*

Proof Clearly $A + F$ has nonempty, closed and convex values. Also, for every finite-dimensional subspace $V \subseteq X$, $(A + F)|_V$ is usc into X_w.

We need to verify condition $[a_3]$ in Definition 2.10.1 (a). So, let $\{(x_n, x_n^*)\}_{n\geqslant 1} \subseteq \mathrm{Gr}(A + F)$ such that

$$x_n \xrightarrow{w} x \text{ in } X \text{ and } \limsup_{n\to\infty} \langle x_n^*, x_n - x\rangle \leqslant 0. \tag{2.89}$$

We have

$$x_n^* = u_n^* + v_n^* \text{ with } u_n^* \in A(x_n),\ \ v_n^* \in F(x_n) \text{ for all } n \geqslant 1.$$

From (2.89) we have

$$\limsup_{n\to\infty} \left[\langle u_n^*, x_n - x\rangle + \langle v_n^*, x_n - x\rangle\right] \leqslant 0. \tag{2.90}$$

We claim that (2.90) implies

$$\limsup_{n\to\infty} \langle u_n^*, x_n - x\rangle \leqslant 0 \text{ and } \limsup_{n\to\infty} \langle v_n^*, x_n - x\rangle \leqslant 0. \tag{2.91}$$

Arguing by contradiction, suppose that one of the limits in (2.91) does not hold. To fix things assume that

$$\limsup_{n\to\infty} \langle u_n^*, x_n - x\rangle > 0.$$

By passing to a suitable subsequence if necessary, we can say that

$$\lim_{n\to\infty} \langle u_n^*, x_n - x\rangle > 0.$$

Then (2.90) implies that

$$\limsup_{n\to\infty} \langle v_n^*, x_n - x\rangle \leqslant -\xi < 0. \tag{2.92}$$

The pseudomonotonicity of F implies that for every $h \in X$, we can find $v^*(h) \in A(x)$ such that

$$\langle v^*(h), x - h\rangle \leqslant \liminf_{n\to\infty} \langle v_n^*, x_n - h\rangle.$$

Choose $h = x$. Then

$$0 \leqslant \liminf_{n\to\infty} \langle v_n^*, x_n - x\rangle. \tag{2.93}$$

Comparing (2.92) and (2.93), we reach a contradiction. This proves (2.91). Then from the pseudomonotonicity of A and F, we know that given $h \in X$, we can find $u^*(h) \in A(x)$ and $v^*(h) \in F(x)$ such that

$$\begin{aligned}\left\langle u^*(h), x-h\right\rangle &\leqslant \liminf_{n\to\infty}\left\langle u_n^*, x_n-h\right\rangle \text{ and}\\ &\left\langle v^*(h), x-h\right\rangle \leqslant \liminf_{n\to\infty}\left\langle v_n^*, x_n-h\right\rangle.\end{aligned} \tag{2.94}$$

Let $x^*(h) = u^*(h) + v^*(h) \in (A + F)(x)$. Then

$$\left\langle x^*(h), x-h\right\rangle \leqslant \liminf_{n\to\infty}\left\langle x^*(h), x_n-h\right\rangle \text{ (see (2.94))}$$

and this proves that $A + F$ is pseudomonotone. □

The importance of pseudomonotone operators is due to their remarkable surjectivity properties. In fact, as for maximal monotone maps (see Corollary 2.8.7), we can show that pseudomonotonicity and strong coercivity imply surjectivity. To do this we need some auxiliary results. The first auxiliary result shows that in a reflexive Banach space without any separability assumption, the weak closure of a bounded set can be completely characterized by weakly convergent sequence. In general, even for Hilbert spaces, it is impossible to characterize the weak closure of each unbounded set by weakly convergent sequences. Consider the following counterexample attributed to von Neumann.

Example 2.10.7 Let $H = l^2$ and let C be the set of vectors $x_{mn} \in l^2, m, n \geqslant 1$, which have coordinates

$$x_{mn}(k) = \begin{cases} 1 & \text{if } k = m \\ m & \text{if } k = n \\ 0 & \text{otherwise.}\end{cases}$$

Then $0 \in \overline{C}^w$ but there is no sequence of vectors in C weakly converging to 0. Indeed, if we can find $\{x_{m_kn_k}\}_{k\geqslant 1} \subseteq C$ such that $x_{m_k,n_k} \xrightarrow{w} 0$, then $(u, x_{m_kn_k})_{l^2} \to 0$ as $k \to \infty$ for all $u \in l^2$. Choosing $u = \left(\frac{1}{k}\right)_{k\geqslant 1} \in l^2$, we see that this cannot happen. On the other hand, if $U = \{x \in l^2 : (u, x)_{l^2} < \epsilon\}$ for some $u \in l^2$ and $\epsilon > 0$ (a basic weak neighborhood of the origin in l^2), then we have that $x_{mn} \in U$, if we take $m \geqslant 1$ such that $|y_m| < \frac{\epsilon}{2}$ and then $n \geqslant 1$ such that $|y_n| < \frac{\epsilon}{2m}$.

Proposition 2.10.8 *If X is a reflexive Banach space, $C \subseteq X$ is bounded and $x \in \overline{C}^w$, then there exists a sequence $\{x_n\}_{n\geqslant 1} \subseteq C$ such that $x_n \xrightarrow{w} x$ in X.*

Proof First we produce a countable set $D \subseteq C$ such that $0 \in \overline{D}^w$. We fix $m, n \in \mathbb{N}$ and consider $(\overline{B_1}^*)^m$ (=the product of m-copies of $\overline{B_1}^* = \{x^* \in X^* : ||x^*||_* \leqslant 1\}$). Since $x \in \overline{C}^w$, for every $(x_k^*)_{k=1}^m \in (\overline{B}_1^x)^m$ we can find $u \in C$ such that

$$|\left\langle x_k^*, u-x\right\rangle| < \frac{1}{n} \text{ for all } k = 1, \ldots, m. \tag{2.95}$$

For every $u \in C$, we introduce the set

$$U_{mn}(u) = \{(x_k^*)_{k=1}^m \in (X^*)^m : (2.95) \text{ holds}\}.$$

Clearly, $U_{mn}(u)$ is w-open in the reflexive Banach space $(X^*)^m$. By Tychonov and Alaoglu's theorems, $(\overline{B}_1^*)^m$ is a w-compact subset of $(X^*)^m$. We have

$$(\overline{B}_1^*)^m \subseteq \bigcup_{\mathrm{u} \in \mathrm{C}} U_{mn}(u).$$

So, we can find a finite subset $F_{mn} \subseteq C$ such that

$$(\overline{B}_1^*)^m \subseteq \bigcup_{\mathrm{u} \in \mathrm{F_{mn}}} U_{mn}(u). \tag{2.96}$$

Let $D = \bigcup_{\mathrm{m,n} \geqslant 1} F_{mn}$. Then D is a countable subset of C.

We show that $x \in \overline{D}^w$. Let V be a weak neighborhood of x. Then there exist $m, n \in \mathbb{N}$ and $(x_k^*)_{k=1}^m \subseteq (\overline{B}_1^*)^m$ such that every $u \in X$ which satisfies (2.95) belongs to V. By (2.96) there exists a $u \in F_{mn}$ such that $(x_k^*)_{k=1}^m \in U_{mn}(u)$. So, u satisfies (2.95), hence $u \in V$. Therefore $V \cap D \neq \emptyset$ and this implies that $x \in \overline{D}^w$.

Let $V = \overline{\text{span}}\, D$. Then V is a separable closed linear subspace of X. The weak topology of V is the restriction on V of the weak topology of X. The weak topology of $\overline{D}^w \subseteq V$ is metrizable. Since $x \in \overline{D}^w$, we can find $\{x_n\}_{n \geqslant 1} \subseteq D \subseteq X$ such that $x_n \xrightarrow{w} x$ in X. $\square$

The second auxiliary result is the following surjectivity result for usc multifunctions on finite-dimensional Banach spaces.

Proposition 2.10.9 *If X is a finite-dimensional Banach space and $F : X \to P_{kc}(X^*)$ is usc and strongly coercive, then F is surjective (that is, $R(F) = X^*$).*

Proof Let $h^* \in X^*$ and let $\hat{F}(x) = F(x) - h^*$ for all $x \in X$. We introduce the following homotopy

$$\hat{F}_t(x) = t\hat{F}(x) + (1-t)x \text{ for all } t \in [0, 1], \text{ all } x \in X.$$

The strong coercivity of $F(\cdot)$ implies that we can find $r > 0$ such that

$$\langle \hat{F}(x), x\rangle > 0 \text{ for all } ||x|| \geqslant r. \tag{2.97}$$

Then we have

$$\langle \hat{F}_t(x), x\rangle = t\langle \hat{F}(x), x\rangle + (1-t)||x||^2 > 0 \text{ for all } ||x|| = r \text{ and all } t \in [0, 1].$$

So, the homotopy invariance property of Brouwer's degree for multifunctions (see Sect. 3.1), implies that

$$\begin{aligned}
&\hat{d}_B(F, B_r, 0) = \hat{d}_B(I, B_r, 0) = 1 \text{ (here } B_r = \{x \in X : ||x|| < r\})\\
\Rightarrow\ & 0 \in \hat{F}x \text{ has a solution } x \in B_r\\
\Rightarrow\ & h^* \in F(x).
\end{aligned}$$

Since $h^* \in X^*$ is arbitrary, we conclude that F is surjective. □

Now we are ready for the surjectivity theorem for pseudomonotone maps.

Theorem 2.10.10 *If X is a reflexive Banach space and $A : X \to 2^{X^*}$ is pseudomonotone and strongly coercive, then A is surjective (that is, $R(A) = X^*$).*

Proof Again the proof is based on Galerkin approximations.

Let $\mathscr{F}$ be the family of finite-dimensional subspace of X partially ordered by inclusion. Let $V \in \mathscr{F}$ and let $i_v \in \mathscr{L}(V, X)$ be the inclusion (embedding) map of V into X. Then $i_V^* \in \mathscr{L}(X^*, V^*)$ is the corresponding projection operator of X^* onto V^*. We consider the following finite-dimensional (Galerkin) approximations of A:

$$A_V = i_V^* \circ A \circ i_V : V \to 2^{V^*}.$$

Evidently, A_V has nonempty, convex, compact values, it is usc and also strongly coercive.

Given any $u^* \in X^*$, we consider the map $A_{u^*}(x) = A(x) - u^*$. This map has the same properties as A. So, it suffices to show that $0 \in R(A)$.

Applying Proposition 2.10.9, for every $V \in \mathscr{F}$, we can find $x_V \in V$ such that $0 \in A_V(x_V)$. Then

$$0 = i_V^* x_V^* \text{ for same } x_V^* \in A(x_V).$$

The strong coercivity of A implies that $\{x_V\}_{V \in \mathscr{F}}$ is bounded.

Let $V \in \mathscr{F}$ and introduce $S_V = \bigcup\limits_{\substack{V' \in \mathscr{F} \\ V' \supseteq V}} \{x_{V'}\}$. Then $S_V \subseteq \overline{B}_M$ for some large $M > 0$.

The reflexivity of X and the finite intersection property imply

$$\bigcap_{V \in \mathscr{F}} \overline{S_V}^w \neq \emptyset.$$

Let $x_0 \in \bigcap\limits_{V \in \mathscr{F}} \overline{S_V}^w$ and let $u \in X$. Choose $V \in \mathscr{F}$ such that $\{x_0, u\} \in V$. Then by virtue of Proposition 2.10.8 we can find $\{x_{V_k} = x_k\}_{k \geqslant 1} \subseteq S_V$ such that $x_k \xrightarrow{w} x_0$ in X. We have

$$\langle x_k^*, x_k - u \rangle = 0 \text{ and } x_k^* \in A(x_k) \text{ for all } k \geqslant 1.$$

The pseudomonotonicity of A implies that we can find $x^*(u) \in A(x_0)$ such that

$$\langle x^*(u), x_0 - u \rangle \leqslant \liminf \langle x_k^*, x_k - u \rangle = 0. \tag{2.98}$$

If $0 \notin A(x_0)$, then by the strong separation theorem, we can find $u \in X$ such that

$$0 < \inf \left\{ \left\langle x^*, x_0 - u \right\rangle : x^* \in A(x_0) \right\}. \tag{2.99}$$

Comparing (2.98) and (2.99), we reach a contradiction. Hence $0 \in A(x_0)$ and so $0 \in R(A)$. This proves the surjectivity of A. $\square$

We introduce some additional monotonicity type conditions and some other notions that are useful in connection with operators of monotone type.

Definition 2.10.11 Let X be a reflexive Banach space, $D \subseteq X$ nonempty and $A : D \to X^*$.

(a) We say that A is an $(S)_+$-map if for $\{x_n, x\}_{n\geqslant 1} \subseteq D$ we have

$$\text{“}x_n \xrightarrow{w} x \text{ and } \limsup_{n\to\infty} \langle A(x_n), x_n - x\rangle \leqslant 0 \Rightarrow x_n \to x \text{ in } X.\text{''}$$

(b) We say that A is a (P)-map if for $\{x_n, x\}_{n\geqslant 1} \subseteq D$ we have

$$\text{“}x_n \xrightarrow{w} x \Rightarrow \limsup_{n\to\infty} \langle A(x_n), x_n - x\rangle \geqslant 0.\text{''}$$

(c) We say that A is an (M)-map if for $\{x_n, x\}_{n\geqslant 1} \subseteq D$ we have

$$\text{“}x_n \xrightarrow{w} x \text{ in } X,\ A(x_n) \xrightarrow{w} x^* \text{ in } X^* \text{ and } \limsup_{n\to\infty} \langle A(x_n), x_n\rangle \leqslant \left\langle x^*, x\right\rangle \Rightarrow A(x) = x^*.\text{''}$$

(d) We say that A is demicontinuous if for $\{x_n, x\}_{n\geqslant 1} \subseteq D$ we have

$$\text{“}x_n \to x \text{ in } X \Rightarrow A(x_n) \xrightarrow{w} A(x) \text{ in } X^*.\text{''}$$

Remark 2.10.12 Sometimes, we say that A is an $(S)_+$ (resp. (P), (M))-map on D in order to emphasize that the property holds only on D and not on the whole space X. Evidently, demicontinuity is sequential continuity from X into X^*_w, which we already encountered in previous sections.

Proposition 2.10.13 *If X is reflexive Banach and $A, F : X \to X^*$, then*

(a) when A is demicontinuous and $(S)_+$, it is pseudomonotone;
(b) when A is monotone and F is completely continuous, $A + F$ is a (P)-map.

Proof (a) Let $x_n \xrightarrow{w} x$ in X and assume that $\limsup\limits_{n\to\infty} \langle A(x_n), x_n - x\rangle \leqslant 0$. Since A is an $(S)_+$-map we have $x_n \to x$ in X. Then the demicontinuity of A implies that $A(x_n) \xrightarrow{w} A(x)$ in X^*. So, for every $u \in X$, we have

$$\begin{aligned} &\langle A(x), x - u\rangle = \lim_{n\to\infty} \langle A(x_n), x_n - u\rangle \\ &\Rightarrow A \text{ is pseudomonotone (see Definition 2.10.1 } (a)). \end{aligned}$$

(*b*) Suppose that $x_n \xrightarrow{w} x$ in X. The complete continuity of F implies $F(x_n) \to F(x)$ in X^*. Also, from the monotonicity of A, we have

$$\begin{aligned} & \langle A(x_n), x_n - x\rangle \geqslant \langle A(x), x_n - x\rangle \text{ for all } n \geqslant 1 \\ \Rightarrow\ & \liminf_{n\to\infty} \langle A(x_n), x_n - x\rangle \geqslant 0. \end{aligned}$$

So, finally $\limsup_{n\to\infty} \langle A(x_n) + F(x_n), x_n - x\rangle \geqslant 0$ and we conclude that $A + F$ is a (P)-map. □

Proposition 2.10.14 *If X is a locally uniformly convex reflexive Banach space with a strictly convex dual X^*, then the duality map $J : X \to X^*$ is $(S)_+$.*

Proof Since X^* is strictly convex, J is single-valued (see Proposition 2.7.27). Suppose

$$x_n \xrightarrow{w} x \text{ in } X \text{ and } \limsup_{n\to\infty} \langle J(x_n), x_n - x\rangle \leqslant 0. \tag{2.100}$$

Using Proposition 2.7.23, we have

$$\begin{aligned} & (||x_n|| - ||x||)^2 \leqslant \langle J(x_n) - J(x), x_n - x\rangle \text{ for all } n \geqslant 1 \\ \Rightarrow\ & ||x_n|| \to ||x|| \text{ (see (2.100))} \\ \Rightarrow\ & x_n \to x \text{ in } X \text{ (from the Kadec–Klee property, see Remark 2.7.30).} \end{aligned}$$

The proof is now complete. □

A compact map need not be an (M)-map.

Example 2.10.15 Let $X = l^2$ and let $A(x) = \big((2 - ||x||)\delta_{k,1}\big)_{k\geqslant 1}$ (recall that $\delta_{k,1} = \begin{cases} 1 \text{ if } k = 1 \\ 0 \text{ if } k \neq 1 \end{cases}$, the Kronecker symbol) for all $x \in l^2$. Clearly $A : l^2 \to l^2$ is compact (in fact it is finite rank). Let $x_n = (\delta_{k,n})_{k\geqslant 1} \in l^2$. We have $A(x_n) = (\delta_{k,1})_{k\geqslant 1}$ and $x_n \to 0$, $A(x_n) \to (\delta_{k,1})_{k\geqslant 1}$ in l^2 and $\lim_{n\to\infty} (A(x_n), x_n)_{l^2} = 0$. Since $A(0) = (2\delta_{k,1})_{k\geqslant 1}$, the map A is not an (M)-map (see Definition 2.10.11 (*c*)).

In particular, the above example shows that the sum of an (M)-map and a compact map need not be an (M)-map (just note the zero map is an (M)-map).

Proposition 2.10.16 *If X is a reflexive Banach space, $A : X \to X^*$ is a demicontinuous $(S)_+$-map and $K : X \to X^*$ is compact, then $A + K$ is an (M)-map.*

Proof Suppose that $x_n \xrightarrow{w} x$ in X, $(A + K)(x_n) \xrightarrow{w} x^*$ in X^* and

$$\limsup_{n\to\infty} \langle (A + K)(x_n), x_n\rangle \leqslant \langle x^*, x\rangle.$$

The compactness of K implies that we can find a subsequence $\{x_{n_k}\}_{k\geqslant 1}$ of $\{x_n\}_{n\geqslant 1}$ and $v^* \in X^*$ such that $K(x_{n_k}) \to v^*$ in X^*. Then we have

$$\limsup_{k\to\infty} \left\langle A(x_{n_k}), x_{n_k} - x\right\rangle \leqslant 0$$
$$\Rightarrow x_{n_k} \to x \text{ in } X \text{ (since } A \text{ is } (S)_+).$$

The demicontinuity of A implies $A(x_{n_k}) \xrightarrow{w} A(x)$ in X^*, while the continuity of K implies $K(x_{n_k}) \to K(x) = v^*$ in X^*. So,

$$x^* = (A + K)(x)$$

and this proves that $A + K$ is an (M)-map. □

Corollary 2.10.17 *If X is a locally uniformly convex reflexive Banach space with strictly convex dual X^* and $K : X \to X^*$ is compact, then $J + K : X \to X^*$ is an (M)-map; in particular, if $X = H =$ a Hilbert space and $K : H \to H$ is compact, then $x \mapsto x + K(x)$ is an (M)-map.*

In general, $(S)_+$-maps exhibit remarkable stability properties under perturbations.

Proposition 2.10.18 *If X is a reflexive Banach space and $A, F : D \subseteq X \to X^*$ are demicontinuous maps, then*

(a) if A is an $(S)_+$-map on D and F is a (P)-map on D, then $x \mapsto (A + F)(x)$ is an $(S)_+$-map on D;
(b) if A and F are both $(S)_+$-maps on $D = X$, then $A + F$ is an $(S)_+$-map;
(c) if A is an $(S)_+$ map on D and F is compact on D, then $A + F$ is an $(S)_+$-map.

Proof (a) This is immediate from Definitions 2.10.11 (a) and (b).

(b) From Proposition 2.10.13 (a), F is pseudomonotone. Then directly from the definition of pseudomonotonicity, we have that F is a (P)-map. So, again we can apply part (a) and conclude that $A + F$ is an $(S)_+$-map.

(c) A compact map on D is clearly a (P)-map. So, again we can apply part (a). □

Proposition 2.10.19 *If X is a reflexive Banach space, $A : X \to X^*$ is an (M)-map and $F : X \to X^*$ is sequentially continuous from X_w into X_w^* and $x \to \langle F(x), x\rangle$ is sequentially weakly lower semicontinuous, then $A + F : X \to X^*$ is an (M)-map.*

Proof Let $x_n \xrightarrow{w} x$ in X, $(A + F)(x_n) \xrightarrow{w} x^*$ in X^* and $\limsup_{n\to\infty} \langle (A + F)(x_n), x_n\rangle \leqslant \langle x^*, x\rangle$. Then $F(x_n) \xrightarrow{w} F(x)$ in X^* and $A(x_n) \xrightarrow{w} x^* - F(x)$ in X^*. Hence

$$\begin{aligned}
\limsup_{n\to\infty} \langle A(x_n, x_n)\rangle &\leqslant \limsup_{n\to\infty} \langle (A+F)(x_n), x_n\rangle - \liminf_{n\to\infty} \langle F(x_n), x_n\rangle \\
&\leqslant \left\langle x^* - F(x), x\right\rangle \\
\Rightarrow \qquad & A(x) = x^* - F(x) \quad \text{(since } A \text{ is an } (M)\text{-map)} \\
\Rightarrow \qquad & A(x) + F(x) = x^* \quad \text{and so } A + F \text{ is an } (M)\text{-map.}
\end{aligned}$$

The proof is now complete. □

Remark 2.10.20 It is easy to check that if $F : X \to X^*$ is monotone and sequentially continuous from X_w into X^*_w, then $x \mapsto \langle F(x), x\rangle$ is sequentially weakly lower semicontinuous.

The following notion will be useful in the degree theory of $(S)_+$-maps which we develop in Chap. 3.

Definition 2.10.21 Let X be a reflexive Banach space and $\{A_t\}_{t\in[0,1]}$ a family of maps $A_t : X \to X^*$. We say that $\{A_t\}_{t\in[0,1]}$ is an $(S)^t_+$-family if the following is true: "for every $\{x_n\}_{n\geqslant 1} \subseteq D$ and $\{t_n\}_{n\geqslant 1} \subseteq [0, 1]$ satisfying $x_n \xrightarrow{w} x$ in X, $t_n \to t$ and $\limsup_{n\to\infty} \left\langle A_{t_n}(x_n), x_n - x\right\rangle \leqslant 0$, we have $x_n \to x$ in X."

We show that affine homotopies of demicontinuous $(S)_+$-maps satisfy this property.

Proposition 2.10.22 *If X is a reflexive Banach space and $A_0, A_1 : X \to X^*$ are demicontinuous $(S)_+$-maps, then $A_t(x) = tA_0(x) + (1-t)A_1(x)$ for all $x \in X$ and all $t \in [0, 1]$ is an $(S)^t_+$-family.*

Proof Suppose that $\{x_n\}_{n\geqslant 1} \subseteq X$ and $\{t_n\}_{n\geqslant 1} \subseteq [0, 1]$ such that

$$x_n \xrightarrow{w} x \text{ in } X,\ t_n \to t \text{ and } \limsup_{n\to\infty} \left\langle A_{t_n}(x_n), x_n - x\right\rangle \leqslant 0. \tag{2.101}$$

Recall that A_0 and A_1 are pseudomonotone (see Proposition 2.10.13 (a)). So, from the inequality in (2.101) and reasoning as in the proof of Proposition 2.10.6 (see (2.90) and (2.91)) we obtain

$$\begin{aligned}
&\limsup_{n\to\infty} \langle A_0(x_n), x_n - x\rangle \leqslant 0 \text{ and } \limsup_{n\to\infty} \langle A_1(x_n), x_n - x\rangle \leqslant 0 \\
\Rightarrow\ & x_n \to x \text{ in } X.
\end{aligned}$$

The proof is now complete. □

2.11 Remarks

2.1: The systematic study of compact maps was initiated by Schauder [372–374]. However, for linear operators, the notion can be traced back to the works of Hilbert [IV] [208] and Riesz [354]. As we already mentioned in Remark 2.1.2, the terminology is not uniform in the literature. See, for example, Granas and Dugundji [197,

p. 112], where the use of the terms compact and completely continuous is different from ours. Our terminology here is consistent with that of linear operators. Speaking of linear operators, we should mention that another term for complete continuity when dealing with linear operators between Banach spaces is "Dunford–Pettis operators" (see Megginson [295, p. 336]). Theorem 2.1.7 is the reason for the many powerful properties of compact maps and it is due to Schauder [374]. It can be extended to locally convex spaces (see Leray [264] and Nagumo [314]) and to vector spaces that need not be locally convex (see Klee [237]). Proposition 2.1.9 and its trivial consequence Proposition 2.1.10 are due to Dugundji [149]. It is a remarkable generalization of the well-known Tietze extension theorem for $\mathbb{R}$-valued continuous functions. For more results of this type, we refer to the books of Bessaga and Pelczynski [49] and Hu [216]. We mention two such extension results.

Proposition 2.11.1 *If X is a paracompact (resp. normal) topological space, $A \subseteq X$ is closed, Y is a Banach space (resp., a separable Banach space) and $f : A \to Y$ is continuous, then there exists a continuous map $\hat{f} : X \to Y$ such that $\hat{f}|_A = f$.*

2.2: Proper mappings are discussed in detail in Bourbaki [59, Section I.10]. Note that in Bourbaki [59] the definition of properness is more general (see Proposition 6, p. 104). Proper maps are also considered by Berger [44]. Additional results on gradient maps can be found in the survey paper of Rothe [364] and in the book of Krasnoselskii [250].

2.3: As we already mentioned, the first to consider linear compact operators were Hilbert [208] and Riesz [354]. Theorem 2.3.6 is due to Schauder [374]. Another result relating properties of the adjoint of an operator A with the compactness of A is the following (see Dunford and Schwartz [151] (Theorem 6, p. 486)).

Proposition 2.11.2 *If X, Y are Banach spaces, $A \in \mathscr{L}(X, Y)$ and $A^* : Y^* \to X^*$ is weak*-to-norm continuous, then $A \in \mathscr{L}_c(X, Y)$.*

The notion of Schauder basis was introduced (of course) by Schauder [372]. In his classical monograph Banach [31, p. 111] asked whether every infinite-dimensional separable Banach space has a Schauder basis. This is known in Banach space theory as the "basis problem". This and the approximation problem initially asked by T. Hildebrandt (1931) (see Remark 2.3.8) were answered in the negative by Enflo [162]. We also have the following result.

Proposition 2.11.3 *Every Banach space with a Schauder basis has the approximation property.*

Remark 2.11.4 The converse is not in general true (see Szarek [396]).

2.4: The term "spectrum" is due to Hilbert, who made major contributions to spectral theory in particular and to functional analysis in general, in a series of six papers that appeared in the journal of the Academy of Sciences of Göttingen between 1904 and 1910. They appeared in book form in 1912. The book was reprinted in 1953, see Hilbert [208]. Note that Hilbert defined the spectrum of A as the set of λ for which $I - \lambda A$ is not invertible. So, Hilbert's spectrum consists of the reciprocals of the elements of $\sigma(A)$ introduced in Definition 2.4.1. A more complete version of the Fredholm alternative theorem (see Theorem 2.4.15) is given below:

Theorem 2.11.5 *If X is a Banach space and $A \in \mathscr{L}_c(X)$, then*

(a) $N(I - A) = \ker(I - A)$ *is finite-dimensional;*
(b) $R(I - A)$ *is closed and* $R(I - A) = N(I - A^*)^{\perp}$*;*
(c) $N(I - A) = \{0\}$ *if and only if* $R(I - A) = X$*;*
(d) $\dim N(I - A) = \dim N(I - A^*)$.

Remark 2.11.6 Note that if X, Y are Banach spaces and $A \in \mathscr{L}(X, Y)$, then

(a) if either X or Y is finite-dimensional, we have

$$A \text{ is surjective if and only if } A^* \text{ is injective;}$$

$$A \text{ is injective if and only if } A^* \text{ is surjective;}$$

(b) if both X and Y are infinite-dimensional, we have

$$A \text{ is surjective } \Rightarrow A^* \text{ is injective;}$$

$$A^* \text{ is surjective } \Rightarrow A \text{ is injective.}$$

The description of the spectrum of a linear compact operator (see Theorem 2.4.20 and Corollary 2.4.21) is due to Riesz [354].

More on linear compact operators and their spectral properties can be found in the books of Akhiezer and Glazman [7], Brezis [65], Kato [229], Reed and Simon [352] and Yosida [418].

2.5: Multifunctions are an important tool in many applications, such as the calculus of variations, optimal control, optimization and mathematical economics. The notions introduced in Definition 2.5.1 are topological in the sense that we can introduce topologies on the hyperspace 2^Y which correspond to the notions of upper and lower semicontinuity and to Vietoris continuity. This was done by Michael [1951]. Theorem 2.5.14 is due to Michael [298]. The approximate continuous selection theorem for usc multifunctions stated in Theorem 2.5.19 is due to Cellina [112]. This result is the starting point for the extension of degree theory to multifunctions (see Chap. 3). Dugundji's extension theorem (see Proposition 2.1.9) has an analog for usc multifunctions. This result is due to Ma [285].

Proposition 2.11.7 *If X is a metric space, Y is a locally convex space, $A \subseteq X$ is nonempty closed and $F : A \to P_{kc}(Y)$ is usc, then there exists a usc multifunction $\hat{F} : X \to P_{kc}(Y)$ such that $\hat{F}|_A = F$ and*

$$\hat{F}(x) \subseteq \operatorname{conv} F(A) \text{ for all } x \in X.$$

For multifunctions with values in the Lebesgue–Bochner space $L^1(\Omega, X)$, we have a continuous selection theorem in which the convexity of the values is replaced by the notion of decomposability.

Definition 2.11.8 Let $(\Omega.\Sigma, \mu)$ be a σ-finite measure space and X a separable Banach space. A set $C \subseteq L^1(\Omega, X)$ is said to be decomposable if for every triple $(A, u_1, u_2) \in \Sigma \times \ C \times C$, we have $\chi_A u_1 + \chi_{A^c} u_2 \in C$.

Remark 2.11.9 Since $\chi_{A^c} = 1 - \chi_A$, we see that the concept of decomposability formally looks like that of convexity. Only now the coefficients in the "convex combination" are characteristic functions. Decomposable sets are up to closure the set of $L^1(\Omega, X)$-selectors of a measurable multifunction with closed values.

The "decomposable" version of Theorem 2.5.14 is due to Bressan and Colombo [60] and Fryszkowski [174].

Theorem 2.11.10 *If E is a separable metric space and $F : E \to P_f(L^1(\Omega, X))$ is an lsc multifunction with decomposable values, then there exists a continuous map $f : E \to L^1(\Omega, X)$ such that*

$$f(e) \in F(e) \text{ for all } e \in E.$$

For a detailed study of decomposable sets and their relation to multifunctions we refer to Fryszkowski [175].

Theorem 2.5.33 is due to Kuratowski and Ryll-Nardzewski [254]. Theorem 2.5.35, modelled after the corresponding result for continuous selections of an lsc multifunction (see Proposition 2.5.26, due to Michael [298]), was proved by Castaing [106]. Theorem 2.5.38 as stated can be found in Saint-Beuve [369]. But earlier more restricted versions were obtained by Yankov [417], von Neumann [317] and Aumann [24]. The result is also related to the so-called "Yankov–von Neumann–Aumann projection theorem".

Theorem 2.11.11 *If (Ω, Σ) is a measurable space, X is a Souslin space and $E \in \Sigma \times B(X)$ (recall that $B(X)$ is the Borel σ-field of X), then* $\mathrm{proj}_\Omega E \in \hat{\Sigma}$.

More about multifunctions (both in the topological and measure theoretic directions) can be found in the books of Aliprantis and Border [9], Aubin and Frankowska [22], Castaing and Valadier [108], Denkowski, Migorski and Papageorgiou [143], Hu and Papageorgiou [218], Kisielewicz [236], Klein and Thompson [238] and Papageorgiou and Kyritsi [329].

2.6: Monotone maps are rooted in the calculus of variations and were introduced in the early sixties in order to provide a framework broader than that of compact operators, for the study of nonlinear functional equations. The first use of the monotonicity condition in Hilbert spaces can be traced back to Golomb [193, pp. 66–72], in the study of nonlinear Hammerstein integral equations. Kachurovski [224] extended the concept of monotonicity to maps from a Banach space into its dual (see Definition 2.6.1 (a)). The systematic study of monotone type maps started in the mid-sixties with the works of Browder [78, 79, 81, 84] and Minty [300, 301]. The local boundedness of monotone maps was first proved by Kato [229]. A detailed study of maximal monotone maps and their generalizations can be found in the books of Barbu [32], Bauschke and Combettes [39], Brezis [64], Denkowski, Migorski and Papageorgiou [143], Gasinski and Papageorgiou [182], Hu and Papageorgiou [218], Papageorgiou and Kyritsi [329], and Pascali and Sburlan [331].

2.7: Although affine continuous minorants of convex functions were considered earlier, the systematic study of the subdifferential multifunction started with the works of Moreau [303, 304] and Rockafellar [356, 357]. Also the duality theory of convex functions started with the work of Fenchel [167], who extended the Legendre transform to nondifferentiable functions on $\mathbb{R}^N$. Extensions to general locally convex spaces were produced by Brondsted [71], Moreau [304] and Rockafellar [356]. Detailed expositions of this duality theory can be found in the books of Barbu and Precupanu [33], Ekeland and Temam [161], Gasinski and Papageorgiou [182], Ioffe and Tichomirov [221], Laurent [260] and Rockafellar [358, 360]. The differentiability properties of convex functions and their interplay with the geometry of Banach spaces can be found in the books of Giles [188] and Phelps [341].

The notion of a duality map (see Definition 2.7.21) was first introduced by Beurling and Livingston [50] and was studied in detail by Asplund [18], Browder [82], [87] and Kachurovski [225]. Also, extensive discussions of the duality map can be found in Cioranescu [125], Gasinski and Papageorgiou [182] and Zeidler [427]. For a detailed discussion of the various geometric properties of the unit ball in a Banach space, we refer to the books of Day [138] and Megginson [295].

Theorem 2.7.36 is a consequence of a more general result of Troyanski [405] and of the Asplund averaging process (see Asplund [19]). See also the books of Cioranescu [125, pp. 98, 108] and Diestel [147, pp. 111, 164]. In fact in Diestel [147, p. 164] one can find the more general version of the Troyanski theorem mentioned above.

2.8: Proposition 2.8.3 is due to Debrunner and Floer [140] and for this reason in some books it appears under the name "Debrunner–Floer Lemma". Theorem 2.8.5 and 2.8.9 were first proved by Minty [300] for pivot Hilbert spaces H (that is, $H = H^*$). Theorem 2.8.6 is due to Browder [84]. The maximal monotonicity of the subdifferential map (see Theorem 2.8.10), was proved by Rockafellar [359]. Also, Rockafellar [355, 359] obtained the characterization of the subdifferential as a maximal cyclically monotone map (see Theorem 2.8.14). More on the sum of maximal monotone maps can be found in the work of Brezis and Nirenberg [68].

2.9: The resolvent and the Yosida approximation are basic to the study of accretive and m-accretive maps (see Brezis, Crandall and Pazy [67]). Here we restrict ourselves to pivot Hilbert spaces, in which case m-accretivity coincides with maximal monotonicity. Our presentation is based on the works of Brezis [63], [64]. Similarly, for linear maximal monotone operators, we refer to Brezis [64].

2.10: The notion of pseudomonotonicity was introduced by Brezis [62, pp. 123–124], but using nets instead of sequences. Here, in defining pseudomonotone and generalized pseudomonotone maps (see Definition 2.10.1), we follow Browder and Hess [91], who were the first to conduct a systematic study of these classes of maps. Proposition 2.10.8 is attributed to Browder (see, for example, Dal Maso [134,

p. 95]) and indeed it can be found in the book of Browder [86, p. 81]. However, it is a particular case of a general result from functional analysis concerning angelic spaces (see Floret [171, p. 30]). A Banach space with the weak topology is angelic (see also Edwards [154, p. 549]). Theorem 2.10.10 is due to Browder and Hess [91]. Calvert and Webb [100] call (P)-maps quasimonotone. Here we follow the terminology of Hess [206] and Zeidler [426, p. 586]. In fact, property (P) is equivalent to saying that $x_n \xrightarrow{w} x$ in $X \Rightarrow \liminf_{n\to\infty} \langle A(x_n), x_n - x\rangle \geqslant 0$. We mention that Skrypnik [382, p. 35] calls $(S)_+$-maps maps of class $\alpha(D)$. The $(S)_+$-property was introduced by Browder [85]. More details on these operators of monotone type can be found in Pascali and Sburlan [331].

Chapter 3
Degree Theories

...the primary question was not What we know, but How do we know it.

Aristotle (384–322 BCE)

Degree theory deals with an abstract equation of the form $\varphi(u) = \xi$. In many situations, we are interested not only in the existence of solutions for the equation, but also in their multiplicity. This is, for example, the case in bifurcation theory, where we want to determine the values of the parameter λ, where the number of solutions of an equation of the form $\varphi(u, \lambda) = \xi$ changes.

Given the equation $\varphi(u) = \xi$, with $\varphi : \Omega \subseteq X \to X$ defined on a set Ω in an appropriate space X and $\xi \in X$, the idea of topological degree is to assign to each triple (φ, Ω, ξ) an integer $d(\varphi, \Omega, \xi)$ which will be a measure of the number of solutions of the equation. In particular, the equation will have solutions in Ω whenever $d(\varphi, \Omega, \xi) \neq 0$. Also, we want the integer $d(\varphi, \Omega, \xi)$ to be stable with respect to small perturbations of φ and ξ. The development of degree theory (which starts in 1912, with the seminal work of Brouwer [75], see also the Remarks) revealed that it is useful to deal with certain fundamental properties that a "topological degree" should have and from which the other useful properties that lead to interesting applications can be deduced. This leads to an axiomatic scheme which firmly defines the notion of "topological degree" and also justifies any new extensions.

We will start with the finite-dimensional theory, which is known as *Brouwer's degree theory*.

N. S. Papageorgiou et al., *Nonlinear Analysis–Theory and Methods*,
Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-030-03430-6_3

3.1 Brouwer Degree

To develop the finite-dimensional degree theory (originally due to Brouwer), we will follow the analytic approach developed by Heinz [205].

In what follows, $\Omega \subseteq \mathbb{R}^N$ is a bounded open set, $\partial\Omega$ denotes the boundary of Ω, $\varphi : \Omega \to \mathbb{R}^N$ and $\xi \in \mathbb{R}^N$. If $\varphi \in C^1(\Omega, \mathbb{R}^N)$, then by $J_\varphi(x)$ we denote the Jacobian of φ at $x \in \Omega$, that is,

$$J_\varphi(x) = \det \varphi'(x) = \det \left[\frac{\partial \varphi^k}{\partial x_i}(x)\right]_{i,k=1}^N .$$

Definition 3.1.1 Suppose $\varphi \in C^1(\Omega, \mathbb{R}^N) \cap C(\overline{\Omega}, \mathbb{R}^N)$ and $\xi \notin \varphi(\partial\Omega)$. Let $0 < \epsilon < d(\xi, \varphi(\partial\Omega))$ and $\vartheta \in (0, +\infty)$ a function with compact support contained in $(0, \epsilon)$ such that

$$\int_{\mathbb{R}^N} \vartheta(||x||)dx = 1.$$

The Brouwer degree for the triple (φ, Ω, ξ) is defined by

$$d(\varphi, \Omega, \xi) = \int_\Omega \vartheta(||\varphi(x) - \xi||) J_\varphi(x)dx.$$

In order for this notion to be well-defined, we need to show that it is independent of $\epsilon > 0$ and $\vartheta \in C(0, \infty)$.

To show that Definition 3.1.1 makes sense, we will need a series of technical lemmata.

Lemma 3.1.2 *If $f : \Omega \to \mathbb{R}^{N-1}$ is a C^2-function and we set*

$$D_k = \det\left[\partial_1 f, \dots, \partial_{k-1} f, \partial_{k+1} f, \dots, \partial_N f\right]$$

where $\partial_i f = \frac{\partial f}{\partial x_i}$, then $\sum\limits_{\mathrm{k}=1}^N (-1)^k \partial_k D_k = 0$.

Proof For $1 \leqslant k \leqslant N$ we set $E_{kk} = 0$, for $i < k$ we set

$$E_{ki} = \det\left[\partial_1 f, \dots, \partial_{i-1} f, \partial_{ik} f, \partial_{i+1} f, \dots, \partial_{k-1} f, \partial_{k+1} f, \dots, \partial_N f\right],$$

where $\partial_{ik} f = \frac{\partial f}{\partial x_i \partial x_k}$ and for $k < i$ we set

$$E_{ki} = \det\left[\partial_1 f, \dots, \partial_{k-1} f, \partial_{k+1} f, \dots, \partial_{i-1} f, \partial_{ik} f, \partial_{i+1} f, \dots, \partial_N f\right]$$

Then we have $\partial_k D_k = \sum\limits_{\mathrm{i}=1}^N E_{ki}$. Let $\beta = \sum\limits_{\mathrm{k}=1}^N (-1)^k \partial_k D_k$. Then

$$\beta = \sum_{\mathrm{k,i=1}}^{N} (-1)^k E_{ki}.$$

A simple calculation involving the properties of determinants, shows that $E_{ki} = (-1)^{k+i-1} E_{ik}$. So, it follows that

$$\beta = \sum_{\mathrm{k,i=1}}^{N} (-1)^k E_{ki} = \sum_{\mathrm{k,i=1}}^{N} (-1)^{i-1} E_{ik} = -\beta$$
$$\Rightarrow \ \beta = 0.$$

The proof is now complete. $\square$

Lemma 3.1.3 *If $h : \Omega \to \mathbb{R}^N$ is a C^2-function and $C_{ki}(x)$ is the cofactor of $\partial_k h_i(x)$ in $J_h(x)$, then for every fixed $i \in \{1, \dots, N\}$ we have*

$$\sum_{\mathrm{k=1}}^{N} \partial_k C_{ki} = 0.$$

Proof Recall that

$$C_{ki}(x) = (-1)^{k+1} \det[\partial_m h_j]_{m \neq k, j \neq i}.$$

Fix $i \in \{1, \dots, N\}$ and let $f : \Omega \to \mathbb{R}^{N-1}$ be defined by $f = (h_1, \dots, h_{i-1}, h_{i+1}, \dots, h_N)$. Evidently f is C^2 and we can apply Lemma 3.1.2 to conclude that

$$\sum_{\mathrm{k=1}}^{N} \partial_k C_{ki} = 0.$$

The proof is now complete. $\square$

Lemma 3.1.4 *If $\varphi \in C^1(\Omega, \mathbb{R}^N) \cap C(\overline{\Omega}, \mathbb{R}^N)$, $0 \notin \varphi(\partial\Omega)$, $\vartheta : \mathbb{R}_+ \to \mathbb{R}$ is continuous,* supp $\vartheta \subseteq [0, \epsilon]$ *where $0 < \epsilon < d(0, \varphi(\partial\Omega))$ and $\int_0^\infty r^{N-1} \vartheta(r) dr = 0$, then $\int_\Omega \vartheta(||\varphi(z)||) J_\varphi(z) dz = 0$.*

Proof Using regularization, we may assume that $\varphi \in C^2(\Omega, \mathbb{R}^N)$. We introduce the function $\eta : \mathbb{R}_+ \to \mathbb{R}$ defined by

$$\eta(0) = 0 \text{ and } \eta(r) = \frac{1}{r^N} \int_0^r t^{N-1} \vartheta(t) dt.$$

Evidently, $\eta \in C^1(0, \infty)$ and has compact support. Also

$$r\eta' + N\eta = \vartheta. \tag{3.1}$$

Let $\sigma(x) = \eta(||x||)x$ for all $x \in \mathbb{R}^N$. Then

$$\operatorname{div}\sigma(x) = ||x||\eta'(||x||) + N\eta(||x||) = \vartheta(||x||) \text{ (see (3.1))}. \tag{3.2}$$

For $z \in \Omega$, using Lemma 3.1.3, we have

$$\begin{aligned}
&\sum_{k=1}^{N}\partial_k\sum_{i=1}^{N}C_{ki}(z)\sigma_i(\varphi(z)) = \\
&\sum_{k=1}^{N}\sum_{i=1}^{N}(\partial_k C_{ki}(z)\sigma_i(\varphi(z))) + \sum_{k=1}^{N}\sum_{i=1}^{N}C_{ki}(z)\partial_k(\sigma_k(\varphi(z))) = \\
&\sum_{k=1}^{N}\sum_{i=1}^{N}C_{ki}(z)\sum_{m=1}^{N}\partial_k\varphi_k(z)\frac{\partial\sigma_i}{\partial x_k}(\varphi(z)) = \\
&\sum_{i=1}^{N}\sum_{m=1}^{N}\left(\sum_{k=1}^{N}C_{ki}(z)\partial_k\varphi_m(z)\right)\frac{\partial\sigma_i}{\partial x_k}(\varphi(z)).
\end{aligned} \tag{3.3}$$

By Cramer's rule, we have

$$\sum_{k=1}^{N}C_{ki}(z)\partial_k\varphi_m(z) = \delta_{mi}J_\varphi(z) \ (\delta_{mi} \text{ being the Kronecker symbol}). \tag{3.4}$$

Using (3.4) in (3.3), we obtain

$$\sum_{k=1}^{N}\partial_k\sum_{i=1}^{N}C_{ki}(z)\sigma_i(\varphi(z)) = J_\varphi(z)\operatorname{div}\sigma(\varphi(z)) \text{ for all } z \in \Omega. \tag{3.5}$$

Therefore we have

$$\begin{aligned}
\vartheta(||\varphi(z)||)J_\varphi(z) &= J_\varphi(z)\left[r\eta'(r) + N\eta\right]\Big|_{r=\varphi(z)} \text{ (see (3.1))} \\
&= J_\varphi(z)\operatorname{div}\sigma(\varphi(z)) \\
&= \sum_{k=1}^{N}\partial_k\sum_{i=1}^{N}C_{ki}(z)\sigma_i(\varphi(z)) \text{ (see (3.5))} \\
\Rightarrow \int_\Omega \vartheta(||\varphi(z)||)J_\varphi(z)dz &= \sum_{k=1}^{N}\int_\Omega \partial_k\left(\sum_{i=1}^{N}C_{ki}(z)\sigma_i(\varphi(z))\right)dz.
\end{aligned}$$

Note that $\sigma_i(\varphi(z)) = 0$ for all z in a neighborhood for $\partial\Omega$. So, by integration by parts we obtain

$$\int_\Omega \vartheta(||\varphi(z)||)J_\varphi(z)dz = 0,$$

which completes the proof ∎

Now we can show that the definition of degree given in Definition 3.1.1 is actually independent of $\epsilon \in (0, d(\xi, \varphi(\partial\Omega))$ and ϑ, hence the notion is well-defined.

Proposition 3.1.5 *If the triple* (φ, Ω, ξ) *and* $\epsilon > 0,\ \vartheta \in C(0, \infty)$ *are as in Definition 3.1.1, then* $d(\varphi, \Omega, \xi)$ *is independent of* $\epsilon > 0$ *(provided* $\epsilon < d(\xi, \varphi(\partial\Omega))$*) and of the function* $\vartheta \in C(0, +\infty)$.

Proof Let $\widehat{\epsilon} = d(\xi, \varphi(\partial\Omega))$ and let $\epsilon_1, \epsilon_2 \in (0, \widehat{\epsilon})$. Suppose that ϑ_1, ϑ_2 are two continuous functions with supports in $(0, \epsilon_1)$ and $(0, \epsilon_2)$ respectively and such that

$$\int_{\mathbb{R}^N} \vartheta_1(||x||)dx = \int_{\mathbb{R}^N} \vartheta_2(||x||)dx = 1.$$

We set $\vartheta = \vartheta_1 - \vartheta_2$. Then $\int_0^\infty r^{N-1}\vartheta(r)dr = 0$ and so we can apply Lemma 3.1.4 with $z \to \varphi(z) - b$, to obtain

$$\begin{aligned} & \int_\Omega \vartheta(||\varphi(z) - \xi||)J_\varphi(z)dz = 0 \\ \Rightarrow\ & \int_\Omega \vartheta_1(||\varphi(z) - \xi||)J_\varphi(z)dz = \int_\Omega \vartheta_2(||\varphi(z) - \xi||)J_\varphi(z)dz. \end{aligned}$$

The proof is now complete. $\square$

We can also show the stability of $\varphi \to d(\varphi, \Omega, \xi)$.

Proposition 3.1.6 *If* $\varphi_1, \varphi_2 \in C^1(\Omega, \mathbb{R}^N) \cap C(\overline{\Omega}, \mathbb{R}^N),\ \xi \notin \varphi_1(\partial\Omega) \cup \varphi_2(\partial\Omega)$ *and* $\epsilon > 0$ *satisfies*

$$\epsilon < \frac{1}{4}d(\xi, \varphi_1(\partial\Omega) \cup \varphi_2(\partial\Omega)),\ \ ||\varphi_1 - \varphi_2||_\infty < \epsilon,$$

then $d(\varphi_1, \Omega, \xi) = d(\varphi_2, \Omega, \xi)$.

Proof It is clear from Definition 3.1.1 that $d(\varphi, \Omega, \xi) = d(\varphi - \xi, \Omega, 0)$ and so without any loss of generality, we may assume that $\xi = 0$. Let $\gamma : \mathbb{R}_+ \to [0, 1]$ be a C^∞-function such that

$$\gamma(r) = \begin{cases} 1 \text{ if } r \in [0, \epsilon] \\ 0 \text{ if } r \geqslant 2\epsilon. \end{cases}$$

Let $\varphi_3(z) = (1 - \gamma(||\varphi_1(z)||)\varphi_1(z) + \gamma(||\varphi_1(z)||)\varphi_2(z)$ for all $z \in \Omega$. Then $||\varphi_k - \varphi_i||_\infty < \epsilon$ for $k, i \in \{1, 2, 3\}$, $||\varphi_k(z)|| > 3\epsilon$ for all $z \in \partial\Omega$, all $k \in \{1, 2, 3\}$ and

$$\varphi_3(z) = \begin{cases} \varphi_1(z) \text{ if } ||\varphi_1(z)|| > 2\epsilon \\ \varphi_2(z) \text{ if } ||\varphi_1(z)|| < \epsilon. \end{cases}$$

According to Definition 3.1.1 and Proposition 3.1.5, we can find $\vartheta_1, \vartheta_2 \in C(0, +\infty)$ such that

$$\operatorname{supp} \vartheta_1 \subseteq (2\epsilon, 3\epsilon), \int_{\mathbb{R}^N} \vartheta_1(||x||)dx = 1, \; d(\varphi_1.\Omega, 0) = \int_\Omega \vartheta_1(||\varphi_1(z)||)J_{\varphi_1}(z)dz,$$

$$\operatorname{supp} \vartheta_2 \subseteq (0, \epsilon), \int_{\mathbb{R}^N} \vartheta_2(||x||)dx = 1, \; d(\varphi_2, \Omega, 0) = \int_\Omega \vartheta_2(||\varphi_2(z)||)J_{\varphi_2}(z)dz.$$

From these choices and the definition of φ_3, we have

$$\vartheta_1(||\varphi_3(z)||)J_{\varphi_3}(z) = \vartheta_1(||\varphi_1(z)||)J_{\varphi_1}(z),$$
$$\vartheta_2(||\varphi_3(z)||)J_{\varphi_3}(z) = \vartheta_2(||\varphi_2(z)||)J_{\varphi_2}(z) \text{ for all } z \in \Omega.$$

So, we conclude that

$$d(\varphi_3, \Omega, 0) = d(\varphi_1, \Omega, 0) = d(\varphi_2, \Omega, 0).$$

The proof is now complete. □

This stability property of the degree $d(\varphi, \Omega, \xi)$ with respect to the function $\varphi \in C^1(\Omega, \mathbb{R}^N) \cap C(\overline{\Omega}, \mathbb{R}^N)$ permits the extension of the notion of topological degree to functions $\varphi \in C(\overline{\Omega}, \mathbb{R}^N)$. Indeed, recall that according to Proposition 1.1.3(a), given $\varphi \in C(\overline{\Omega}, \mathbb{R}^N)$, we can find $\varphi_\epsilon \in C^1(\mathbb{R}^N, \mathbb{R}^N)$ such that $||\varphi_\epsilon - \varphi||_\infty < \epsilon$. So, we can find $\{\varphi_n\}_{n\geqslant 1} \subseteq C^1(\Omega, \mathbb{R}^N) \cap C(\overline{\Omega}, \mathbb{R}^N)$ such that $||\varphi_n - \varphi||_\infty \to 0$ as $n \to \infty$.

Proposition 3.1.7 *If $\varphi, \varphi_n : \overline{\Omega} \to \mathbb{R}^N$, $n \geqslant 1$, are as above and $\xi \notin \varphi(\partial\Omega)$, then we can find $n_0 \geqslant 1$ such that for all $n, m \geqslant n_0$, $d(\varphi_n, \Omega, \xi) = d(\varphi_m, \Omega, \xi)$.*

Proof Since $\xi \notin \varphi(\partial\Omega)$ and $\varphi_n \to \varphi$ in $C(\overline{\Omega}, \mathbb{R}^N)$, we can find $\widehat{n} \geqslant 1$ such that $\xi \notin \varphi_n(\partial\Omega)$ for all $n \geqslant \hat{n}$. So $d(\varphi_n, \Omega, \xi)$ is well-defined. We can find $n_0 \geqslant \widehat{n} \geqslant 1$ such that

$$||\varphi_n - \varphi_m||_\infty < \frac{1}{4} d(\xi, \varphi_n(\partial\Omega) \cup \varphi_m(\partial\Omega)) \text{ for all } n, m \geqslant n_0.$$

Invoking Proposition 3.1.6, we have

$$d(\varphi_n, \Omega, \xi) = d(\varphi_m, \Omega, \xi) \text{ for all } n, m \geqslant n_0,$$

which completes the proof. □

Remark 3.1.8 It is routine to see that this stabilized value is in fact independent of the approximating family $\{\varphi_n\}_{n\geqslant 1}$. This leads to the following extension of the notion of the Brouwer topological degree to continuous functions.

Definition 3.1.9 Let $\Omega \subseteq \mathbb{R}^N$ be a bounded open set, $\varphi \in C(\overline{\Omega}, \mathbb{R}^N)$ and $\xi \notin \varphi(\partial\Omega)$. The Brouwer topological degree of φ on Ω at ξ, is defined by

$$d(\varphi, \Omega, \xi) = \lim_{n\to\infty} d(\varphi_n, \Omega, \xi),$$

where $\{\varphi_n\}_{n\geqslant 1} \subseteq C^1(\Omega, \mathbb{R}^N) \cap C(\overline{\Omega}, \mathbb{R}^N)$ satisfies $||\varphi_n - \varphi||_\infty \to 0$ as $n \to \infty$.

Now, we will derive the main properties of this topological degree. We start with an obvious one which results from Proposition 3.1.6 and Definition 3.1.9.

Proposition 3.1.10 *If* $\varphi_1, \varphi_2 \in C(\overline{\Omega}, \mathbb{R}^N)$, $\xi \notin \varphi_1(\partial\Omega) \cup \varphi_2(\partial\Omega)$ *and*

$$||\varphi_1 - \varphi_2||_\infty < \frac{1}{4} d(\xi, \varphi_1(\partial\Omega) \cup \varphi_2(\partial\Omega)),$$

then $d(\varphi_1, \Omega, \xi) = d(\varphi_2, \Omega, \xi)$.

A similar stability holds with respect to the reference point $\xi \in \mathbb{R}^N$.

Proposition 3.1.11 *If* $\varphi \in C(\overline{\Omega}, \mathbb{R}^N)$*, then* $d(\varphi, \Omega, \cdot)$ *is constant on the connected components of* $\mathbb{R}^N \backslash \varphi(\partial\Omega)$*, that is, if* ξ_1, ξ_2 *belong to the same connected component of* $\mathbb{R}^N \backslash \varphi(\partial\Omega)$*, then* $d(\varphi, \Omega, \xi_1) = d(\varphi, \Omega, \xi_2)$.

Proof Recall that $d(\varphi, \Omega, \xi) = d(\varphi - \xi, \Omega, 0)$. So, if $\epsilon > 0$ satisfies $\epsilon < \frac{1}{4} d(\xi, \varphi(\partial\Omega))$, then by Proposition 3.1.10, we see that then when $||\xi_1 - \xi_2|| < \epsilon$ we have

$$d(\varphi, \Omega, \xi_1) = d(\varphi - \xi_1, \Omega, 0) = d(\varphi - \xi_2, \Omega, 0) = d(\varphi, \Omega, \xi_2).$$

Clearly then for all ξ_1, ξ_2 in the same connected component of $\mathbb{R}^N \backslash \varphi(\partial\Omega)$, we have $d(\varphi, \Omega, \xi_1) = d(\varphi, \Omega, \xi_2)$. □

Proposition 3.1.12 *If* $\varphi \in C(\overline{\Omega}, \mathbb{R}^N)$, $\xi \notin \varphi(\partial\Omega)$ *and* $y \in \mathbb{R}^N$*, then* $d(\varphi - y, \Omega, \xi - y) = d(\varphi, \Omega, \xi)$.

To formulate the next property of the topological degree, we need to recall the following notion from topology.

Definition 3.1.13 Let X, Y be Hausdorff topological spaces and assume that $\varphi, \psi \in C(X, Y)$. We say that φ and ψ are homotopic if there exists a continuous function $h : [0, 1] \times X \to Y$ such that $h(0, x) = \varphi(x)$ and $h(1, x) = \psi(x)$. The function $h(\cdot, \cdot)$ is a homotopy between φ and ψ.

The next property is arguably the most important property of the topological degree. It says that the Brouwer degree is homotopy invariant. It allows the computation of the topological degree of a function by performing the computation of the degree of another function which is simple and easier.

Proposition 3.1.14 *If* $h : [0, 1] \times \overline{\Omega} \to \mathbb{R}^N$ *is continuous and* $\xi \notin h([0, 1] \times \partial\Omega)$*, then* $d(h(t, \cdot), \Omega, \xi) = d(h(0, \cdot), \Omega, \xi)$ *for all* $t \in [0, 1]$.

Proof Let $\epsilon = \frac{1}{4} d(\xi, h([0, 1] \times \partial\Omega))$ and note that h is uniformly continuous on $[0, 1] \times \overline{\Omega}$. So, we can find $\delta = \delta(\epsilon) > 0$ such that

$$|t_1 - t_2| < \delta \Rightarrow |h(t_1, x) - h(t_2, x)| < \epsilon \text{ for all } x \in \mathbb{R}^N.$$

Hence by virtue of Proposition 3.1.10, we have

$$d(h(t_1,\cdot),\Omega,\xi) = d(h(t_2,\cdot),\Omega,\xi) \text{ for all } t_1,t_2 \in [0,1] \text{ such that } |t_1 - t_2| < \delta.$$

Since $[0, 1]$ is compact, we can cover it by a finite number of integrals of length $\delta > 0$. So, the homotopy invariance property of the degree follows. □

As a consequence of the homotopy invariance property of the degree, we have the next result which says that in degree theory what matters is the boundary behavior of the functions.

Proposition 3.1.15 *If* $\varphi, \psi \in C(\overline{\Omega}, \mathbb{R}^N)$, $\varphi|_{\partial\Omega} = \psi|_{\partial\Omega}$ *and* $\xi \notin \varphi(\partial\Omega)$, *then* $d(\varphi, \Omega, \xi) = d(\psi, \Omega, \xi)$.

Proof Consider the homotopy $h : [0, 1] \times \mathbb{R}^N \to \mathbb{R}^N$ defined by

$$h(t,x) = (1-t)\varphi(x) + t\psi(x) \text{ for all } t \in [0,1], \text{ all } x \in \mathbb{R}^N.$$

Since $\varphi|_{\partial\Omega} = \psi|_{\partial\Omega}$, we have $\xi \notin h([0,1] \times \partial\Omega)$ and so by Proposition 3.1.14, we have

$$\begin{aligned} & d(h(0,\cdot),\Omega,\xi) = d(h(1,\cdot),\Omega,\xi) \\ \Rightarrow\ & d(\varphi,\Omega,\xi) = d(\psi,\Omega,\xi). \end{aligned}$$

The proof is now complete. □

Next, we want to examine the dependence of the degree on the domain Ω. To do this we recall the following theorem due to Sard [370], which says that given $\varphi \in C^1(\Omega, \mathbb{R}^N)$ a generic point $\xi \in \varphi(\Omega)$ should be a regular value.

Theorem 3.1.16 *If* $\Omega \subseteq \mathbb{R}^N$ *is bounded open,* $\varphi \in C^1(\Omega, \mathbb{R}^N)$ *and* $S_\varphi = \{z \in \Omega : J_\varphi(z) = 0\}$, *then* $\varphi(S_\varphi)$ *is Lebesgue-null.*

Proof Consider a closed cube C contained in Ω of side a. We subdivide C into k^N ($k \geqslant 1$ an integer) subcubes each of side $\frac{a}{k}$. Then $x \to \varphi'(x)$ is uniformly continuous on C and so, given $\epsilon > 0$, we can find $\delta = \delta(\epsilon) > 0$ such that

$$\text{“for all } z, y \in C \text{ with } ||z - y|| < \delta, \text{ we have } ||\varphi'(z) - \varphi'(y)|| < \epsilon\text{”}. \tag{3.6}$$

We choose $k \geqslant 1$ such that $\sqrt{N}\frac{a}{k} < \delta$ (note that $\sqrt{N}\frac{a}{k}$ is the diameter of each subcube). From the mean value theorem we have

$$||\varphi(z) - \varphi(y)|| \leqslant \sup_{u \in C} ||\varphi'(u)||\ ||z - y||.$$

Let $M = \sup\limits_{u \in C} ||\varphi'(u)||$. Let $z \in C \cap S_\varphi$. Then we can find a subcube $\widehat{C}$ of C which contains z. We have

$$||\varphi(z) - \varphi(y)|| \leqslant M\sqrt{N}\frac{a}{k} \text{ for all } y \in \widehat{C}$$
$$\Rightarrow \varphi(\widehat{C}) \subseteq \overline{B}_{M\sqrt{N}\frac{a}{k}}(\varphi(z)).$$

Also, for all $y \in \widehat{C}$, we have

$$\begin{aligned} ||\varphi(y) - \varphi(z) - \varphi'(z)(y-z)|| &= \left\| \int_0^1 \left[\varphi'(z + t(y-z)) - \varphi'(z)\right](y-z)dt \right\| \\ &\leqslant \epsilon||y-z|| \text{ (see (3.6)).} \end{aligned}$$

Since $z \in C \cap S_\varphi$, we have $J_\varphi(z) = 0$ and so $\varphi'(z)$ is not invertible. This means that $\varphi'(z)(\mathbb{R}^N)$ is contained in a subspace E of $\mathbb{R}^N$ of dimension $N-1$. Then

$$d(\varphi(y), \varphi(x) + E) \leqslant \epsilon\sqrt{N}\frac{a}{k} \text{ for all } y \in \widehat{C}.$$

So, $\varphi(\widehat{C})$ is in a cuboid of center $\varphi(x)$ and volume $2\epsilon\left(\sqrt{N}\frac{a}{k}\right)^N$. Hence

$$\begin{aligned} &\lambda^*(\varphi(\widehat{C})) \leqslant 2^N M^{N-1} N^{\frac{N}{2}} \frac{a^N}{k^N}\epsilon \quad (\lambda^* \text{ is the Lebesgue outer measure}) \\ \Rightarrow\ &\lambda^*(\varphi(C \cap S_\varphi)) \leqslant \sum_{\widehat{C}\cap S_\varphi \neq \emptyset} \lambda^*(\varphi(\widehat{C} \cap S_\varphi)) \leqslant \sum_{\widehat{C}\cap S_\varphi \neq \emptyset} \lambda^*(\varphi(\widehat{C})) \leqslant \\ &\left(2^N M^{N-1} N^{\frac{N}{2}} a^N\right)\epsilon = \\ &\beta(N, C)\epsilon. \end{aligned} \tag{3.7}$$

Since $\epsilon > 0$ is arbitrary, from (3.7) it follows that $\varphi(C \cap S_\varphi)$ is Lebesgue-null. Since S_φ can be covered by a finite family of sets of the form $C \cap S_\varphi$, we infer that $\varphi(S_\varphi)$ is Lebesgue-null. $\square$

Now we can establish an additive property with respect to the domain $\Omega \subseteq \mathbb{R}^N$.

Proposition 3.1.17 *If $\Omega_1, \Omega_2 \subseteq \mathbb{R}^N$ are bounded open sets, $\Omega_1 \cap \Omega_2 = \emptyset$, $\Omega = \Omega_1 \cup \Omega_2$, $\varphi \in C(\overline{\Omega}, \mathbb{R}^N)$, and $\xi \notin \varphi(\partial\Omega_1) \cup \varphi(\partial\Omega_2) = \varphi(\partial\Omega)$, then $d(\varphi, \Omega, \xi) = d(\varphi, \Omega_1, \xi) + d(\varphi, \Omega_2, \xi)$.*

Proof Let $r = d(\xi, \varphi(\partial\Omega)) > 0$. Then $r \leqslant d(\xi, \varphi(\partial\Omega_k))$ for $k = 1, 2$. Let $\psi \in C^1(\overline{\Omega}, \mathbb{R}^N)$ such that $||\psi - \varphi||_\infty < \frac{r}{2}$. Then

$$d(\psi, \Omega, \xi) = d(\varphi, \Omega, \xi) \text{ and } d(\psi, \Omega_k, \xi) = d(\varphi, \Omega_k, \xi) \text{ for } k = 1, 2. \tag{3.8}$$

The open ball with center ξ and radius $\frac{1}{2}d(\xi, \varphi(\partial\Omega))$ contains an $\eta \notin \psi(S_\psi)$ (see Theorem 3.1.16). Evidently ξ and η are in the same connected component of $\mathbb{R}^N \backslash \psi(\partial\Omega)$ and of $\mathbb{R}^N \backslash \psi(\partial\Omega_k)$ $k = 1, 2$. Then Proposition 3.1.11 implies that

$$d(\psi, \Omega, \xi) = d(\psi, \Omega, \eta) \text{ and } d(\psi, \Omega_k, \xi) = d(\psi, \Omega_k, \eta) \text{ for } k = 1, 2. \quad (3.9)$$

Then since η is a regular value of ψ (recall that $\eta \in \psi(S_\psi)$), from Definition 3.1.1 (see also Lemma 3.1.23 below), we have

$$\begin{aligned} d(\psi, \Omega, \eta) = \sum_{z\in\psi^{-1}(\eta)} \operatorname{sgn} J_\psi(z) = & \sum_{z\in\psi^{-1}(\eta)\cap\Omega_1} \operatorname{sgn} J_\psi(z) \\ & + \sum_{z\in\psi^{-1}(\eta)\cap\Omega_2} \operatorname{sgn} J_\psi(z) \end{aligned}$$

$$\Rightarrow d(\psi, \Omega, \eta) = d(\psi, \Omega_1, \eta) + d(\psi, \Omega_2, \eta)$$
$$\Rightarrow d(\varphi, \Omega, \xi) = d(\varphi, \Omega_1, \xi) + d(\varphi, \Omega_2, \xi) \text{ (see (3.8), (3.9))}.$$

The proof is now complete. □

An easy consequence of the above additivity property is the so-called excision property.

Corollary 3.1.18 *If $\Omega \subseteq \mathbb{R}^N$ is bounded open, $\varphi \in C(\overline{\Omega}, \mathbb{R}^N)$, $K \subseteq \overline{\Omega}$ is compact and $\xi \notin \varphi(K) \cup \varphi(\partial\Omega)$, then $d(\varphi, \Omega, \xi) = d(\varphi, \Omega\backslash K, \xi)$.*

The next property is very useful in establishing the existence of solutions for the equation $\varphi(x) = \xi$. For this reason this property is often called the "existence property".

Proposition 3.1.19 *If $\Omega \subseteq \mathbb{R}^N$ is bounded open, $\varphi \in C(\overline{\Omega}, \mathbb{R}^N)$, $\xi \notin \varphi(\partial\Omega)$ and $d(\varphi, \Omega, \xi) \neq 0$, then the equation $\varphi(x) = \xi$ admits at least one solution.*

Proof It suffices to show that if $\varphi^{-1}(\xi) = \emptyset$, then $d(\varphi, \Omega, \xi) = 0$. By virtue of Definition 3.1.9, we may assume that $\varphi \in C^1(\overline{\Omega}, \mathbb{R}^N)$. Let $\epsilon \in (0, d(\xi, \varphi(\overline{\Omega})))$ and let $\vartheta \in C(0, +\infty)$ be as in Definition 3.1.1 with $\operatorname{supp} \vartheta \subseteq (0, \epsilon)$. Then by that definition, we have $d(\varphi, \Omega, \xi) = 0$. □

Remark 3.1.20 So, as a by-product of the above proof, we can equivalently reformulate the existence property as follows:

$$\text{"If } \xi \notin \varphi(\overline{\Omega}), \text{ then } d(\varphi, \Omega, \xi) = 0\text{"}. \quad (3.10)$$

Applying Proposition 3.1.19 to the identity map $i(x) = x$ for all $x \in \overline{\Omega}$, we have (see also (3.10)).

Corollary 3.1.21 *If $\Omega \subseteq \mathbb{R}^N$ is bounded open and $\xi \in \mathbb{R}^N$, then*

$$d(i, \Omega, \xi) = \begin{cases} 1 \text{ if } \xi \in \Omega \\ 0 \text{ if } \xi \notin \overline{\Omega} \end{cases} \quad \text{and} \quad d(-i, \Omega, \xi) = \begin{cases} (-1)^N & \text{if } \xi \in \Omega \\ 0 & \text{if } \xi \notin \overline{\Omega}. \end{cases}$$

Corollary 3.1.22 *If $\Omega \subseteq \mathbb{R}^N$ is bounded open, $\varphi \in C(\overline{\Omega}, \mathbb{R}^N)$ and $\varphi(\overline{\Omega}) \subseteq H =$ hyperplane, then for all $\xi \notin \varphi(\partial\Omega)$ we have $d(\varphi, \Omega, \xi) = 0$.*

Finally, we show that the degree map is $\mathbb{Z}$-valued. To this end, first we produce a convenient expression of the degree for a C^1-map (see Definition 3.1.1) when $\xi \in \mathbb{R}^N$ is a regular value of φ. This is an easy consequence of Definition 3.1.1.

Lemma 3.1.23 *If $\Omega \subseteq \mathbb{R}^N$ is bounded open, $\varphi \in C^1(\Omega, \mathbb{R}^N) \cap C(\overline{\Omega}, \mathbb{R}^N)$, $S_\varphi = \{z \in \Omega : J_\varphi(z) = 0\}$ and $\xi \notin \varphi(\partial\Omega) \cup \varphi(S_\varphi)$, then*

$$d(\varphi, \Omega, \xi) = \sum_{z \in \varphi^{-1}(\xi)} \operatorname{sgn} J_\varphi(z) \in \mathbb{Z}.$$

Proof By virtue of (3.10), we assume that $\varphi^{-1}(\xi) \neq \emptyset$. Then the inverse function theorem implies that $\varphi^{-1}(\xi) = \{z_k\}_{k=1}^m$. Choose $r > 0$ small such that $B_r(\xi) \cap [\varphi(\partial\Omega) \cup \varphi(S_\varphi)] = \emptyset$. From the inverse function theorem we know that for every $k \in \{1, \dots, m\}$, we can find a neighborhood U_k of z_k such that $\varphi|_{U_k}$ is a diffeomorphism and $\varphi^{-1}(B_r(\xi)) = \bigcup_{\mathrm{k}=1}^{N} U_k$. We choose $\epsilon > 0$ small such that $\epsilon < \min\{r, d(\xi, \varphi(\partial\Omega) \cup \varphi(S_\varphi))\}$ and a function $\vartheta \in C(0, +\infty)$ such that $\operatorname{supp} \vartheta \subseteq (0, \epsilon)$ and $\int_{\mathbb{R}^N} \vartheta(||x||)dx = 1$. Then according to Definition 3.1.1, we have

$$\begin{aligned} d(\varphi, \Omega, \xi) &= \int_\Omega \vartheta(||\varphi(z) - \xi||) J_\varphi(z) dz \\ &= \sum_{\mathrm{k}=1}^{N} \int_{U_k} \vartheta(||\varphi(z) - \xi||) J_\varphi(z) dz. \end{aligned} \tag{3.11}$$

We can perform a change of variables $x = \varphi(z) - \xi$ since $J_\varphi|_{U_k} \neq 0$. We have

$$dx = |J_\varphi(z)| dz = \operatorname{sgn}(J_\varphi(z)) J_\varphi(z) dz.$$

Therefore, we have

$$\int_{U_k} \vartheta(||\varphi(z) - \xi||) J_\varphi(z) dz = \operatorname{sgn} J_\varphi(z_k) \int_{B_r(0)} \vartheta(||y||) dy = \operatorname{sgn} J_\varphi(z_k)$$
$$\text{for all } k \in \{1, \dots, m\}.$$

Then from (3.11) it follows that

$$d(\varphi, \Omega, \xi) = \sum_{\mathrm{k}=1}^{N} \operatorname{sgn} J_\varphi(z_k) = \sum_{z \in \varphi^{-1}(\xi)} \operatorname{sgn} J_\varphi(z).$$

The proof is now complete. □

Proposition 3.1.24 *If $\Omega \subseteq \mathbb{R}^N$ is bounded open, $\varphi \in C(\overline{\Omega}, \mathbb{R}^N)$ and $\xi \notin \varphi(\partial\Omega)$, then $d(\varphi, \Omega, \xi) \in \mathbb{Z}$.*

Proof According to Definition 3.1.9, we can find $\psi \in C^1(\Omega, \mathbb{R}^N) \cap C(\overline{\Omega}, \mathbb{R}^N)$ such that

$$d(\varphi, \Omega, \xi) = d(\psi, \Omega, \xi). \tag{3.12}$$

Let $S_\psi = \{z \in \Omega : J_\psi(z) = 0\}$ (the singular set of ψ). Using Theorem 3.1.16, we see that $\mathbb{R}^N \backslash (\psi(\partial\Omega) \cup \psi(S_\psi))$ is dense in $\mathbb{R}^N \backslash \psi(\partial\Omega)$. So, we can find $\xi' \in \psi(\partial\Omega) \cup \psi(S_\psi)$ such that ξ and ξ' are in the same connected component of $\mathbb{R}^N \backslash \psi(\partial\Omega)$. Invoking Proposition 3.1.11 we have

$$d(\psi, \Omega, \xi) = d(\psi, \Omega, \xi'). \tag{3.13}$$

Moreover, Lemma 3.1.23 implies that

$$d(\psi, \Omega, \xi') \in \mathbb{Z}. \tag{3.14}$$

Finally from (3.12), (3.13) and (3.14), we conclude that $d(\varphi, \Omega, \xi) \in \mathbb{Z}$. □

So, summarizing the properties of Brouwer's degree, we can formulate the following theorem.

Theorem 3.1.25 *If*

$$\tau = \left\{(\varphi, \Omega, \xi) : \Omega \subseteq \mathbb{R}^N \text{ is bounded open}, \ \varphi \in C(\overline{\Omega}, \mathbb{R}^N), \ \xi \notin \varphi(\partial\Omega)\right\},$$

then there exists a map $d : \tau \to \mathbb{Z}$, known as the Brouwer degree, such that

(a) Normalization: $d(i, \Omega, \xi) = 1$ provided $\xi \in \Omega$.
(b) Domain Additivity: $d(\varphi, \Omega, \xi) = d(\varphi, \Omega_1, \xi) + d(\varphi, \Omega_2, \xi)$ with Ω_1, Ω_2 disjoint open subsets of Ω and $\xi \notin \varphi(\overline{\Omega} \backslash (\Omega_1 \cup \Omega_2))$.
(c) Homotopy Invariance: $d(h(t, \cdot), \Omega, \xi(t))$ is independent of $t \in [0, 1]$ whenever $h \in C([0, 1] \times \overline{\Omega}, \mathbb{R}^N)$, $\xi \in C([0, 1], \mathbb{R}^N)$ and $\xi(t) \notin h(t, \partial\Omega)$ for all $t \in [0, 1]$.
(d) Solution Property: $d(\varphi, \Omega, \xi) \neq 0$ implies $\varphi^{-1}(\xi) \neq \emptyset$.
(e) Continuity in (φ, ξ): $\varphi \to d(\varphi, \Omega, \xi)$ is constant on $B_\epsilon^{C(\overline{\Omega})}(\varphi) = \{\psi \in C(\overline{\Omega}) : \|\psi - \varphi\|_\infty < \epsilon\}$ where $\epsilon = d(\xi, \varphi(\partial\Omega)) > 0$; $\xi \to d(\varphi, \Omega, \xi)$ is constant on every connected component of $\mathbb{R}^N \backslash \varphi(\partial\Omega)$.
(f) Dependence on Boundary Values: $d(\varphi, \Omega, \xi) = d(\psi, \Omega, \xi)$ for every $\psi \in C(\overline{\Omega}, \mathbb{R}^N)$ with $\varphi|_{\partial\Omega} = \psi|_{\partial\Omega}$.
(g) Excision Property: $d(\varphi, \Omega, \xi) = d(\varphi, \Omega_1, \xi)$ for every open set $\Omega_1 \subseteq \Omega$ such that $\xi \notin \varphi(\overline{\Omega} \backslash \Omega_1)$.

Remark 3.1.26 Here we have stated the homotopy invariance property in a slightly more general form since the reference point depends also on $t \in [0, 1]$. However, we can easily see that this makes no difference if we recall that

$$d(h(t, \cdot), \Omega, \xi(t)) = d(h(t, \cdot) - \xi(t), \Omega, 0) \ \text{ for all } \ t \in [0, 1].$$

We also have another form of the homotopy invariance property, in which the domain is also t-dependent (see Lloyd [283, p. 28]).

Proposition 3.1.27 *If $\Omega^* \subseteq [0,1] \times \mathbb{R}^N$ is bounded open, $\Omega_t = \{z \in \mathbb{R}^N : (t,z) \in \Omega^*\}$ for all $t \in [0,1]$, $\varphi \in C(\overline{\Omega}^*, \mathbb{R}^N)$, for every $t \in [0,1]$ φ_t is the map $x \to \varphi(t,x)$ and $\xi \in C([0,1], \mathbb{R}^N)$ such that $\xi(t) \notin \varphi_t(\partial\Omega_t)$ for all $t \in [0,1]$, then $d(\varphi_t, \Omega_t, \xi(t))$ is independent of $t \in [0,1]$.*

In the above construction of Brouwer's degree, we used the natural basis $\{e_k\}_{k=1}^N$ of $\mathbb{R}^N$, where $e_k = (e_{ki} = \delta_{ki})_{i=1}^N$. The natural basis is ordered. We obtain the same degree function if instead we consider a different ordered basis $\{\hat{e}_k\}_{k=1}^N$. This can be easily verified using the transition matrix which corresponds to the change of basis $(e_k)_{k=1}^N \to (\hat{e}_k)_{k=1}^N$ and then the chain rule. So, if $\Omega \subseteq \mathbb{R}^N$ is bounded open and $\varphi \in C^1(\Omega, \mathbb{R}^N) \cap C(\overline{\Omega}, \mathbb{R}^N)$, then we define $\hat{z} = A(z)$, where A is the transition matrix, $\widehat{\Omega} = A(\Omega)$ and $\widehat{\varphi}(\hat{z}) = A\varphi(A^{-1}(\hat{z}))$ for all $\hat{z} \in \widehat{\Omega}$. The chain rule yields.

$$\begin{aligned} &J_{\widehat{\varphi}}(\hat{z}) = \det A\, J_\varphi(A^{-1}(\hat{z})) \det A^{-1} = J_\varphi(A^{-1}(\hat{z})) = J_\varphi(z) \\ &\Rightarrow \operatorname{sgn} J_{\widehat{\varphi}}(\hat{z}) = \operatorname{sgn} J_\varphi(z). \end{aligned} \tag{3.15}$$

So, if $\xi \notin \varphi(\partial\Omega \cup S_\varphi)$, then

$$d(\varphi, \Omega, \xi) = d(A\varphi A^{-1}, A(\Omega), A(\xi)) = d(\widehat{\varphi}, \widehat{\Omega}, A(\xi)) \tag{3.16}$$

(see Lemma 3.1.23 and (3.15)).

The requirement that $\xi \notin \varphi(S_\varphi)$ can be removed using Sard's theorem (see Theorem 3.1.16). Finally the passage to continuous functions can be achieved by approximating a continuous function by smooth functions.

Arguing similarly, we can replace $\mathbb{R}^N$ by X, an N-dimensional normed (hence Banach) space. So, we fix an ordered basis $(x_k)_{k=1}^N$ of X. For every $x \in \mathbb{R}^N$ we have

$$x = \sum_{\mathrm{k}=1}^N c_k(x) x_k, \text{ with } c_k(x) \in \mathbb{R} \text{ (the } k^{th} \text{ coordinate of } x).$$

Then, the map $\sigma : X \to \mathbb{R}^N$ defined by $\sigma(x) = (c_k(x))_{k=1}^N$ is a homeomorphism. Given $\Omega \subseteq X$ bounded open, $\varphi \in C(\overline{\Omega}, X)$ and $\xi \notin \varphi(\partial\Omega)$. Since σ is a homeomorphism if we set $\psi = \sigma\varphi\sigma^{-1} : \sigma(\overline{\Omega}) \subseteq \mathbb{R}^N \to \mathbb{R}^N$, then $d(\psi, \sigma(\Omega), \sigma(\xi))$ is well-defined. So, we can define

$$d(\varphi, \Omega, \xi) = d(\psi, \sigma(\Omega), \sigma(\xi)). \tag{3.17}$$

For this definition to be valid, it should not depend on the homeomorphism σ (that is, on the original choice of basis on X). So, suppose that $(\hat{x}_k)_{k=1}^N$ is another ordered basis of X and let $\widehat{\sigma}$ be the corresponding homeomorphism. Let A be the transition matrix passing from $(\widehat{x}_k)_{k=1}^N$ to $(x_k)_{k=1}^N$. Then $\sigma = A\widehat{\sigma}$ and so

$$\begin{aligned} d(\psi, \sigma(\Omega), \sigma(\xi)) &= d(\sigma\varphi\sigma^{-1}, \sigma(\Omega), \sigma(\xi)) \\ &= d(A\widehat{\sigma}\varphi\widehat{\sigma}^{-1}A^{-1}, A\widehat{\sigma}(\Omega), A\widehat{\sigma}(\xi)) \\ &= d(\widehat{\sigma}\varphi\widehat{\sigma}^{-1}, \widehat{\sigma}(\Omega), \widehat{\sigma}(\xi)) \text{ (see (3.16)).} \end{aligned}$$

This proves that (3.17) is well defined.

Therefore, we can make the following definition:

Definition 3.1.28 Let X be an N-dimensional Banach space and let

$$\widehat{\tau} = \{(\varphi, \Omega, \xi) : \Omega \subseteq X \text{ is bounded open, } \varphi \in C(\overline{\Omega}, X) \text{ and } \xi \notin \varphi(\partial\Omega)\}.$$

Suppose that $(x_k)_{k=1}^N$ is an ordered basis of X and $(e_k)_{k=1}^N$ is the natural basis of $\mathbb{R}^N$.

Then we define the Brouwer degree on $\widehat{\tau}$ by

$$d(\varphi, \Omega, \xi) = d(\sigma\varphi\sigma^{-1}, \sigma(\Omega), \sigma(\xi)),$$

where $\sigma : X \to \mathbb{R}^N$ is the homeomorphism defined by $\sigma(x_k) = e_k$ for all $k \in \{1, \ldots, N\}$.

The previous discussion leads to the following theorem.

Theorem 3.1.29 *The degree function $d : \widehat{\tau} \to \mathbb{Z}$ defined in Definition 3.1.28 has all the properties listed in Theorem 3.1.25.*

Finally, suppose that X and Y are both real N-dimensional Banach spaces, $\Omega \subseteq X$ is bounded, open, $\varphi : \overline{\Omega} \to Y$ is continuous and $\xi \in Y \backslash \varphi(\partial\Omega)$. We consider two bases $\{e_k\}_{k=1}^N \subseteq X$ and $\{h_k\}_{k=1}^N \subseteq Y$ and the corresponding isomorphisms $\sigma : X \to \mathbb{R}^N$ and $\sigma : Y \to \mathbb{R}^N$. We set $\psi = \widehat{\sigma}\varphi\sigma^{-1}$ and we may define

$$d(\varphi, \Omega, \xi) = d(\psi, \sigma(\Omega), \widehat{\sigma}(\xi)).$$

Suppose we change the bases on X and Y. Then $\sigma = A\widehat{\sigma}$, $\widehat{\sigma} = B\sigma^*$ and the new function is $B^{-1}\psi A$. So, we have

$$d(B^{-1}\psi A, \widehat{\sigma}(\Omega), \sigma^*(\xi)) = \operatorname{sgn}(\det A \det B)\, d(\psi, \sigma(\Omega), \widehat{\sigma}(\xi)).$$

So, we see that the definition of the degree we gave earlier depends on the choice of the bases. We say that two bases $\{e_k\}_{k=1}^N \subseteq X$ and $\{h_k\}_{k=1}^N$ have the same orientation if the matrix A defined by

$$Ae_k = h_k \text{ for all } k \in \{1, \ldots, N\}$$

has $\det A > 0$. This introduces an equivalence relation with exactly two equivalence classes. We call X oriented if we have chosen one class and ignore the other one. Therefore, if X and Y are oriented, then $\det A > 0$, $\det B > 0$ and so the Brouwer degree given above is well-defined.

In the rest of this section, we present some important application of Brouwer's degree.

We start by recalling the following important notion from topology.

Definition 3.1.30 Let X be a Hausdorff topological space and $C \subseteq X$. We say that C is a retract of X if there is a continuous map $r : X \to C$ such that $r|_C$ =identity (that is, the identity map $i : C \to C$ admits a continuous extension $r : X \to C$ and this extension is called a retraction).

Remark 3.1.31 Every retract is closed. By Proposition 2.1.9 every closed convex set of a normal space, is a retract. The notion of retract is topologically invariant, that is, if $h : X \to Y$ is a homeomorphism and $C \subseteq X$, then $h(C)$ is a retract of Y if and only if C is a retract of X. Evidently, $\partial\overline{B}_1 = \{x \in \mathbb{R}^N : ||x|| = 1\}$ is a retract of $\overline{B}_1\backslash\{0\}$ with a retraction given by $x \to \frac{x}{||x||}$. This is no longer true if $\overline{B}_1\backslash\{0\}$ is replaced by $\overline{B}_1 = \{x \in \mathbb{R}^N : ||x|| \leqslant 1\}$.

Proposition 3.1.32 *$\partial B_1 = S^{N-1}$ is not a retract of $\overline{B}_1$.*

Proof Arguing by contradiction suppose that ∂B_1 is a retract of $\overline{B}_1$ and let $r : \overline{B}_1 \to \partial B_1$ be a retraction. Then from Theorem 3.1.25(f) and Corollary 3.1.21, we have

$$d(r, B_1, 0) = d(i, B_1, 0) = 1.$$

Invoking Theorem 3.1.25(d), we can find $x \in B_1$ such that $r(x) = 0$, which is a contradiction to the fact that $r(\cdot)$ has values in ∂B_1. $\square$

Remark 3.1.33 In contrast, in an infinite-dimensional normed space X, $\partial B_1 = \{x \in X; ||x|| = 1\}$ is a retract of $\overline{B}_1 = \{x \in X : ||x|| \leqslant 1\}$. To see this, let V be a dense linear subspace and let $D = \partial B_1 \cap V$. Then D is dense ∂B_1 and $i|_D$ has a continuous extension $g : \overline{B}_1 \cup \text{conv}\, D \subseteq \partial B_1 \cup (\overline{B}_1 \cap V) \subset \overline{B}_1$ (see Proposition 2.1.9 and its proof). Now let $x_0 \in \overline{B}_{1/3} - g(\overline{B}_1)$ and define $\hat{r} : \overline{B}_1\backslash\{x_0\} \to \partial B_1$, by

$$\hat{r}(x) = \begin{cases} \dfrac{\frac{1}{2}(x - x_0) + ||x - x_0||x_0}{||\frac{1}{2}(x - x_0) + ||x - x_0||x_0||} & \text{if } 0 < ||x - x_0|| < \frac{1}{2} \\ \dfrac{x}{||x||} & \text{if } ||x - x_0|| \geqslant \frac{1}{2}. \end{cases}$$

Evidently, $\hat{r}g : \overline{B}_1 \to \partial B_1$ is continuous and $\hat{r}g|_{\partial B_1}$ = identity.

The next result is the celebrated Brouwer's fixed point theorem.

Theorem 3.1.34 *If $\varphi : \overline{B}_1 = \{x \in \mathbb{R}^N : ||x|| \leqslant 1\} \to \overline{B}_1$ is continuous, then there exists an $x \in \overline{B}_1$ such that $\varphi(x) = x$ (that is, φ admits a fixed point).*

Proof If there exists an $x \in \partial B_1$ such that $\varphi(x) = x$, then we are done. Therefore, we may assume that $x - \varphi(x) \neq 0$ for all $x \in \partial B_1$. We consider the homotopy

$$h(t, x) = x - t\varphi(x) \text{ for all } t \in [0, 1], \text{ all } x \in \overline{B}_1.$$

Note that $0 \notin h(1, \partial B_1)$. Suppose that for some $t \in [0, 1)$ and some $x \in \partial B_1$, we have

$$\begin{aligned} & x - t\varphi(x) = 0 \\ \Rightarrow\ & t||\varphi(x)|| = 1, \text{ a contraction.} \end{aligned}$$

Therefore $0 \notin h(t, \partial B_1)$ for all $t \in [0, 1]$ and so Theorem 3.1.25(c) and Corollary 3.1.21 imply

$$d(i - \varphi, B_1, 0) = d(i, B_1, 0) = 1.$$

So, by virtue of Theorem 3.1.25(d), there exists an $x \in \overline{B}_1$ such that $\varphi(x) = x$. □

Remark 3.1.35 In fact, in the above theorem, we can replace $\overline{B}_1$ by any convex and compact subset of $\mathbb{R}^N$ (see also Sect. 4.1).

Proposition 3.1.36 *Theorem 3.1.34 and Proposition 3.1.32 are equivalent.*

Proof Theorem 3.1.34 $\Rightarrow$ *Proposition* 3.1.32.

Suppose that ∂B_1 is a retract of $\overline{B}_1$. Then we can find $r : \overline{B}_1 \to \partial B_1$ continuous such that $r|_{\partial B_1} = i$. Consider the continuous map $\varphi : \overline{B}_1 \to \overline{B}_1$ defined by $\varphi(x) = -r(x)$. Clearly, $\varphi(\cdot)$ has no fixed point, a contradiction to Theorem 3.1.34.

Proposition 3.1.32 $\Rightarrow$ *Theorem* 3.1.34.

Suppose we can find a continuous function $\varphi : \overline{B}_1 \to \overline{B}_1$ such that $\varphi(x) \neq x$ for all $x \in \overline{B}_1$. For every $x \in \overline{B}_1$ let $t(x) > 0$ such that $(1 - t(x))\varphi(x) + t(x)x \in \partial B_1$. It is easy to see that $x \to t(x)$ is continuous on $\overline{B}_1$. Let $\psi(x) = \varphi(x) + t(x)(x - \varphi(x))$. Then $\psi \in C(\overline{B}_1, \partial B_1)$ and $\psi|_{\partial B_1} = i$, which contradicts Proposition 3.1.32. □

Another important result of Nonlinear Analysis which is equivalent to Theorem 3.1.25 (Brouwer's fixed point theorem), is the following Hartman–Stampacchia theorem concerning variational inequalities.

Theorem 3.1.37 *If $C \subseteq \mathbb{R}^N$ is nonempty, convex compact and $\varphi : C \to \mathbb{R}^N$ is continuous, then there exists a $\hat{u} \in C$ such that $(\varphi(\hat{u}), u - \hat{u})_{\mathbb{R}^N} \geqslant 0$ for all $u \in C$.*

Proposition 3.1.38 *Theorems 3.1.34 and 3.1.37 are equivalent.*

Proof Theorem 3.1.34 $\Rightarrow$ *Theorem* 3.1.37

Let $\sigma : C \to C$ be the continuous map defined by $\sigma(u) = \mathrm{proj}_C\,(-\varphi(u) + u)$. Then according to Theorem 3.1.34 (see also Remark 3.1.35), we can find $\hat{u} \in C$ such that $\sigma(\hat{u}) = \hat{u}$. Hence

$$\begin{aligned} & (-\varphi(\hat{u}) + \hat{u} - \hat{u}, u - \hat{u})_{\mathbb{R}^N} \leqslant 0 \\ \Rightarrow\ & (\varphi(\hat{u}), u - \hat{u})_{\mathbb{R}^N} \geqslant 0 \text{ for all } u \in C. \end{aligned}$$

Theorem 3.1.37 $\Rightarrow$ *Theorem* 3.1.34

Consider a continuous map $\varphi : \overline{B}_1 \to \overline{B}_1$ and let $\psi = i - \varphi$ (i = identity map). We apply Theorem 3.1.37 on ψ and so can find $\hat{u} \in \overline{B}_1$ such that

$$(\hat{u} - \varphi(\hat{u}), u - \hat{u})_{\mathbb{R}^N} \geqslant 0 \text{ for all } u \in \overline{B}_1. \tag{3.18}$$

If $\hat{u} \in B_1$, then we can find $\epsilon > 0$ such that $B_\epsilon(\hat{u}) \subseteq \overline{B}_1$. So, in (3.18) we choose $u = \hat{u} + \epsilon h$ for $h \in \partial B_1$. Then

$$\begin{aligned} &(\hat{u} - \varphi(\hat{u}), h)_{\mathbb{R}^N} \geqslant 0 \text{ for all } h \in \partial B_1 \\ \Rightarrow\ &\hat{u} = \varphi(\hat{u}). \end{aligned}$$

If $\hat{u} \in \partial B_1$, then from $(\varphi(\hat{u}) - \hat{u}, u - \hat{u})_{\mathbb{R}^N} \leqslant 0$ for all $u \in \overline{B}_1$, we infer that $\hat{u} = \mathrm{proj}_{\overline{B}_1}\varphi(\hat{u})$ and so $\varphi(\hat{u}) = \lambda\hat{u}$ for some $\lambda \geqslant 1$. Therefore

$$\begin{aligned} &\hat{u} - \varphi(\hat{u}) = \mu\hat{u} \text{ for some } \mu \leqslant 0 \\ \Rightarrow\ &\varphi(\hat{u}) = (1 - \mu)\hat{u}. \end{aligned}$$

Since φ has values in $\overline{B}_1$, we have $|1 - \mu| \leqslant 1 \Rightarrow -1 \leqslant 1 - \mu \leqslant 1 \Rightarrow \mu \geqslant 0 \Rightarrow \mu = 0$ and so finally we have $\varphi(\hat{u}) = \hat{u}$. ☐

Proposition 3.1.39 *If $\varphi, \psi \in C(\overline{\Omega}, \mathbb{R}^N)$, $h \in C([0, 1] \times \partial\Omega, \mathbb{R}^N)$, $\xi \notin h(t, \partial\Omega)$ for all $t \in [0, 1]$ and $h(0, \cdot) = \varphi|_{\partial\Omega}$, $h(1, \cdot) = \psi|_{\partial\Omega}$, then $d(\varphi, \Omega, \xi) = d(\psi, \Omega, \xi)$.*

Proof By the Tietze extension theorem, we can find $\hat{h} \in C([0, 1] \times \overline{\Omega}, \mathbb{R}^N)$ such that $\hat{h}|_{[0,1]\times\partial\Omega} = h$. Let $\widehat{\varphi} = h(0, \cdot)$ and $\widehat{\psi} = h(1, \cdot)$. Evidently, $\varphi|_{\partial\Omega} = \widehat{\varphi}|_{\partial\Omega}$, $\psi|_{\partial\Omega} = \widehat{\psi}|_{\partial\Omega}$. So, from Theorem 3.1.25(f), we have

$$d(\varphi, \Omega, \xi) = d(\widehat{\varphi}, \Omega, \xi) \text{ and } d(\psi, \Omega, \xi) = d(\widehat{\psi}, \Omega, \xi). \tag{3.19}$$

On the other hand, the homotopy invariance property (see Theorem 3.1.25(c)) implies that

$$d(\widehat{\varphi}, \Omega, \xi) = d(\widehat{\psi}, \Omega, \xi). \tag{3.20}$$

From (3.19) and (3.20), we conclude that $d(\varphi, \Omega, \xi) = d(\psi, \Omega, \xi)$. ☐

Proposition 3.1.40 *If $\varphi \in C(\mathbb{R}^N, \mathbb{R}^N)$ and $\lim\limits_{||u||\to\infty} \frac{(\varphi(u),u)_{\mathbb{R}^N}}{||u||} = +\infty$, then φ is surjective.*

Proof Let $y \in \mathbb{R}^N$ and consider the homotopy

$$h(t, u) = tu + (1 - t)\varphi(u) - y \text{ for all } (t, u) \in [0, 1] \times \mathbb{R}^N.$$

For $||u|| = r$, we have

$$(h(t,u),u)_{\mathbb{R}^N} \geqslant r\left[tr + (1-t)\frac{(\varphi(u),u)_{\mathbb{R}^N}}{||u||} - ||y||\right] > 0$$

for all $t \in [0, 1]$ and for large $r > ||y||$. So, from the homotopy invariance property (see Theorem 3.1.25(c)), we have

$$d(\varphi, B_r, y) = d(\varphi - y, B_r, 0) = d(i, B_r, 0) = 1 \text{ (see Corollary 3.1.21).}$$

So, by the solution property (see Theorem 3.1.25(d)), we can find $\hat{u} \in \mathbb{R}^N$ such that

$$\varphi(\hat{u}) = y.$$

Since $y \in \mathbb{R}^N$ is arbitrary, we conclude that φ is surjective. □

Proposition 3.1.41 *If $d(\varphi, \Omega, \xi) \neq 0$, then $\varphi(\Omega)$ is a neighborhood of ξ.*

Proof From Theorem 3.1.25(d), we can find $u_0 \in \Omega$ such that $\varphi(u_0) = \xi$. Let U_ξ be the connected component of $\mathbb{R}^N \setminus \varphi(\partial\Omega)$ containing ξ. Then from Theorem 3.1.25(e), we have

$$\begin{aligned} & 0 \neq d(\varphi, \Omega, \xi) = d(\varphi, \Omega, y) \text{ for all } y \in U_\xi \\ \Rightarrow\ & U_\xi \subseteq \varphi(\Omega) \text{ (again by Theorem 3.1.25(d))} \\ \Rightarrow\ & \varphi(\Omega) \text{ is a neighborhood of } \xi. \end{aligned}$$

The proof is now complete. □

Corollary 3.1.42 *If $\varphi(\Omega)$ is contained in a proper linear subspace of $\mathbb{R}^N$, then $d(\varphi, \Omega, \xi) = 0$*

The next result is useful in extending the notion of Brouwer's degree to infinite dimensions.

Proposition 3.1.43 *If $\Omega \subseteq \mathbb{R}^N$ is bounded open, $\varphi \in C(\overline{\Omega}, \mathbb{R}^N)$, $\psi = i + \varphi$, $m \leqslant N$, $\xi \in \mathbb{R}^m$ and $\xi \notin \psi(\partial\Omega)$, then $d(\psi, \Omega, \xi) = d(\psi|_{\overline{\Omega \cap \mathbb{R}^m}}, \Omega \cap \mathbb{R}^m, \xi)$.*

Proof Note that $\psi(\overline{\Omega} \cap \mathbb{R}^m) \subseteq \mathbb{R}^m$. Also, by virtue of Definition 3.1.9 and Theorem 3.1.16 (Sard's theorem), we may assume that $\varphi \in C^1(\Omega, \mathbb{R}^n) \cap C(\overline{\Omega}, \mathbb{R}^N)$ and $\xi \notin \psi(S_\psi)$. Suppose that $\psi(u) = \xi$ for some $u \in \Omega \cap \mathbb{R}^m$. Let $\psi_m = \psi|_{\overline{\Omega \cap \mathbb{R}^m}}$. We have

$$\begin{aligned} J_{\psi_m}(u) &= \det\left[I_m - (\partial_k \varphi_i(u))_{k,i=1}^m\right] \ (I_m = m \times m \text{ identity matrix}) \\ \text{and} \quad J_\psi(u) &= \det\left[\begin{array}{c|c} I_m - (\partial_k \varphi_i(u))_{k,i=1}^m & -(\partial_k \varphi_i(u))_{k,i=1}^m \\ \hline 0 & I_{N-m} \end{array}\right]. \end{aligned}$$

Developing with respect to the last $N - m$ rows, we obtain $J_\psi(u) = J_{\psi_m}(u)$ and so

$$d(\psi, \Omega, \xi) = d(\psi_m, \Omega \cap \mathbb{R}^m, \xi) \text{ (see Lemma 3.1.23).}$$

The proof is now complete. □

According to Theorem 3.1.25(d), in order to show the solvability of the equation $\varphi(u) = \xi$, it suffices to show that $d(\varphi, \Omega, \xi) \neq 0$. In this direction, the so-called Borsuk's theorem is very useful.

Theorem 3.1.44 *If $\Omega \subseteq \mathbb{R}^N$ is bounded, open, symmetric (that is, $\Omega = -\Omega$), $0 \in \Omega$, $\varphi \in C(\overline{\Omega}, \mathbb{R}^N)$ is odd and $0 \notin \varphi(\partial\Omega)$, then $d(\varphi, \Omega, 0)$ is odd.*

Proof We claim that we may assume that $\varphi \in C^1(\Omega, \mathbb{R}^N) \cap C(\overline{\Omega}, \mathbb{R}^N)$ and $J_\varphi(0) \neq 0$. To this end we approximate uniformly φ by $\psi_1 \in C^1(\Omega, \mathbb{R}^N) \cap C(\overline{\Omega}, \mathbb{R}^N)$ and then consider the odd part $\psi_2(u) = \frac{1}{2}[\psi_1(u) - \psi_1(-u)]$. We choose ϑ which is not an eigenvalue of $\psi_2'(0)$ and set $\psi = \psi_2 - \vartheta i \in C^1(\Omega, \mathbb{R}^N) \cap C(\overline{\Omega}, \mathbb{R}^N)$ (i=identity map). Evidently ψ is odd with $J_\psi(0) \neq 0$ and it is uniformly close to φ, if ϑ and $||\psi_1 - \varphi||_\infty$ are sufficiently small. We have $d(\varphi, \Omega, 0) = d(\psi, \Omega, 0)$.

So, we assume that $\varphi \in C^1(\Omega, \mathbb{R}^N) \cap C(\overline{\Omega}, \mathbb{R}^N)$ and $J_\varphi(0) \neq 0$. We will produce an odd map $\psi \in C^1(\Omega, \mathbb{R}^N) \cap C(\overline{\Omega}, \mathbb{R}^N)$, sufficiently close to φ such that $0 \notin \psi(S_\psi)$. This will be done by induction.

Let $\Omega_k = \{u \in \Omega : u_i \neq 0 \text{ for some } i \leqslant k\}$ and let $\eta \in C^1(\mathbb{R})$ be odd such that $\eta'(0) = 0$ and $\eta(t) = 0$ if and only if $t = 0$. Let $\widehat{\varphi}(u) = \frac{\varphi(u)}{\eta(u_1)}$ for all $u \in \Omega_1$. By Theorem 3.1.16, we can find a $v' \notin \widehat{\varphi}(S_{\widehat{\varphi}})$ with $||v^1||$ as small as necessary for what follows. Then 0 is a regular value for $\psi_1(u) = \varphi(u) - \eta(u_1)v^1$ for all $u \in \Omega_1$ (note that $\psi_1'(u) = \eta(u_1)\widehat{\psi}(u)$ for all $u \in \Omega_1$ such that $\psi_1(u) = 0$). Now suppose that we have already produced an odd function $\psi_k \in C^1(\Omega, \mathbb{R}^N) \cap C(\overline{\Omega}, \mathbb{R}^N)$ close to φ on $\overline{\Omega}$ such that $0 \notin \psi_k(S_{\psi_k})$ for some $k < n$ and set

$$\psi_{k+1}(u) = \psi_k(u) - \eta(u_{k+1})v^{k+1}$$

with $||v^{k+1}||$ sufficiently small such that 0 is a regular value of ψ_{k+1} on $\{u \in \Omega : u_{k+1} \neq 0\}$. Note that $\psi_{k+1} \in C^1(\Omega, \mathbb{R}^N) \cap C(\overline{\Omega}, \mathbb{R}^N)$ is odd and uniformly close to φ on $\overline{\Omega}$. If $u \in \Omega_{k+1}$ and $u_{k+1} = 0$, then $u \in \Omega_k$, $\psi_{k+1}(u) = \psi_k(u)$ and $\psi_{k+1}'(u) = \psi_k'(u)$, hence $J_{\psi_{k+1}}(u) \neq 0$. It follows $0 \notin \psi_{k+1}(S_{\psi_{k+1}})$. Therefore $\psi = \psi_N$ is odd, uniformly close to $\widehat{\varphi}$ on $\overline{\Omega}$ and $0 \notin \psi(S_\psi)$ on $\Omega \backslash \{0\}$ (note $\Omega_N = \Omega\backslash\{0\}$). But by the induction hypothesis we know that $\psi'(0) = \psi_1'(0) = \varphi'(0)$ and so $0 \notin \psi(S_\psi)$.

Therefore we have reduced the theorem to the case $\varphi \in C^1(\Omega, \mathbb{R}^N) \cap C(\overline{\Omega}, \mathbb{R}^N)$, $J_\varphi(0) \neq 0$ and φ is odd. We have

$$\varphi^{-1}(0) = \{0\} \text{ or } \varphi^{-1}(0) = \{0\} \cup \{(u_{k'} - u_k)\}_{k=1}^m \text{ with } u_k \in \Omega,\ u_k \neq 0.$$

In the first case we have $d(\varphi, \Omega, 0) = \pm 1$ and in the second case we have

$$\begin{aligned} d(\varphi, \Omega, 0) &= \pm 1 + \sum_{\mathrm{k}=1}^{m} \left[\operatorname{sgn} J_{\varphi}(u_k) + \operatorname{sgn} J_{\varphi}(-u_k)\right] \text{ (see Lemma 3.1.23)} \\ &= \pm 1 + 2\sum_{\mathrm{k}=1}^{m} \operatorname{sgn} J_{\varphi}(u_k) \text{ (since } J_{\varphi} \text{ is even).} \end{aligned}$$

Therefore in both cases we see that $d(\varphi, \Omega, 0)$ is odd. □

The following result is know as the "Borsuk–Ulam theorem".

Theorem 3.1.45 *If $\Omega \subseteq \mathbb{R}^N$ is bounded, open, symmetric with $0 \in \Omega$, $\varphi \in C(\partial\Omega, \mathbb{R}^N)$ is odd and $\varphi(\partial\Omega)$ is contained in a hyperplane H of $\mathbb{R}^N$, then there exists a $u \in \partial\Omega$ such that $\varphi(u) = 0$.*

Proof From the Tietze extension theorem, we can find $\widehat{\varphi} \in C(\overline{\Omega}, H)$ such that $\widehat{\varphi}|_{\partial\Omega} = \varphi$. Let

$$\varphi_0(u) = \frac{\widehat{\varphi}(u) - \widehat{\varphi}(-u)}{2}.$$

Then $\varphi_0 \in C(\overline{\Omega}, H)$ is odd. If $0 \notin \varphi_0(\partial\Omega)$, then from Theorem 3.1.44 and Proposition 3.1.41, we have that $\varphi_0(\Omega)$ is a neighborhood of 0, which is a contradiction to the fact that $\varphi_0(\overline{\Omega}) \subseteq H$. □

Corollary 3.1.46 *If $\Omega \subseteq \mathbb{R}^N$ is bounded, open, symmetric with $0 \in \Omega$, $\varphi \in C(\partial\Omega, \mathbb{R}^N)$ and $\varphi(\partial\Omega)$ is contained in a hyperplane H of $\mathbb{R}^N$, then there exists a $\hat{u} \in \partial\Omega$ such that $\varphi(\hat{u}) = \varphi(-\hat{u})$.*

Proof Let $\psi(u) = \frac{\varphi(u)-\varphi(-u)}{2}$ for all $u \in \overline{\Omega}$. Then ψ is odd and so we can apply Theorem 3.1.45 and obtain $\hat{u} \in \partial\Omega$ such that $\psi(\hat{u}) = 0$. It follows that $\varphi(\hat{u}) = \varphi(-\hat{u})$. □

Theorem 3.1.45 permits us to distinguish between the spheres of finite-dimensional Banach spaces. This is a consequence of Theorem 3.1.45, since every continuous odd map from $\partial B_1^N = \{u \in \mathbb{R}^N : ||u|| = 1\}$ into $\partial B_1^m = \{u \in \mathbb{R}^m : ||u|| = 1\}$ $(m < N)$ must vanish somewhere.

Corollary 3.1.47 *If $N > m$, then there is no continuous odd map from ∂B_1^N into ∂B_1^m.*

The next result, also a consequence of Theorem 3.1.44, is known as the "invariance of domain theorem". Recall that a map $\varphi : \mathbb{R}^N \to \mathbb{R}^N$ is locally one-to-one if for every $x \in \mathbb{R}^N$, there exists a neighborhood $U(x)$ such that $\varphi|_{U(x)}$ is one-to-one.

Theorem 3.1.48 *If $\Omega \subseteq \mathbb{R}^N$ is open and $\varphi : \Omega \to \mathbb{R}^N$ is continuous and locally one-to-one, then φ is an open map.*

Proof Let $z_0 \in \Omega$. By replacing Ω by $\Omega - z_0$ and φ by $\widehat{\varphi}(z) = \varphi(z + z_0) - \varphi(z_0)$ for all $z \in \Omega\backslash\{z_0\}$ if necessary, we may assume that $z_0 = 0$ and $\varphi(0) = 0$.

Since φ is locally one-to-one, we can find $r > 0$ such that $\varphi|_{\bar{B}_r}$ is one-to-one. We consider the homotopy

$$h(t,u) = \varphi\left(\frac{u}{1+t}\right) - \varphi\left(-\frac{tu}{1+t}\right) \text{ for all } t \in [0,1], \text{ all } u \in \bar{B}_r.$$

Clearly, h is continuous and $h(0,\cdot) = \varphi$, $h(1,u) = \varphi(\frac{1}{2}u) - \varphi(-\frac{1}{2}u)$. Hence $h(1,\cdot)$ is odd. If $h(t,u) = 0$ for some $t \in [0,1]$ and some $u \in \partial B_r$, then

$$\begin{aligned} \frac{u}{1+t} &= -\frac{tu}{1+t} \text{ (since } \varphi|_{\bar{B}_r} \text{ is one-to-one)} \\ \Rightarrow u &= 0, \text{ a contradiction.} \end{aligned}$$

Then by virtue of the homotopy invariance property of the degree (see Theorem 3.1.25(c)), we have

$$d(\varphi, B_r, \xi) = d(h(1,\cdot), B_r, 0) \neq 0 \text{ (see Theorem 3.1.44)}$$

for every $\xi \in B_s$, for some $x > 0$ and so $B_s \subseteq \varphi(B_r)$. $\square$

Remark 3.1.49 In fact, one can show that if $\Omega \subseteq \mathbb{R}^N$ is bounded open and $\varphi \in C(\overline{\Omega}, \mathbb{R}^N)$ is one-to-one (injective), then for $\xi \in \varphi(\Omega)$ we have $d(\varphi, \Omega, \xi) = \pm 1$ (see Lloyd [283, p. 51]).

We conclude this section with two nice results. The first essentially says that you cannot comb a coconut.

Proposition 3.1.50 *If $\Omega \subseteq \mathbb{R}^N$ is bounded open with $0 \in \Omega$, N is odd, and $\varphi \in C(\partial\Omega, \mathbb{R}^N \setminus \{0\})$, then there exist $u \in \partial\Omega$ and $\lambda \neq 0$ such that $\varphi(u) = \lambda u$.*

Proof Thanks to the Tietze extension theorem, without any loss of generality we may assume that $\varphi \in C(\overline{\Omega}, \mathbb{R}^N)$. From Corollary 3.1.21 and since N is odd, we have

$$d(-i, \Omega, 0) = -1 \text{ (recall } i \text{is the identity map).}$$

If $d(\varphi, \Omega, 0) \neq -1$, then the homotopy

$$h(t,u) = (1-t)\varphi(u) - tu \text{ for all } t \in [0,1], \text{ all } u \in \overline{\Omega}$$

must have a zero $(\hat{t}, \hat{u}) \in (0,1) \times \partial\Omega$. Then $\varphi(\hat{u}) = \frac{t\hat{u}}{1-t}$ and so the result follows with $\lambda = \frac{t}{1-t} > 0$.

If $d(\varphi, \Omega, 0) = -1$, then we apply same argument to the homotopy

$$h(t,u) = (1-t)\varphi(u) + tu \text{ for all } t \in [0,1], \text{ all } u \in \overline{\Omega}.$$

The proof is now complete. $\square$

The second result is known in the literature as the Ljusternik–Schnirelmann–Borsuk theorem.

Proposition 3.1.51 *If* $\Omega \subseteq \mathbb{R}^N$ *is bounded open and symmetric with* $0 \in \Omega$ *and* $\{C_k\}_{k=1}^m$ *is a closed covering of* $\partial\Omega$ *such that* $C_k \cap (-C_k) = \emptyset$ *for all* $k \in \{1, \ldots, m\}$, *then* $m \geqslant N + 1$.

Proof Arguing by contradiction, suppose that $m \leqslant N$. Let $\varphi_k|_{C_k} = 1$ $\varphi_k|_{-C_k} = 1$ for $k \in \{1, \ldots, m-1\}$ and $\varphi_k|_{\overline{\Omega}} = 1$ for $k \in \{m, \ldots, N\}$. For $k \in \{1, \ldots, n\}$, extend φ_k continuously to $\overline{\Omega}$ and then set $\varphi = (\varphi_k)_{k=1}^N$ a continuous map from $\overline{\Omega}$ into $\mathbb{R}^N$. We claim that

$$\varphi(-u) \neq \lambda\varphi(u) \text{ for all } u \in \partial\Omega, \text{ all } \lambda \geqslant 0.$$

If this is not true, then we can find $u_0 \in \partial\Omega$ and $\lambda_0 > 0$ such that $\varphi(-u_0) = \lambda_0\varphi(u_0)$ (note that $0 \notin \varphi(\partial\Omega)$). Also $u_0 \notin C_k \cup (-C_k)$ for all $k \in \{1, \ldots, m-1\}$ since $\varphi_k(-x) = -\varphi_k(x)$. So, $u_0 \in C_m$. Then $u_0 \notin -C_m$ and so $-u_0 \in C_k$ for some $k \in \{1, \ldots, m-1\}$, hence $u_0 \in -C_k$, a contradiction. Therefore $\varphi(-u) \neq \lambda\varphi(u)$ for all $u \in \partial\Omega$ and all $\lambda \geqslant 0$. Consider the homotopy

$$h(t, u) = \varphi(u) - t\varphi(-u) \text{ for all } (t, u) \in [0, 1] \times \overline{\Omega}.$$

By the homotopy invariance property and Theorem 3.1.44 we have

$$d(\varphi, \Omega, 0) = d(\varphi_0, \Omega, 0) \neq 0 \text{ where } \varphi_0(u) = \varphi(u) - \varphi(-u) \text{ odd}.$$

So, we can find $u \in \Omega$ such that $\varphi(u) = 0$, a contradiction. □

3.2 The Leray–Schauder Degree

In most applications where we want to use degree theoretic techniques, the ambient space is infinite-dimensional. So, we need to extend Brouwer's degree to continuous functions defined on an infinite-dimensional Banach space. The next example illustrates that this is not possible.

Example 3.2.1 We consider the Hilbert space l^2 and let $\overline{B}_1$ be the closed unit ball of l^2. We define $\varphi : l^2 \to l^2$ by

$$\varphi(u) = \left(\sqrt{1 - ||u||^2}, u_1, u_2, \ldots, u_n, \ldots\right) \text{ for all } u = (u_k)_{k\geqslant 1} \subseteq l^2.$$

Evidently, φ is continuous and $\varphi(\overline{B}_1) \subseteq \overline{B}_1$. If Brouwer's degree admitted an infinite-dimensional extension, then φ would have had a fixed point. So, suppose

$$\varphi(\hat{u}) = \hat{u} \text{ for some } \hat{u} = (\hat{u}_k)_{k\geqslant 1} \in l^2.$$

Then for $n \geqslant 1$, $\hat{u}_{n+1} = \hat{u}_n$ and $\hat{u}_1 = \sqrt{1 - ||u||^2}$. Note that $1 = ||\varphi(\hat{u})|| = ||\hat{u}||$ hence $\hat{u}_1 = 0$ and so $\hat{u}_n = 0$ for all $n \geqslant 1$. Therefore $\hat{u} = 0$, a contradiction to the fact that $||\hat{u}|| = 1$.

However, we can have a degree function exhibiting all the main properties listed in Theorem 3.1.25 if we limit ourselves to maps of the form.

$$\varphi = i - f,$$

with i the identity map and f a compact map (see Definition 2.1.1(a)). This family of maps is a reasonable candidate for the extension of Brouwer's theory, due to the possibility of approximating the compact map f by finite rank maps (see Theorem 2.1.7).

The Leray–Schauder degree will be defined on triples (φ, Ω, ξ), where

$$\begin{aligned} &\varphi = i - f \text{ with } f : \overline{\Omega} \to X \text{ compact,} \\ &\Omega \subseteq X \text{ bounded open and } \xi \in X \text{ such that } \xi \notin \varphi(\partial\Omega). \end{aligned}$$

Note that since X is an infinite-dimensional Banach space, $\overline{\Omega}$ is never compact. So, if we set $r = d(\xi, \varphi(\partial\Omega))$, then it is not immediately clear that $r > 0$. Indeed, suppose that $r = 0$. Then we can find $\{u_n\}_{n\geqslant 1} \subseteq \partial\Omega$ such that $\varphi(u_n) \to \xi$ in X. Since f is compact, we have that $\{f(u_n)\}_{n\geqslant 1} \subseteq X$ is relatively compact (see Definition 2.1.1(a)). So, by the Eberlein–Smulian theorem and by passing to a subsequence if necessary, we may assume that

$$f(u_n) \to y \text{ in } X.$$

We have $y \in f(\overline{\Omega})$ and

$$u_n = f(u_n) + \varphi(u_n) \to y + \xi \text{ in } X \text{ as } n \to \infty. \tag{3.21}$$

Since $\{u_n\}_{n\geqslant 1} \subseteq \partial\Omega$ and the latter is closed, we have $y + \xi \in \partial\Omega$. The continuity of f implies

$$\begin{aligned} &y = \lim_{n\to\infty} f(u_n) = f(y + \xi) \text{ (see (3.21))} \\ \Rightarrow\ &\varphi(y + \xi) = \xi, \text{ hence } \xi \in \varphi(\partial\Omega), \text{ a contradiction.} \end{aligned}$$

Now let $\epsilon \in (0, r)$. By virtue of Theorem 2.1.7, we can find a finite rank map $f_\epsilon : \overline{\Omega} \to X$ such that

$$||f(x) - f_\epsilon(x)|| < \epsilon \text{ for all } x \in \overline{\Omega}.$$

Let $X_\epsilon = \text{span}\,\{f_\epsilon(\overline{\Omega}), \xi\}$, $\Omega_\epsilon = \Omega \cap X_\epsilon$ and $\varphi_\epsilon(u) = u - f_\epsilon(u)$ for all $u \in \overline{\Omega}$. Then $\Omega_\epsilon \subseteq X_\epsilon$ is bounded, open and $\partial_{X_\epsilon}\Omega_\epsilon \subseteq \partial\Omega$. Evidently $\varphi_\epsilon(\overline{\Omega}_\epsilon) \subseteq X_\epsilon$ and for $u \in \partial\Omega$ we have

$$||u - f_\epsilon(u) - \xi|| \geqslant ||u - f(u) - \xi|| - ||f(u) - f_\epsilon(u)|| > r - \epsilon > 0.$$

Therefore $d(\varphi_\epsilon, \Omega_\epsilon, \xi)$ is well-defined (if $\Omega_\epsilon = \emptyset$, then $d(\varphi_\epsilon, \Omega_\epsilon, \xi) = 0$).

Lemma 3.2.2 *For* $\epsilon \in (0, r)$, $d(\varphi_\epsilon, \Omega_\epsilon, \xi)$ *is independent of* ϵ.

Proof Let $\epsilon, \delta \in (0, r)$. Let $\widehat{X} = \mathrm{span}\{X_\epsilon, X_\delta\}$ and set $\widehat{\Omega} = \Omega \cap \widehat{X}$. Proposition 3.1.43 implies

$$d(\varphi_\epsilon, \Omega_\epsilon, \xi) = d(\varphi_\epsilon, \widehat{\Omega}, \xi) \text{ and } d(\varphi_\delta, \Omega_\delta, \xi) = d(\varphi_\delta, \widehat{\Omega}, \xi). \tag{3.22}$$

We consider the homotopy

$$h(t, u) = t\varphi_\epsilon(u) + (1 - t)\varphi_\delta(u) \text{ for all } (t, u) \in [0, 1] \times \overline{\widehat{\Omega}}.$$

We have

$$\begin{aligned} ||h(t, u) - \varphi(u)|| &\leqslant t||\varphi_\epsilon(u) - \varphi(u)|| + (1 - t)||\varphi_\delta(u) - \varphi(u)|| \\ &< t\epsilon + (1 - t)\delta < r. \end{aligned} \tag{3.23}$$

So, for $u \in \partial\widehat{\Omega}$ we have

$$||h(t, u) - \xi|| \geqslant ||\varphi(u) - \xi|| - ||h(t, u) - \varphi(u)|| > r - r = 0 \text{ (see (3.23))}.$$

By the homotopy invariance property of Brouwer's degree, we have

$$d(\varphi_\epsilon, \Omega_\epsilon, \xi) = d(\varphi_\delta, \Omega_\delta, \xi). \tag{3.24}$$

Then the result follows from (3.22) and (3.24). ☐

For any finite-dimensional subspace Y of X such that $X_\epsilon \subseteq Y$ $(0 < \epsilon < r)$, we let $\Omega_Y = \Omega \cap Y$ and from Proposition 3.1.43, we have

$$d(\varphi_\epsilon, \Omega_Y, \xi) = d(\varphi_\epsilon, \Omega, \xi).$$

So, we are led to the following definition.

Definition 3.2.3 Let $\Omega \subseteq X$ be bounded open, $\varphi = i - f$ with $f : \overline{\Omega} \to X$ compact and $\xi \notin \varphi(\partial\Omega)$. Let $\hat{f} : \overline{\Omega} \to X$ be a finite rank map such that

$$||f(u) - \hat{f}(u)|| < d(u, \varphi(\partial\Omega)) \text{ for all } u \in \overline{\Omega}.$$

Choose Y a finite-dimensional subspace of X containing $\hat{f}(\overline{\Omega})$ and ξ. We set $\Omega_Y = \Omega \cap Y$ and then define the Leray–Schauder degree of (φ, Ω, ξ) to be

$$d_{LS}(\varphi, \Omega, \xi) = d(\widehat{\varphi}, \Omega_Y, \xi),$$

where $\widehat{\varphi} = i - \hat{f}$.

Remark 3.2.4 A careful reading of the above definition reveals that $\Omega \subseteq X$ need not be bounded. It is enough to assume that for every finite-dimensional subspace Y of X, $\Omega \cap Y$ is bounded. Such sets are usually called "finitely bounded".

Then the properties of the Leray–Schauder degree can be derived from the above definition and the corresponding properties of the Brouwer degree.

In what follows, X is a Banach space, $\Omega \subseteq X$ is bounded open, $\varphi = i - f$ with i being the identity map on X and $f : \overline{\Omega} \to X$ is a compact map and $\xi \notin \varphi(\partial\Omega)$.

Proposition 3.2.5 *(a) $d_{LS}(i, \Omega, \xi) = 1$ for all $\xi \in \Omega$. (b) $d_{LS}(i, \Omega, \xi) = 0$ for all $\xi \notin \overline{\Omega}$.*

Proof (a) Let $f_\epsilon(u) = 0$ for all $u \in \overline{\Omega}$, $X_\epsilon = \text{span}\,\{\xi\}$ and $\Omega_\epsilon = \Omega \cap X_\epsilon$. Then according to Definition 3.2.3, we have

$$d_{LS}(i, \Omega, \xi) = d(i, \Omega_\epsilon, \xi). \tag{3.25}$$

Since $\xi \in \Omega$, we have $\xi \in \Omega_\epsilon$ and then Corollary 3.1.21 implies that

$$\begin{aligned} &d_{LS}(i, \Omega_\epsilon, \xi) = 1 \\ \Rightarrow\ &d_{LS}(i, \Omega, \xi) = 1 \text{ (see (3.25)).} \end{aligned}$$

(b) Similarly if $\xi \notin \overline{\Omega}$, using Remark 3.1.20. □

Proposition 3.2.6 *If $d_{LS}(\varphi, \Omega, \xi) \neq 0$, then there exists a $\hat{u} \in \Omega$ such that $\varphi(\hat{u}) = \xi$.*

Proof For every $n > \frac{1}{d(\xi, \varphi(\partial\Omega))} > 0$, we can find a finite rank map $f_n : \overline{\Omega} \to X$ such that

$$||f_n(x) - f(x)|| < \frac{1}{n} \text{ for all } x \in \overline{\Omega}.$$

Let $X_n = \text{span}\,\{\varphi_n(\overline{\Omega}), \xi\}$. Then $\dim X_n < \infty$ and $\Omega_n = \Omega \cap X_n$, and from Definition 3.2.3 we have

$$0 \neq d_{LS}(\varphi, \Omega, \xi) = d(i - f_n, \Omega_n, \xi) \text{ for every } n.$$

Theorem 3.1.25(a) implies that for every n we can find $u_n \in \Omega_n$ such that

$$u_n - f_n(u_n) = \xi. \tag{3.26}$$

Since $u_n \in \overline{\Omega}$ and $f : \overline{\Omega} \to X$ is compact, by passing to a subsequence if necessary we may assume that $f(u_n) \to y$ in X as $n \to \infty$. Then from (3.26), we have

$$u_n = \xi + f_n(u_n) \to \xi + y \text{ in } X \text{ as } n \to \infty$$
$$\Rightarrow \varphi(u_n) \to \varphi(\xi + y) = \xi \text{ (see (3.26))}.$$

Since $\xi \notin \varphi(\partial\Omega)$, we have $\xi + y \in \Omega$. So, the equation $\varphi(u) = \xi$, has a solution $\hat{u} = \xi + y$. □

To state the homotopy invariance property of the Leray–Schauder degree, we need to introduce the family of admissible homotopies.

Definition 3.2.7 Let $D \subseteq X$ and suppose that $h : [0, 1] \to K(D, X) = \{$family of compact maps from D into $X\}$ (see Definition 2.1.1(a)). We say that h is a "homotopy of compact maps on D" if given $\epsilon > 0$ and $B \subseteq D$ bounded, we can find $\delta = \delta(\epsilon, B) > 0$ such that

$$||h(t)(u) - h(s)(u)|| < \epsilon \text{ for all } u \in B \text{ and all } |t - s| < \delta.$$

Proposition 3.2.8 *If* $\{h(t)\}_{t\in[0,1]}$ *is a homotopy of compact maps on* $\overline{\Omega}$, $\varphi_t = i - h(t)$ *for all* $t \in [0, 1]$ *and* $\xi \notin \varphi_t(\partial\Omega)$ *for all* $t \in [0, 1]$, *then* $d_{LS}(\varphi_t, \Omega, \xi)$ *is independent of* $t \in [0, 1]$.

Proof We claim that there exists an $r > 0$ such that

$$||\varphi_t(u) - \xi|| \geqslant r \text{ for all } u \in \partial\Omega \text{ and all } t \in [0, 1]. \tag{3.27}$$

Arguing by contradiction, suppose that (3.27) is not true. Then we can find $\{u_n\}_{n\geqslant 1} \subseteq \partial\Omega$ and $\{t_n\}_{n\geqslant 1} \subseteq [0, 1]$ such that

$$||\varphi_{t_n}(u_n) - \xi|| < \frac{1}{n} \text{ for all } n \geqslant 1. \tag{3.28}$$

We may assume that $t_n \to t \in [0, 1]$. Also, since $h(t) \in K(\overline{\Omega}, X)$, we may assume that

$$h(t)(u_n) \to y \text{ in } X \text{ as } n \to \infty. \tag{3.29}$$

From Definition 3.2.7, we have

$$||h(t)(u_n) - h(t_n)(u_n)|| \to 0 \text{ as } n \to \infty. \tag{3.30}$$

Then

$$\begin{aligned} &||h(t_n)(u_n) - y|| \leqslant ||h(t_n)(u_n) - h(t)(u_n)|| + ||h(t)(u_n) - y|| \to 0 \\ &\qquad\qquad \text{as } n \to \infty \text{ (see (3.29) and (3.30))} \\ &\Rightarrow h_{t_n}(u_n) \to y \text{ as } n \to \infty \qquad (3.31) \\ &\Rightarrow u_n = \varphi_{t_n}(u_n) - \xi + h(t_n)(u_n) + \xi \to y + \xi \text{ (see (3.28))}. \end{aligned}$$

Since $\{u_n\}_{n\geqslant 1} \subseteq \partial\Omega$, we have that $y + \xi \in \partial\Omega$ and

$$\varphi_t(y+\xi) = y + \xi - \lim_{n\to\infty} h(t_n)(u_n) = \xi \text{ (see (3.31))}$$
$$\Rightarrow \xi \in \varphi_t(\partial\Omega), \text{ a contradiction.}$$

This proves (3.27).
On [0, 1] we introduce the relation

$$t \sim s \text{ if and only if } d_{LS}(\varphi_t, \Omega, \xi) = d_{LS}(\varphi_s, \Omega, \xi).$$

Evidently, $\sim$ is an equivalence relation on [0, 1]. We show that the equivalence classes are open sets in [0, 1]. To this end, let $\tau \in [0, 1]$ and $\epsilon \in (0, \frac{r}{4})$ (see (3.27)). We can find $h_\epsilon(\tau) \in K_f(\overline{\Omega}, X)$ such that

$$||h_\epsilon(\tau)(u) - h(\tau)(u)|| < \frac{r}{4} \text{ for all } u \in \overline{\Omega}. \tag{3.32}$$

We can find $\delta > 0$ such that

$$|t-\tau| < \delta \Rightarrow ||h(t)(u) - h(\tau)(u)|| < \frac{r}{4} \text{ for all } u \in \overline{\Omega} \tag{3.33}$$
$$\text{(see Definition 3.2.7).}$$

From (3.32) and (3.33) it follows that

$$||h(t)(u) - h_\epsilon(\tau)(u)|| < \frac{r}{2} \text{ for } |t-\tau| < \delta, \text{ all } u \in \overline{\Omega}$$
$$\Rightarrow d_{LS}(\varphi_t, \Omega, \xi) = d(i - h_\epsilon(\tau), \Omega \cap V, \xi) = d(\varphi_\tau, \Omega, \xi),$$

where V is a finite-dimensional subspace of X such that $h_\epsilon(\tau)(\overline{\Omega}) \subseteq V$ (see Definition 3.2.3). Therefore $t \sim \tau$ if $|t-\tau| < \delta$ and so the equivalence classes are open sets. This means that there is only one equivalence class, the whole interval [0, 1], and so

$$d_{LS}(\varphi_t, \Omega, \xi) \text{ is independent of } t \in [0, 1].$$

The proof is now complete. □

Proposition 3.2.9 *If $\varphi = i - f$, $\psi = i - g$ with i the identity map on X and $f, g \in K(\overline{\Omega}, X)$ such that*

$$f|_{\partial\Omega} = g|_{\partial\Omega} \text{ and } \xi \notin \varphi(\partial\Omega),$$

then $d_{LS}(\varphi, \Omega, \xi) = d_{LS}(\psi, \Omega, \xi)$.

Proof Let $h(t, u) = (1-t)f(u) + tg(u)$ for all $(t, u) \in [0, 1] \times \overline{\Omega}$. Evidently, $\{h(t, \cdot)\}_{t\in[0,1]}$ is a homotopy of compact maps on $\overline{\Omega}$ and $\xi \notin (i - h(t, \cdot))(\partial\Omega)$ for all $t \in [0, 1]$. So, Proposition 3.2.8 implies that

$$d_{LS}(\varphi, \Omega, \xi) = d_{LS}(\psi, \Omega, \xi).$$

The proof is now complete. □

Proposition 3.2.10 *If $\varphi = i - f$ with i the identity map on X, $f \in K(\overline{\Omega}, X)$, $\xi \notin \varphi(\partial\Omega)$ and $\eta \in X$, then $d_{LS}(\varphi, \Omega, \xi) = d_{LS}(\varphi - \eta, \Omega, \xi - \eta)$.*

Proof Note that $\varphi - \eta = i - (f + \eta)$ and $f + \eta \in K(\overline{\Omega}, X)$. Let $g \in K_f(\overline{\Omega}, X)$ such that if $\psi = i - g$ then

$$d_{LS}(\varphi, \Omega, \xi) = d(\psi, \Omega \cap V, \xi), \tag{3.34}$$

where $V \subseteq X$ is a finite-dimensional linear subspace such that $g(\overline{\Omega}) \subseteq V$ (see Definition 3.2.3). Let $g_1 = g + \eta \in K_f(\overline{\Omega}, X)$. Then

$$\begin{aligned}
&||g_1(u) - (f + \eta)(u)|| < r = d(\xi - \eta, (i - (f + \eta))(\partial\Omega)) \\
\Rightarrow\ & d_{LS}(\varphi - \eta, \Omega, \xi - \eta) = d(\psi - \eta, \Omega \cap V, \xi - \eta) = d(\psi, \Omega \cap V, \xi) \\
&\qquad\qquad\qquad\qquad \text{(see Proposition 3.1.12)} \\
&= d_{LS}(\varphi, \Omega, \xi) \text{ (see (3.34))}.
\end{aligned}$$

The proof is now complete. □

Proposition 3.2.11 *If $\varphi = i - f$, $\psi = i - g$ with i the identity map on X, $f, g \in K(\overline{\Omega}, X)$, $\xi \notin \varphi(\partial\Omega)$, and $||\varphi(u) - \psi(u)|| < d(\xi, \varphi(\partial\Omega)) = r$ for all $u \in \overline{\Omega}$, then $\xi \notin \psi(\partial\Omega)$ and $d_{LS}(\varphi, \Omega, \xi) = d_{LS}(\psi, \Omega, \xi)$.*

Proof Let $h(t, u) = (1 - t)f(u) + tg(u)$ for all $(t, u) \in [0, 1] \times \overline{\Omega}$. Evidently, $\{h(t, \cdot)\}_{t \in [0,1]}$ is a homotopy of compact transformations (see Definition 3.2.7). Let $\sigma_t(u) = u - h(t, u)$ for all $(t, u) \in [0, 1] \times \overline{\Omega}$. For every $u \in \partial\Omega$ we have

$$\begin{aligned}
&||\xi - \sigma_t(u)|| \geqslant ||\xi - \varphi(u)|| - t||\varphi(u) - \psi(u)|| > \\
&||\xi - \varphi(u)|| - tr \geqslant r - tr = (1 - t)r \\
\Rightarrow\ & \xi \notin \sigma_t(\partial\Omega) \text{ for all } t \in [0, 1] \\
\Rightarrow\ & d_{LS}(\varphi, \Omega, \xi) = d_{LS}(\psi, \Omega, \xi) \text{ (see Proposition 3.2.8)}.
\end{aligned}$$

The proof is now complete. □

Proposition 3.2.12 *If $\varphi = i - f$ with i the identity map on X and $f \in K(\overline{\Omega}, X)$, then $d_{LS}(\varphi, \Omega, \cdot)$ is constant on every connected component of $X \backslash \varphi(\partial\Omega)$.*

Proof Let U_X be a connected component of $X \backslash \varphi(\partial\Omega)$ and consider the map $\gamma : u \to \varphi(\partial\Omega)$. It suffices to show that γ is continuous on U.

Let $\xi \in U$ and let $r = d(\xi, \varphi(\partial\Omega) > 0)$. For $\eta \in U$, let $\varphi_\eta : \overline{\Omega} \to X$ be defined by

$$\varphi_\eta(u) = \varphi(u) - (\eta - \xi) \text{ for all } u \in \overline{\Omega}.$$

Proposition 3.2.10 implies that

$$d_{LS}(\varphi, \Omega, \eta) = d_{LS}(\varphi - (\eta - \xi), \Omega, \eta - (\eta - \xi)) = d(\varphi_\eta, \Omega, \xi). \quad (3.35)$$

If $||\xi - \eta|| < r$, then $||\varphi(u) - \varphi_\eta(u)|| < r$. Hence Proposition 3.2.11 implies that

$$\begin{aligned} &d_{LS}(\varphi_\eta, \Omega, \xi) = d_{LS}(\varphi, \Omega, \xi) \\ \Rightarrow\ &d_{LS}(\varphi, \Omega, \eta) = d_{LS}(\varphi, \Omega, \xi) \text{ (see (3.35))}, \end{aligned}$$

which completes the proof. $\square$

Proposition 3.2.13 *If $\Omega_1, \Omega_2 \subseteq X$ are bounded, open, $\Omega \cap \Omega_2 = \emptyset$, $\varphi = i - f$ with i the identity map on X, $f \in K(\overline{\Omega}_1 \cup \overline{\Omega}_2, X)$ and $\xi \notin \varphi(\partial\Omega_1 \cup \partial\Omega_2)$, then $d_{LS}(\varphi, \Omega_1 \cup \Omega_2, \xi) = d_{LS}(\varphi, \Omega_1, \xi) + d_{LS}(\varphi, \Omega_2, \xi)$.*

Proof Let $g \in K_f(\overline{\Omega}_1 \cup \overline{\Omega}_2, X)$ such that

$$||g(u) - f(u)|| < d(\xi, \varphi(\partial\Omega_1 \cup \partial\Omega_2)) \text{ for all } u \in \overline{\Omega}.$$

Then for $\psi = i - g$, we have

$$d_{LS}(\varphi, \Omega_1 \cup \Omega_2, \xi) = d(\psi, (\Omega_1 \cup \Omega_2) \cap V, \xi)$$

with $V \subseteq X$ a finite-dimensional subspace of X such that $g(\overline{\Omega}_1 \cup \overline{\Omega}_2) \subseteq V$ (see Definition 3.2.7). But from Theorem 3.1.25(b), we have

$$\begin{aligned} &d(\psi, (\Omega_1 \cup \Omega_2) \cap V, \xi) = d(\psi, \Omega_1 \cap V, \xi) + d(\psi, \Omega_2 \cap V, \xi) \\ \Rightarrow\ &d_{LS}(\varphi, \Omega_1 \cup \Omega, \xi) = d_{LS}(\varphi, \Omega_1, \xi) + d_{LS}(\varphi, \Omega_2, \xi). \end{aligned}$$

The proof is now complete. $\square$

In a similar fashion, we prove the excision property for the Leray–Schauder degree.

Proposition 3.2.14 *If $C \subseteq X$ is compact, $\varphi = i - f$ with i the identity map on X and $f \in K(\overline{\Omega}, X)$ and $\xi \notin \varphi(C)$, then $d_{LS}(\varphi, \Omega, \xi) = d_{LS}(\varphi, \Omega \backslash C, \xi)$.*

So, we can state the following theorem summarizing the main properties of the Leray–Schauder degree.

Theorem 3.2.15 *If $\tau_{LS} = \{(\varphi, \Omega, \xi) : \Omega \subseteq X$ bounded open, $\varphi = i - f$ with i the identity map on X and $f \in K(\overline{\Omega}, X)$ and $\xi \notin \varphi(\partial\Omega)\}$, then there exists a map $d_{LS} : \tau_{LS} \to \mathbb{Z}$, known as the Leray–Schauder degree, such that the following properties hold:*

(a) Normalization: $d_{LS}(i, \Omega, \xi) = 1$ provided $\xi \in \overline{\Omega}$.
(b) Domain Additivity: $d_{LS}(\varphi, \Omega, \xi) = d_{LS}(\varphi, \Omega_1, \xi) + d_{LS}(\varphi, \Omega_2, \xi)$ with Ω_1, Ω_2 disjoint open subsets of Ω and $\xi \in \varphi(\overline{\Omega} \backslash (\Omega_1 \cup \Omega_2))$.

(c) *Homotopy Invariance:* $d_{LS}(i - h(t,\cdot), \Omega, \xi(t))$ *is independent of* $t \in [0,1]$ *when* $\{h(t,\cdot)\}_{t\in[0,1]}$ *is a homotopy of compact maps,* $0 \notin (i - h(t,\cdot))(\partial\Omega)$*, and* $\xi \in C([0,1], X)$.

(d) *Solution Property:* $d_{LS}(\varphi, \Omega, \xi) \neq 0$ *implies* $\varphi^{-1}(\xi) \neq \emptyset$.

(e) *Continuity in* (φ, ξ)*:* $d(\varphi, \Omega, \xi) = d(\psi, \Omega, \xi)$ *for all* $\varphi = i - f$, $\psi = i - g$ *with* i *the identity map on* X, $f, g \in K(\overline{\Omega}, X)$ *with*

$$||f(u) - g(u)|| < d(\xi, \varphi(\partial\Omega)) \text{ and } \xi \notin \psi(\partial\Omega);$$

also $d_{LS}(\varphi, \Omega, \cdot)$ *is constant in every connected component of* $X\setminus\varphi(\partial\Omega)$.

(f) *Dependence on Boundary Values:* $d_{LS}(\varphi, \Omega, \xi) = d_{LS}(\psi, \Omega, \xi)$ *where* $\varphi, \psi \in \tau_{LS}$ *with* $\varphi = i - f$, $\psi = i - g$ *and* $\varphi|_{\partial\Omega} = \psi|_{\partial\Omega}$.

(g) *Excision Property:* $d_{LS}(\varphi, \Omega, \xi) = d_{LS}(\varphi, \Omega_1, \xi)$ *for every open set* $\Omega_1 \subseteq \Omega$ *such that* $\xi \notin \varphi(\Omega\setminus\Omega_1)$.

Borsuk's theorem (see Theorem 3.1.44) remains valid for the Leray–Schauder degree.

Theorem 3.2.16 (Borsuk) *If* $\Omega \subseteq X$ *is bounded, open, symmetric with* $0 \in \Omega$, $\varphi = i - f$ *with* i *the identity map on* X, $f \in K(\overline{\Omega}, X)$ *and it is odd on* $\partial\Omega$ *and* $0 \notin \varphi(\partial\Omega)$ *then* $d_{LS}(\varphi, \Omega, 0)$ *is odd.*

Proof Evidently by considering the odd part of φ (see the beginning of the proof of Theorem 3.1.44), we may assume that φ is odd on $\overline{\Omega}$. Then $K = \overline{f(\overline{\Omega})}$ is symmetric and compact. We can find a finite-dimensional subspace V of X and $g \in C(K, V)$ such that

$$||u - g(u)|| \leqslant \frac{r}{2} \text{ for all } u \in X \text{ with } r = d(0, \varphi(\partial\Omega)).$$

Let $g_0(u) = \frac{1}{2}[g(u) - g(-u)]$ for all $u \in K$. Then

$$||u - g_0(u)|| \leqslant \frac{r}{2} \text{ for all } u \in K.$$

If we set $\varphi_1 = i - g_0 \circ f$, then φ_1 is odd and

$$\begin{aligned} & d(\varphi, \Omega, 0) = d(\varphi_1|_{\overline{\Omega}\cap V}, \Omega \cap V, 0) \neq 0 \text{ (see Proposition 3.1.43)} \\ \Rightarrow\ & d(\varphi, \Omega, 0) \text{ is odd.} \end{aligned}$$

The proof is now complete. □

Then, a straightforward modification of the proof of Theorem 3.1.48 leads to the following infinite-dimensional version of the "invariance of domain theorem".

Theorem 3.2.17 *If* $\Omega \subseteq X$ *is open,* $f : \Omega \to X$ *is compact and* $\varphi = i - f$ *is locally one-to-one, then* $i - f$ *is open.*

Corollary 3.2.18 *If $\Omega \subseteq X$ is bounded open, $\varphi = i - f$ with i the identity map on X and $f \in K(\overline{\Omega}, X)$, φ is one-to-one and $\xi \in \varphi(\Omega)$, then $d_{LS}(\varphi, \Omega, \xi) = \pm 1$.*

Now let us see some topological applications of the Leray–Schauder degree. In Proposition 3.1.32 we saw that in a finite-dimensional space the boundary of the unit ball is not a retract of the unit ball. We also mentioned (see Remark 3.1.33) that in contrast in an infinite-dimensional Banach space X, $\partial B_1 = \{x \in X : ||x|| = 1\}$ is always a retract of $B_1 = \{x \in X : ||x|| < 1\}$. However, if we limit ourselves to compact perturbations of the identity, this is no longer true.

Proposition 3.2.19 *There is no $\varphi \in C(\overline{B}_1, \partial B_1)$ of the form $\varphi = i - f$ with i the identity map of X and $f \in K(\overline{B}_1, X)$ such that $\varphi|_{\partial B_1} = i|_{\partial B_1}$.*

Proof If such a φ exists, then $d(\varphi, B_1, 0) = 1$ (see Theorem 3.2.15(a) and (f)). Then by the solution property (see Theorem 3.2.15(d)), we can find $\hat{u} \in B_1$ such that $\varphi(\hat{u}) = 0$. But by hypothesis φ has values in ∂B_1, a contradiction. □

Brouwer's fixed point theorem (see Theorem 3.1.34 and Remark 3.1.35) has an infinite-dimensional analogue, known as "Schauder's fixed point theorem".

Theorem 3.2.20 *(Schauder) If X is Banach space, $C \subseteq X$ is nonempty bounded, closed, convex and $f : C \to C$ is compact, then there exists a $\hat{u} \in C$ such that $f(\hat{u}) = \hat{u}$.*

Proof Since C is bounded, we can find a large $\rho > 0$ such that $C \subseteq B_\rho = \{u \in X : ||u|| < \rho\}$. By virtue of Proposition 2.1.10, we can find a compact extension $\hat{f} : \overline{B}_\rho \to C \subseteq \overline{B}_\rho$ of f (that is, $\hat{f}|_C = f$). From Proposition 2.1.9, we know that $\overline{B}_\rho$ is a retract of X. Let $r : X \to \overline{B}_\rho$ be a corresponding retraction and define $h(t, u) = t\hat{f}(r(u))$ for all $(t, u) \in [0, 1] \times X$. Then $\{h(t, \cdot)\}_{t \in [0,1]}$ is a homotopy of compact maps. We may assume that $\hat{f}(u) \neq u$ for all $u \in \partial B_\rho$, or otherwise there is nothing more to prove. It follows that $h(t, u) \neq u$ for all $(t, u) \in [0, 1] \times \partial B_\rho$. Then the homotopy invariance property of the Leray–Schauder degree (see Theorem 3.2.15(c)) implies that $d(i - \hat{f}, B_\rho, 0) = 1$ and so we can find $\hat{u} \in B_\rho$ such that

$$\hat{u} = \hat{f}(\hat{u}) \in C.$$

The proof is now complete. □

Remark 3.2.21 A careful reading of the above proof reveals that the above fixed point theorem remains valid if C is only homeomorphic to a bounded, closed and convex subset of X. We can give a proof of the theorem which avoids the use of degree theory and instead uses Theorem 2.1.7 (the approximation of f by finite rank maps). So, for every $n \geqslant 1$, we can find a finite rank map $f_n : C \to C$ such that $||f_n(u) - f(u)|| < \frac{1}{n}$ for all $u \in C$ (see Theorem 2.1.7). By the Brouwer fixed point theorem (see Theorem 3.1.34 and Remark 3.1.35), we can find $u_n \in C$ such that $f_n(u_n) = u_n$. Due to compactness, we may assume that $f_n(u_n) \to \hat{u}$ in X. Then $||f_n(\hat{u}_n) - f(u_n)|| = ||u_n - f(u_n)|| \leqslant \frac{1}{n}$ for all $n \geqslant 1$ and so $f(\hat{u}) = \hat{u}$.

The next result, known as the "Schaefer fixed point theorem" (or sometimes as the Leray–Schauder alternative theorem), is useful in establishing the existence of solutions for nonlinear boundary value problems. Roughly speaking, it says that a priori estimates lead to the existence of solutions.

Proposition 3.2.22 *If $f : X \to X$ is compact and*

$$D_f = \{u \in X : \textit{there exists a } t \in (0, 1) \textit{ such that } u = tf(u)\},$$

then the following alternative holds:

(a) *D_f is unbounded*
or
(b) *f has a fixed point.*

Proof If D_f is bounded, then $D_f \subseteq B_\rho$ for $\rho > 0$ large. Then $tf|_{\overline{B}_r} : \overline{B}_r \to X$ $t \in (0, 1)$ is a compact map with no fixed points on ∂B_ρ. The homotopy invariance property of the Leray–Schauder degree (see Proposition 3.1.15(c)) implies

$$d_{LS}(i - f, B_\rho, 0) = 1.$$

So, f must have a fixed point (see Theorem 3.2.15(d)). □

We present two more fixed point results involving compact maps.

Proposition 3.2.23 *If $\Omega \subseteq X$ is bounded, open convex, $f : \overline{\Omega} \to X$ is compact and*

$$f(\partial\Omega) \subseteq \overline{\Omega},$$

then f has a fixed point.

Proof If f has a fixed point on $\partial\Omega$, then there is nothing to prove. So, we assume that $0 \notin (i - f)(\partial\Omega)$ with i being the identity map on X. Fix $u_0 \in \Omega$ and let

$$h(t, u) = tf(u) + (1 - t)u_0 \text{ for all } (t, u) \in [0, 1] \times \overline{\Omega}.$$

We claim that $h(t, u) \neq u$ for all $t \in [0, 1]$ and all $u \in \partial\Omega$. Indeed, if we can find $t \in [0, 1]$ and $u \in \partial\Omega$ such that $h(t, u) = u$, then

$$\begin{aligned} tf(u) &= tu_0 + (u - u_0) \\ \Rightarrow t &\neq 0, t \neq 1. \end{aligned}$$

Then we have

$$f(u) = u_0 + \frac{1}{t}(u - u_0) \notin \overline{\Omega} \text{ since } \frac{1}{t} > 1,$$

a contradiction to the hypothesis that $f(\partial\Omega) \subseteq \overline{\Omega}$. Since $\{h(t,\cdot)\}_{t\in[0,1]}$ is a homotopy of compact maps, from the homotopy invariance property of the Leray–Schauder degree (see Proposition 3.1.15(c)), we have

$$d_{LS}(i - f, \Omega, 0) = d_{LS}(i + u_0, \Omega, 0) = d_{LS}(i, \Omega, u_0) = 1.$$

So, there exists a $\hat{u} \in \Omega$ such that $\hat{u} = f(\hat{u})$. □

When $X = H =$ a Hilbert space, we have the following variant of the above fixed point theorem.

Proposition 3.2.24 *If H is a Hilbert space, $\Omega \subseteq H$ is bounded open with $0 \notin \partial\Omega$, $f : \overline{\Omega} \to H$ is compact and $(f(u), u)_H \leqslant ||u||^2$ for all $u \in \partial\Omega$ (by $(\cdot,\cdot)_H$ we denote the inner product of H), then f has a fixed point.*

Proof Arguing by contradiction, suppose that f has no fixed point on $\overline{\Omega}$. Then

$$d_{LS}(i - f, \Omega, 0) = 0$$

with i being the identity map on H. Hence, we can find $t_0 \in (0, 1]$ and $u_0 \in \partial\Omega$ such that $u_0 = t_0 f(u_0)$. Then $||u_0||^2 = t_0(f(u_0), u_0)_H \leqslant t_0||u_0||^2$, a contradiction. □

Remark 3.2.25 The hypothesis in the above fixed point result implies that u and $u - f(u)$ make an acute angle. For this reason, the result is sometimes called the "acute angle fixed point principle".

Extending this result to Banach spaces, we have:

Proposition 3.2.26 *If X is a Banach space, $\Omega \subseteq X$ is bounded open convex with $0 \in \Omega$, $f : \overline{\Omega} \to X$ is compact and*

$$||u - f(u)||^2 \geqslant ||f(u)||^2 - ||u||^2 \text{ for all } u \in \partial\Omega, \tag{3.36}$$

then f has a fixed point.

Proof Let $h(t,u) = tf(u)$ for all $t \in [0,1]$ and all $u \in \overline{\Omega}$. Then $\{h(t,\cdot)\}_{t\in[0,1]}$ is a homotopy of compact maps on $\overline{\Omega}$. Suppose that

$$h(t_0, u_0) = u_0 \text{ for some } t_0 \in [0,1] \text{ for some } u_0 \in \partial\Omega.$$

Evidently, $t_0 \neq 0$ and so $f(u_0) = \frac{1}{t_0}u_0$. Then

$$||f(u_0) - u_0||^2 = (1 - t_0)^2||f(u_0)||^2 = \frac{(1-t_0)^2}{t_0^2}||u_0||^2, \tag{3.37}$$

$$||f(u_0)^2|| - ||u_0||^2 = ||u_0||^2\left(\frac{1}{t_0^2} - 1\right) = ||u_0||^2\frac{1-t_0^2}{t_0^2}. \tag{3.38}$$

Using (3.37) and (3.38) in (3.36), we obtain

$$
\begin{aligned}
&||u_0||^2 \frac{(1-t_0)^2}{t_0^2} \geqslant ||u_0||^2 \frac{1-t_0^2}{t_0^2} \\
\Rightarrow\ &(1-t_0)^2 \geqslant 1-t_0^2 \\
\Rightarrow\ &t_0 \geqslant 1, \ \text{hence } t_0 = 1.
\end{aligned}
$$

Therefore $f(u_0) = u_0$ and so we have a fixed point.
If $h(t, u) \neq u$ for all $t \in [0, 1]$, all $u \in \partial\Omega$, then

$$
\begin{aligned}
&d_{LS}(i - f, \Omega, 0) = d_{LS}(i, \Omega, 0) = 1 \\
\Rightarrow\ &f \text{ has a fixed point.}
\end{aligned}
$$

The proof is now complete. □

Remark 3.2.27 Additional fixed point theorems of a topological nature will be proved in Sect. 4.2.

Finally, we mention a result on the degree of gradient maps, due to Amann [14], where the interested reader can find its proof.

So, suppose H is a Hilbert space, $U \subseteq H$ an open set, $\varphi \in C^1(U, \mathbb{R})$ and $\nabla\varphi = i - f$ with $f : \overline{U} \to H$ a compact map.

Proposition 3.2.28 *If, for some $\beta \in \mathbb{R}$, the set $V = \varphi^{-1}(-\infty, \beta)$ is bounded, $\overline{V} \subseteq U$ and there exist numbers $\alpha < \beta$ and $r > 0$ and a point $u_0 \in U$ such that*

$$\varphi^{-1}(-\infty, a] \subseteq \overline{B}_r(x_0) \subseteq V$$

and $\nabla\varphi(u) \neq 0$ for all $u \in \varphi^{-1}[\alpha, \beta]$, then $d_{LS}(\nabla\varphi, V, 0) = 1$.

3.3 Degree for Multifunctions

In this section, we extend the Leray–Schauder degree to multifunctions which are upper semicontinuous and compact and have nonempty, closed and convex values.

We will consider multifunctions of the form $i - F$ with i being the identity map on the ambient Banach space X and F being a multifunction which belongs to the following family:

Definition 3.3.1 Let X be a Banach space, $D \subseteq X$ and $F : D \to 2^X \setminus \{\emptyset\}$ a multifunction. We say that $F(\cdot)$ is "compact" if it has values in $P_{fc}(X) = \{C \subseteq X :$ nonempty, closed, convex$\}$, it is usc and maps bounded sets to relatively compact sets.

From Theorem 2.5.19, we know that a usc multifunction with nonempty and convex values admits an approximate continuous selection. Since this result will lead to the extension of the Leray–Schauder degree to multifunctions, we recall it here.

Proposition 3.3.2 *If Z is a metric space, X is a Banach space and $F : Z \to 2^X \backslash \{\emptyset\}$ is a usc multifunction with nonempty and convex values, then given $\epsilon > 0$, we can find a continuous map $f_\epsilon : Z \to X$ such that*

$$f_\epsilon(z) \in \operatorname{conv} F(Z) \textit{ for all } z \in Z$$

and for every $z \in Z$, there exist $(u, y) \in \operatorname{Gr} F$ such that

$$d_Z(z, u) < \epsilon \textit{ and } ||f_\epsilon(z) - y|| < \epsilon.$$

Then the degree of the multiplication $u \to u - F(u)$ will be defined in terms of these approximations. The next proposition is the crucial step in this direction.

Proposition 3.3.3 *If X is a Banach space, $\Omega \subseteq X$ is bounded open, $F : \overline{\Omega} \to P_{f_c}(X)$ is a compact multifunction and $\xi \notin (i - F)(\partial\Omega)$ with i the identity map on X, then there exists an $\epsilon_0 > 0$ such that $\xi \notin (i - f_\epsilon)(\partial\Omega)$ for all $\epsilon \in (0, \epsilon_0)$ with f_ϵ being the approximate continuous selection of F produced in Proposition 3.3.2.*

Proof We argue by contradiction. So, suppose that the proposition is not true. Then we can find $\epsilon_n \to 0^+$ and $\{u_n\}_{n \geqslant 1} \subseteq \partial\Omega$ such that $\xi = u_n - f_{\epsilon_n}(u_n)$ for all $n \geqslant 1$. From Proposition 3.3.2 we know that there exist $(v_n, y_n) \in \operatorname{Gr} F$, $n \geqslant 1$, such that

$$||u_n - v_n|| < \epsilon_n \text{ and } ||f_{\epsilon_n}(u_n) - y_n|| < \epsilon_n \text{ for all } n \geqslant 1. \tag{3.39}$$

Since F is compact (see Definition 3.3.1), by passing to a subsequence if necessary, we may assume that $y_n \to y$ in X. Then from (3.39) we have

$$f_{\epsilon_n}(u_n) \to y \text{ and } v_n \to \xi + y \in \partial\Omega \text{ in } X. \tag{3.40}$$

Recall that $(v_n, y_n) \in \operatorname{Gr} F$ for all $n \geqslant 1$. Proposition 2.5.8 and (3.40) imply

$$\begin{aligned} &(\xi + y, y) \in \operatorname{Gr} F \\ \Rightarrow\ & y \in F(\xi + y), \text{ hence } \xi \in (i - F)(\partial\Omega), \text{ a contradiction.} \end{aligned}$$

This proves the proposition. □

Evidently each f_ϵ is compact. So, using this proposition, we can extend the Leray–Schauder degree to multifunctions.

Definition 3.3.4 Let X be a Banach space, $\Omega \subseteq X$ a bounded open set, and $F : \overline{\Omega} \to P_{f_c}(X)$ a compact multifunction such that $\xi \notin (i - F)(\partial\Omega)$ with i the identity map on X. Then we define

$$\hat{d}_{LS}(i - F, \Omega, \xi) = \lim_{\epsilon \to 0^+} d_{LS}(i - f_\epsilon, \Omega, \xi)$$

with f_ϵ the single-valued compact map on $\overline{\Omega}$ defined as in Proposition 3.3.3.

Remark 3.3.5 First of all note that Proposition 3.3.2 guarantees that $d_{LS}(i - f_\epsilon, \Omega, \xi)$ is well-defined for all $\epsilon \in (0, \epsilon_0)$. Next, we show that $\{d_{LS}(i - f_\epsilon, \Omega, \xi)\}_{\epsilon \in (0, \epsilon_0)}$ evidently stabilizes. We consider the homotopy of compact maps

$$h_{\epsilon,\delta}(t, u) = (1 - t) f_\epsilon(u) + t f_\delta(u) \text{ for all } (t, u) \in [0, 1] \times \overline{\Omega}.$$

We claim that there exists an $\epsilon_1 \in (0, \epsilon_0]$ such that $\xi \notin (i - h_{\epsilon,\delta}(t, \cdot))(\partial\Omega)$ for all $t \in [0, 1]$ and all $\epsilon, \delta \in (0, \epsilon_1)$. Indeed, if this is not true, then we can find $\{t_n\}_{n \geqslant 1} \subseteq [0, 1]$, $\epsilon_n, \delta_n \to 0^+$ and $\{u_n\}_{n \geqslant 1} \subseteq \partial\Omega$ such that

$$\xi = u_n - (1 - t_n) f_{\epsilon_n}(u_n) - t_n f_{\epsilon_n}(u_n) \text{ for all } n \geqslant 1. \tag{3.41}$$

The compactness of $F(\cdot)$ implies that $\overline{\text{conv}}\, F(\overline{\Omega})$ is compact in X. Since $\{f_{\epsilon_n}(u_n)\}_{n \geqslant 1}$, $\{f_{\delta_n}(u_n)\}_{n \geqslant 1} \subseteq \overline{\text{conv}}\, F(\overline{\Omega})$ (see Proposition 3.3.2), by passing to a suitable subsequence if necessary, we may assume that

$$t_n \to t \in [0, 1] \text{ and } f_{\epsilon_n}(u_n) \to y_1, \ \ f_{\delta_n}(u_n) \to y_2 \text{ in } X. \tag{3.42}$$

So, from (3.41) and (3.42), it follows that

$$u_n \to \xi + (1 - t) y_1 + t y_2 = \hat{u} \in \partial\Omega.$$

Note that $y_1, y_2 \in \hat{u} - F(\hat{u}) \subseteq (i - F)(\partial\Omega)$ and so

$$\xi = (1 - t)(\hat{u} - y_1) + t(\hat{u} - y_2) \in (1 - t)(i - F)(\hat{u}) + t(i - F)(\hat{u}).$$

Since F has convex values, we infer that $\xi \in (i - F)(\partial\Omega)$, a contradiction. So, our initial claim is true and by virtue of the homotopy invariance property of the Leray–Schauder degree, we have

$$d_{LS}(i - f_\epsilon, \Omega, \xi) = d_{LS}(i - f_\delta, \Omega, \xi) \text{ for all } \epsilon, \delta \in (0, \epsilon_1).$$

This argument also shows that Definition 3.3.4 is in fact independent of the particular choice of the approximate continuous selection f_ϵ.

Using this definition, we can transfer the main properties of the Leray–Schauder degree to multifunctions. The proofs are straightforward and so they are omitted.

Theorem 3.3.6 *If* $\widehat{\tau}_{LS} = \{(i - F, \Omega, \xi) : \Omega \subseteq X$ *is bounded open,* i *is the identity map on* X, $F : \overline{\Omega} \to P_{f_{c}}(X)$ *is a compact multifunction and* $\xi \notin (i - F)(\partial\Omega)\big\}$, *then there exists a map* $\widehat{d}_{LS} : \widehat{\tau}_{LS} \to \mathbb{Z}$ *such that*

(a) *Normalization:* $\hat{d}_{LS}(i, \Omega, \xi) = 1$ *provided* $\xi \in \overline{\Omega}$.
(b) *Domain Additivity:* $\hat{d}_{LS}(i - F, \Omega, \xi) = \hat{d}_{LS}(i - F, \Omega_1, \xi) + \hat{d}_{LS}(i - F, \Omega_2, \xi)$ *with* Ω_1, Ω_2 *disjoint open subsets of* Ω *and* $\xi \notin (i - F)(\overline{\Omega}\backslash(\Omega_1 \cup \Omega_2))$.
(c) *Homotopy Invariance: If* $h : [0, 1] \times \overline{\Omega} \to P_{f_c}(X)$ *is a compact multifunction and* $\xi \notin (i - h(t, \cdot))(\partial\Omega)$ *for all* $t \in [0, 1]$*, then* $\hat{d}_{LS}(i - h(t, \cdot), \Omega, \xi)$ *is independent of* $t \in [0, 1]$.
(d) *Solution Property:* $\hat{d}_{LS}(i - F, \Omega, \xi) \neq 0$ *implies that there exists a* $\hat{u} \in \Omega$ *such that* $\xi \in \hat{u} - F(\hat{u})$.
(e) *Dependence on Boundary Values: If* $(i - F, \Omega, \xi)$, $(i - G, \Omega, \xi) \in \hat{\tau}_{LS}$ *and* $F|_{\partial\Omega} = G|_{\partial\Omega}$*, then* $\hat{d}_{LS}(i - F, \Omega, \xi) = \hat{d}_{LS}(i - G, \Omega, \xi)$.
(f) *Excision Property:* $\hat{d}_{LS}(i - F, \Omega, \xi) = \hat{d}_{LS}(i - F, \Omega_1, \xi)$ *for every open set* $\Omega_1 \subseteq \Omega$ *such that* $\xi \notin (i - F)(\Omega\backslash\Omega_1)$.

We also have Borsuk's theorem.

Proposition 3.3.7 *If* $\Omega \subseteq X$ *is bounded, open, symmetric with* $0 \in \Omega$, $F : \overline{\Omega} \to P_{f_c}(X)$ *is a compact multifunction such that* $F(-u) = -F(u)$ *for all* $u \in \partial\Omega$ *and*

$$0 \notin (1 - F)(\partial\Omega),$$

then $\hat{d}_{LS}(i - F, \Omega, 0)$ *is odd.*

Finally, we also have the reduction property.

Proposition 3.3.8 *If* $(i - F, \Omega, \xi) \in \widehat{\tau}_{LS}$ *and* $F(\cdot)$ *has values in a closed linear subspace* V *of* X*,* $\xi \in V$*, then* $\hat{d}_{LS}(i - F, \Omega, \xi) = \hat{d}_{LS}(i - F, \Omega \cap V, \xi)$.

3.4 Degree for $(S)_+$-Maps

In this section, we present a topological degree for $(S)_+$-maps (see Definition 2.10.11(a)). So, let X be a reflexive Banach space. By X^* we denote its topological dual and by $\langle \cdot, \cdot \rangle$ the duality brackets for the pair (X^*, X). Let $\Omega \subseteq X$ be bounded open and $\varphi : \overline{\Omega} \to X^*$ be an $(S)_+$-map. To simplify things, without any loss of generality, we will assume that the reference point $\xi = 0$. Otherwise, we can replace φ by $\varphi - \xi$, which remains an $(S)_+$-map.

First we fix our terminology.

Definition 3.4.1 Given a map $\varphi : \overline{\Omega} \to X^*$, we say:

(a) φ is demicontinuous if whenever $u_n \to u$ in $\overline{\Omega}$, then $\varphi(u_n) \xrightarrow{w} \varphi(u)$ in X^* (that is, φ is sequentially continuous from X with the strong topology into X^* with the weak topology, denoted by X^*_w).
(b) φ is an $(S)_+$-map if whenever $\{u_n, u\}_{n \geqslant 1} \subseteq \overline{\Omega}$, $u_n \xrightarrow{w} u$ in X and

$$\limsup_{n\to\infty} \langle \varphi(u_n), u_n - u \rangle \leqslant 0,$$

then $u_n \to u$ in X (see also Definition 2.10.11(a)).

If $\{\varphi_t\}_{t\in[0,1]}$ is a family of maps from $\overline{\Omega}$ into X^*, then

(c) $\{\varphi_t\}_{t\in[0,1]}$ is an $(S)_+^t$-family on $\overline{\Omega}$ if whenever $\{t_n\}_{n\geqslant 1} \subseteq [0,1]$, $\{u_n\}_{n\geqslant 1} \subseteq \overline{\Omega}$, $t_n \to t$ in $[0,1]$, $u_n \xrightarrow{w} u$ in X and $\limsup\limits_{n\to\infty} \langle \varphi_{t_n}(u_n), u_n - u\rangle \leqslant 0$, then $u_n \to u$ in X (see Definition 2.10.21).

In what follows, we denote by $\mathscr{F}$ the family of finite-dimensional subspaces of X.

Proposition 3.4.2 *If $U \subseteq X$ is open, $K \subseteq \overline{U}$ is bounded closed and $\varphi_t : \overline{U} \to X^*$, $t \in [0,1]$, is a family of maps such that*

(i) $\{\varphi_t\}_{t\in[0,1]}$ is an $(S)_+^t$-family on K;

(ii) $t_n \to \hat{t}$ in $[0,1]$, $\{u_n\}_{n\geqslant 1} \subseteq \overline{U}$ and $u_n \to u$ in X imply $\varphi_{t_n}(u_n) \xrightarrow{w} \varphi_{\hat{t}}(u)$ in X^*;

(iii) $\varphi_t(u) \neq 0$ for all $t \in [0,1]$ and $u \in K$,

then there exists an $F_0 \in \mathscr{F}$ such that

$$Z(F_0, F) = \{(t,u) \in [0,1] \times (K \cap F) : \langle \varphi_t(u), u\rangle \leqslant 0,\ \langle \varphi_t(u), y\rangle = 0 \text{ for all } y \in F_0\} = \emptyset$$

for all $F \in \mathscr{F}$ with $F_0 \subseteq F$.

Proof We argue indirectly. So, suppose that given any $F_0 \in \mathscr{F}$ we can find $F_1 \in \mathscr{F}$, $F_1 \supseteq F_0$, such that $Z(F_0, F_1) \neq \emptyset$. We set

$$\Gamma(F_0) = \bigcup_{F_0 \subseteq F \in \mathscr{F}} Z(F_0, F). \tag{3.43}$$

Evidently $\Gamma(F_0) \neq \emptyset$ for every $F_0 \in \mathscr{F}$.

We consider the family $\{\overline{\Gamma(F)}^w : F \in \mathscr{F}\}$ (here $\overline{\Gamma(F)}^w$ denotes the weak closure of $\Gamma(F)$). Let $\{F_k\}_{k=1}^m \subseteq \mathscr{F}$ and let $\hat{F} = \sum\limits_{k=1}^m F_k \in \mathscr{F}$. Then

$$Z(F_k, F) \supseteq Z(\hat{F}, F) \text{ for all } k \in \{1,\dots,m\},\ \text{ all } F \in \mathscr{F},\ \hat{F} \subseteq F. \tag{3.44}$$

From (3.43) and (3.44), we have

$$\begin{aligned} &\Gamma(\hat{F}) \subseteq \Gamma(F_k) \text{ for all } k \in \{1,\dots,m\} \\ \Rightarrow\ &\Gamma(\hat{F}) \subseteq \bigcap_{k=1}^m \Gamma(F_k) \subseteq \bigcap_{k=1}^m \overline{\Gamma(F_k)}^w \text{ and } \Gamma(\hat{F}) \neq \emptyset. \end{aligned}$$

So, the family $\{\overline{\Gamma(F)}^w : F \in \mathscr{F}\}$ has the finite intersection property.

Since K is bounded and X is reflexive, it follows that $\overline{\Gamma(F)}^w$ is w-compact in $\mathbb{R} \times X$, which is reflexive. Then the finite intersection property implies that

$$\bigcap_{F \in \mathscr{F}} \overline{\Gamma(F)}^w \neq \emptyset.$$

Let $(t_0, u_0) \in \bigcap_{F \in \mathscr{F}} \overline{\Gamma(F)}^w$.

Claim 3. $u_0 \in K$ and $\varphi_{t_0}(u_0) = 0$.

Let $x \in X$ and let $F_x \in \mathscr{F}$ such that $u_0, x \in F_x$. We have $(t_0, u_0) \in \overline{\Gamma(F_x)}^w$. Since $\Gamma(F_x) \subseteq \mathbb{R} \times X$ is bounded, by virtue of Proposition 2.10.8, we can find a sequence $\{(t_n, u_n)\}_{n \geqslant 1} \subseteq \Gamma(F_x)$ such that

$$t_n \to t_0 \text{ in } [0, 1] \text{ and } u_n \xrightarrow{w} u_0 \text{ in } X.$$

Recalling that $Z(F_x, F) \subseteq Z(F_x, \hat{F})$ for all $\hat{F} \in \mathscr{F}$, $F \subseteq \hat{F}$, we see that we can find a sequence $\{F_x^n\}_{n \geqslant 1} \subseteq \mathscr{F}$, $F_x \subseteq F_x^n$, such that

$$(t_n, u_n) \in Z(F_x, F_x^n) \text{ for all } n \geqslant 1.$$

We have

$$u_n \in K \cap F_x^n \text{ and } \left\langle \varphi_{t_n}(u_n), u_n \right\rangle \leqslant 0, \ \left\langle \varphi_{t_n}(u_n), y \right\rangle = 0 \quad \text{for all } y \in F_x \text{ and all } n \geqslant 1. \tag{3.45}$$

Recall that $u_0 \in F_x$. Then from (3.45) we have

$$\langle \varphi_t(u_n), u_n - u_0 \rangle \leqslant 0 \text{ for all } n \geqslant 1. \tag{3.46}$$

Since by hypothesis $\{\varphi_t\}_{t \in [0,1]}$ is an $(S)_+^t$-family on K, from (3.46), we obtain $u_n \to u_0$ in X (see Definition 3.4.1(c)). Hence $u_0 \in K$ (since K is closed). From hypothesis (ii) we have $\varphi_{t_n}(u_n) \xrightarrow{w} \varphi_{t_0}(u_0)$ in X. Then

$$\begin{aligned} & 0 = \left\langle \varphi_{t_n}(u_n), y \right\rangle \to \left\langle \varphi_{t_0}(u_0), y \right\rangle \text{ as } n \to \infty \text{ for all } y \in F_x \\ \Rightarrow & \left\langle \varphi_{t_0}(u_0), y \right\rangle = 0 \text{ for all } y \in F_x, \text{ in particular for } y = x. \end{aligned}$$

Because $x \in X$ is arbitrary, we conclude that $\varphi_{t_0}(u_0) = 0$. This proves the claim.

Since $(t_0, u_0) \in [0, 1] \times K$, the claim contradicts hypothesis (iii). This proves the proposition. □

Now let $\varphi : \overline{U} \subseteq X \to X^*$, $F \in \mathscr{F}$ and let $\{e_k\}_{k=1}^m$ be an ordered basis of F. We consider the finite-dimensional map $\varphi_F : \overline{U} \cap F \to F$ defined by

$$\varphi_F(u) = \sum_{\mathrm{k}=1}^{m} \langle \varphi(u), e_k \rangle \, e_k \text{ for all } u \in \overline{U} \cap F. \tag{3.47}$$

Note that $\varphi_F : F \to F$, while φ maps X into its dual X^*. So, we cannot say that φ_F is a finite-dimensional approximation of φ.

If φ is demicontinuous (see Definition 3.4.1(a)), then φ_F is continuous and so we can consider the Brouwer degree of φ_F at the origin with respect to the set $\overline{U} \cap F$, provided $U \cap F$ is bounded and $\varphi_F(u) \neq 0$ for all $u \in \partial_F U$, with $\partial_F U$ being the boundary of $U \cap F$ in the relative topology of F. Note that $\partial_F U \subseteq \partial U \cap F$ with ∂U being the boundary of U in X.

From Proposition 3.4.2 we infer at once the following result.

Corollary 3.4.3 *If $\Omega \subseteq X$ is bounded open and $\varphi_t : \overline{\Omega} \to X^*$, $t \in [0, 1]$, is a family of maps which satisfy hypotheses (i), (ii), (iii) of Proposition 3.4.2 with $K = \partial\Omega$, then there exists an $F_0 \in \mathscr{F}$ such that $(\varphi_t)_F \neq 0$ for all $(t, u) \in [0, 1] \times (\partial\Omega \cap F)$ for any $F \in \mathscr{F}$ with $F_0 \subseteq F$.*

This corollary together with the proposition that follows pave the way for the introduction of a degree map for demicontinuous (S_+)-maps.

Proposition 3.4.4 *If $\Omega \subseteq X$ is bounded open, $\varphi : \overline{\Omega} \to X^*$ is a demicontinuous, (S_+)-map, $\varphi(u) \neq 0$ for all $u \in \partial\Omega$ and $F_0 \in \mathscr{F}$ is as in Proposition 3.4.2, then $d(\varphi_F, \overline{\Omega} \cap F, 0) = d(\varphi_{F_0}, \overline{\Omega} \cap F_0, 0)$ for all $F \in \mathscr{F}$, $F_0 \subseteq F$.*

Proof Let $\{e_k\}_{k=1}^m$ be an ordered basis of F_0 and let $\{e_k\}_{k=1}^m \cup \{v_k\}_{k=m+1}^n$ be a basis of $F \supseteq F_0$. Let $u_k^* \in X^x$ such that $u_k^*|_{F_0} = 0$ and $\langle u_k^*, v_i \rangle = \delta_{ki}$ for all $i \in \{m + 1, \dots, n\}$. On $\overline{\Omega} \cap F$ we consider the following continuous maps.

$$\varphi_F(u) = \sum_{\text{k=1}}^{m} \langle \varphi(u), e_k \rangle\, e_k + \sum_{\text{k=m+1}}^{n} \langle \varphi(u), v_k \rangle\, v_k,$$
$$\psi_F(u) = \sum_{\text{k=1}}^{m} \langle \varphi(u), e_k \rangle\, e_k + \sum_{\text{k=m+1}}^{n} \langle u_k^*, u \rangle\, v_k.$$

Claim 4. $d(\varphi_F, \overline{\Omega} \cap F, 0) = d(\psi_F, \overline{\Omega} \cap F, 0) = d(\varphi_{F_0}, \overline{\Omega} \cap F_0, 0)$.
To show the first equivalent, we consider the homotopy

$$h_F(t, u) = t\varphi_F(u) + (1 - t)\psi_F(u) \text{ for all } (t, u) \in [0, 1] \times \overline{\Omega}.$$

We will show that

$$h_F(t, u) \neq 0 \text{ for all } (t, u) \in [0, 1] \times (\partial U \cap F). \tag{3.48}$$

If (3.48) is not true, then we can find $t_0 \in [0, 1]$ and $u_0 \in \partial\Omega \cap F$ such that

$$\langle \varphi(u_0), e_k \rangle = 0 \text{ for all } k \in \{1, \dots, m\} \text{ and } t_0 \langle \varphi(u_0), v_k \rangle + (1 - t_0) \langle u_k^*, u_0 \rangle \quad \text{for all } k \in \{m + 1, \dots, n\}. \tag{3.49}$$

It follows that $\langle \varphi(u_0), y \rangle = 0$ for all $y \in F_0$.

Suppose $t_0 = 1$. Then $\langle \varphi(u_0), y\rangle = 0$ for all $y \in F$. Since $u_0 \in \partial\Omega \cap F$, this contradicts Proposition 3.4.2 with $K = \partial\Omega \cap F$, because then $\emptyset \neq Z(F, F) \subseteq Z(F_0, F) = \emptyset$.

Therefore $t_0 \in (0, 1]$. Let

$$u_0 = \sum_{\mathrm{k=1}}^{m} \vartheta_k e_k + \sum_{\mathrm{k=m+1}}^{n} \eta_k v_k.$$

Note that $\eta_k = \left\langle u_k^*, u_0\right\rangle$ for all $k \in \{m+1, \ldots, n\}$. We have

$$\begin{aligned}
\frac{t_0}{1-t_0}\langle \varphi(u_0), u_0\rangle &= \frac{t_0}{1-t_0}\left\langle \varphi(u_0), \sum_{\mathrm{k=1}}^{n} \vartheta_k e_k + \sum_{\mathrm{k=m+1}}^{n} \eta_k v_k \right\rangle \\
&= \frac{t_0}{1-t_0}\sum_{\mathrm{k=1}}^{n} \vartheta_k \langle \varphi(u_0), e_k\rangle + \frac{t_0}{1-t_0}\sum_{\mathrm{k=m+1}}^{n} \eta_k \langle \varphi(u_0), v_k\rangle \\
&= -\sum_{\mathrm{k=m+1}}^{n} \left\langle u_k^x, u_0\right\rangle^2 \leqslant 0 \text{ (see (3.49))} \\
&\Rightarrow \langle \varphi(u_0), u_0\rangle \leqslant 0 \text{ and } \langle \varphi(u_0), y\rangle = 0 \text{ for all } y \in F_0 \\
&\Rightarrow Z(F_0, F) \neq \emptyset, \text{ which contradicts Proposition 3.4.2.}
\end{aligned}$$

This proves (3.48) and then the homotopy invariance of the Brouwer degree implies

$$d(\varphi_F, \overline{\Omega} \cap F, 0) = d(\psi_F, \overline{\Omega} \cap F, 0).$$

Next, let $p_{F_0} : F \to F_0$ be the projection map. We have $\varphi_{F_0} = p_{F_0} \circ \varphi_F$ and

$$\psi_F = i_F + p_{F_0} \circ (\varphi_F - i_F)$$

with i_F being the identity map on F. Proposition 3.1.43 implies

$$d(\psi_F, \overline{\Omega} \cap F, 0) = d(\varphi_{F_0}, \overline{\Omega} \cap F_0, 0).$$

This proves the claim and so the proposition, too. □

Remark 3.4.5 Note that it does not matter how we complete the basis of F_0. Moreover, if X is separable and reflexive, then the degree does not change if we change the basis on F_0. This is a consequence of the uniqueness of the degree map for demicontinuous $(S)_+$-maps defined on a separable reflexive Banach space, established by Brouwer [75] and Berkovitz and Mustonen [46]. In general, it makes sense to speak about a degree map when the space F_0 is fixed together with an ordered basis on it. This then leads to the following definition.

Definition 3.4.6 Let $\Omega \subseteq X$ be bounded open, $\varphi : \overline{\Omega} \to X^*$ a demicontinuous $(S)_+$-map, $0 \notin \varphi(\partial\Omega)$, $\varphi_F : \overline{\Omega} \cap F \to F$ with $F \in \mathscr{F}$ be defined by (3.47) and $F_0 \in \mathscr{F}$ be as in Corollary 3.4.3. Then we define

$$d_{(S)_+}(\varphi, \Omega, 0) = d(\varphi_F, \Omega \cap F, 0) \text{ for } F_0 \subseteq F.$$

Remark 3.4.7 If $u^* \notin \varphi(\partial\Omega)$, then $d_{(S)_+}(\varphi, \Omega, u^*) = d(S)_+(\varphi - u^*, \Omega, 0)$ since $u \to \varphi(u) - u^*$ is still a demicontinuous $(S)_+$-map in $\overline{\Omega}$. From the above construction of $d_{(S)_+}$ it is clear that it suffices to assume that $\varphi : \overline{\Omega} \to X^*$ is $(S)_+$ on $\partial\Omega$.

Next, using Definition 3.4.6, we will establish some classical properties for the degree map $d_{(S)_+}$. We will start with the homotopy invariance property. To state this property, we need to introduce the family of admissible homotopies.

Definition 3.4.8 Let $\Omega \subseteq X$ be bounded open and $h_t : \overline{\Omega} \to X^*$, $t \in [0, 1]$, be a one-parameter family of maps such that

(a) $h_t(u) \neq 0$ for all $t \in [0, 1]$, all $u \in \partial\Omega$;
(b) $\{h_t\}_{t\in[0,1]}$ is an $(S)_+^t$-family (see Definition 3.4.1(c));
(c) $t_n \to t_0$ and $u_n \to u_0$ in X imply $h_{t_n}(u_n) \xrightarrow{w} h_{t_0}(u_0)$ in X^*.

Then the family $\{h_t\}_{t\in[0,1]}$ is an "admissible $(S)_+$-homotopy".

Proposition 3.4.9 *If $\varphi, \psi : \overline{\Omega} \to X^*$ are demicontinuous $(S)_+$-maps, $0 \notin \varphi(\partial\Omega)$, $\psi(\partial\Omega)$ and*

$$h_t(u) = t\varphi(u) + (1-t)\psi(u) \text{ for all } (t, u) \in [0, 1] \times \overline{\Omega},$$

then h is an admissible $(S)_+$-homotopy.

Proof Evidently, we only need to show property (b) in Definition 3.4.8.

So, let $t_n \to t_0$ and $\{u_n, u_0\}_{n\geqslant 1} \subseteq \overline{\Omega}$ with $u_n \xrightarrow{w} u_0$ which satisfy

$$\limsup_{n\to\infty} \left\langle h_{t_n}(u_n), u_n - u_0 \right\rangle \leqslant 0. \tag{3.50}$$

From Proposition 2.10.13 we know that φ and ψ are pseudomonotone. So, from the proof of Proposition 2.10.6 and (3.50), we have

$$\begin{aligned}
&\limsup_{n\to\infty} \langle \varphi(u_n), u_n - u_0 \rangle \leqslant 0 \text{ or } \limsup_{n\to\infty} \langle \psi(u_n), u_n - u_0 \rangle \leqslant 0 \\
\Rightarrow\ & u_n \to u_0 \text{ in} \to\ X \text{ (since both } \varphi, \psi \text{ are } (S)_+ - \text{maps)} \\
\Rightarrow\ & \{h_t\}_{t\in[0,1]} \text{ is an } (S)_+^t\text{-family.}
\end{aligned}$$

Therefore h is an admissible $(S)_+$-homotopy. □

Now we can formulate the homotopy invariance property for $d_{(S)_+}$.

Proposition 3.4.10 *If* $h_t : \overline{\Omega} \to X^*$, $t \in [0, 1]$ *is an admissible* $(S)_+$*-homotopy and*

$$0 \notin h_t(\partial\Omega) \text{ for all } t \in [0, 1],$$

then $d_{(S)_+}(h_t, \Omega, 0)$ *is independent of* $t \in [0, 1]$.

Proof By virtue of Proposition 3.4.2, we can find $F_0 \in \mathscr{F}$ such that

$$0 \notin (\varphi_t)_F(\partial\Omega \cap F) \text{ for all } t \in [0, 1] \text{ and all } F \in \mathscr{F} \text{ with } F_0 \subseteq F.$$

Then the homotopy invariance property of Brouwer's degree implies that

$$d((h_t)_F, \Omega \cap F, 0) \text{ is independent of } t \in [0, 1] \text{ for any } F \in \mathscr{F} \text{ with } F_0 \subseteq F.$$

Then according to Definition 3.4.8, we have

$$d_{(S)_+}(h_t, \Omega, 0) \text{ is independent of } t \in [0, 1].$$

The proof is now complete. □

Next, we establish the solution property of the degree.

Proposition 3.4.11 *If* $\varphi : \overline{\Omega} \to X^*$ *is a demicontinuous* $(S)_+$*-map and* $\varphi(u) \neq 0$ *for all* $u \in \overline{\Omega}$, *then* $d_{(S)_+}(\varphi, \Omega, 0) = 0$.

Proof From Proposition 3.4.2, we know that there is an $F_0 \in \mathscr{F}$ such that $d(\varphi_F, \Omega \cap F, 0)$ is well-defined for all $F \in \mathscr{F}$ with $F_0 \subseteq F$. Then $\varphi_F(u) \neq 0$ for all $u \in \overline{\Omega} \cap F$ and so $d(\varphi_F, \Omega \cap F, 0) = 0$. Definition 3.4.8 implies that $d_{(S)_+}(\varphi, \Omega, 0) = 0$, which concludes the proof. □

Corollary 3.4.12 *If* $\varphi : \overline{\Omega} \to X^*$ *is a demicontinuous* $(S)_+$*-map,* $0 \notin \varphi(\partial\Omega)$ *and*

$$d_{(S)_+}(\varphi, \Omega, 0) \neq 0,$$

then there exists a $u_0 \in \Omega$ *such that* $\varphi(u_0) = 0$.

We also have the domain additivity property.

Proposition 3.4.13 *If* $\varphi : \overline{\Omega} \to X^*$ *is a demicontinuous* $(S)_+$*-map,* $\Omega_1, \Omega_2 \subseteq \Omega$ *are nonempty disjoint open sets and* $\varphi(u) \neq 0$ *for all* $\overline{\Omega} \setminus (\Omega_1 \cup \Omega_2)$, *then*

$$d_{(S)_+}(\varphi, \Omega, 0) = d_{(S)_+}(\varphi, \Omega_1, 0) + d_{(S)_+}(\varphi, \Omega_2, 0).$$

Proof Note that $\overline{\Omega} \setminus (\Omega_1 \cup \Omega_2)$ is bounded and closed. So, by Remark 3.2.4 we can find $F_0 \in \mathscr{F}$ such that

$$\varphi_F(u) \neq 0 \text{ for all } u \in (\overline{\Omega} \setminus (\Omega_1 \cup \Omega_2)) \cap F$$

whenever $F \in \mathscr{F}$ with $F_0 \subseteq F$. Let $\Omega_1^F = \Omega_1 \cap F$, $\Omega_2^F = \Omega_2 \cap F$. We have

$$\begin{aligned}
&(\overline{\Omega} \setminus (\Omega_1 \cup \Omega_2)) \cap F = \overline{\Omega} \cap (\Omega_1 \cup \Omega_2)^c \cap F = \\
&(\overline{\Omega} \cap F) \cap [(\Omega_1 \cup \Omega_2)^c \cup F^c] = \\
&(\overline{\Omega} \cap F) \cap ((\Omega_1 \cup \Omega_2) \cap F)^c = \\
&(\overline{\Omega} \cap F) \setminus (\Omega_1^F \cup \Omega_2^F).
\end{aligned}$$

The sets Ω_1^F, Ω_2^F are open and $\overline{\Omega} \cap F$ is relatively closed. So, the additivity property of Brouwer's degree implies

$$\begin{aligned}
&d(\varphi_F, \Omega \cap F, 0) = d(\varphi_F, \Omega_1^F, 0) + d(\varphi_F, \Omega_2^F, 0) \\
\Rightarrow\ & d_{(S)_+}(\varphi, \Omega, 0) = d_{(S)_+}(\varphi, \Omega_1, 0) + d_{(S)_+}(\varphi_2, \Omega_2, 0) \text{ (see Definition 3.4.8).}
\end{aligned}$$

The proof is now complete. □

The normalization property is satisfied by the duality map, which depends on the geometry of the Banach space X and uses the properties established so far.

Proposition 3.4.14 *If J is the duality map of X generated by a norm on X which is locally uniformly convex and its dual norm is also locally uniformly convex (see Theorem 2.7.36) and $\xi \in J(\Omega)$, then $d_{(S)_+}(J, \Omega, \xi)$.*

Proof Recall that $J : X \to X^*$ is a homeomorphism (Proposition 2.7.33) and of type $(S)_+$ (Proposition 2.10.14). So, it suffices to show that $d_{(S)_+}(J, \Omega, 0) = 1$ if $0 \in \Omega$.

Since J is strictly monotone, and $\langle J(u), u \rangle = ||u||^2$ for all $u \in X$, any nontrivial finite-dimensional subspace $F_0 \subseteq X$ satisfies Proposition 3.4.2. Let $F \in \mathscr{F}$, let $\{v_k\}_{k=1}^n$ be a basis of F and let J_F be defined as in (3.47). We consider the homotopy

$$h_t^F(u) = t i_F(u) + (1-t) J_F(u) \text{ for all } (t, u) \in [0, 1] \times (\overline{\Omega} \cap F)$$

with i_F being the identity map on F. Suppose that we can find $t \in [0, 1]$ and $u \in \partial\Omega \cap F$ such that $0 = h_t^F(u)$. Then

$$0 = t \sum_{\mathrm{k}=1}^{n} \left\langle v_k^*, u \right\rangle v_k + (1-t) \sum_{\mathrm{k}=1}^{n} \left\langle J(u), v_k \right\rangle v_k \tag{3.51}$$

with $v_k^* \in X^*$ such that $\left\langle v_k^*, v_i \right\rangle = \delta_{ki}$ for all $k, i \in \{1, \ldots, n\}$ (recall that for every $u \in X$ we have $u = \sum\limits_{\mathrm{k}=1}^{n} \left\langle v_k^*, u \right\rangle v_k$). Since $\{v_k\}_{k=1}^n$ are linearly independent, from (3.51) we infer that

$$t \left\langle v_k^*, u \right\rangle = -(1-t) \left\langle J(u), v_k \right\rangle \text{ for all } k \in \{1, \ldots, n\}.$$

Therefore we have

$$\begin{aligned} 0 \leqslant t\sum_{k=1}^{n} \left\langle v_k^*, u\right\rangle^2 &= -(1-t)\sum_{k=1}^{n} \langle J(u), v_k\rangle \left\langle v_k^*, u\right\rangle \\ &= -(1-t)\left\langle J(u), \sum_{k=1}^{n} \left\langle v_k^*, u\right\rangle v_k \right\rangle \\ &= -(1-t)\langle J(u), u\rangle \\ &= -(1-t)||u||^2 \leqslant 0, \end{aligned}$$

hence $u = 0$ (recall that $t \in (0, 1]$), a contradiction since $u \in \partial\Omega$.

So, we have

$$h_t^F(u) \neq 0 \text{ for all } t \in [0, 1] \text{ and all } u \in \partial\Omega \cap F.$$

The homotopy invariance property of Brouwer's degree implies that

$$\begin{aligned} &d(J_F, \Omega \cap F, 0) = d(i_F, \Omega \cap F, 0) = 1 \\ \Rightarrow\ &d_{(S)_+}(J, \Omega, 0) = 1 \text{ (see Definition 3.4.6).} \end{aligned}$$

The proof is now complete. □

The next theorem summarizes the situation for demicontinuous and $(S)_+$ maps from a reflexive Banach space X into its dual X^*.

Theorem 3.4.15 *If* $\tau_{(S)_+} = \{(\varphi, \Omega, \xi) : \Omega \subseteq X$ *bounded open,* $\varphi : \overline{\Omega} \to X^*$ *is demicontinuous and* $(S)_+$ *and* $\xi \notin \varphi(\partial\Omega)\}$*, then there exists a map* $d_{(S)_+} : \tau_{(S)_+} \to \mathbb{Z}$ *such that*

(a) Normalization: $d_{(S)_+}(J, \Omega, \xi) = 1$ *provided* $\xi \in J(\Omega)$.
(b) Domain Additivity: $d_{(S)_+}(J, \Omega, \xi) = d_{(S)_+}(J, \Omega_1, \xi) + d_{(S)_+}(J, \Omega_2, \xi)$ *with* $\Omega_1, \Omega_2 \subseteq \Omega$ *disjoint open and* $\xi \notin \varphi(\overline{\Omega}\setminus(\Omega_1 \cup \Omega_2))$.
(c) Homotopy Invariance: $d_{(S)_+}(h_t, \Omega, \xi)$ *is independent of* $t \in [0, 1]$ *when* $\{h_t\}_{t\in[0,1]}$ *is an* $(S)_+$*-homotopy (see Definition 3.4.8).*
(d) Solution Property: $d_{(S)_+}(\varphi, \Omega, \xi) \neq 0$ *implies* $\varphi^{-1}(\xi) \neq \emptyset$.

Remark 3.4.16 Of course this degree map has other properties too, such as the dependence on boundary values (that is, $d_{(S)_+}(\varphi, \Omega, \xi) = d_{(S)_+}(\psi, \Omega, \xi)$ for every $\psi : \overline{\Omega} \to X^*$ demicontinuous and $(S)_+$, provided $\varphi|_{\partial\Omega} = \psi|_{\partial\Omega}$) and the excision property (that is, $d_{(S)_+}(\varphi, \Omega, \xi) = d_{(S)_+}(\varphi, \Omega_1, \xi)$ for every open set $\Omega_1 \subseteq \Omega$ such that $\xi \notin \varphi(\overline{\Omega}\setminus\Omega_1)$).

As a direct consequence of Definition 3.4.6 and of Theorem 3.1.44 (Borsuk's Theorem), we have:

Proposition 3.4.17 *If* $\Omega \subseteq X$ *is bounded, open, symmetric,* $0 \in \Omega$, $\varphi : \Omega \to X^*$ *is demicontinuous,* $(S)_+$ *and odd and* $\xi \notin \varphi(\partial\Omega)$*, then* $d_{(S)_+}(\varphi, \Omega, \xi)$ *is odd.*

Similarly, Definition 3.4.6 and Proposition 3.2.28 lead to the following result.

Proposition 3.4.18 *If X is a reflexive Banach space, $U \subseteq X$ is open, $\Phi : U \to \mathbb{R}$ is Gateaux differentiable with $\varphi = \Phi' : U \to X^*$ demicontinuous, $(S)_+$ and there exist $\alpha, \beta \in \mathbb{R}$, $\alpha < \beta$ and $u_0 \in U$ such that*

(i) $V = \{\Phi < \beta\}$ *is bounded and* $\overline{V} \subseteq U$*;*
(ii) if $u \in \{\Phi \leqslant \alpha\}$*, then* $tu + (1-t)u_0 \in V$ *for all* $t \in [0, 1]$*;*
(iii) $\varphi(u) = \Phi'(u) \neq 0$ *for all* $u \in \{\alpha \leqslant \Phi \leqslant \beta\}$*,*

then $d_{(S)_+}(\varphi, V, 0) = 1$.

This result has some noteworthy consequences.

Proposition 3.4.19 *If X is a Banach space, $\Phi : X \to \mathbb{R}$ is bounded, Gâteaux differentiable such that $\varphi = \Phi' : X \to X^*$ is demicontinuous, $(S)_+$, $\Phi(x) \to +\infty$ as $||x|| \to +\infty$ and there exists an $r_0 > 0$ such that $\varphi(x) \neq 0$ for all $||x|| \geqslant r_0$, then there exists an $r_1 \geqslant r_0$ such that $d_{(S)_+}(\varphi, B_r, 0) = 1$ for all $r \geqslant r_1$.*

Proof Let $\alpha = \sup[\Phi(u) : u \in B_{r_0}]$ and $r_1 = \sup[||u|| : u \in \{\Phi \leqslant \alpha\}]$. Moreover, given $r \geqslant r_1$ let $\beta > \sup[\Phi(u) : u \in B_r]$. Then the result follows from Proposition 3.4.18 with $u_0 = 0$ and from the excision property of $d_{(S)_+}$ (see Remark 3.4.16). □

Proposition 3.4.20 *If X is a reflexive Banach space, $U \subseteq X$ is open and convex, $\Phi \in C^1(U)$ such that $\varphi = \Phi' : U \to X^*$ is an $(S)_+$-map, u_0 is a local minimum of φ and an isolated critical point of Φ, then $d_{(S)_+}(\varphi, B_r, 0) = 1$ for some $r > 0$.*

Proof We start by showing that Φ is sequentially weakly lower semicontinuous on U. We argue indirectly. So, suppose that Φ is not sequentially weakly lower semicontinuous. Then we can find $\{u_n, u\}_{n \geqslant 1} \subseteq U$ such that

$$u_n \xrightarrow{w} u \text{ in } X \text{ and } \liminf_{n\to\infty} \Phi(u_n) < \Phi(u). \tag{3.52}$$

We can find a subsequence $\{u_{n_k}\}_{k \geqslant 1}$ of $\{u_n\}_{n \geqslant 1}$ such that

$$\lim_{k\to\infty} \Phi(u_{n_k}) = \liminf_{k\to+\infty} \Phi(u_n). \tag{3.53}$$

The mean value theorem implies that we can find $t_{n_k} \in (0, 1)$ such that

$$\Phi(u_{n_k}) - \Phi(u) = \left\langle \varphi(u + t_{n_k}(u_{n_k} - u)), u_{n_k} - u \right\rangle \text{ for all } k \geqslant 1 \tag{3.54}$$

$$\begin{aligned} \Rightarrow & \limsup_{k\to\infty} [t_{n_k}(\Phi(u_{n_k}) - \Phi(u))] \\ & = \limsup_{k\to+\infty} \left\langle \varphi(u + t_{n_k}(u_{n_k} - u))u + t_{n_k}(u_{n_k} - u) - u \right\rangle. \end{aligned} \tag{3.55}$$

From (3.52) and (3.53), we infer that

$$\limsup_{k\to\infty} [t_{n_k}(\Phi(u_{n_k}) - \Phi(u))] \leqslant 0. \tag{3.56}$$

We have $u + t_{n_k}(u_{n_k} - u) \xrightarrow{w} u$ in X (see (3.52)). So, from (3.55), (3.56) and since by hypothesis $\varphi = \Phi'$ is an $(S)_+$-map, we have

$$\begin{aligned} &u + t_{n_k}(u_{n_k} - u) \to u \text{ in } X \\ \Rightarrow\ &\varphi(u + t_{n_k}(u_{n_k} - u)) \to \varphi(u) = 0 \text{ in } X^* \text{ (recall } \Phi \in C^1(U)). \end{aligned} \tag{3.57}$$

Therefore from (3.54) and (3.57) we have

$$\Phi(u_{n_k}) - \Phi(u) \to 0, \text{ a contradiction to (3.52).}$$

So, indeed Φ is sequentially weakly lower semicontinuous.

By hypothesis, u_0 is a local minimizer of Φ and so an isolated critical point. So, we can find $r_0 > 0$ such that

$$\Phi(u_0) < \Phi(u) \text{ and } \varphi(u) = \Phi'(u) \neq 0 \text{ for all } u \in \overline{B}_{r_0}(u_0) \backslash \{u_0\}. \tag{3.58}$$

We show that for all $r \in (0, r_0)$, we have

$$\Phi(u_0) < \inf\{\Phi(u) : u \in \overline{B}_{r_0}(u_0) \backslash B_r(u_0)\}. \tag{3.59}$$

Again we proceed by contradiction. So, suppose that (3.59) is not true. Then we can find $r > 0$ and $\{u_n\}_{n \geqslant 1} \subseteq \overline{B}_{r_0}(u_0) \backslash B_r(u_0)$ such that

$$\Phi(u_n) \downarrow \Phi(u_0) \text{ as } n \to \infty. \tag{3.60}$$

Since $\{u_n\}_{n \geqslant 1}$ is bounded in a reflexive Banach space, by passing to a suitable subsequence if necessary, we may assume that $u_n \xrightarrow{w} u$ in X. The sequential weak lower semicontinuity of Φ established in the beginning of the proof implies that

$$\begin{aligned} &\Phi(u) \leqslant \liminf_{n \to \infty} \Phi(u_n) = \Phi(u_0) \\ \Rightarrow\ &u = u_0. \end{aligned}$$

By the mean value theorem we have

$$\Phi(u_n) - \Phi\left(\frac{u_n + u_0}{2}\right) = \left\langle \varphi(t_n u_n + (1 - t_n)\frac{u_n + u_0}{2}), \frac{u_n - u_0}{2} \right\rangle$$

with $t_n \in (0, 1)$ for all $n \geqslant 1$.

Passing to the limit as $n \to \infty$ and using (3.60) and that $u_n \xrightarrow{w} u_0$ in X, we have

$$\begin{aligned}
&\limsup_{n\to\infty}\left\langle \varphi(t_nu_n+(1-t_n)\frac{u_n+u_0}{2}),\ \frac{u_n-u_0}{2}\right\rangle \leqslant 0\\
\Rightarrow &\limsup_{n\to\infty}\left\langle \varphi(t_nu_n+(1-t_n)\frac{u_n+u_0}{2}),\ t_nu_n+(1-t_n)\frac{u_n+u_0}{2}-u_0\right\rangle =\\
&\limsup_{n\to\infty}\left[\frac{1+t_n}{2}\left\langle \varphi(t_nu_n+(1-t_n)\frac{u_n+u_0}{2}),\ u_n-u_0\right\rangle\right] \leqslant 0. \qquad (3.61)
\end{aligned}$$

Since φ is an $(S)_+$-map, from (3.61) we infer that

$$t_nu_n+(1-t_n)\frac{u_n+u_0}{2}\to u_0 \text{ in } X. \qquad (3.62)$$

But note that

$$||t_nu_n+(1-t_n)\frac{u_n+u_0}{2}-u_0|| = (1+t_n)\left\|\frac{u_n-u_0}{2}\right\| \geqslant \frac{r}{2} \text{ for all } n\geqslant 1,$$

contradicting (3.62). Therefore (3.59) holds.

Let $\beta=\inf[\Phi(u) : u\in \overline{B}_{r_0}(u_0)\backslash B_{r_0/2}(u_0)]-\Phi(u_0)$. From (3.59) we have $\beta>0$. We set

$$V=\{u\in B_{r_0/2}(u_0) : \Phi(u)-\Phi(u_0)<\beta\}.$$

The set V is nonempty open. Fix $r\in(0,\frac{r_0}{2})$ such that $\overline{B}_r(u_0)\subseteq V$ and choose $0<\alpha<\inf\{\Phi(u) : u\in B_{r_0}(u_0)\backslash B_r(u_0)\}-\Phi(u_0)$. We have

$$\{u\in B_{r_0}(u_0)\backslash B_r(u_0) : \Phi(u)-\Phi(u_0)<\alpha\}\subseteq B_r(u_0)\subseteq \overline{B}_r(u_0)\subseteq V\subseteq \overline{V}\subseteq B_{r_0}(u_0).$$

We apply Proposition 3.4.18 with $U=B_{r_0}(u_0)$, Φ replaced by $\Phi|_{B_{r_0}(u_0)}-\Phi(u_0)$ and $\alpha,\beta\in\mathbb{R}$ as above. Then

$$d_{(S)_+}(\varphi, V, 0)=1.$$

By the excision property (see Remark 3.4.16), we obtain

$$d_{(S)_+}(\varphi, B_r(u_0), 0)=1,$$

which concludes the proof. □

3.5 Degree for Maximal Monotone Perturbation of $(S)_+$-Maps

Let X be a reflexive Banach space and $\Omega\subseteq X$ a bounded open set. In this section, we define a degree for maps $\varphi+A$, where $\varphi:\overline{\Omega}\to X^*$ is bounded, demicontinuous,

$(S)_+$ and $A : X \to 2^{X^*}$ is maximal monotone. We will assume that X is equipped with a norm which is locally uniformly convex together with its dual (see Theorem 2.7.36).

In Sect. 2.9, in the context of Hilbert spaces we introduced an approximation of the identity (the resolvent) and a single-valued approximation of a maximal monotone map (the Yosida approximation). In Remark 2.9.12 we mentioned that analogous notions can be defined in the more general framework of a Banach space and its dual. However, the results are not as precise and strong as in the Hilbert space case. Nevertheless they can be used to produce a degree for maps $u \to \varphi(u) + A(u)$ as described above. For this reason, we have brief look at these notions when the ambient space is a reflexive Banach space which need not be Hilbert.

Recall that our hypotheses on X and its dual X^* imply that the duality map $J : X \to X^*$ is a homeomorphism and J^{-1} is the duality map of X^* (identifying X with X^{**}). We consider the following operator inclusion

$$0 \in J_\lambda(y - u) + \lambda A(y). \tag{3.63}$$

Theorem 2.8.5 implies that for every fixed $u \in X$ and $\lambda > 0$, problem (3.63) has a solution $(u_\lambda, u_\lambda^*) \in \mathrm{Gr}\, A$ and this solution is unique due to the strict monotonicity of J (recall that X and X^* are both locally uniformly convex). We have

$$\begin{aligned} & u_\lambda^* \in A(u_\lambda) \text{ and } \lambda u_\lambda^* \in J(u - u_\lambda) \\ \Rightarrow\ & u - u_\lambda = J^{-1}(\lambda u_\lambda^*) \\ \Rightarrow\ & u = (A^{-1} + \lambda J^{-1})(u_\lambda^*) \\ \Rightarrow\ & u_\lambda^* = (A^{-1} + \lambda J^{-1})^{-1}(u). \end{aligned}$$

We make the following definition.

Definition 3.5.1 Let $A : X \to 2^{X^*}$ be a maximal monotone map. The Yosida approximation of A is defined by

$$A_\lambda = (\lambda J^{-1} + A^{-1})^{-1} \text{ for every } \lambda > 0$$

and the resolvent of A is defined by

$$J_\lambda^A = I - \lambda J^{-1} A_\lambda \text{ for every } \lambda > 0.$$

Remark 3.5.2 So, according to the previous discussion, for every $u \in X$ and $\lambda > 0$, $A_\lambda(u)$ is the unique solution of problem (3.63). It is easy to see that for every $\lambda > 0$, $A_\lambda : X \to X^*$ is maximal monotone. If $X = H$ is a Hilbert space and $H = H^*$ (pivot Hilbert space), then $J = i$=the identity map on H. So

$$A_\lambda = \frac{1}{\lambda}(I - J_\lambda^A) \text{ and } J_\lambda^A = (I + \lambda A)^{-1} \text{ for all } \lambda > 0 \tag{3.64}$$

and we recover Definition 2.9.1. For all $u \in X$, we have the splitting

$$u = J_\lambda^A(u) + \lambda J^{-1}(A_\lambda(u)).$$

Also we have $A_\lambda(u) \in A(J_\lambda^A(u))$. Let $(y, y^*) \in \operatorname{Gr} A$. Then from the monotonicity of A, we have

$$\begin{aligned}
\left\langle y^*, J_\lambda^A(u) - y \right\rangle &\leqslant \left\langle A_\lambda(u), J_\lambda^A(u) - y \right\rangle \\
&= -\frac{1}{\lambda}\left\langle J(J_\lambda^A(u) - u), J_\lambda^A(u) - y \right\rangle \text{ (see (3.63))} \\
&= -\frac{1}{\lambda}\left\langle J(J_\lambda^A(u) - u), J_\lambda^A(u) - y \right\rangle \\
&- \frac{1}{\lambda}\left\langle J(J_\lambda^A(u) - u), u - y \right\rangle \\
&\Rightarrow ||J_\lambda^A(u) - u||^2 \leqslant -\lambda\left\langle y^*, J_\lambda^A(u) - y \right\rangle - \left\langle J(J_\lambda^A(u) - u), u - y \right\rangle. \quad (3.65)
\end{aligned}$$

Choosing $y = u$, we have

$$\begin{aligned}
&||J_\lambda^A(u) - u|| \leqslant \lambda||y^*||_* \quad (3.66)\\
\Rightarrow\ &||J_\lambda^A(u)|| \leqslant \lambda||y^*||_* + ||u|| \text{ for all } u \in D(A).
\end{aligned}$$

Moreover, from (3.64) we also have

$$||A_\lambda(u)||_* \leqslant ||y^*||_* \text{ for all } u \in D(A). \quad (3.67)$$

Now, given a maximal monotone map $A : X \to 2^{X^*}$, from Proposition 2.6.5 we know that A has closed and convex values. Because X^* is locally uniformly convex, we can define the single-valued map $u \to A^0(u)$ (also known as the "minimal section of A"), by setting

$$||A^0(u)||_* = \min\left\{ ||u^*||_* : u^* \in A(u)\right\}.$$

Proposition 3.5.3 *If $A : X \to X^*$ is a maximal monotone map, then*

(a) $J_\lambda^A(u) \to u$ *in* X *as* $\lambda \to 0^+$ *for all* $u \in \overline{\operatorname{conv}}\, D(A)$*;*
(b) $A_\lambda(u) \to A^0(u)$ *as* $\lambda \to 0^+$ *for all* $u \in D(A)$.

Proof (a) Let $\{\lambda_n\}_{n\geqslant 1}$ be a sequence such that $\lambda_n \to 0^+$. From (3.66) we see that $\{J_{\lambda_n}^A(u) - u\}_{n\geqslant 1} \subseteq X$ is bounded, hence $\{J(J_{\lambda_n}^A(u) - u)\}_{n\geqslant 1} \subseteq X^*$ is bounded. The reflexivity of X^* implies that by passing to a subsequence if necessary, we may assume that $J(J_{\lambda_n}^A(u) - u) \xrightarrow{w} v^*$ in X^*. From (3.65) we have

$$\limsup_{n\to\infty} ||J_{\lambda_n}^A(u) - u||^2 \leqslant \left\langle v^*, y - u\right\rangle \text{ for all } y \in D(A).$$

In fact, this holds for every $y \in \overline{\text{conv}}\, D(A)$. So, taking $y = u$, we have

$$J^A_{\lambda_n}(u) \to u \text{ in } X.$$

(b) From (3.67) we know that

$$\begin{aligned} &||A_\lambda(u)||_* \leqslant ||u^*||_* \text{ for all } u^* \in A(u) \\ \Rightarrow\ &||A_\lambda(u)||_* \leqslant ||A^0(u)||_* \text{ for all } \lambda > 0. \end{aligned} \tag{3.68}$$

The reflexivity of X^* implies that for a sequence $\lambda_n \to 0^+$ we have

$$\begin{aligned} &A_{\lambda_n}(u) \xrightarrow{w} v^* \text{ in } X^* \\ \Rightarrow\ &||v^*||_* \leqslant ||A^0(u)||_* \text{ (see (3.68)).} \end{aligned} \tag{3.69}$$

The monotonicity of A implies

$$0 \leqslant \left\langle y^* - A_{\lambda_n}(u), y - J^A_{\lambda_n}(u)\right\rangle \text{ for all } (y, y^*) \in \text{Gr}\, A \text{ and all } n \geqslant 1$$

(recall that $A_{\lambda_n}(u) \in A(J^A_{\lambda_n}(u)),\ n \geqslant 1$). Passing to the limit as $n \to \infty$ and since $J^A_{\lambda_n}(u) \to u$ in X (see part (a)), we obtain

$$0 \leqslant \left\langle y^* - v^*, y - u\right\rangle \text{ for all } (y, y^*) \in \text{Gr}\, A.$$

The maximality of A implies that $v^* \in A(u)$ and so

$$||A^0(u)||_* \leqslant ||v^*||_*. \tag{3.70}$$

From (3.69) and (3.70) and since $v^* \in A(u)$, we have $v^* = A^0(u)$. Therefore

$$A_{\lambda_n}(u) \xrightarrow{w} A^0(u) \text{ in } X^* \text{ for all } u \in D(A).$$

The proof is now complete. □

Proposition 3.5.4 *If $A : X \to 2^{X^*}$ is maximal monotone, $\lambda \to \lambda_0 > 0$ and $u \in D(A)$, then $J^A_\lambda(u) \to J^A_{\lambda_0}(u)$ and $A_\lambda(u) \to A_{\lambda_0}(u)$.*

Proof Note that

$$(||J^A_\lambda(u) - u|| - ||J^A_{\lambda_0}(u) - u||)^2 \leqslant \tag{3.71}$$
$$\langle J(J^A_\lambda(u) - u) - J(J^A_{\lambda_0}(u) - u), J^A_\lambda(u) - J^A_{\lambda_0}(u)\rangle. \tag{3.72}$$

Also using the equation

$$A_\lambda(u) = \frac{1}{\lambda} J(u - J^A_\lambda(u)) \text{ and } A_\lambda(u) \in A(J^A_\lambda(u)) \text{ for all } u \in X \text{ and all } x \in X,$$

we obtain

$$\begin{aligned}\left\langle J(J_\lambda^A(u)-u)-J(J_{\lambda_0}^A(u)-u),\, J_\lambda^A(u)-J_{\lambda_0}^A(u)\right\rangle &\leqslant \\ \frac{\lambda-\lambda_0}{\lambda_0}\left\langle J(J_{\lambda_0}^A(u)-u),\, J_\lambda^A(u)-J_{\lambda_0}^A(u)\right\rangle&. \qquad (3.73)\end{aligned}$$

Combining (3.71) and (3.73), we see that

$$\lim_{\lambda\to\lambda_0}||J_\lambda^A(u)-u||=||J_{\lambda_0}^A(u)-u||.$$

It follows that

$$\lim_{\lambda\to\lambda_0}\left\langle J(J_{\lambda_0}^A(u)-u),\, J_\lambda^A(u)-u\right\rangle=||J_{\lambda_0}^A(u)-u||^2.$$

Since X and X^* are both locally uniformly convex, from the Kadec–Klee property we infer that

$$\lim_{\lambda\to\lambda_0}A_\lambda(u)=A_{\lambda_0}(u) \text{ in } X^* \text{ and } \lim_{\lambda\to\lambda_0}J_\lambda^A(u)=J_{\lambda_0}^A(u) \text{ in } X \text{ for all } u\in X.$$

The proof is now complete. □

Let $\Omega\subseteq X$ be bounded open, $\varphi:\overline{\Omega}\to X^*$ bounded, demicontinuous, $(S)_+$ and $A:X\to 2^{X^*}$ be maximal monotone. We will define a degree for the sum $\varphi+A$ of such maps.

Lemma 3.5.5 *If (Ω,φ,A) are as above, $\Omega\cap D(A)\neq\emptyset$ and $0\notin(\varphi+A)(\partial\Omega\cap D(A))$, then there exists a $\lambda_0>0$ such that $0\notin(\varphi+A_\lambda)(\partial\Omega)$ for all $\lambda\in(0,\lambda_0)$.*

Proof We argue by contradiction. So, suppose that the lemma is not true. Then we can find $\lambda_n\to 0^+$ and $u_n\in\partial\Omega$ such that $\varphi(u_n)+A_{\lambda_n}(u_n)=0$ for all $n\geqslant 1$. We may assume that

$$u_n\xrightarrow{w}u \text{ in } X \text{ and } \varphi(u_n)\xrightarrow{w}u^* \text{ in } X^*. \qquad (3.74)$$

We have

$$\left\langle\varphi(u_n)+A_{\lambda_n}(u_n),\, u_n-u\right\rangle=0 \text{ for all } n\geqslant 1. \qquad (3.75)$$

From the monotonicity of A_λ (see Definition 3.5.1 and Remark 3.5.2), we have

$$\begin{aligned}&\left\langle A_{\lambda_n}(u_n),\, u_n-u\right\rangle\geqslant\left\langle A_{\lambda_n}(u),\, u_n-u\right\rangle \text{ for all } n\geqslant 1\\ \Rightarrow\ &\liminf_{n\to\infty}\left\langle A_{\lambda_n}(u_n),\, u_n-u\right\rangle\geqslant 0.\end{aligned}$$

So, from (3.75) we have

$$\begin{aligned}&\limsup_{n\to\infty} \langle \varphi(u_n), u_n - u\rangle \leqslant 0\\ &\Rightarrow u_n \to u \text{ in } X \text{ (since } \varphi \text{ is an } (S)_+ - \text{map) and so } u \in \partial\Omega.\end{aligned} \tag{3.76}$$

Also, from the monotonicity of A and since $A_{\lambda_n}(u_n) \in A(J^A_{\lambda_n}(u_n))$ for all $n \geqslant 1$, we have

$$\left\langle y^* - A_{\lambda_n}(u_n), y - J^A_{\lambda_n}(u_n)\right\rangle \geqslant 0 \text{ for all } n \geqslant 1. \tag{3.77}$$

Note that $A_{\lambda_n}(u_n) \xrightarrow{w} -u^*$ on X^* (see (3.74) and recall that by hypothesis $\varphi(u_n) = -A_{\lambda_n}(u_n)$ for all $n \geqslant 1$). Also

$$J^A_{\lambda_n}(u_n) = u_n - \lambda_n J^{-1}(A_{\lambda_n}(u_n)).$$

Since $\{A_{\lambda_n}(u_n)\}_{n\geqslant 1} \subseteq X^*$ is bounded, so is $\{J^{-1}(A_{\lambda_n}(u_n))\}_{n\geqslant 1} \subseteq X$. Hence

$$\begin{aligned}&\lambda_n J^{-1}(A_{\lambda_n}(u_n)) \to 0 \text{ in } X\\ &\Rightarrow J^A_{\lambda_n}(u_n) \to u \text{ in } X \text{ (see (3.76))}.\end{aligned}$$

Therefore, if in (3.77) we pass to the limit as $n \to \infty$, then

$$\begin{aligned}&\left\langle y^* + u^*, y - u\right\rangle \geqslant 0 \text{ for all } (y, y^*) \in \operatorname{Gr} A\\ &\Rightarrow (u, -u^*) \in \operatorname{Gr} A \text{ (since } A \text{ is maximal monotone)}.\end{aligned}$$

Hence $u \in D(A)$ and so $u \in \partial\Omega \cap D(A)$ (see (3.76)) and we have

$$0 \in \varphi(u) + A(u),$$

a contradiction. □

This lemma implies that

$$0 \notin (\varphi + A_\lambda)(\partial\Omega \cap D(A)) \text{ for all } \lambda \in (0, \lambda_0).$$

Also note that if $u_n \to u$ in X and

$$\begin{aligned}&\limsup_{n\to\infty} \langle \varphi(u_n) + A_\lambda(u_n), u_n - u\rangle \leqslant 0\\ &\Rightarrow \limsup_{n\to\infty} \langle \varphi(u_n) + A_\lambda(u), u_n - u\rangle \leqslant 0 \text{ (due to the monotonicity of } A_\lambda)\\ &\Rightarrow \limsup \langle \varphi(u_n), u_n - u\rangle \leqslant 0\\ &\Rightarrow u_n \to u \text{ in } X \text{ (since } \varphi \text{ is an } (S)_+\text{-map)}\\ &\Rightarrow u \to (\varphi + A_\lambda)(u) \text{ is an } (S)_+\text{-map which is demicontinuous.}\end{aligned}$$

Therefore $d_{(S)_+}(\varphi + A_\lambda, \Omega, 0)$ is well-defined for all $\lambda \in (0, \lambda_0)$. Moreover, using Proposition 3.5.4, we see that for $\lambda_1, \lambda_2 \in (0, \lambda_0)$

$$\{h_t = \varphi + A_{t\lambda_1+(1-t)\lambda_2}\}_{t\in[0,1]}$$

is an $(S)_+$-homotopy (see Definition 3.4.8). So, Theorem 3.4.15(c) implies that $d_{(S)_+}(\varphi + A_\lambda, \Omega, 0)$ is independent of $\lambda \in (0, \lambda_0)$. Hence the following definition makes sense.

Definition 3.5.6 Let X be a reflexive Banach space such that both X and its dual X^* are locally uniformly convex. Let $\Omega \subseteq X$ be bounded open, $\varphi : \overline{\Omega} \to X^*$ a bounded demicontinuous $(S)_+$-map and $A : X \to 2^{X^*}$ a maximal monotone map such that $\Omega \cap D(A) \neq \emptyset$. Suppose that $0 \notin (\varphi + A)(\partial\Omega \cap D(A))$. Then we define

$$d_M(\varphi + A, \Omega, 0) = \lim_{\lambda\to 0^+} d_{(S)_+}(\varphi + A_\lambda, \Omega, 0).$$

An immediate consequence of this definition and of the results in Sect. 3.4 is the following proposition.

Proposition 3.5.7 *If $d_M(\varphi + A, \Omega, 0)$ is as in Definition 3.5.6, then*

(a) $d_M(J, \Omega, 0) = 1$ *provided* $0 \in J(\Omega)$.
(b) $d_M(\varphi + A, \Omega, 0) = d_M(\varphi + A, \Omega_1, 0) + d_M(\varphi + A_1, \Omega_2, 0)$ *with* $\Omega_1, \Omega_2 \subseteq \Omega$ *open disjoint and* $0 \notin (\varphi + A)(\overline{\Omega}\backslash(\Omega_1 \cup \Omega_2))$.
(c) $d_M(\varphi_t + A, \Omega, 0)$ *is independent of t when* $\{\varphi_t\}_{t\in[0,1]}$ *is an $(S)_+$-homotopy with each φ_t bounded.*
(d) $d_M(\varphi + A, \Omega, 0) \neq 0$ *implies that there exists a $u \in \Omega \cap D(A)$ such that* $0 \in (\varphi + A)(u)$.

Of course, the homotopies employed in part (c) of the above proposition are not the most general ones. We can do better and consider homotopies which also involve A.

Definition 3.5.8 Let $\{A_t\}_{t\in[0,1]}$ be a family of maximal monotone maps from X into 2^{X^*}. We say that $\{A_t\}_{t\in[0,1]}$ is a "pseudomonotone homotopy" of monotone maps if it satisfies the following mutually equivalent conditions:

(a) If $t_n \to t$ in $[0, 1]$, $u_n \xrightarrow{w} u$ in X, $u_n^* \xrightarrow{w} u^*$ in X^*, $(u_n, u_n^*) \in \mathrm{Gr}\, A_{t_n}$ for all $n \geqslant 1$ and

$$\limsup_{n\to\infty} \langle u_n^*, u_n\rangle \leqslant \langle u^*, u\rangle$$

then $(u, u^*) \in \mathrm{Gr}\, A$ and $\langle u_n^*, u_n\rangle \to \langle u^*, u\rangle$.
(b) For every $\lambda > 0$ (equivalently for some $\lambda > 0$), then the map $(t, u) \to J_\lambda^{A_t}(u)$ is continuous from $[0, 1] \times X$ into X (X is equipped with the strong topology).

(c) For every $\lambda > 0$ (equivalently for some $\lambda > 0$) and every $u \in X$, the map $t \to J_\lambda^{A_t}(u)$ is continuous from $[0, 1]$ into X (X is equipped with the strong topology).
(d) If $t_n \to t$ in $[0, 1]$ and $(u, u^*) \in \operatorname{Gr} A_t$, then there exist sequences $\{u_n\}_{n\geqslant 1} \subseteq X$ and $\{u_n^*\}_{n\geqslant 1} \subseteq X^*$ such that $(u_n.u_n^*) \in \operatorname{Gr} A_{t_n}$ for all $n \geqslant 1, u_n \to u$ in X, $u_n^* \to u^*$ in X^*.

Remark 3.5.9 In general affine homotopies need not be pseudomonotone homotopies. That is, if $A_1, A_2 : X \to 2^{X^*}$ are maximal monotone maps, then $A_t = tA_1 + (1 - t)A_2$, $t \in [0, 1]$, need not be a pseudomonotone homotopy. However, if one of A_1 or A_2 is continuous, everywhere defined, then $\{A_t\}_{t\in[0,1]}$ is a pseudomonotone homotopy.

Theorem 3.5.10 *If* $\tau_M = \big\{(\varphi + A, \Omega, 0) : \Omega \subseteq X$ *bounded, open,* $\varphi : \overline{\Omega} \to X^*$ *bounded, demicontinuous,* $(S)_+$, $A : X \to 2^{X^*}$ *maximal monotone such that* $\Omega \cap D(A) \neq \emptyset$ *and* $\xi \notin (\varphi + A)(\partial\Omega \cap D(A))\big\}$, *then there exists a map* $d_M : \tau_M \to \mathbb{Z}$ *such that*

(a) Normalization: $d_M(J, \Omega, 0) = 1$ *provided* $\xi \in J(\Omega)$.
(b) Domain Additivity: $d_M(\varphi + A, \Omega, \xi) = d_M(\varphi + A, \Omega_1, \xi) + d_M(\varphi + A, \Omega_2, \xi)$ *with* $\Omega_1, \Omega_2 \subseteq \Omega$ *disjoint open and* $\xi \in (\varphi + A)(\overline{\Omega}\backslash(\Omega_1 \cup \Omega_2))$.
(c) Homotopy Invariance: $d_M(\varphi_t + A_t, \Omega, \xi)$ *is independent of* $t \in [0, 1]$ *when* $\{\varphi_t\}_{t\in[0,1]}$ *is an* $(S)_+$*-homotopy with each* φ_t *bounded,* $\{A_t\}_{t\in[0,1]}$ *is a pseudomonotone homotopy of maximal monotone maps and* $\xi \notin (\varphi_t + A_t)(\partial\Omega \cap D(A_t))$ *for all* $t \in [0, 1]$.
(d) Solution Property: $d_M(\varphi + A, \Omega, 0) \neq 0$ *implies that there exists a* $u \in \Omega \cap D(A)$ *such that* $0 \in \varphi(u) + A(u)$.

Remark 3.5.11 In the homotopy invariance property, we can replace $\xi \in X^*$ by a continuous map $\xi : [0, 1] \to X^*$ such that $\xi(t) \notin (\varphi_t + A_t)(\partial\Omega \cap D(A_t))$ for all $t \in [0, 1]$. Of course, we can have additional properties such as the excision property and the dependence on boundary value problems.

3.6 Degree for Subdifferential Operators

In this section we construct a degree for maps of the form $\partial\varphi + F$, where $\partial\varphi$ is the subdifferential of a lower semicontinuous convex function φ defined on a separable Hilbert space H into $\mathbb{R}_+ = [0, +\infty)$ and F is a multifunction on H.

So, let H be a separable pivot Hilbert space (that is, $H = H^*$). For the subdifferential term $\partial\varphi$, we employ the following class of functions.

Definition 3.6.1 Let $\Gamma_c(H)$ denote the family of lower semicontinuous and convex functions $\varphi : H \to \overline{\mathbb{R}}_+ = [0, +\infty]$ which satisfy the following conditions:

(a) $\varphi(0) = 0$.

(b) For every $\eta \in (0, +\infty)$, the level set

$$\{u \in H : \varphi(u) + ||u||^2 \leqslant \eta\}$$

is compact in H.

Remark 3.6.2 Condition (a) implies that $0 \in \partial\varphi(0)$. If for $\lambda = 0$, $J_\lambda^{\partial\varphi} = (I + \lambda\partial\varphi)^{-1}$ is the resolvent map and $(\partial\varphi)_\lambda = \partial\varphi_\lambda = \frac{1}{\lambda}(I - J_\lambda^{\partial\varphi})$ is the Yosida approximation map (see Definition 2.9.1 and Proposition 2.9.13), then $J_\lambda^{\partial\varphi}(0) = 0$ and $(\partial\varphi)_\lambda(0) = \partial\varphi_\lambda(0) = 0$. Hence $||J_\lambda^{\partial\varphi}(u)|| \leqslant ||u||$ and $||\partial\varphi_\lambda(u)|| \leqslant \frac{1}{\lambda}||u||$ (see Corollary 2.9.3 and Theorem 2.9.11).

For the family $\Gamma_c(H)$ we will consider the following two homotopies.

Definition 3.6.3 (a) Let $\Gamma_c^{h,1}(H)$ be the family $\{\varphi_t\}_{t\in[0,1]}$ such that for each $t \in [0, 1]$, $\varphi_t : H \to \overline{\mathbb{R}}_+ = [0, +\infty]$ is lower semicontinuous, convex, $\varphi_t(0) = 0$ for all $t \in [0, 1]$ and

(1) for every $\eta \in (0, +\infty)$, the set

$$\bigcup_{t\in[0,1]} \{u \in H : \varphi_t(u) + ||u||^2 \leqslant \eta\}$$

is relatively compact in H;

(2) $\{\partial\varphi_t\}_{t\in[0,1]}$ is a pseudomonotone homotopy of maximal monotone maps (see Definition 3.5.8).

(b) Let $\Gamma_c^{h,2}(H)$ be the family $\{\varphi_t\}_{t\in[0,1]}$ such that for each $t \in [0, 1]$, $\varphi_t : H \to \overline{\mathbb{R}}_+ = [0, +\infty]$ is lower semicontinuous, convex, $\varphi_t(0) = 0$ for all $t \in [0, 1]$ and

(1) the same as in (a);
(2) if $t_n \to t$ in $[0, 1]$ and $u_n \xrightarrow{w} u$ in H, then

$$\varphi_t(u) \leqslant \liminf_{n\to\infty} \varphi_{t_n}(u_n)$$

and for every sequence $t_n \to t$ in $[0, 1]$ and every $y \in \mathrm{dom}\varphi_t$, there exists a sequence $\{y_n\}_{n\geqslant 1} \subseteq H$ such that

$$y_n \xrightarrow{w} y \text{ in } H \text{ and } \varphi_{t_n}(y_n) \to \varphi_t(y).$$

Evidently the two homotopies differ in the second condition. The next proposition explains how these two conditions are related.

Proposition 3.6.4 $\Gamma_c^{h,2}(H) \subseteq \Gamma_c^{h,1}(H)$.

Proof We will check condition (a) in Definition 3.5.8. So, let $t_n \to t$ in $[0, 1]$, $u_n \xrightarrow{w} u$ in X, $u_n^* \xrightarrow{w} u^*$ in X^*, $(u_n, u_n^*) \in \mathrm{Gr}\,\partial\varphi_{t_n}$ and

$$\limsup_{n\to\infty}(u_n^*, u_n)_H \leqslant (u^*, u)_H. \tag{3.78}$$

Let $y \in \operatorname{dom}\varphi_t$. Condition (2) of $\Gamma_c^{h,2}(H)$ implies that we can find $\{y_n\}_{n\geqslant 1} \subseteq H$ such that

$$y_n \xrightarrow{w} y \text{ in } H \text{ and } \varphi_{t_n}(y_n) \to \varphi_t(y). \tag{3.79}$$

By virtue of the common condition (1), we have that

$$y_n \to y \text{ in } H. \tag{3.80}$$

Since $(u_n, u_n^*) \in \operatorname{Gr}\partial\varphi_{t_n}$, we have

$$\begin{aligned}
&(u_n^*, y_n - u_n)_H \leqslant \varphi_{t_n}(y_n) - \varphi_{t_n}(u_n) \text{ for all } n \geqslant 1\\
\Rightarrow\ &\limsup_{n\to\infty}(u_n^*, y_n - u_n)_H \leqslant \varphi_t(y) - \varphi_t(u)\\
&\qquad\qquad\qquad\qquad\qquad \text{(see (3.79) and Definition 3.6.3(b)(2))}\\
\Rightarrow\ &(u^*, y - u)_H \leqslant \varphi_t(y) - \varphi_t(u) \text{ (see (3.78) and (3.80))}.
\end{aligned}$$

Since $y \in \operatorname{dom}\varphi_t$ is arbitrary, it follows that $(u, u^*) \in \operatorname{Gr}\partial\varphi_t$. Choosing $y = u$ we have $y_n \to u$ in H and

$$\begin{aligned}
&0 \geqslant (u^*, u)_H - \liminf_{n\to\infty}(u_n^*, u_n)_H \geqslant (u^*, u)_H - \limsup_{n\to\infty}(u_n^*, u_n)_H \geqslant 0 \text{(see (3.78))}\\
\Rightarrow\ &(u_n^*, u_n)_H \to (u^*, u)_H.
\end{aligned}$$

Therefore we have Definition 3.5.8(a) and so we obtain

$$\Gamma_c^{h,2}(H) \subseteq \Gamma_c^{h,1}(H),$$

which concludes the proof. □

Next we look at the multivalued perturbation F. We consider two classes of such perturbations. In what follows,

$$S = \{\vartheta : \mathbb{R}_+ = [0, +\infty) \to \mathbb{R}_+ : \vartheta \text{ is monotone increasing}\}.$$

Definition 3.6.5 Let $\Omega \subseteq H$ be bounded open and $\varphi \in \Gamma_c(H)$.

(a) We denote by $M_1(\varphi, \Omega)$ the family of all multivalued maps $F : H \to 2^H$ which satisfy the following conditions:

(1) $F(u) \in P_{f_c}(H)$ for all $u \in \overline{\Omega} \cap D(\partial\varphi)$;
(2) $\operatorname{Gr} F$ is sequentially closed in $H \times H_w$;
(3) there exist $c_1 \in (0, 1),\ r \in (0, 2)$ and $\vartheta_1 \in S$ such that

$$||u^*||^2 \leqslant c_1||v^*||^2 + \vartheta_1(||u||)(\varphi(u)^r + 1)$$

for all $u \in \overline{\Omega} \cap D(\partial\varphi)$, all $v^* \in \partial\varphi(u)$ and all $u^* \in F(u)$.

(b) We denote by $M_2(\varphi, \Omega)$ the family of all multivalued map $F : H \to 2^H$ which satisfy the following conditions:

(1) the same as $(a)(1)$;
(2) the same as $(a)(2)$;
(3) there exist $c_2 \in (0, 1)$ and $\vartheta_1, \vartheta_2 \in S$ such that

$$||u^*||^2 \leqslant c_2||v^*||^2 + \vartheta_2(\varphi(u)) + \vartheta_3(||u||)$$

for all $u \in \overline{\Omega} \cap D(\partial\varphi)$, all $v^* \in \partial\varphi(u)$ and all $u^* \in F(u)$;
(4) there exist $c_3 \in (0, 1)$ and $\vartheta_4 \in S$ such that

$$-(u^*, u)_H \leqslant c_3\varphi(u) + \vartheta_4(||u||)$$

for all $u \in \overline{\Omega} \cap D(\partial\varphi)$, $v^* \in \partial\varphi(u)$ and $u^* \in F(u)$.

For these two families of multivalued maps, we introduce two classes of homotopies which are similar to each other.

Definition 3.6.6 Let $\Omega \subseteq H$ be bounded open and $\{\varphi_t\}_{t\in[0,1]} \in \Gamma_c^{h,1}(H)$ (resp, $\{\varphi_t\}_{t\in[0,1]} \in \Gamma_c^{h,2}(H)$). By $M_1(\{\varphi_t\}, \Omega)$ (resp. $M_2(\{\varphi_t\}, \Omega)$) we denote the collection of all one-parameter families $\{F_t\}_{t\in[0,1]}$ of multivalued maps in $M_1(\varphi_t, \Omega)$ (resp. in $M_2(\varphi_t, \Omega)$) which satisfy

(1) $t_n \to t$ in $[0, 1]$, $\{u_n\}_{n\geqslant 1} \subseteq \overline{\Omega} \cap D(\partial\varphi_{t_n})$, $u_n \to u$ in H, $u_n^* \in F_{t_n}(u_n)$ $n \geqslant 1$ and $u_n^* \xrightarrow{w} u^*$ in H, then $u^* \in F_t(u)$;
(2) condition (3) in Definition 3.6.5(a) (resp. conditions (3), (4) in Definition 3.6.5(b)) hold uniformly in t (that is, the constants $c_1, c_2, c_3, r > 0$ and the functions $\vartheta_1, \vartheta_2, \vartheta_3, \vartheta_4 \in S$ can be chosen independent of $t \in [0, 1]$).

Before proceeding with the construction of the degree, let us produce sufficient conditions for a one parameter family of functions $\{\varphi_t\}_{t\in[0,1]}$ to belong in $\Gamma_c^{h,1}(H)$ or $\Gamma_c^{h,2}(H)$.

Proposition 3.6.7 *If $\{\varphi_t\}_{t\in[0,1]}$ is a family of functions in $\Gamma_c(H)$ such that for every $\eta \in (0, +\infty)$ the set*

$$\bigcup_{t\in[0,1]} \{u \in H : \varphi_t(u) + ||u||^2 \leqslant \eta\}$$

is relatively compact in H, then the following conditions are equivalent:

(a) $\{\varphi_t\}_{t\in[0,1]} \in \Gamma_c^{h,1}(H)$.
(b) If $t_n \to t$ in $[0, 1]$ and $(u, v^*) \in \mathrm{Gr}\,\partial\varphi_t$, then we can find $(u_n, v_n^*) \in \mathrm{Gr}\,\partial\varphi_{t_n}$ for all $n \geqslant 1$ such that $u_n \to u$ in H, $v_n^* \xrightarrow{w} v^*$ in H.

(c) *If* $t_n \to t$ *in* $[0, 1]$, $(u_n, v_n^*) \in \operatorname{Gr} \partial\varphi_{t_n}$ $n \geqslant 1$, $u_n \to u$ *in* H, $v_n^* \xrightarrow{w} v^*$ *in* H, *then* $(u, v^*) \in \operatorname{Gr} \partial\varphi_t$.

(d) *For every* $u \in H$ *and every* $\lambda > 0$ *(equivalently for some* $\lambda > 0$*) the map* $t \to J_\lambda^{\partial\varphi_t}(u)$ *is continuous from* $[0, 1]$ *into* H.

Proof $(a) \Rightarrow (b)$: See Definition 3.5.8(d) and use the hypothesis to conclude that if $v_n^* \xrightarrow{w} v^*$ in H, then $v_n^* \to v^*$ in H.

$(b) \Rightarrow (c)$: Let $t_n \to t$ in $[0, 1]$, $(u_n, v_n^*) \in \operatorname{Gr} \partial\varphi_{t_n}$ $n \geqslant 1$, $u_n \to u$ in H and $v_n^* \xrightarrow{w} v^*$ in H. Let $(y, y^*) \in \operatorname{Gr} \partial\varphi_t$. Since (b) holds, we can find $(\hat{u}_n, \hat{v}_n^*) \in \operatorname{Gr} \partial\varphi_{t_n}$, $n \geqslant 1$, such that $\hat{u}_n \to y$ in H and $\hat{v}_n^* \xrightarrow{w} y^*$ in H. Since $\partial\varphi_{t_n}$, $n \geqslant 1$, is monotone, we have

$$\begin{aligned} &0 \leqslant (v_n^* - \hat{v}_n^*, u_n - \hat{u}_n)_H \text{ for all } n \geqslant 1 \\ \Rightarrow\ &0 \leqslant (v^* - y^*, u - y)_H. \end{aligned}$$

Since $(y, y^*) \in \operatorname{Gr} \partial\varphi_t$ is arbitrary, from the last inequality we infer that $(u, v^*) \in \operatorname{Gr} \partial\varphi_t$.

$(c) \Rightarrow (d)$ (for all $\lambda > 0$): Let $\lambda > 0$ and suppose that $t_n \to t$ in $[0, 1]$. We set

$$z_n = J_\lambda^{\partial\varphi_{t_n}}(u) \text{ and } z_n^* = (\partial\varphi_{t_n})_\lambda(u) \text{ for all } n \geqslant 1.$$

We know that $(z_n, z_n^*) \in \operatorname{Gr} \partial\varphi_{t_n}$ for all $n \geqslant 1$. Since $\varphi_{t_n}(0) = 0$ and $(0, 0) \in \operatorname{Gr} \partial\varphi_{t_n}$ $n \geqslant 1$, we have

$$\begin{aligned} &\varphi_{t_n}(z_n) \leqslant (z_n^*, z_n)_H \leqslant \frac{1}{\lambda}||u||^2 \\ \Rightarrow\ &\varphi_{t_n}(z_n) + ||z_n||^2 \leqslant (\frac{1}{\lambda} + 1)||u||^2. \end{aligned}$$

Then our hypothesis implies that $\{z_n\}_{n\geqslant 1} \subseteq H$ is relatively compact. So, we can find a subsequence $\{z_{n_k}\}_{k\geqslant 1}$ of $\{z_n\}_{n\geqslant 1}$ such that $z_{n_k} \to z$. Then

$$z_{n_k}^* = \frac{1}{\lambda}(u - z_{n_k}) \to \frac{1}{\lambda}(u - z) = z^* \text{ in } H \text{ (see Definition 2.9.1).}$$

From (c) it follows that $(z, z^*) \in \operatorname{Gr} \partial\varphi_t$. Evidently

$$z = J_\lambda^{\partial\varphi_t}(u) \text{ and } z^* = (\partial\varphi_t)_\lambda(u).$$

From Urysohn's criterion for the convergence of a sequence, for the original sequence we have

$$\begin{aligned} &z_n = J_\lambda^{\partial\varphi_{t_n}}(u) \to z = J_\lambda^{\partial\varphi_t}(u) \text{ in } H \\ \Rightarrow\ &t \to J_\lambda^{\partial\varphi_t}(u) \text{ is continuous.} \end{aligned}$$

(d) (for some $\lambda > 0$) $\Rightarrow$ (a): See Definition 3.5.8. □

Proposition 3.6.8 *If* $\varphi_1, \varphi_2 : H \to \overline{\mathbb{R}}_+ = [0, +\infty]$ *are lower semicontinuous, convex functions such that* $\varphi_1(0) = \varphi_2(0) = 0$, $\overline{\operatorname{dom}\varphi_1} = \overline{\operatorname{dom}\varphi_2}$ *and there exist functions* $\vartheta_1, \vartheta_2 : (0, 1] \to [0, +\infty)$ *such that* $\lim\limits_{\lambda\to 0^+} \vartheta_i(\lambda) = 0$ $(i = 1, 2)$ *and*

$$\varphi_1(J_\lambda^{\partial\varphi_2}(u)) \leqslant (1 + \vartheta_1(\lambda))\varphi_1(u) + \vartheta_2(\lambda) \text{ for all } u \in \operatorname{dom}\varphi_1, \text{ all } \lambda > 0,$$
$$\varphi_2(J_\lambda^{\partial\varphi_1}(u)) \leqslant (1 + \vartheta_2(\lambda))\varphi_2(u) + \vartheta_2(\lambda) \text{ for all } u \in \operatorname{dom}\varphi_2, \text{ all } \lambda > 0,$$

then for $\varphi_t = (1 - t)\varphi_1 + t\varphi_2$ $t \in [0, 1]$, $\{\partial\varphi_t\}_{t\in[0,1]}$ *is a pseudomonotone homotopy.*

Proof Evidently, if $t_n \to t$ in $[0, 1]$ and $u_n \xrightarrow{w} u$ in H, then

$$\varphi_t(u) \leqslant \liminf_{n\to\infty} \varphi_{t_n}(u_n).$$

Next, we will verify the second part of Definition 3.6.3(b) (2) and then invoking Proposition 3.6.4, we will conclude the proof.

So, suppose $t_n \to t$ in $[0, 1]$ and let $y \in \operatorname{dom}\varphi_t$. We may assume that $t_n \in (0, 1)$ for all $n \geqslant 1$. Since the hypotheses of the proposition are symmetric with respect to φ_1, φ_2 we may assume that $t \neq 1$.

If $t \in (0, 1)$, then $\operatorname{dom}\varphi_t = \operatorname{dom}\varphi_1 \cap \operatorname{dom}\varphi_2 = \operatorname{dom}\varphi_{t_n}$ and so we can take $y_n = y$ for all $n \geqslant 1$ (see Definition 3.6.3(b) (2)).

If $t = 0$ and $y \in \operatorname{dom}\varphi_1$, then let $y_n = J_{\sqrt{t_n}}^{\partial\varphi_2}(y)$ $n \geqslant 1$ and we have $y_n \to y$ in H. It remains to show that $\varphi_{t_n}(y_n) \to \varphi_t(y)$. We have

$$\begin{aligned}
&\varphi_1(y_n) \leqslant (1 + \vartheta_1(\sqrt{t_n}))\varphi_1(y) + \vartheta_2(\sqrt{t_n}) \\
\Rightarrow\ &\limsup_{n\to\infty} \varphi_1(y_n) \leqslant \varphi_1(y) \\
\Rightarrow\ &\varphi_1(y_n) \to \varphi_1(y) \text{ (since } \varphi_1 \text{is lower semicontinuous).}
\end{aligned} \tag{3.81}$$

Since $(J_{\lambda_n}^{\partial\varphi_2})(y), (\partial\varphi_2)_{\lambda_n}(y) \in \operatorname{Gr}\partial\varphi_2$ and $\varphi_2(0) = 0$, $(0, 0) \in \operatorname{Gr}\partial\varphi_2$, we have

$$\begin{aligned}
t_n\varphi_2(y_n) &\leqslant t_n(\partial\varphi_{\lambda_n}^2(y), J_{\lambda_n}^{\partial\varphi_2}(y))_H \\
&\leqslant t_n\frac{2}{\sqrt{t_n}}||y||^2 \to 0 \text{ as } n \to \infty.
\end{aligned}$$

So, finally

$$\varphi_{t_n}(y_n) = (1 - t_n)\varphi_1(y_n) + t_n\varphi_2(y_n) \to \varphi_1(y)$$
$$\text{(see (3.81) and recall that } t_n \to 0).$$

The proof is now complete. □

As a direct consequence of this proposition, we have:

Corollary 3.6.9 *If* $\varphi_1, \varphi_2 \in \Gamma_c(H)$, $\overline{\operatorname{dom}\varphi_1} = \overline{\operatorname{dom}\varphi_2}$ *and there exists a* $c > 0$ *such that* $\varphi_1(J_\lambda^{\partial\varphi_2}(u)) \leqslant \varphi_1(u) + c\lambda$ *for all* $u \in \operatorname{dom}\varphi_1$ *and all* $\lambda > 0$*, then* $\{\varphi_t = (1-t)\varphi_1 + t\varphi_2\}_{t\in[0,1]} \in \Gamma_c^{h,2}(H)$.

Remark 3.6.10 We mention that the hypothesis in the above proposition is equivalent to

$$-\hat{c} \leqslant ((\partial\varphi_1)_\lambda(u), (\partial\varphi_2)_\mu(u))_H \text{ for all } u \in H \text{ and all } \lambda, \mu > 0$$

(see Otani [324]). We stress that this condition is symmetric in φ_1 and φ_2 and guarantees the maximal monotonicity of $(1-t)\partial\varphi_1 + t\partial\varphi_2$ and that $\partial\varphi_1 = (1-t)\partial\varphi_1 + t\partial\varphi_2$ for all $t \in [0, 1]$.

Now we can start with the construction of the degree for $\partial\varphi + F$. The idea is to approximate $\partial\varphi$ by its Yosida approximation and F by compact multifunctions.

Recall that H is separable. So, we can find an increasing sequence $\{H_n\}_{n\geqslant 1}$ of finite-dimensional subspaces such that $H = \overline{\bigcup_{n\geqslant 1} H_n}$. By $p_n \in \mathscr{L}(H, H_n)$ we denote the orthogonal projection from H onto H_n.

Let $\{\varphi_t\}_{t\in[0,1]} \in \Gamma_c^{h,1}(H)$ and $\{F_t\}_{t\in[0,1]} \in M_1(\{\varphi_t\}, \Omega)$. For $k \in \mathbb{N}$, $\lambda > 0$ and $t \in [0, 1]$, we define

$$F_t^{k,\lambda} = p_k \circ F_t \circ J_\lambda^{\partial\varphi_t}. \tag{3.82}$$

Proposition 3.6.11 *If* $0 < \lambda_0 < \lambda_1$ *and* $k \in \mathbb{N}$*, then* $\left\{J_\lambda^{\partial\varphi_t}\right\}_{t\in[0,1]}$, $\left\{J_\lambda^{\partial\varphi_t}\right\}_{\lambda\in[\lambda_0,\lambda_1]}$, $\left\{F_t^{k,\lambda}\right\}_{t\in[0,1]}$, $\left\{F_t^{k,\lambda}\right\}_{\lambda\in[\lambda_0,\lambda_1]}$ *are all compact homotopies.*

Proof Let $u_n \xrightarrow{w} u$ in H, $t_n \to t$ in $[0, 1]$ and $\lambda_n \to \lambda > 0$. Then as in the proof of Proposition 3.6.7 (see $(c) \Rightarrow (d)$), we have

$$\begin{aligned}
&J_{\lambda_n}^{\partial\varphi_{t_n}}(u_n) \to J_\lambda^{\partial\varphi_t}(u) \text{ in } H \text{and } (\partial\varphi_{t_n})_{\lambda_n}(u_n) \xrightarrow{w} (\partial\varphi_t)_\lambda(u) \text{ in } H \\
\Rightarrow\ &(u, t, \lambda) \to J_\lambda^{\partial\varphi_t}(u) \text{ is compact} \\
\Rightarrow\ &\{J_\lambda^{\partial\varphi_t}\}_{t\in[0,1]} \text{ and } \{J_\lambda^{\partial\varphi_t}\}_{\lambda\in[\lambda_0,\lambda_1]} \text{ are compact homotopies.}
\end{aligned}$$

Next we show the compactness of $(u, t, \lambda) \to F_t^{k,\lambda}(u)$.

So, let $u_n \xrightarrow{w} u$ in H, $t_n \to t$ in $[0, 1]$, $\lambda_n \to \lambda > 0$ and $u_n^* \in \Gamma_{t_n}^{k,\lambda_n}(u_n)$, $n \geqslant 1$. We have $u_n^* = p_k(y_n)$ with $y_n \in F_{t_n}\left(J_{\lambda_n}^{\partial\varphi_{t_n}}(u_n)\right)$ for all $n \geqslant 1$ (see (3.82)). From Definition 3.6.6 and Definition 3.6.5(a)(3), we have

$$\begin{aligned}
||y_n||^2 &\leqslant c_1||(\partial\varphi_{t_n})_{\lambda_n}(u_n)||^2 + \vartheta_1(||J_{\lambda_n}^{\partial\varphi_{t_n}}(u_n)||)(\varphi(J_{\lambda_n}^{\partial\varphi_{t_n}}(u))^r + 1) \\
&\leqslant c_1\left(\frac{1}{\lambda_n}||u_n||\right)^2 + \vartheta_1(||u_n||)\left[\left(\frac{1}{\lambda_n}||u_n||^2\right)^r + 1\right] \\
&\Rightarrow \{y_n\}_{n\geqslant 1} \subseteq H \text{ is bounded.}
\end{aligned}$$

So, by passing to a suitable subsequence if necessary, we may assume that $y_n \xrightarrow{w} y$ in H. Then $u_n^* = p_k(y_n) \to p_k(y) = u^*$ in H (recall that H_k is finite-dimensional). On the other hand, from the first part of the proof we know that

$$J_{\lambda_n}^{\partial\varphi_{t_n}}(u_n) \to J_{\lambda}^{\partial\varphi_t}(u) \text{ in } H.$$

From Definition 3.6.6 we have $y \in F_t(J_\lambda^{\partial\varphi_t}(u))$, hence $u^* \in F_t^{k,\lambda}(u)$. This proves the compactness of the homotopies

$$\{F_t^{k,\lambda}\}_{t\in[0,1]} \text{ and } \{F_t^{k,\lambda}\}_{\lambda\in[\lambda_0,\lambda_1]}\,.$$

The proof is now complete. $\square$

Remark 3.6.12 In fact the above proof established the compactness of the mapping $(u, t, \lambda) \mapsto F_t^\lambda(u) = F_t \circ J_\lambda^{\partial\varphi_t}(u)$.

The next proposition is the crucial step in the direction of introducing a degree for maps of the form $\partial\varphi + F$.

Proposition 3.6.13 *If* $\{\varphi_t\}_{t\in[0,1]} \in \Gamma_c^{h,1}(H)$, $\{F_t\}_{t\in[0,1]} \in M_1(\{\varphi_t\}, \Omega)$, $\xi : [0, 1] \to H$ *is continuous and* $\xi(t) \notin (\partial\varphi_t + F_t)(\partial\Omega)$ *for all* $t \in [0, 1]$, *then there exist* $\lambda_0 > 0$ *and* $k_0 \in \mathbb{N}$ *such that*

$$\lambda\xi(t) \notin (I - J_\lambda^{\partial\varphi_t} + \lambda((1 - s)F_t^{k,\lambda} + sF_t^{i,\lambda}))(\partial\Omega)$$

for all $\lambda \in (0, \lambda_0)$, *all* $k, i \geqslant k_0$ *and all* $s, t \in [0, 1]$

Proof We argue by contradiction. So, suppose that the proposition is false. This means that we can find $\lambda_n \to 0^+$, $k_n, i_n \to +\infty$, $s_n, t_n \in [0, 1]$, $u_n \in \partial\Omega$ and $y_n, w_n \in F_{t_n}(J_{\lambda_n}^{\partial\varphi_{t_n}}(u_n))$ for all $n \geqslant 1$ such that

$$\begin{aligned}
&\lambda_n\xi(t_n) = u_n - J_{\lambda_n}^{\partial\varphi_{t_n}}(u_n) + \lambda_n((1 - s_n)p_{k_n}(y_n) + s_n p_{i_n}(w_n)) \\
&\hspace{20em}\text{for all } n \geqslant 1 \\
&\Rightarrow \xi(t_n) = v_n^* + (1 - s_n)p_{k_n}(y_n) + s_n p_{i_n}(w_n) \text{ for all } n \geqslant 1
\end{aligned} \tag{3.83}$$

with $v_n^* = (\partial\varphi_{t_n})_{\lambda_n}(u_n)$, see Definition 2.9.1.

From (3.83) we have

$$||v_n^*||^2 = (v_n^*, \xi(t_n))_H - (1 - s_n)(v_n^*, p_{k_n}(y_n))_H - s_n(v_n^*, p_{i_n}(w_n))_H.$$

Using Young's inequality, given $\epsilon > 0$, we can find $c_\epsilon > 0$ such that

$$||v_n^*||^2 \leqslant \epsilon||v_n^*||^2 + c_3 + \frac{1}{2}(1-s_n)||p_{k_n}(y_n)||^2 + \frac{1}{2}s_n||p_{i_n}(w_n)||^2 + \frac{1}{2}||v_n^*||^2.$$

Using Definition 3.6.5(a) (3), we have

$$(\frac{1}{2}-\epsilon)||v_n^*|| \leqslant c_3 + \frac{1}{2}[c_1||v_n^*||^2 + c_4(\varphi_{t_n}(J_{\lambda_n}^{\partial\varphi_{t_n}}(u_n))^r + 1)]$$

for some $c_4 > 0$. We choose $\epsilon > 0$ such that $\frac{1}{2} - \epsilon > \frac{c_1}{2}$. Then

$$||v_n^*|| \leqslant c_5(1 + \varphi_{t_n}(J_{\lambda_n}^{\partial\varphi_{t_n}}(u_n))^r) \text{ for some } c_5 > 0 \text{ and all } n \geqslant 1. \tag{3.84}$$

It follows that for all $n \geqslant 1$ we have

$$\begin{aligned}\varphi_{t_n}(J_{\lambda_n}^{\partial\varphi_{t_n}}(u_n)) &\leqslant (v_n^*, J_{\lambda_n}^{\partial\varphi_{t_n}}(u_n))_H \\ &\leqslant c_6||v_n^*|| \text{ for some } c_6 > 0 \\ &\leqslant c_7(1 + \varphi_{t_n}(J_{\lambda_n}^{\partial\varphi_{t_n}}(u_n))^r) \text{ with } c_7 = c_5c_6 > 0. \end{aligned} \tag{3.85}$$

Since $r > 2$, from (3.85) we infer that $\left\{\varphi_{t_n}\left(J_{\lambda_n}^{\partial\varphi_{t_n}}(u_n)\right)\right\}_{n\geqslant 1} \subseteq \mathbb{R}_+$ is bounded. Definition 3.6.3(a)(1) implies that $\{J_{\lambda_n}^{\partial\varphi_{t_n}}(u_n)\}_{n\geqslant 1} \subseteq H$ is relatively compact and so we may assume that

$$J_{\lambda_n}^{\partial\varphi_{t_n}}(u_n) \to u \text{ in } H.$$

Then (3.84) implies that $\{v_n^*\}_{n\geqslant 1} \subseteq H$ is bounded. Also, $\{y_n\}_{n\geqslant 1}, \{w_n\}_{n\geqslant 1} \subseteq H$ are both bounded (see Remark 3.6.12). Therefore, we may assume that

$$\begin{aligned}&v_n^* \xrightarrow{w} v^*,\ y_n \xrightarrow{w} y,\ w_n \xrightarrow{w} w \text{ in } H,\\ &s_n \to s,\ t_n \to t \text{ in } [0,1].\end{aligned}$$

Proposition 3.6.7(c) implies that $(u, v^*) \in \operatorname{Gr}\partial\varphi_t$. Also, Definition 3.6.6 implies that $(1-s)y + sw \in F_t(u)$. Since

$$\begin{aligned}&||J_{\lambda_n}^{\partial\varphi_{t_n}}(u_n) - u_n|| = \lambda_n||v_n^*|| \to 0\\ \Rightarrow\ &u_n \to u \text{ in } H \text{ and so } u \in \partial\Omega.\end{aligned}$$

Passing to the limit as $n \to \infty$ in (3.83), we obtain

$$\begin{aligned}&\xi(t) = v^* + (1-s)y + sw\\ \Rightarrow\ &\xi(t) \in (\partial\varphi_t + F_t)(\partial\Omega),\ \text{a contradiction.}\end{aligned}$$

The proof is now complete. $\square$

Now, let $\varphi \in \Gamma_c(H)$ and $F \in M_1(\varphi, \Omega)$. Assume that

$$\xi \notin (\partial\varphi + F)(\partial\Omega).$$

Applying Proposition 3.6.13 to $\varphi_t = \varphi,\ F_t = F$ and $\xi(t) = \xi$ for all $t \in [0, 1]$, we see that for $\lambda > 0$ small and for $k \in \mathbb{N}$ sufficiently big, we can define

$$d_{LS}(I - J_\lambda^{\partial\varphi} + \lambda F^{k,\lambda}, \Omega, \lambda\xi) = d_{LS}(\lambda(\partial\varphi_\lambda + F^{k,\lambda}), \Omega, \lambda\xi)$$

(see also Proposition 3.6.11) and in fact this degree stabilizes. So, we can make the following definition.

Definition 3.6.14 With $(\varphi, F, \Omega, \xi)$ as above, we define

$$d_{S_1}(\partial\varphi + F, \Omega, \xi) = \lim_{\substack{\lambda\to 0^+ \\ k\to+\infty}} d_{LS}(I - J_\lambda^{\partial\varphi} + \lambda F^{k,\lambda}, \Omega, \lambda\xi).$$

This degree has the main properties.

Theorem 3.6.15 *If H is a separable pivot Hilbert space and*

$$\tau_{S_1} = \{(\varphi + F, \Omega, \xi) : \varphi \in \Gamma_c(H),\ F \in M_1(\varphi, \Omega),\ \Omega \subseteq H \\ \textit{is bounded open and } \xi \notin (\partial\varphi + F)(\partial\Omega)\}$$

then there exists a map $d_{S_1} : \tau_{S_1} \to \mathbb{Z}$ such that

(a) Normalization: $d_{S_1}(\partial\varphi, \Omega, \xi) = 1$ provided $\xi \notin \partial\varphi(\Omega)\setminus\partial\varphi(\partial\Omega)$.
(b) Domain Additivity: $d_{S_1}(\partial\varphi + F, \Omega, \xi) = d_{S_1}(\partial\varphi + F, \Omega_1, \xi) + d_{S_1}(\partial\varphi + F, \Omega_2, \xi)$, with $\Omega_1, \Omega_2 \subseteq \Omega$ disjoint open and $\xi \notin (\partial\varphi + F)(\overline{\Omega}\setminus(\Omega_1 \cup \Omega_2))$.
(c) Homotopy Invariance: $d_{S_1}(\partial\varphi_t + F_t, \Omega, \xi(t))$ is independent of $t \in [0, 1]$, when $\{\varphi_t\}_{t\in[0,1]} \in \Gamma_c^{h,1}(H)$, $\{F_t\}_{t\in[0,1]} \in M_1(\{\varphi_t\}, \Omega)$ and $\xi : [0, 1] \to H$ is continuous such that $\xi(t) \notin (\partial\varphi_t + F_t)(\partial\Omega)$ for all $t \in [0, 1]$.
(d) Solution Property: $d_{S_1}(\partial\varphi + F, \Omega, \xi) \neq 0$ implies that there exists a $u \in \Omega$ such that $\xi \in \partial\varphi(u) + F(u)$.

Proof Properties (b), (c) and (d) follow from Definition 3.6.14, the corresponding properties of the Leray–Schauder degree and from Proposition 3.6.13.

So, it remains to check the normalization property (a).

Let $\varphi \in \Gamma_c(H)$, $\Omega \subseteq H$ be bounded open and $\xi \in \partial\varphi(\Omega)\setminus\partial\varphi(\partial\Omega)$. We can find $u_0 \in \Omega$ such that $\xi \in \partial\varphi(u_0)$. Let $u_\lambda = u_0 + \lambda\xi$. Then $(\partial\varphi)_\lambda(u_\lambda) = \partial\varphi_\lambda(u_\lambda) = \xi$. Fix $\lambda > 0$ small such that $u_\lambda \in \Omega$, $\lambda\xi \notin (I - J_\lambda^{\partial\varphi})(\partial\Omega)$ and $d_{S_1}(\partial\varphi, \Omega, \xi) = d_{LS}(I - J_\lambda^{\partial\varphi}, \Omega, \lambda\xi)$ (see Definition 3.6.14).

Choose $r > 0$ such that $\Omega \subseteq B_r$. We know that $I - J_\lambda^{\partial\varphi} = \lambda\partial\varphi_\lambda$ and so $I - J_\lambda^{\partial\varphi}$ is maximal monotone. Hence $(I - J_\lambda^{\partial\varphi})^{-1}$ is maximal monotone too and from

Proposition 2.6.5 it follows that $(I - J_\lambda^{\partial\varphi})^{-1}(\lambda\xi)$ is closed and convex. It follows that

$$\lambda\xi \notin (I - J_\lambda^{\partial\varphi})(\overline{B}_r \backslash \Omega)$$

and then the excision property of the Leray–Schauder degree implies that

$$d_{LS}(I - J_\lambda^{\partial\varphi}, \Omega, \lambda\xi) = d_{LS}(I - J_\lambda^{\partial\varphi}, B_r, \lambda\xi). \tag{3.86}$$

We consider the homotopy $h_t(u) = u - tJ_\lambda^{\partial\varphi}(u)$ for all $t \in [0, 1]$, all $u \in H$ and claim that

$$t\lambda\xi \notin (I - tJ_\lambda^{\partial\varphi})(\partial B_r) \text{ for all } t \in [0, 1]. \tag{3.87}$$

Suppose that (3.87) is not true. Then we can find $t \in [0, 1]$ and $u \in \partial B_r$ such that

$$t\lambda\xi = u - tJ_\lambda^{\partial\varphi}(u) = (1 - t)u + t\partial\varphi_\lambda(u). \tag{3.88}$$

If $t = 1$, then from the first equality in (3.88) we contradict the fact that

$$\lambda\xi \notin (I - J_\lambda^{\partial\varphi})(\overline{B}_r \backslash \Omega).$$

So, $t \neq 1$. On (3.88) we act with $u_\lambda - u$ and obtain

$$t\lambda(\xi - \partial\varphi_\lambda(u), u_\lambda - u)_H = (1 - t)(u, u_\lambda - u)_H. \tag{3.89}$$

Recall that $\xi = \partial\varphi_\lambda(u_\lambda)$. So, the monotonicity of $\partial\varphi_\lambda(\cdot)$ implies

$$t\lambda(\xi - \partial\varphi_\lambda(u), u_\lambda - u)_H \geqslant 0. \tag{3.90}$$

From (3.89) and (3.90), we have

$$(1 - t)(u, u_\lambda - u)_H \geqslant 0. \tag{3.91}$$

On the other hand, since $t < 1$ and $u_\lambda \in B_r$, we have

$$(1 - t)(u, u_\lambda - u) \leqslant (1 - t)(r||u_\lambda|| - r^2) < 0 \text{ (recall that } u \in \partial B_r). \tag{3.92}$$

Comparing (3.91) and (3.92), we reach a contradiction. So, (3.87) holds and from the homotopy invariance of the Leray–Schauder degree, we have

$$\begin{aligned} &d_{LS}(I - J_\lambda^{\partial\varphi}, B_r, \lambda\xi) = 1 \\ \Rightarrow\ &d_{LS}(I - J_\lambda^{\partial\varphi}, \Omega, \lambda\xi) = 1 \text{ (see (3.86))} \\ \Rightarrow\ &d_S(\partial\varphi, \Omega, \xi) = 1. \end{aligned}$$

The proof is now complete. $\square$

Next, we consider the case where $F \in M_2(\varphi, \Omega)$.

Let $\varphi \in \Gamma_c(H)$ and $F \in M_2(\varphi, \Omega)$. Without loss of generality, we may assume that the function $\vartheta_2 \in S$ in Definition 3.6.5(b)(3) is C^1 and $\vartheta_2(0) = 0$. For $\epsilon > 0$, we set

$$\varphi_\epsilon(u) = \varphi(u) + \epsilon\vartheta_2(\varphi(u)). \tag{3.93}$$

Proposition 3.6.16 *If φ_ϵ is defined by (3.93), then*

(a) $\{\varphi_\epsilon\}_{\epsilon\in[\epsilon_0,\epsilon_1]} \in \Gamma_c^{h,1}(H)$ *for* $0 < \epsilon_0 < \epsilon_1$*;*
(b) *if* $\{\varphi_t\}_{t\in[0,1]} \in \Gamma_c^{h,2}(H)$ *and* $\epsilon > 0$*, then* $\{(\varphi_t)_\epsilon\}_{t\in[0,1]} \in \Gamma_c^{h,1}(H)$*;*
(c) $\partial\varphi_\epsilon(\cdot) = (1 + \epsilon\vartheta_2'(\varphi(\cdot)))\partial\varphi(\cdot)$ *for all* $\epsilon > 0$.

Proof (a) and (b) are straightforward consequences of Definition 3.6.3 and Proposition 3.6.4.

(c) From (3.93) we have

$$(1 + \epsilon\vartheta_2'(\varphi(\cdot)))\partial\varphi(\cdot) \subseteq \partial\varphi_\epsilon(\cdot).$$

So, in order to prove equality, we need to show that $(1 + \epsilon\vartheta_2'(\varphi(\cdot)))\,\partial\varphi(\cdot)$ is maximal monotone. To this end, let $h \in H$ and consider the function

$$\gamma_h(\lambda) = \epsilon\vartheta_2'(\varphi(J_{1+\lambda}(h))).$$

Suppose that λ_0 is a fixed point of γ_h. Then for $u_0 = J_{1+\lambda_0}^{\partial\varphi}(h)$ we have $u_0 + (1 + \epsilon\vartheta_2'(\varphi(u_0)))\partial\varphi(u_0) \ni h$ and this implies the maximality of $(1 + \epsilon\vartheta_2'(\varphi(\cdot)))\,\partial\varphi(\cdot)$.

For $0 < \mu < \lambda$, we have

$$\begin{aligned}
&||J_\lambda(h) - J_\mu(h)|| \leqslant (\lambda - \mu)(\varphi(J_\mu(h)) - \varphi(J_\lambda(h))),\\
&|\varphi(J_\mu(h)) - \varphi(J_\lambda(h))| \leqslant (||\partial\varphi_\lambda(h)|| + ||\partial\varphi_\mu(h)||)||J_\lambda(h) - J_\mu(h)||\,.
\end{aligned}$$

From these inequalities it follows that $\lambda \to \varphi(J_{1+\lambda}(h))$ is continuous and monotone decreasing on $\mathbb{R}_+ = [0, +\infty)$. Then so is γ_h (recall ϑ_2' is monotone increasing). Therefore, γ_h has a fixed point. $\square$

From Proposition 3.6.16(c), we deduce the following result.

Corollary 3.6.17 *If* $\{\varphi_t\}_{t\in[0,1]} \in \Gamma_c^{h,2}(H)$ *and* $\{F_t\}_{t\in[0,1]} \in M_2(\{\varphi_t\}, \Omega)$*, then*

(a) $\{F_t\}_{t\in[0,1]} \in M_1(\{(\varphi_t)_\epsilon\}, \Omega)$ *for all* $\epsilon > 0$*;*
(b) *for* $0 > \epsilon_0 < \epsilon_1$ *and* $t \in [0, 1]$*, set* $(F_t)_\epsilon = F_t$ *for all* $\epsilon \in [\epsilon_0, \epsilon_1]$*, then*

$$\{(F_t)_\epsilon\}_{\epsilon\in[\epsilon_0,\epsilon_1]} \in M_1(\{(\varphi_t)\epsilon\}_{\epsilon\in[\epsilon_0,\epsilon_1]}, \Omega).$$

The next result realizes the crucial step in the direction of defining the degree when $F \in M_2(\varphi, \Omega)$.

Proposition 3.6.18 *If* $\{\varphi_t\}_{t\in[0,1]} \in \Gamma_c^{h,2}(H)$, $\{F_t\}_{t\in[0,1]} \in M_2(\{\varphi_t\}, \Omega)$, $\xi : [0,1] \to H$ *is continuous and* $\xi(t) \notin (\partial\varphi_t + F_t)(\partial\Omega)$ *for all* $t \in [0,1]$, *then there exists an* $\epsilon_0 > 0$ *such that*

$$\xi(t) \notin (\partial(\varphi_t)_\epsilon + F_t)(\partial\Omega) \textit{ for all } \epsilon \in (0, \epsilon_0] \textit{ and all } t \in [0,1].$$

Proof We argue by contradiction. So, suppose we can find

$$\begin{aligned}&\epsilon_n \to 0,\ t_n \in [0,1],\ u_n \in \partial\Omega \cap D(\partial\varphi_{t_n}),\ v_n^* \in \partial\varphi_{t_n}(u_n),\ y_n \in F_{t_n}(u_n) \text{ such that}\\ &\quad \xi(t_n) = (1 + \epsilon_n \vartheta_2'(\varphi_{t_n}(u_n)))u_n + y_n \text{ for all } n \geqslant 1.\end{aligned} \tag{3.94}$$

From Definition 3.6.5(b), (4), we have

$$\begin{aligned}\varphi_{t_n}(u_n) &\leqslant (\varphi_{t_n})_{\epsilon_n}(u_n)\\ &\leqslant ((1 + \epsilon_n \vartheta_2'(\varphi_{t_n}(u_n)))v_n^*, u_n)_H\\ &= (\xi(t_n) - y_n, u_n)_H\\ &\leqslant c_3\varphi_{t_n}(u_n) + \hat{c} \text{ with } c_3 \in (0,1), \hat{c} > 0\\ &\Rightarrow \{\varphi_{t_n}(u_n)\}_{n\geqslant 1} \text{ is bounded.}\end{aligned}$$

Then from Definition 3.6.3(b)(1), we infer that $\{u_n\}_{n\geqslant 1} \subseteq H$ is relatively compact. So, we may assume that $u_n \to u$ in H, hence $u \in \partial\Omega$. From Definition 3.6.5(b)(3) we have

$$\begin{aligned}||v_n^*||^2 &\leqslant ||\xi(t_n) - y_n||^2\\ &\leqslant (1+\beta)||y_n||^2 + c_\beta\\ &\leqslant (1+\beta)c_2||v_n^*||^2 + \hat{c} - c_\beta,\end{aligned} \tag{3.93'}$$

where $\beta > 0$ and $c_\beta > 0$ is a constant depending only on β. Choosing $\beta > 0$ such that $(1+\beta)c_2 < 1$ (recall $c_2 \in (0,1)$), from (3.93′) it follows that $\{v_n^*\}_{n\geqslant 1} \subseteq H$ is bounded. By virtue of Definition 3.6.5(b)(3) it follows that $\{y_n\}_{n\geqslant 1} \subseteq H$ is bounded. So, we may assume that

$$v_n^* \xrightarrow{w} v^* \text{ and } y_n \xrightarrow{w} y \text{ in } H.$$

Proposition 3.6.7(iii) implies that $(u, v^*) \in \operatorname{Gr} \partial\varphi_t$ and Definition 3.6.6 implies $y \in F_t(u)$. Therefore

$$\xi(t) \in \partial\varphi_t(u) + F_t(u) \text{ with } u \in \partial\Omega,$$

a contradiction. □

This proposition suggests that for $\varphi \in \Gamma_c(H)$ and $F \in M_2(\varphi, \Omega)$, we can define

$$d_{S_2}(\partial\varphi + F, \Omega, \xi) = \lim_{\epsilon\to 0^+} d_{S_1}(\partial\varphi_\epsilon + F, \Omega, \xi), \tag{3.95}$$

where the degree in the right-hand side is defined in the sense of Definition 3.6.14. To do this, we need to show that this definition is independent of the choice of the function ϑ_2.

So, let $\vartheta_2, \widehat{\vartheta}_2$ be two convex C^1-functions satisfying Definition 3.6.5(b)(3). We set

$$\varphi_\epsilon(u) = \varphi(u) + \epsilon\vartheta_2(u) \text{ and } \widehat{\varphi}_\epsilon(u) = \varphi(u) + \epsilon\widehat{\vartheta}_2(u).$$

We need to show that

$$d_{S_2}(\partial\varphi_\epsilon + F, \Omega, \xi) = d_{S_2}(\partial\widehat{\varphi}_\epsilon + F, \Omega, \xi) \text{ for all } \epsilon > 0 \text{ small.}$$

Let $\widetilde{\varphi}_t(u) = (1-t)\varphi_\epsilon(u) + t\widehat{\varphi}_\epsilon(u)$ for all $t \in [0, 1]$. Then we can easily check that $\{\widetilde{\varphi}_t\} \subseteq \Gamma_c^{h,2}(H) \subseteq \Gamma_c^{h,1}(H)$ (see Proposition 3.6.4). Since $\widetilde{\varphi}_0 = \varphi_\epsilon$, $\widetilde{\varphi}_1 = \widehat{\varphi}_\epsilon$, we have

$$\begin{aligned} & d_{S_1}(\partial\varphi_\epsilon + F, \Omega, \xi) = d_{S_1}(\partial\widehat{\varphi}_\epsilon + F, \Omega, \xi) \text{ (see Theorem 3.6.15(c))} \\ \Rightarrow\ & d_{S_2}(\partial\varphi + F, \Omega, \xi) \text{ is independent of } \vartheta_2. \end{aligned}$$

So, we can make the following definition.

Definition 3.6.19 For $\varphi \in \Gamma_c(H)$, $F \in M_2(\varphi, \Omega)$ and $\xi \notin (\partial\varphi + F)(\partial\Omega)$, we set

$$d_{S_2}(\partial\varphi + F, \Omega, 0) = \lim_{\epsilon\to 0^+} d_{S_1}(\partial\varphi_\epsilon + F, \Omega, 0).$$

This degree exhibits all the main properties.

Proposition 3.6.20 *If H is a separable pivot Hilbert space*

$$\begin{aligned} \tau_{S_2} = \{(\varphi + F, \Omega, \xi) : \varphi \in \Gamma_c(H),\ F \in M_2(\varphi, \Omega),\ \Omega \subseteq H \\ \text{is bounded open and } \xi \notin (\partial\varphi + F)(\partial\Omega)\ \} \end{aligned}$$

then there exists a map $d_{S_2} : \tau_{S_2} \to \mathbb{Z}$ such that

(a) Normalization: $d_{S_2}(\partial\varphi, \Omega, \xi) = 1$ provided $\xi \in \partial\varphi(\Omega)\backslash\partial\varphi(\partial\Omega)$.
(b) Domain Additivity: $d_{S_2}(\partial\varphi + F, \Omega, \xi) = d_{S_2}(\partial\varphi + F, \Omega_1, \xi) + d_{S_2}(\partial\varphi + F, \Omega_2, \xi)$ with $\Omega_1, \Omega_2 \subseteq \Omega$ disjoint open and $\xi \notin (\partial\varphi + F)(\overline{\Omega}\backslash(\Omega_1 \cup \Omega_2))$.
(c) Homotopy Invariance: $d_{S_2}(\partial\varphi_t + F_t, \Omega, \xi(t))$ is independent of $t \in [0, 1]$ when $\{\varphi_t\}_{t\in[0,1]} \in \Gamma_c^{h,2}(H)$, $\{F_t\}_{t\in[0,1]} \in M_2(\{\varphi_t\}, \Omega)$ and $\xi : [0, 1] \to H$ is continuous such that $\xi(t) \notin (\partial\varphi_t + F_t)(\partial\Omega)$ for all $t \in [0, 1]$.
(d) Solution Property: $d_{S_2}(\partial\varphi + F, \Omega, \xi) \neq 0$ implies that there exists a $u \in \Omega$ such that $\xi \in \partial\varphi(u) + F(u)$.

3.7 Some Generalizations

In this section we present some useful extensions of the degrees discussed so far.

In Sect. 3.2 we presented an infinite-dimensional extension of Brouwer's degree to maps of the of the form $I - f$ with f compact. However, in many applications f is not compact. So, it is natural to ask what kind of maps can replace compact maps and still have an infinite-dimensional extension of Brouwer's degree. We have already seen that simple continuity of f does not lead to a degree (see Example 3.2.1). It turns out that we can have a degree theory when f is a set condensing map. This brings on stage the notion of a measure of noncompactness.

Definition 3.7.1 Let X be a Banach space and $\mathscr{B}$ the family of bounded subsets of X.

(a) The "Kuratowski measure of noncompactness" is the map $\alpha : \mathscr{B} \to \mathbb{R}_+$ defined by

$$\alpha(B) = \inf\{d > 0 : \; B \text{ admits a finite cover by sets of diameter} \leqslant d\}.$$

(b) The "Hausdorff (or ball) measure of noncompactness" is the map $\beta : \mathscr{B} \to \mathbb{R}_+$ defined by

$$\beta(B) = \inf\{r > 0 : B \text{ admits a finite cover by balls of radius } r\}.$$

Remark 3.7.2 Note that in the definition of $\alpha(B)$, the covering sets may be taken to be subsets of B since $\operatorname{diam}(D_k \cap B) \leqslant \operatorname{diam} D_k$ and $B = \bigcup\limits_{\mathrm{k}=1}^{n} (D_k \cap B)$. However, in the definition of $\beta(B)$ it is important to specify the set in which the centers of the covering balls of B are located. If, for some reason, we want those centers to belong in some particular $E \supseteq B$, then we should write $\beta_E(B)$ and evidently we have $\beta(B) \leqslant \beta_E(B)$, with strict inequality possible.

The next result is a straightforward consequence of Definition 3.7.1.

Proposition 3.7.3 *For every $B \in \mathscr{B}$ we have $\beta(B) \leqslant \alpha(B) \leqslant 2\beta(B)$.*

On $\mathscr{B}$ we can define a distance.

Definition 3.7.4 The Hausdorff distance between $B, C \in \mathscr{B}$ is defined by

$$h(B, C) = \max\{\sup_{b \in B} d(b, C), \sup_{c \in C} d(c, B)\}.$$

Remark 3.7.5 If

$$P_{bf}(X) = \{B \subseteq X : B \text{ is bounded and closed}\},$$

then $(P_{bf}(X), h)$ is a metric space. This metric space is complete if and only if X is complete (see, for example, Hu and Papageorgiou [217]).

Proposition 3.7.6 *If $\gamma = \alpha$ or β and $B, C \in \mathscr{B}$, then*

(a) $\gamma(B) = 0$ *if and only if* $\overline{B} \subseteq X$ *is compact.*
(b) $\gamma(B) = \gamma(\overline{B})$.
(c) *if* $B \subseteq C$, *then* $\gamma(B) \leqslant \gamma(C)$.
(d) $\gamma(B \cup C) \leqslant \max\{\gamma(B), \gamma(C)\}$ *and* $\gamma(B \cap C) \leqslant \min\{\gamma(B), \gamma(C)\}$.
(e) γ *is a seminorm, that is,*

$$\gamma \textit{ is positively homogeneous, namely } \gamma(\lambda A) = |\lambda|\gamma(A) \textit{ for all } \lambda \in \mathbb{R};$$
$$\gamma \textit{ is subadditive, namely } \gamma(B + C) \leqslant \gamma(B) + \gamma(C).$$

(f) $|\gamma(B) - \gamma(C)| \leqslant 2h(B, C)$.
(g) $\gamma(\text{conv } B) = \gamma(B)$.

Proof (a)–(e) are straightforward consequences of Definition 3.7.1.

We only need to prove (f) and (g).

(f) To fix things, we assume that $\gamma = \alpha$ (the proof is similar if $\gamma = \beta$). Given $\epsilon > 0$ we can find $D_1, \ldots, D_n$ subsets of B such that

$$\text{diam } D_k \leqslant \alpha(B) + \epsilon \text{ for all } k \in \{1, \ldots, n\} \text{ and } \bigcup_{\text{k}=1}^{n} D_k = B.$$

Let $\xi = h(B, C) + \epsilon$ and $C_k = \{u \in C : \text{ there exists an } x \in D_k,\ ||x - u|| \leqslant \xi\}$ for all $k \in \{1, \ldots, n\}$. Since $h(B, C) < \xi$, we see that $C = \bigcup_{\text{k}=1}^{n} C_k$. Also

$$\begin{aligned} &\text{diam } C_k \leqslant 2\xi + \text{diam } D_k \leqslant 2h(B, C) + \alpha(B) + 3\epsilon \\ \Rightarrow\ &\alpha(C) - \alpha(B) \leqslant 2h(B, C) \text{ (letting } \epsilon \downarrow 0). \end{aligned}$$

Reversing the roles of B and C, we conclude that

$$|\alpha(C) - \alpha(B)| \leqslant 2h(B, C).$$

(g) Again we fix $\gamma = \alpha$, the proof being similar if $\gamma = \beta$.

From (c) we know that $\alpha(B) \leqslant \alpha(\text{conv } B)$. Given $\epsilon > 0$, we can find $D_1, \ldots, D_n$ subsets of B such that $\text{diam } D_k \leqslant \alpha(B) + \epsilon$ for all $k \in \{1, \ldots, n\}$. We may assume that each D_k is convex, since $\text{diam}(\text{conv } D_k) = \text{diam } D_k$. Let

$$\Delta = \{\widehat{\lambda} = (\lambda_k)_{k=1}^{n} : \lambda_k \geqslant 0 \text{ for all } k \in \{1, \ldots, n\}, \sum_{\text{k}=1}^{n} \lambda_k = 1\}$$

and define

$$D(\widehat{\lambda}) = \sum_{\text{k}=1}^{n} \lambda_k\ D_\lambda \text{ for every } \widehat{\lambda} \in \Delta.$$

We have

$$\alpha(D(\widehat{\lambda})) \leqslant \alpha(B) + \epsilon \text{ for every } \widehat{\lambda} \in \Delta \text{ (see (e)).} \tag{3.96}$$

We show that $\bigcup\limits_{\widehat{\lambda}\in\Delta} D(\widehat{\lambda})$ is convex. Let $\widehat{\lambda}, \widehat{\eta} \in \Delta$ and

$$x = \sum_{k=1}^{n} \lambda_k x_k, \; u = \sum_{k=1}^{n} \eta_k u_k \text{ with } x_k, u_k \in D_k \text{ for all } k \in \{1, \ldots, n\}.$$

Then for $t \in [0, 1]$ we have

$$tx + (1-t)u = \sum_{k=1}^{n} (t\lambda_k + (1-t)\eta_k) \left[\frac{t\lambda_k}{t\lambda_k + (1-t)\eta_k} x_k + \frac{(1-t)\eta_k}{t\lambda_k + (1-t)\eta_k} u_k \right]$$

$$\Rightarrow \bigcup_{\widehat{\lambda}\in\Delta} S(\widehat{\lambda}) \text{ is convex.}$$

So, it follows that

$$\text{conv } B = \text{conv } (\bigcup_{k=1}^{n} D_k) \subseteq \bigcup_{\widehat{\lambda}\in\Delta} D(\widehat{\lambda}). \tag{3.97}$$

Evidently, the set $\Delta \subseteq \mathbb{R}^n$ is compact. So, we can find $\{\widehat{\lambda}_k\}_{k=1}^{m} \subseteq \Delta$ such that

$$\begin{aligned} &\bigcup_{\widehat{\lambda}\in\Delta} D(\widehat{\lambda}) \subseteq \bigcup_{k=1}^{m} D(\widehat{\lambda}_k) + \epsilon B_1 \\ &\Rightarrow \text{conv } B \subseteq \bigcup_{k=1}^{m} D(\widehat{\lambda}_k) + \epsilon B_1 \text{ (see (3.96))} \\ &\Rightarrow \alpha(\text{conv } B) \leqslant \alpha(B) + \epsilon \text{ (see (3.95 and (d))).} \end{aligned}$$

Let $\epsilon \downarrow 0$ and conclude that $\alpha(\text{conv } B) = \alpha(B)$. □

Proposition 3.7.7 *If X is an infinite-dimensional Banach space, then $\alpha(B_1) = 2$ and $\beta(B_1) = 1$ (recall that $B_1 = \{x \in X : ||x|| < 1\}$).*

Proof Evidently, $\alpha(B_1) \leqslant 2$. Suppose $\alpha(B_1) < 2$. Then we can find $D_1, \ldots, D_n$ subsets of B_1 such that diam $D_k < 2$ for all $k \in \{1, \ldots, n\}$ and $B_1 = \bigcup\limits_{k=1}^{n} D_k$. Let X_n be an n-dimensional subspace of X and set $B_1^n = B_1 \cap X_n$, $D_k^n = D_k \cap X_n$ for all $k \in \{1, \ldots, n\}$. Then

$$\text{diam } D_k^n < 2 \text{ for all } k \in \{1, \ldots, n\} \text{ and } B_1^n = \bigcup_{k=1}^{n} D_k^n,$$

which contradicts Proposition 3.1.15. Therefore $\alpha(B_1) = 2$.

Then from the above fact and Proposition 3.7.3, we conclude that $\beta(B_1) = 1$. $\square$

Definition 3.7.8 Let X be a Banach space, $D \subseteq X$, $\varphi : D \to X$ continuous and $\gamma = \alpha$ or β.

(a) We say that φ is a γ-Lipschitz map if

$$\gamma(\varphi(B)) \leqslant k\gamma(B) \text{ for some } k > 0 \text{ and all bounded } B \subseteq D.$$

(b) We say that φ is a γ-contraction if φ is γ-Lipschitz with $k \in (0, 1)$.

(c) We say that φ is a γ-condensing map if

$$\gamma(\varphi(B)) < \gamma(B) \text{ for all } B \subseteq D \text{ bounded with } \gamma(B) > 0.$$

Remark 3.7.9 If $\varphi = A \in \mathscr{L}(X, X)$, then φ is γ-Lipschitz with $k = ||A||_{\mathscr{L}}$. Also, if $\varphi : D \subseteq X \to X$ is Lipschitz continuous with Lipschitz constant $k > 0$, then φ is also γ-Lipschitz with the same constant. Another important map which is γ-Lipschitz is provided by the next proposition.

Proposition 3.7.10 *If X is a Banach space, $\overline{B}_1 = \{x \in X : ||x|| \leqslant 1\}$ and $r : X \to \overline{B}_1$ is the radial retraction, that is, $r(x) = \begin{cases} x & \text{if } x \in \overline{B}_1 \\ \frac{x}{||x||} & \text{if } ||x|| > 1 \end{cases}$ then r is γ-Lipschitz with $k = 1$.*

Proof Let $B \in \mathscr{B}$. Then $\gamma(r(B)) \leqslant \gamma(\text{conv}\,(\{0\} \cup B)) = \gamma(\{0\} \cup B) = \gamma(B)$ (see Proposition 3.7.6). $\square$

Proposition 3.7.11 *If X is an infinite-dimensional Banach space, $\vartheta : [0, 1] \to [0, 1]$ is a continuous and strictly decreasing function and $\varphi : \overline{B}_1 \to \overline{B}_1$ is defined by*

$$\varphi(u) = \vartheta(||u||)u \text{ for all } u \in \overline{B}_1,$$

then φ is γ-condensing but not a γ-contraction.

Proof Let $B \subseteq \overline{B}_1$ with $\gamma(B) > 0$. Let $r \in (0, \frac{\gamma(B)}{2})$ and let

$$C_1 = B \cap \overline{B}_r \text{ and } C_2 = B \backslash \overline{B}_r.$$

Then $B = C_1 \cup C_2$ and so $\varphi(B) = \varphi(C_1) \cup \varphi(C_2)$. So, we have

$$\gamma(\varphi(B)) = \gamma(\varphi(C_1) \cup \varphi(C_2)) \leqslant \max[\gamma(\varphi(C_1)), \gamma(\varphi(C_2))]. \tag{3.98}$$

Also, we have

$$\gamma(\varphi(C_1)) \leqslant \gamma(\text{conv}\,(\{0\} \cup C_1)) = \gamma(C_1) \leqslant \text{diam}\, C_1 \leqslant 2r < \gamma(B) \tag{3.99}$$

and $\varphi(C_2) \subseteq \text{conv}(\{0\} \cup \vartheta(r)B)$. Hence

$$\gamma(\varphi(C_2)) \leqslant \gamma(\vartheta(r)B) < \gamma(B) \text{ (see Proposition 3.7.6(e))}. \tag{3.100}$$

From (3.98), (3.99) and (3.100), we conclude that φ is γ-condensing.
Note that $\partial B_{\lambda\vartheta(\lambda)} \subseteq \varphi(\overline{B}_\lambda)$ for all $\lambda \in [0, 1]$ and so

$$\gamma(\varphi(\overline{B}_\lambda)) \geqslant \gamma(\partial B_{\lambda\vartheta(\lambda)}) = 2\lambda\vartheta(\lambda) = \vartheta(\lambda)\alpha(\overline{B}_\lambda)$$
$$\text{(see Proposition 3.7.7)}.$$

Since $\vartheta(\lambda) \to 1$ as $\lambda \to 0^+$, we see that φ cannot be a γ-contraction. $\square$

The next proposition is crucial in the definition of the degree which will extend the Leray–Schauder degree.

Proposition 3.7.12 *If X is a Banach space, $\Omega \subseteq X$ is bounded, $\varphi : \overline{\Omega} \to X$ is γ-condensing, and $F = \{x \in \overline{\Omega} : \varphi(x) = x\}$, then there exists a compact, convex C such that*

(a) $F \subseteq C$;
(b) if $x_0 \in \overline{\text{conv}}\,[C \cup \{\varphi(x_0)\}]$, then $x_0 \in C$;
(c) $C = \overline{\text{conv}}\,\varphi(C \cap \overline{\Omega})$.

Proof Let

$$\mathscr{S} = \{K : F \subseteq K, \text{ closed convex}, \ \varphi(\overline{\Omega} \cap K) \subseteq K \text{ and } (b) \text{ holds for } K\}.$$

Note that $\overline{\text{conv}}\,\varphi(\overline{\Omega}) \in \mathscr{S}$. Hence $\mathscr{S} \neq \emptyset$. We set $C = \bigcap_{k \in \mathscr{S}} K$. Evidently C satisfies (a), (b), (c) and it is closed convex. It remains to show that it is compact.

Suppose that C is not compact. Then there exists a sequence $C_1 = \{x_n\}_{n \geqslant 1} \subseteq C$ with no convergent subsequence. By (c) we have $C = \overline{\text{conv}}\,\varphi(\overline{\Omega} \cap C)$. So, there is a countable set $E_1 \subseteq \overline{\Omega} \cap C$ such that $C_1 \subseteq \overline{\text{conv}}\,\varphi(E_1)$. Therefore, $S_1 = \overline{\text{conv}}\,\varphi(\overline{\Omega} \cup C_1)$ is separable and so is $\overline{\Omega} \cap S_1$. It follows that we can find countable sets $D_1 \subseteq S_1$ and $D_1^* \subseteq \overline{\Omega} \cap S_1$ such that $\overline{D}_1 = S_1$, $\overline{D}_1^* = \overline{\Omega} \cap S_1$. Let $C_2 = C_1 \cup E_1 \cup D_1 \cup D_1^*$. Then

$$\begin{aligned}
&C_1 \subseteq C_2, \\
&\overline{\text{conv}}\,\varphi(\overline{\Omega} \cap C_1) \subseteq \overline{C}_2, \\
&\overline{\text{conv}}\,\varphi(\overline{\Omega} \cap C_1) \cap \overline{\Omega} \subseteq \overline{\overline{\Omega} \cap C_2}.
\end{aligned}$$

Using these relations and induction, we produce a sequence $\{C_n\}_{n \geqslant 1}$ of subsets of C such that for all $n \geqslant 1$ we have

$$
\begin{aligned}
&C_n \subseteq C_{n+1}, \\
&\overline{\mathrm{conv}}\, \varphi(\overline{\Omega} \cap C_n) \subseteq \overline{C}_{n+1}, \\
&\overline{\mathrm{conv}}\, \varphi(\overline{\Omega} \cap C_n) \cap \overline{\Omega} \subseteq \overline{\overline{\Omega} \cap C_{n+1}}.
\end{aligned}
$$

Set $M = \bigcup\limits_{n \geqslant 1} C_n$. Then

$$
\begin{aligned}
&M \subseteq \overline{\mathrm{conv}}\, \varphi(\overline{\Omega} \cap M) \\
\Rightarrow\ &\gamma(M) \leqslant \gamma(\overline{\mathrm{conv}}\, \varphi(\overline{\Omega} \cap M)) < \gamma(\overline{\Omega} \cap M),
\end{aligned}
$$

a contradiction. This proves the compactness of C and completes the proof of the proposition. □

Using this proposition, we can define a degree for γ-condensing maps.

The setting is the following. Let X be a Banach space and $\Omega \subseteq X$ bounded open. We consider a γ-condensing map $\varphi : \overline{\Omega} \to X$ such that $0 \notin (I - \varphi)(\partial\Omega)$. If $0 \notin (I - \varphi)(\Omega)$, then we set for the new degree map d_C

$$d_C(I - \varphi, \Omega, 0) = 0.$$

Therefore, we need to consider the case $0 \in (I - \varphi)(\Omega)$. In this case the set $F = \{x \in \Omega : \varphi(x) = x\}$ is nonempty. Let C be the nonempty compact convex set produced in Proposition 3.7.12. From part (c) of that proposition, we have $\varphi : \overline{\Omega} \cap C \to C$. By virtue of Dugundji's extension theorem (see Proposition 2.1.9) C is a retract of X. So, let $r : X \to C$ be a retraction. Then $u \to (\varphi \circ r)(u)$ is a compact map and $r^{-1}(\Omega)$ is open. We have $0 \notin (I - \varphi \circ r)(\partial(\Omega \cap r^{-1}(\Omega)))$. So, we can define $d_{LS}(I - \varphi \circ r, \Omega \cap r^{-1}(\Omega), 0)$. We set

$$d_C(I - \varphi, \Omega, 0) = d_{LS}(I - \varphi \circ r, \Omega \cap r^{-1}(\Omega), 0). \tag{3.101}$$

Of course, in principle this definition depends on the choice of the retraction r and on C from Proposition 3.7.11. We will show that this is not the case.

So, let $r_1, r_2 : X \to C$ be two retractions onto C. Let

$$h_t(x) = tr_1(x) + (1 - t)r_2(x) \text{ for all } (t, x) \in [0, 1] \times X.$$

Clearly, for each $t \in [0, 1]$, $h_t(\cdot)$ is a retraction of X onto C. Also, $x = (\varphi \circ h_t)(x)$ for all $(t, x) \in [0, 1] \times \partial(\Omega \cap r_1^{-}(\Omega) \cap r_2^{-}(\Omega))$. So, from the homotopy invariance property, we have

$$
\begin{aligned}
&d_{LS}(I - \varphi \circ r_1, \Omega \cap r_1^{-1}(\Omega) \cap r_2^{-1}(\Omega), 0) = \\
&d_{LS}(I - \varphi \circ r_2, \Omega \cap r_1^{-1}(\Omega) \cap r_2^{-1}(\Omega), 0).
\end{aligned} \tag{3.102}
$$

We can easily check that

$$0 \notin \overline{\Omega \cap r_1^{-1}(\Omega)} \setminus (\Omega \cap r_1^{-1}(\Omega) \cap r_2^{-1}(\Omega)) \text{ and } 0 \notin \overline{\Omega \cap r_2^{-1}(\Omega)} \setminus (\Omega \cap r_1^{-1}(\Omega) \cap r_2^{-1}(\Omega)).$$

Then the excision property of the Leray–Schauder degree implies that

$$d_{LS}(I - \varphi \circ r_1, \Omega \cap r_1^{-1}(\Omega), 0) = d_{LS}(I - \varphi \circ r_1, \Omega \cap r_1^{-1}(\Omega) \cap r_2^{-1}(\Omega), 0), \tag{3.103}$$
$$d_{LS}(I - \varphi \circ r_2, \Omega \cap r_2^{-1}(\Omega), 0) = d_{LS}(I - \varphi \circ r_2, \Omega \cap r_1^{-1}(\Omega) \cap r_2^{-1}(\Omega), 0). \tag{3.104}$$

From (3.102), (3.103) and (3.104), we infer that

$$d_{LS}(I - \varphi \circ r_1, \Omega \cap r_1^{-1}(\Omega), 0) = d_{LS}(I - \varphi \circ r_2, \Omega \cap r_2^{-1}(\Omega), 0).$$

Moreover, the excision property of the Leray–Schauder degree shows that $d_C(I - \varphi, \Omega, 0)$ in (3.101) is also independent of C as in Proposition 3.7.12.

Therefore the following definition makes sense.

Definition 3.7.13 Let X be a Banach space, $\Omega \subseteq X$ be bounded open and $\varphi : \overline{\Omega} \to X$ be γ-condensing. Suppose that $0 \notin (I - \varphi)(\partial\Omega)$ and let $C \subseteq X$ be a compact convex set as in Proposition 3.7.12 and $r : X \to C$ a retraction onto C. We define

$$d_C(I - \varphi, \Omega, 0) = d_{LS}(I - \varphi \circ r, \Omega \cap r^{-1}(\Omega), 0).$$

This new degree has the usual properties.

Theorem 3.7.14 *If X is a Banach space and*

$$\tau_C = \big\{(I - \varphi, \Omega, 0) : \Omega \subseteq X \textit{ is bounded open}, \ \varphi : \overline{\Omega} \to X \\ \textit{is } \gamma\textit{-condensing and } 0 \notin (I - \varphi)(\partial\Omega)\big\},$$

then there exists a map $d_C : \tau_C \to \mathbb{Z}$ such that

(a) Normalization: $d_C(I, \Omega, 0) = 1$ provided $0 \in \Omega$.
(b) Domain Additivity: $d_C(I - \varphi, \Omega, 0) = d_C(I - \varphi, \Omega_1, 0) + d_C(I - \varphi, \Omega_2, 0)$ with $\Omega_1, \Omega_2 \subseteq \Omega$ disjoint open and $0 \notin (I - \varphi)(\overline{\Omega} \setminus (\Omega_1 \cup \Omega_2))$.
(c) Homotopy Invariance: $d_C(I - h_t, \Omega, 0)$ is independent of $t \in [0, 1]$ when $h : [0, 1] \times \overline{\Omega} \to X$ is a γ-condensing map and $h_t(u) \neq u$ for all $(t, u) \in [0, 1] \times \partial\Omega$.
(d) Solution Property: $d_C(I - \varphi, \Omega, 0) \neq 0$ implies that there exists a $u \in \Omega$ such that $\varphi(u) = u$.

Proof Properties (a), (b) and (d) follow from Definition 3.7.13 and the corresponding properties of the Leray–Schauder degree.

It remains to prove (c) (the homotopy invariance property). To this end let

$$C_0 = \overline{\operatorname{conv}}\, h([0, 1] \times \overline{\Omega}) \text{ and } C_n = \overline{\operatorname{conv}}\, h([0, 1] \times (\overline{\Omega} \cap C_{n+1})) \text{ for all } n \geqslant 1.$$

Then $C = \bigcap_{n \geqslant 1} C_n$ is nonempty, compact, convex (see Proposition 3.9.10) and $h : [0, 1] \times C \to C$. Let $r : X \to C$ be a retraction. Then $h_t(r(x)) \neq x$ for all $x \in \partial(\Omega \cap r^{-1}(\Omega \cap C))$. So the homotopy invariance property of the Leray–Schauder degree implies that

$$d_{LS}(I - h_t \circ r, \Omega \cap r^{-1}(\Omega \cap C), 0)$$

is independent of $t \in [0, 1]$. Then (c) follows from Definition 3.7.13 and the excision property of the Leray–Schauder degree. $\square$

In the second part of this section, we exploit the Galerkin approximation method to define a generalized degree theory for a large class of maps, called A-proper maps.

Definition 3.7.15 Let X be a Banach space. If there is a sequence $\{X_n\}_{n \geqslant 1}$ of finite-dimensional subspaces of X and a sequence $\{P_n\}_{n \geqslant 1}$ of continuous, linear projections from X onto X_n such that $P_n(x) \to x$ in X as $n \to \infty$ for all $x \in X$, then we say that X has a projection scheme $\{P_n, X_n\}_{n \geqslant 1}$.

Remark 3.7.16 Evidently X is separable, $X = \overline{\bigcup_{n \geqslant 1} X_n}$ and $\sup_{n \geqslant 1} ||P_n||_{\mathscr{L}} = c < +\infty$ (by the uniform boundedness principle).

Proposition 3.7.17 *If X is a Banach space with a Schauder basis $\{e_n\}_{n \geqslant 1}$ and for all $n \geqslant 1$*

$$X_n = \text{span}\,\{e_n\}_{k=1}^n \text{ and } P_n(x) = \sum_{k=1}^n \lambda_k(x) e_k$$

where $x = \sum_{n \geqslant 1} \lambda_n(x) e_n$ (see Definition 2.3.10), then $\{P_n, X_n\}_{n \geqslant 1}$ is a projection scheme.

Remark 3.7.18 In the above case $\{X_n\}_{n \geqslant 1}$ is an increasing sequence of finite-dimensional subspaces of X and $P_n \circ P_m = P_l$ with $l = \min\{m, n\}$. If X is a separable Hilbert space, then we can choose an orthonormal basis $\{e_n\}_{n \geqslant 1}$ and then P_n are the orthogonal projections $P_n(u) = \sum_{k=1}^n (e_k, u)_H e_k$ which satisfy $P_n = P_n^*$ and $||P_n||_{\mathscr{L}} = 1$ for all $n \geqslant 1$.

Proposition 3.7.19 *If X is a reflexive Banach space and $\{P_n, X_n\}_{n \geqslant 1}$ is a projection scheme on X such that $P_n \circ P_m = P_l$ with $l = \min\{m, n\}$, then $\{P_n^*, P_n^*(X^*)\}_{n \geqslant 1}$ is a projection scheme on X^*.*

Proof We have

$$\begin{aligned} &\langle P_n^* P_n^*(u^*), u\rangle = \langle u^*, P_n^2(u)\rangle = \langle P_n^*(u^*), u\rangle \text{ for all } u \in X,\ u^* \in X^* \\ \Rightarrow\ &P_n^* : X^* \to X_n^* = P_n(X^*) \text{ is a projection operator for every } n \geqslant 1. \end{aligned}$$

From Theorem 2.11.5, we know that

$$\dim X_n^* = \dim (I - P_n^*) = \dim (I - P_n) = \dim X_n \text{ for all } n \geqslant 1.$$

So, it remains to show that $P_n^*(u^*) \to u^*$ in X^* as $n \to \infty$.

First note that $X^* = \overline{\bigcup_{n\geqslant 1} X_n^*}$. Indeed, if this is not true we can find $u \in X\backslash\{0\}$ such that $u|_{\overline{\bigcup_{n\geqslant 1} X_n^*}} = 0$. Then $\langle u^*, P_n(u)\rangle = 0$ for all $u^* \in X^*$ and all $n \geqslant 1$, hence passing to the limit as $n \to \infty$, we obtain $\langle u^*, u\rangle = 0$ for all $u^* \in X^*$ and so $u = 0$, a contradiction. Note that for $m \geqslant n$, $P_m \circ P_n = P_n$ and so $P_m^* \circ P_n^* = P_n^*$, which in turn implies that $\{X_n^*\}_{n\geqslant 1}$ is an increasing sequence of subspaces, that is, $X_n^* \subseteq X_{n+1}^*$ for all $n \geqslant 1$. Given $u^* \in X^*$ and $\epsilon > 0$, we can find $n_0 \geqslant 1$ and $y^* \in X_{n_0}^*$ such that $||u^* - y^*||_* < \epsilon$. Then

$$\begin{aligned}
&||P_n^*(u^*) - u^*||_* \leqslant ||P_n^*(u^* - y^*)||_* + ||y^* - u^*||_* \leqslant \epsilon(c+1) \text{ for all } n \geqslant n_0\\
\Rightarrow\ & P_n^*(u^*) \to u^* \text{ in } X^* \text{ as } n \to \infty\\
\Rightarrow\ & \{P_n^*, X_n^* = P_n^*(X)\}_{n\geqslant 1} \text{ is a projection scheme on } X^*.
\end{aligned}$$

The proof is now complete. □

Definition 3.7.20 Let X, Y be Banach spaces with projection schemes

$$\{P_n, X_n\}_{n\geqslant 1} \text{ and } \{Q_n, Y_n\}_{n\geqslant 1} \text{ respectively}$$

and $\dim X_n = \dim Y_n$. We say that $\Pi = \{P_n, X_n; Q_n, Y_n\}$ is an operator projection scheme.

Proposition 3.7.21 *If both Banach spaces X and Y have Schauder bases, then the pair admits an operator projection scheme.*

Remember that our goal is to use Galerkin approximations in order to develop degree theoretic tools which will allow us to solve the operator equation $\varphi(u) = y$. So, suppose Π is an operator projection scheme. We are looking for appropriate ("proper") maps $\varphi : D \subseteq X \to Y$ for which the Galerkin method will work. So, suppose that

$$Q_n(\varphi(u)) = Q_n y$$

has a solution $u_n \in D_n = D \cap X_n$ for all large $n \geqslant 1$ and $\{u_n\}_{n\geqslant 1} \subseteq X$ is bounded. We have

$$Q_n(\varphi(u)) = Q_n y \to y \text{ in } Y.$$

So, we should require that $\{u_n\}_{n\geqslant 1}$ admits a subsequence converging to a solution of $\varphi(u) = y$. This is the starting point of the so-called A-proper maps.

Definition 3.7.22 Let X, Y be Banach spaces and $\Pi = \{P_n, X_n; Q_n, Y_n\}$ an operator projection scheme. Let $\varphi : D \subseteq X \to Y$. We say that φ is "A-proper" with respect to

Π if for any bounded $u_m \in D_m = D \cap X_m, m \geqslant 1$, such that $Q_m(\varphi(u_m)) \to y$ in Y as $m \to \infty$, we can find a subsequence $\{u_{m_k}\}_{k\geqslant 1} \subseteq \{u_m\}_{m\geqslant 1}$ such that $u_{m_k} \to u \in D$ and $\varphi(u) = y$. By $A_\Pi(D, Y)$ we denote the family of all maps $\varphi : D \subseteq X \to Y$ which are A-proper with respect to Π. If $X = Y$, then we write $A_\Pi(D)$.

Remark 3.7.23 The condition of A-properness is very weak since it does not require solvability of the finite-dimensional approximation. Even if those finite-dimensional equations are solvable and the solutions are uniformly bounded, then only a subsequence converges to a solution of the original equation. So, the family of maps introduced in Definition 3.7.22 is large enough to include the operators considered earlier. On the other hand, understandably A-proper maps are not amenable to computations.

Proposition 3.7.24 *If X is a Banach space and $\{P_n, X_n\}$ is a projection schema on X such that $\sup_{n\geqslant 1} ||P_n||_{\mathscr{L}} = c < \infty$, $D \subseteq X$ is closed and $\varphi : D \to X$ is β-Lipschitz with constant $k > 0$, then $\lambda I - \varphi \in A_\Pi(D)$ for all $\lambda > kc$.*

Proof It suffices to show that for every bounded set $B \subseteq X$ we have

$$\beta(B) \leqslant \beta(\bigcup_{n\geqslant 1} P_n(B)) = \lim_{m\to\infty} \beta(\bigcup_{n\geqslant m} P_n(B)) \leqslant \beta(B)c. \tag{3.105}$$

Since $B \subseteq \bigcup_{n\geqslant 1} P_n(B)$, the first inequality in (3.105) is immediate. Note that $\bigcup_{n=1}^{m} P_n(B)$ is relatively compact, being bounded in the finite-dimensional subspace X_m. Therefore

$$\beta(\bigcup_{n\geqslant 1} P_n(B)) = \lim_{m\to\infty} \beta(\bigcup_{n\geqslant m} P_n(B))$$

(see Proposition 3.7.6). Next, suppose that $B \subseteq \bigcup_{k=1}^{m} B_r(x_k)$. Then

$$P_n(B) \subseteq \bigcup_{k=1}^{m} B_{cr+\epsilon}(x_k) \text{ with } \epsilon = \max_{1\leqslant k\leqslant m} ||P_n(x_k) - x_k||$$

and so we conclude that

$$\lim_{m\to\infty} \beta(\bigcup_{n\geqslant m} P_n(B)) \leqslant \beta(B)c.$$

This proves (3.105) and also the proposition. □

Proposition 3.7.25 *If X is a reflexive Banach space with a projection scheme $\{P_n, X_n\}$ such that $P_n \circ P_m = P_l$ with $l = \min\{n, m\}$, $\Pi = \{P_n, X_n; P_n^*, X_n^*\}$ (an* *operator projection scheme by Proposition 3.7.21) and $\varphi : X \to X^*$ is demicontinuous and strongly monotone in the sense that*

$$\langle \varphi(u) - \varphi(y), u - y \rangle \geqslant \vartheta(||u - y||)||u - y||$$

with $\vartheta : \mathbb{R}_+ \to \mathbb{R}_+$ a continuous function such that $\vartheta(0) = 0,\ \vartheta(r) > 0$ for all $r > 0$, then $\varphi \in A_\Pi(X, X^*)$.

Proof Let $\{u_m\}_{m \geqslant 1}$, $u_m \in X_m$ for all $m \geqslant 1$, be a bounded sequence and assume that $P_m^*(\varphi(u_m)) \to y^*$ in X^*. Because of the reflexivity of X and by passing to a suitable subsequence if necessary, we may assume that $u_m \xrightarrow{w} u_0$ in X. We have

$$\begin{aligned}
&\limsup_{m \to \infty} \vartheta(||u_m - u_0||)||u_m - P_m(u_0)|| \\
&\leqslant \limsup_{m \to \infty} \left\langle P_m^*(\varphi(u_m)) - P_m^*(\varphi(P_n(u_0))), u_m - P_n(u_0) \right\rangle \\
&= \left\langle y^* - \varphi(P_n(u_0)), u_0 - P_n(u_0) \right\rangle \to 0 \text{ as } n \to \infty \\
&\Rightarrow u_m \to u_0 \text{ in } X \text{ and } \varphi(u_0) = y^*.
\end{aligned}$$

The proof is now complete. □

Note that $\pm i \in A_\prod(X)$ (i being the identity map on X). But $0 \notin A_\prod(X)$ and so $A_\prod(X)$ is not a linear space. However, it is clear that if $\varphi \in A_\prod(X)$ and $\lambda \neq 0$, then $\lambda\varphi \in A_\prod(X)$. Also, we have

Proposition 3.7.26 *If X, Y are Banach spaces, $D \subset X$, $\prod = \{P_n, X_n; Q_n, Y_n\}$ is an operator projection scheme, $\varphi \in A_\prod(D, Y)$ and $\psi : D \to Y$ is compact, then $\varphi + \psi \in A_\prod(D, Y)$.*

Proof Let $\{u_n\}_{n \geqslant 1} \subseteq D$ be a bounded sequence and $Q_n((\varphi + \psi)(u_n)) \to y$ in Y. Then we can find a subsequence $\{u_m\}$ of $\{u_n\}$ such that $\psi(u_m) \to \hat{y}$ in Y. Then

$$\begin{aligned}
&||Q_m(\psi(u_m)) - \hat{y}|| = ||Q_m(\psi(u_m)) - Q_m(\hat{y})|| + ||Q_m(\hat{y}) - \hat{y}|| \to 0 \\
&\Rightarrow Q_m(\varphi(u_m)) \to y - \hat{y} \text{ in } Y.
\end{aligned}$$

Since φ is A-proper, we can find a subsequence $\{u_{n_k}\}_{k \geqslant 1}$ of $\{u_n\}_{n \geqslant 1}$ such that $u_{n_k} \to$ in X. Then $\varphi(u) = y - \hat{y}$. Also, $\psi(u_{n_k}) \to \psi(u) = \hat{y}$ in Y. Therefore

$$(\varphi + \psi)(u) = y,$$

which completes the proof. □

Proposition 3.7.27 *If X, Y are Banach spaces, $D \subseteq X$ is closed and bounded, $\prod = \{P_n, X_n; Q_n, Y_n\}$ is an operator projection scheme on (X, Y) and $\varphi \in A_\prod(D, Y)$, then φ is closed (that is, maps closed sets to closed sets) and proper (see Definition 2.2.1).*

Proof Let $C \subset D$ be closed and assume that $\varphi(u_n) \to y$ in X with $u_n \in C$ for all $n \geqslant 1$. From Definition 3.7.15, we have $P_m(u_n) \to u_n$ in X as $m \to \infty$ for all $n \geqslant 1$ and $\varphi(P_m(u_n)) \to \varphi(u_n)$ in X as $m \to \infty$ for all $n \geqslant 1$. So, given $n \in \mathbb{N}$, we can find $m_n \geqslant n$ such that

$$||u_n - P_{m_n}(u_n)|| \leqslant \frac{1}{n} \text{ and } ||\varphi(u_n) - \varphi(P_{m_n}(u_n))|| \leqslant \frac{1}{n}. \tag{3.106}$$

Let $v_n = P_{m_n}(u_n)$ for all $n \geqslant 1$. From (3.106) it follows that

$$||u_n - v_n|| \to 0 \text{ and } ||\varphi(u_n) - \varphi(v_n)|| \to 0 \text{ as } n \to \infty. \tag{3.107}$$

Therefore

$$\begin{aligned}||Q_{m_n}(\varphi(v_n)) - y|| = ||Q_{m_n}(\varphi(v_n)) - Q_{m_n}(\varphi(u_n))|| + ||Q_{m_n}(\varphi(u_n)) - Q_{m_n}(y)||\\ +||Q_{m_n}(y) - y|| \to 0 \text{ as } n \to \infty \text{ (see (3.107))}.\end{aligned}$$

Because φ is A-proper we can find a subsequence $\{v_{n_k}\}_{k\geqslant 1}$ of $\{v_n\}_{n\geqslant 1}$ such that $v_{n_k} = P_{m_{n_k}}(u_{n_k}) \to u \in D$ in X and $\varphi(u) = y$. But then (3.107) implies that

$$u_{n_k} \to u \text{ in } X \text{ and } u \in C \text{ (since } C \text{ is closed)}$$

$$\Rightarrow \varphi(u) = y \in \varphi(C) \text{ and so } \varphi \text{ is closed.}$$

Next, we show that φ is proper. So, let $K \subseteq Y$ be compact and $\{u_n\}_{n\geqslant 1} \subset \varphi^{-1}(K)$. Then $\varphi(u_n) \in K$ for all $n \geqslant 1$ and so we may assume that $\varphi(u_n) \to y \in K$ in Y. Then reasoning as in the first part of the proof, we can produce a subsequence $\{u_{n_k}\}_{k\geqslant 1}$ of $\{u_n\}_{n\geqslant 1}$ such that $u_{n_k} \to u \in D$ in X and $\varphi(u) = y$. Therefore $u \in \varphi^{-1}(K)$ and we conclude that $\varphi^{-1}(K)$ is compact, that is, φ is proper. $\square$

Corollary 3.7.28 *If X, Y are Banach spaces, $D \subseteq X$ is closed and bounded, $\prod = \{P_n, X_n; Q_n, Y_n\}$ is an operator projection scheme on (X, Y), $\varphi \in A_{\prod}(D, Y)$, $C \subseteq D$ is closed and $\varphi(u) \neq y$ for all $u \in \partial C$, then there exists a $c > 0$ such that $||\varphi(u) - y|| \geqslant c$ for all $u \in \partial C$.*

Proof From Proposition 3.7.27 we know that $\varphi(\partial C) \subseteq Y$ is closed. Suppose that the result is not true. Then we can find $\{u_n\}_{n\geqslant 1} \subseteq \partial C$ such that $||\varphi(u_n) - y|| < \frac{1}{n}$. So, $\varphi(u_n) \to y$ in Y hence $y \in \partial\Omega$, which contradicts the hypothesis that $y \notin \varphi(\partial C)$. $\square$

The next theorem is an important surjectivity result for A-proper maps and can be viewed as the counterpart of Theorem 2.8.5 (see also Theorem 2.8.6) when maximal monotonicity is replaced by A-properness.

Theorem 3.7.29 *If X, Y are Banach spaces, $\prod = \{P_n, X_n; Q_n, Y_n\}$ is an operator projection scheme on (X, Y) and $\varphi : X \to Y$ is a continuous map such that*

$$||(Q_n \circ \varphi)(u) - (Q_n \circ \varphi)(v)|| \geqslant \vartheta(||u - v||) \textit{ for all } u, v \in X \textit{ and all } n \geqslant 1$$

with $\vartheta : \mathbb{R}_+ \to \mathbb{R}_+$ a continuous function, $\vartheta(0) = 0$, $\vartheta(r) > 0$ for all $r > 0$ and $\vartheta(r) \to +\infty$ as $r \to +\infty$, then $\varphi \in A_{\prod}(X, Y)$ if and only if φ is surjective.

Proof $\underline{\Rightarrow}$: Evidently $Q_n \circ \varphi$ is injective, continuous and $||(Q_n \circ \varphi)(u)|| \to +\infty$ as $||u|| \to \infty$. From Theorem 3.1.48 it follows that $(Q_n \circ \varphi)(X_n)$ is open in X_n. We claim that $(Q_n \circ \varphi)(X_n)$ is also closed in X_n. Indeed, if $(Q_n \circ \varphi)(u_m) \to y$ in X_n as $m \to \infty$, then from the coercivity of $Q_n \circ \varphi$ we have that $\{u_m\}_{m \geqslant 1} \subseteq X_n$ is bounded. So, we may assume that $u_m \to u$ in X_n and then $(Q_n \circ \varphi)(u) = y$. This proves that $(Q_n \circ \varphi)(X_n)$ is closed. The connectedness of X_n implies that $(Q_n \circ \varphi)(X_n) = X_n$ for all $n \geqslant 1$ (that is, $Q_n \circ \varphi$ is surjective for every $n \geqslant 1$).

So, for every $y \in Y$, there exists a unique $u_n \in X_n$ such that

$$\begin{aligned}
&(Q_n \circ \varphi)(u_n) = Q_n(y) \text{ for all } n \geqslant 1 \\
\Rightarrow\ &(Q_n \circ \varphi)(u_n) \to y \text{ in } Y \text{ as } n \to \infty.
\end{aligned}$$

From the hypothesis on $(Q_n \circ \varphi)(\cdot)$, it follows that $\{u_n\}_{n \geqslant 1} \subseteq X$ is bounded. Because φ is A-proper, we can find a subsequence $\{u_{n_k}\}_{k \geqslant 1}$ of $\{u_n\}_{n \geqslant 1}$ such that

$$u_{n_k} \to u \text{ in } X \ \text{ and } \varphi(u) = y.$$

Evidently, $u \in X$ is unique.

$\underline{\Leftarrow}$: Let $\{u_n\}_{n \geqslant 1} \subset X$ be bounded and assume that $(Q_n \circ \varphi)(u_n) \to y$ in Y. Since φ is surjective, there exists a $u \in X$ such that $\varphi(u) = y$. Since $P_n(u) \to$ in X, we have that $(Q_n \circ \varphi)(P_n(u_n)) \to \varphi(u) = y$ in Y. From our hypothesis on $Q_n \circ \varphi$, we have

$$\begin{aligned}
\vartheta(||u_n - P_n(u)||) &\leqslant ||(Q_n \circ \varphi)(u_n) - (Q_n \circ \varphi)(P_n(u))|| \\
&\leqslant ||(Q_n \circ \varphi)(u_n)' - y|| + ||(Q_n \circ \varphi)(P_n(u)) - Q_n(y)|| \\
&\quad + ||Q_n(y) - y|| \to 0
\end{aligned}$$

$\Rightarrow ||u_n - P_n(u)|| \to 0$
$\Rightarrow u_n \to u$ in X and so we conclude that $\varphi \in A_{\prod}(X, Y)$.

This completes the proof. ☐

Corollary 3.7.30 *If X, Y are Banach spaces, $\prod = \{P_n, X_n; Q_n, Y_n\}$ is an operator projection scheme and $A \in \mathscr{L}(X, Y)$, then the following statements are equivalent*

(a) A is injective and A-proper;
(b) $A(X) = Y$ and $||(Q_n \circ A)(u)|| \geqslant c||u||$ for some $c > 0$ and all $n \geqslant 1$.

Proof **(a)**$\Rightarrow$(b): Suppose that the equicoercivity of $\{Q_n \circ \varphi\}_{n \geqslant 1}$ is not true. Then we can find $\{u_n\}_{n \geqslant 1} \subseteq \partial B_1^X = \{u \in X : ||u|| = 1\}$ such that $||A(u_n)|| < \frac{1}{n}$ for all $n \geqslant 1$. Since A is A-proper, we can find a subsequence $\{u_{n_k}\}_{k \geqslant 1}$ of $\{u_n\}_{n \geqslant 1}$ such that $u_{n_k} \to u$ in X and $A(u) = 0$. The injectivity of A implies $u = 0$, a contradiction to the fact that $||u|| = 1$. This proves that there exists a $c > 0$ such that $||(Q_n \circ A)(u)|| \geqslant c||u||$

for all $u \in X$ and all $n \geqslant 1$ (equicoercivity). Invoking Theorem 3.7.29, we conclude that A is surjective

$$(b) \Rightarrow (a): \text{ This follows from Theorem 3.7.29, with } \vartheta(r) = cr.$$

The proof is now complete. □

Proposition 3.7.31 *If X is a reflexive Banach space with a projection scheme $\{P_n, X_n\}_{n\geqslant 1}$, $\Omega \subset X$ is bounded open and $\varphi : \overline{\Omega} \to X^*$ is a bounded, continuous, $(S)_+$-map, then $\varphi \in A_{\prod}(\overline{\Omega}, X^*)$ with $\prod = \{P_n, X_n; P_n^*, X_n^*\}$.*

Proof Let $u_n \in \overline{\Omega} \cap X_n$ and suppose that $(P_n^* \circ \varphi)(u_n) \to u^*$ in X^*. Since $\overline{\Omega}$ is bounded we may assume that $u_n \xrightarrow{w} u$ in X.

For fixed $v \in X$, we have

$$\begin{aligned} |\langle \varphi(u_n) - u^*, u_n - P_n(v)\rangle| &= |\langle \varphi(u_n) - u^*, P_n(u_n) - P_n(v)\rangle| \\ &= |\langle P_n^*(\varphi(u_n)) - P_n^*(u^*), u_n - v\rangle| \\ &\leqslant \hat{c}||P_n^*(\varphi(u_n)) - P_n^*(u^*)||_* \\ &\qquad \text{with } \hat{c} = \sup_{n\geqslant 1}\lceil ||u_n|| + ||v||\rceil . \end{aligned}$$

Note that

$$\begin{aligned} &\langle \varphi(u_n), u_n - v\rangle \to \langle u^*, u - v\rangle \\ \Rightarrow\ &\langle \varphi(u_n), u_n - u\rangle \to 0 \text{ (choosing } v = u) \\ \Rightarrow\ &u_n \to u \text{ in } X \text{ (since } \varphi \text{ is an } (S)_+\text{-map.} \end{aligned}$$

Then $\varphi(u_n) \to \varphi(u)$ in X^* and so

$$\begin{aligned} &||P_n^*(\varphi(u_n)) - \varphi(u)||_* \leqslant ||P_n^*(\varphi(u_n)) - P_n^*(\varphi(u))||_* + ||P_n^*(\varphi(u)) - \varphi(u)||_* \\ \Rightarrow\ &P_n^*(\varphi(u_n)) \to \varphi(u) \text{ in } X^* \text{ and so } \varphi(u) = u^* \\ \Rightarrow\ &\varphi \in A_{\prod}(\overline{\Omega}, X^*). \end{aligned}$$

The proof is now complete. □

Corollary 3.7.32 *If X is a reflexive Banach space with a projection scheme $\{P_n, X_n\}$, $\Omega \subseteq X$ is bounded open and $\varphi : \overline{\Omega} \to X^*$ is bounded continuous and*

$$\langle \varphi(u) - \varphi(v), u - v\rangle \geqslant \vartheta(||u - v||) \text{ for all } u, v \in \overline{\Omega} \tag{3.108}$$

with $\vartheta : \mathbb{R}_+ \to \mathbb{R}_+$ continuous, $\vartheta(0) = 0$, $\vartheta(r) > 0$, for all $r > 0$ and $\vartheta(r) \to +\infty$ as $r \to +\infty$, then $\varphi \in A_{\prod}(\overline{\Omega}, X^*)$ where $\prod = \{P_n, X_n; P_n^*, X_n^*\}$.

Proof Clearly (3.108) implies that φ is $(S)_+$ and so we can apply Proposition 3.7.31. □

Proposition 3.7.33 *If X is a reflexive Banach space normed in such a way that both X and its dual X^* are locally uniformly convex (see Theorem 2.7.36), $\{P_n, X_n\}$ is a projection scheme on X, and $J : X \to X^*$ is the corresponding duality map, then $J \in A_{\prod}(X, X^*)$ with $\prod = \{P_n, X_n; P_n^*, X_n^*\}$.*

Proof Consider a bounded sequence $\{u_n\}_{n \geqslant 1} \subseteq X$ with $u_n \in X_n$ for all $n \geqslant 1$ and assume that

$$(P_n^* \circ J)(u_n) \to u^* \text{ in } X^* \text{ as } n \to \infty. \tag{3.109}$$

We may assume that $u_n \xrightarrow{w} v$ in X as $n \to \infty$. We have

$$\begin{aligned}\langle J(y) - J(v), y - v\rangle =&(||y|| - ||v||)^2 + (||J(y)||_*||v|| - \langle J(y), v\rangle) \\ &+ (||J(v)||_*||y|| - \langle J(v), y\rangle) \\ \geqslant&(||y|| - ||v||)^2 \text{ for all } y, v \in X\end{aligned}$$

$$\Rightarrow \langle J(u_n) - J(P_n(u)), u_n - P_n(u)\rangle \geqslant (||u_n|| - ||P_n(u)||)^2. \tag{3.110}$$

Since $J : X_n \to X_n^*$, we have $P_n^*(J(u_n)) = J(u_n)$ for all $n \geqslant 1$, hence

$$J(u_n) \to u^* \text{ in } X^* \text{ as } n \to \infty \text{ (see (3.109))}. \tag{3.111}$$

Also, from Proposition 2.7.33, we know that J is a homeomorphism. So,

$$J(P_n(u)) \to J(u) \text{ in } X^* \text{ as } n \to \infty. \tag{3.112}$$

If we return to (3.110), pass to the limit and use (3.111) and (3.112), we obtain

$$\begin{aligned}&||u_n|| - ||P_n(u)|| \to 0 \\ \Rightarrow\ &||u_n|| \to ||u||.\end{aligned} \tag{3.113}$$

But X has the Kadec–Klee property (see Remark 2.7.30). So, from (3.113) and since $u_n \xrightarrow{w} u$ in X as $n \to \infty$, we infer that $u_n \to u$ in X. Then $J(u_n) \to J(u)$ in X^* as $n \to \infty$ (see Proposition 2.7.33). Hence

$$\begin{aligned}&||P_n^*(J(u_n)) - J(u)||_* \leqslant ||P_n^*(J(u_n)) - P_n^*(J(u))||_* + ||P_n^*(J(u)) - J(u)||_* \\ \Rightarrow\ &P_n^*(J(u_n)) \to J(u) \text{ in } X^* \text{ as } n \to \infty \\ \Rightarrow\ &J(u) = u^* \text{ (see (3.109))}.\end{aligned}$$

Therefore $J \in A_{\prod}(X, X^*)$. □

The next lemma is important for the definition of a degree for A-proper maps.

Lemma 3.7.34 *If X, Y are Banach spaces, $\prod = \{P_n, X_n; Q_n, Y_n\}$ is an operator projection scheme on (X, Y), $\Omega \subseteq X$ is bounded open, $\varphi \in A_{\prod}(\overline{\Omega}, Y)$ and $y \notin \varphi(\partial\Omega)$, then there exist $n_0 \in \mathbb{N}$ and $c > 0$ such that*

$$||\varphi_n(u) - y|| \geqslant c \textit{ for all } u \in \partial(\Omega \cap X_n) \textit{ and all } n \geqslant n_0.$$

Proof We argue by contradiction. So, suppose that the lemma is not true. Then we can find $u_n \in \partial(\Omega \cap X_n)$ such that $(Q_n \circ \varphi)(u_n) \to y$ in Y. Since φ is A-proper, we can find a subsequence $\{u_{n_k}\}_{k \geqslant 1}$ of $\{u_n\}_{n \geqslant 1}$ such that $u_{n_k} \to u \in \overline{\Omega}$ in X as $k \to \infty$ and $\varphi(u) = y$. Evidently $u \in \partial\Omega$ and so $y \in \varphi(\partial\Omega)$, a contradiction. □

Definition 3.7.35 Let X, Y be Banach spaces, $\prod = \{P_n, X_n; Q_n, Y_n\}$ be an operator projection scheme with X_n, Y_n oriented, $\Omega \subseteq X$ be bounded open, $\varphi \in A_{\prod}(\overline{\Omega}, Y)$ and $\xi \notin \varphi(\partial\Omega)$. We define a multivalued degree on the triple (φ, Ω, ξ) by setting

$$d_A(\varphi, \Omega, \xi) = \{k \in \mathbb{Z} \cup \{\pm\infty\} : d(Q_n \circ \varphi, \Omega \cap X_n, Q_n(\xi)) \to k \text{ as } n \to \infty\}.$$

Remark 3.7.36 Lemma 3.7.34 guarantees that for large $n \geqslant 1$ we have $Q_n(\xi) \notin (Q_n \circ \varphi)(\partial(\Omega \cap X_n))$ and so in the above definition the Brouwer degree $d(Q_n \circ \varphi, \Omega \cap X_n, Q_n(\xi))$ makes sense. Also, if $k \in \mathbb{Z}$, then $d(Q_n \circ \varphi, \Omega \cap X_n, Q_n(\xi)) = k$ for all large $n \geqslant 1$.

The new degree exhibits the usual properties. Only the domain additivity will hold in special cases, since there is no reasonable addition on $\mathbb{Z} \cup \{\pm\infty\}$.

Theorem 3.7.37 *If X, Y are Banach spaces, $\prod = \{P_n, X_m; Q_n, Y_n\}$ is an operator projection scheme on (X, Y) with X_n, Y_n oriented and*

$$\tau_A = \{(\varphi, \Omega, \xi) : \Omega \subseteq X \textit{is bounded}, \varphi \in A_{\prod}(\overline{\Omega}, Y),\ \xi \notin \varphi(\partial\Omega)\},$$

then there exists a map $d_A : \tau_A \to 2^{\mathbb{Z} \cup \{\pm\infty\}}$ such that

(a) Normalization: $d_A(\varphi, \Omega, \xi) \neq \emptyset$.
(b) Homotopy Invariance: if $h : [0, 1] \times \overline{\Omega} \to Y$ is such that $h(\cdot, x)$ is continuous on $[0, 1]$ uniformly with respect to $u \in \overline{\Omega}$ and $h(t, \cdot) \in A_{\prod}(\overline{\Omega}, Y)$ for all $t \in [0, 1]$ then for $\xi \notin h([0, 1] \times \partial\Omega)$

$$d_A(h(t, \cdot), \Omega, \xi)$$

is independent of $t \in [0, 1]$.
(c) Solution Property: if $d_A(\varphi, \Omega, \xi) \neq \{0\}$, then there exists a $u \in \Omega$ such that $\varphi(u) = \xi$.
(d) Odd Maps: if Ω is symmetric with $0 \in \Omega$, φ is odd on $\overline{\Omega}$ and $0 \notin \varphi(\partial\Omega)$, then for all $m \in \mathbb{Z}$, $2m \notin d_A(\varphi, \Omega, 0)$.

Proof (a) Let $n_0 \geqslant 1$ be as postulated by Proposition 2.7.31. We consider the sequence

$$\{d(Q_n \circ \varphi, \Omega \cap X_n, Q_n(\xi))\}_{n \geqslant n_0}.$$

If this sequence is bounded, then there exists a $k \in \mathbb{Z}$ with $k \in d_A(\varphi, \Omega, \xi)$. Otherwise either $+\infty$ or $-\infty$ are in $d_A(\varphi, \Omega, \xi)$.

(b) It suffices to show that

$$Q_n(\xi) \notin (Q_n \circ h)([0,1] \times \partial(\Omega \cap X_n)) \text{ for all large } n \geqslant 1, \tag{3.114}$$

since in this case, by virtue of the homotopy invariance property of the Brouwer degree, $d((Q_n \circ h), \Omega \cap X_n, Q_n(\xi))$ is independent of $t \in [0, 1]$. So, suppose that (3.114) is not true. Then we can find $\{t_n\}_{n\geqslant 1} \subseteq [0, 1]$ and $\{u_n\}_{n\geqslant 1} \subseteq \partial(\Omega \cap X_n)$ such that

$$t_n \to t \text{ and } (Q_n \circ h)(t_n, u_n) = Q_n(\xi) \text{ for all } n \geqslant 1. \tag{3.115}$$

By hypothesis $h(t_n, \cdot) \to h(t, \cdot)$ uniformly on $\overline{\Omega}$, and we have

$$h(t_n, u_n) - h(t, u_n) \to 0 \text{ in } Y \text{ as } n \to \infty$$

$$\Rightarrow ||(Q_n \circ h)(t_n, u_n) - \xi|| \to 0 \text{ as } n \to \infty \text{ (see (3.115))}.$$

Since $h(t, \cdot) \in A_{\prod}(\overline{\Omega}, Y)$, we can find a subsequence $\{u_{n_k}\}_{k\geqslant 1} \subseteq \{u_n\}_{n\geqslant 1}$ such that $u_{n_k} \to u \in \partial\Omega$ in X as $k \to \infty$ and $h(t, u) = \xi$, hence $\xi \in h([0, 1] \times \partial\Omega)$, a contradiction.

(c) If $d_A(\varphi, \Omega, \xi) \neq \{0\}$, then $d(Q_n \circ \varphi, \Omega \cap X_n, Q_n(\xi)) \neq 0$ for some sequence. From the solution property of the Brouwer degree, we can find $u_n \in \Omega \cap X_n$ such that

$$(Q_n \circ \varphi)(u_n) = Q_n(\xi).$$

We have $Q_n(\xi) \to \xi$ in Y as $n \to \infty$. Since $\varphi \in A_{\prod}(\overline{\Omega}, Y)$, we can find a subsequence $\{u_{n_k}\}_{k\geqslant 1}$ of $\{u_n\}_{n\geqslant 1}$ such that $u_{n_k} \to u \in \overline{\Omega}$ and $\varphi(u) = \xi$. Evidently, $u \in \Omega$.

(d) Note that $Q_n \circ \varphi$ is odd. So, by Theorem 3.1.44 $d(Q_n \circ \varphi, \Omega \cap X_n, 0)$ is odd for all $n \geqslant 1$. □

Proposition 3.7.38 *If X is a reflexive Banach space normed in such a way such that both X and X^* are locally uniformly convex (see Theorem 2.7.36), $\{P_n, X_n\}$ is a projection scheme on X and $J : X \to X^*$ is the duality map, then for every $r > 0$ and $u^* \in X^*$, we have*

$$d_A(J, B_r, u^*) \neq \{0\} \textit{ if } ||u^*||_* < r \textit{ and } d_A(J, B_r, u^*) = \{0\} \textit{ if } ||u^*||_* > r.$$

Proof First suppose that $||u^*||_* < r$. Let $h(t, u) = J(u) - tu^*$ for all $(t, u) \in [0, 1] \times \overline{B}_r$. From Proposition 3.7.26 and 3.7.33, we have that $h(t, \cdot) \in A_{\prod}(\overline{B}_r, X^*)$ for all $t \in [0, 1]$ with $\prod = \{P_n, X_n; P_n^*, X_n^*\}$. Moreover, $h(\cdot, u)$ is continuous on $[0, 1]$ uniformly with respect to $u \in \overline{B}_r$. Also, we claim that $h(t, u) \neq 0$ for all $(t, u) \in [0, 1] \times \partial B_r$. Indeed, if $(t_0, u_0) \in [0, 1] \times \partial B_r$ and $h(t_0, u_0) = 0$, then $J(u_0) = t_0 u^*$, hence $||J(u_0)||_* = ||u_0|| = r = t_0||u^*||_* < r$, a contradiction. Therefore, we can apply Theorem 3.7.37(b) and have

$$d_A(J, B_r, 0) = d_A(J - u^*, B_r, 0) = d_A(J, B_r, u^*). \tag{3.116}$$

The duality map J is odd. So, Theorem 3.7.37(d) implies that

$$\begin{aligned}&d_A(J, B_r, 0) \neq \{0\}\\ \Rightarrow\ &d_A(J, B_r, u^*) \neq \{0\} \text{ for all } ||u^*||_* < r.\end{aligned}$$

Next, suppose that $||u^*||_* > r$. So, $J(u) \neq u^*$ for all $u \in \partial B_r$. If $d_A(J, B_r, u^*) \neq \{0\}$, then according to Theorem 3.7.37(c), we can find $u \in \overline{B}_r$ such that $J(u) = u^*$, hence $r = ||J(u)||_* = ||u^*||_* > r$, a contradiction. So, we conclude that

$$d_A(J, B_r, u^*) = \{0\} \text{ for all } ||u^*||_* > r.$$

The proof is now complete. □

Remark 3.7.39 Suppose that X is a Banach space and $\prod = \{P_n, X_n\}$ is a projection scheme on X. Let $\Omega \subseteq X$ be bounded open and $\varphi : \overline{\Omega} \to X$ a compact map. From Proposition 3.7.26 we have that $I - \varphi \in A_{\prod}(\overline{\Omega})$. We claim that

$$d_A(I - \varphi, \overline{\Omega}, \xi) = d_{LS}(I - \varphi, \overline{\Omega}, \xi).$$

Indeed, note that $P_n(\varphi(u)) \to \varphi(u)$ uniformly on $\overline{\Omega}$. Since $\varphi(\Omega)$ is compact, given $\epsilon > 0$, we can find $\{u_k\}_{k=1}^m \subseteq \overline{\Omega}$ such that

$$\varphi(\overline{\Omega}) \subseteq \bigcup_{k=1}^{m} B_\epsilon(\varphi(u_k)).$$

Also, we can find $n_0 = n_0(\epsilon) \geqslant 1$ such that

$$||P_n(\varphi(u_k)) - \varphi(u_k)|| \leqslant \epsilon \text{ for all } n \geqslant n_0,\ \text{ all } k \in \{1, ..., m\}.$$

Then

$$\begin{aligned}&\sup_{u\in\overline{\Omega}} ||P_n(\varphi(u)) - \varphi(u)||\\ \leqslant\ &\sup_{u\in\overline{\Omega}} ||P_n(\varphi(u)) - P_n(\varphi(u_{k_0}))|| + ||P_n(\varphi(u_{k_0})) - \varphi(u_{k_0})||\\ &\qquad\qquad + ||\varphi(u_{k_0}) - \varphi(u)|| \leqslant 3\epsilon\end{aligned}$$

with $k_0 \in \{1, ..., m\}$ such that $\varphi(u) \in B_\epsilon(\varphi(u_{k_0}))$. Then for large $n \geqslant 1$ we have $d(P_n \circ I - P_n \circ \varphi, \Omega \cap X_n, P_n(\xi)) =$ constant $= d_{LS}(I - \varphi, \Omega, \xi)$ (see Proposition 3.1.43).

Therefore, we have seen that d_A extends the Leray–Schauder degree.

3.8 Index of a ξ-Point

In this section, we introduce the notion of the index of a ξ-point $u_0 \in \Omega$.

So, let $\Omega \subseteq \mathbb{R}^N$ be bounded open, $\varphi \in C(\overline{\Omega}, \mathbb{R}^N)$ and let u_0 be an isolated ξ-point, namely $\varphi(u_0) = \xi \in \mathbb{R}^N$ and there exists a neighborhood U of u_0 such that $U \cap \varphi^{-1}(\xi) = \{u_0\}$. Let $\mathscr{S} = \{U : U$ is a neighborhood of u_0 such that $\overline{U}$ contains no other ξ-point of $\varphi\}$.

Definition 3.8.1 We define the index of φ with respect to the pair (u_0, ξ) by

$$i(\varphi, u_0, \xi) = d(\varphi, U, \xi) \text{ for all } u \in \mathscr{S}.$$

Remark 3.8.2 First note that $\xi \notin \varphi(\partial\Omega)$ and so $d(\varphi, U, \xi)$ is well-defined. Also, if $U_1,\ U_2 \in \mathscr{S}$, then $U = U_1 \cup U_2 \in \mathscr{S}$ (recall that $\overline{U} = \overline{U}_1 \cup \overline{U}_2$). Let $K = \overline{U}_1 \cap U_2^c$. Then $K \subseteq \overline{U}$ and so K is compact with $\xi \notin \varphi(K)$. The excision property of the Brouwer degree (see Theorem 3.1.25(g)) implies that

$$\begin{aligned} d(\varphi, U, \xi) &= d(\varphi, U \backslash K, \xi) \\ &= d(\varphi, U \cap (\overline{U}_1^c \cup U_2), \xi) \\ &= d(\varphi, U_2, \xi). \end{aligned} \tag{3.117}$$

By exchanging the roles of U_1, U_2 in the above argument, we also show

$$\begin{aligned} &d(\varphi, U, \xi) = d(\varphi, U_2, \xi) \\ \Rightarrow\ &d(\varphi, U_1, \xi) = d(\varphi, U_2, \xi) \text{ (see (3.117))}. \end{aligned}$$

Hence the definition of the index is independent of the choice of $U \in \mathscr{S}$.

Proposition 3.8.3 *If $\Omega \subseteq \mathbb{R}^N$ is bounded open, $\varphi \in C^1(\overline{\Omega}, \mathbb{R}^N)$, $u_0 \in \Omega$, $\varphi(u_0) = \xi$ and $\varphi'(u_0)$ is invertible, then $i(\varphi, u_0, \xi) = i(\varphi'(u_0) - \varphi'(u_0)u_0, u_0, 0)$.*

Proof Since $\varphi'(u_0)$ is invertible, we can find $c > 0$ such that

$$||\varphi'(u_0)u|| \geqslant c||u|| \text{ for all } u \in X. \tag{3.118}$$

Let U be a neighborhood of u_0 such that

$$||\varphi(u) - \xi - \varphi'(u_0)(u - u_0)|| \leqslant \frac{c}{3}||u - u_0|| \text{ for all } u \in U. \tag{3.119}$$

From (3.118) and (3.119) it is clear that u_0 is an isolated ξ-point of φ. Let $h_t(u) = (1-t)\varphi(u) + t\varphi'(u_0)(u - u_0)$ for all $t \in [0, 1]$, all $u \in \overline{U}$ and $\xi(t) = (1-t)\xi$ for all $t \in [0, 1]$. We have

$$\begin{aligned}||h_t(u) - \xi(t)|| &\geqslant ||\varphi'(u_0)(u - u_0)|| - (1 - t)||\varphi(u) - \xi - \varphi'(u_0)(u - u_0)|| \\ &\geqslant \frac{2c}{3}||u - u_0|| \text{ for all } t \in [0, 1] \text{ (see (3.118), (3.119))} \\ &\Rightarrow \xi(t) \notin h_t(\partial\Omega) \text{ for all } t \in [0, 1].\end{aligned}$$

The homotopy invariance property of the Brouwer degree implies

$$\begin{aligned}&d(\varphi, U, \xi) = d(\psi, U, 0) \text{ where } \psi(u) = \varphi'(u_0)(u - u_0) \\ \Rightarrow\ &i(\varphi, u_0, \xi) = i(\varphi'(u_0) - \varphi'(u_0)u_0, u_0, 0).\end{aligned}$$

The proof is now complete. □

Remark 3.8.4 The importance of the above result comes from the fact that it reduces the calculation of the index of a map φ to that of the index of a linear map, which is easier to compute.

Proposition 3.8.5 *If $\Omega \subseteq \mathbb{R}^N$ is bounded open, $\varphi \in C(\overline{\Omega}, \mathbb{R}^N)$, $\xi \notin \varphi(\partial\Omega)$ and $\varphi^{-1}(\xi)$ is finite, then $d(\varphi, \Omega, \xi) = \sum_{u \in \varphi^{-1}(\xi)} i(\varphi, u, \xi)$.*

Proof Since $\varphi^{-1}(\xi)$ is finite, then every $u \in \varphi^{-1}(\xi)$ is an isolated ξ-point of φ and so $i(\varphi, u, \xi)$ is well-defined (see Definition 3.8.1). Suppose that $\varphi^{-1}(\xi) = \{u_k\}_{k=1}^m$. Let $\{U_k\}_{k=1}^m$ be mutually disjoint open subsets of Ω such that $u_k \in U_k$ for all $k \in \{1, ..., m\}$. Then according to Definition 3.8.1 we have

$$i(\varphi, u_k, \xi) = d(\varphi, U_k, \xi) \text{ for all } k \in \{1, ..., m\}.$$

Then from the domain additivity and excision properties, we have

$$\begin{aligned}\sum_{u \in \varphi^{-1}(\xi)} i(\varphi, u, \xi) &= \sum_{k=1}^m i(\varphi, u_k, \xi) \\ &= \sum_{k=1}^m d(\varphi, U_k, \xi) \\ &= d(\varphi, \bigcup_{k=1}^m U_k, \xi) \text{ (domain additivity)} \\ &= d(\varphi, \Omega \backslash K, \xi) \text{ where } K = \overline{\Omega} \backslash \bigcup_{k=1}^m U_k \text{ (excision property)} \\ &= d(\varphi, \Omega, \xi).\end{aligned}$$

The proof is now complete. □

Proposition 3.8.6 *If $\Omega \subseteq \mathbb{R}^N$ is bounded open, $\varphi \in C^1(\overline{\Omega}, \mathbb{R}^N)$, $\xi \notin \varphi(\partial\Omega)$, $u_0 \in \varphi^{-1}(\xi)$ and $\varphi'(u_0)$ is invertible, then u_0 is an isolated ξ-point of φ and*

$$i(\varphi, u_0, \xi) = (-1)^m,$$

where m is the number of negative eigenvalues of $\varphi'(u_0)$*, counting multiplicities.*

Proof The fact that u_0 is an isolated ξ-point of φ was established in the proof of Proposition 3.8.3 (see (3.118), (3.119)). From Proposition 3.8.3, we have

$$i(\varphi, u_0, \xi) = \operatorname{sgn} J_\varphi(u_0),$$

where $J_\varphi(u_0)$ is the Jacobian of φ at u_0 (that is, $J_\varphi(u_0) = \det \varphi'(u_0)$). Let $\lambda_1, \dots, \lambda_N$ be the eigenvalues of $\varphi'(u_0)$ (counting multiplicities). Then

$$J_\varphi(u_0) = \prod_{k=1}^{N} \lambda_k,$$

where the complex eigenvalues appear in conjugate pairs $\lambda, \overline{\lambda}$ and so $\lambda\overline{\lambda} > 0$. Hence

$$\operatorname{sgn} J_\varphi(u_0) = (-1)^m$$

with m being the number of negative eigenvalues of $\varphi'(u_0)$. □

An immediate consequence of this proposition is the following result (see also Corollary 3.1.21).

Corollary 3.8.7 $d(-i, \Omega, 0) = (-1)^N$ *with* i *being the identity map on* $\mathbb{R}^N$.

Corollary 3.8.8 *If* A *is an invertible* $N \times N$ *matrix (that is,* $A \in GL(\mathbb{R}^N)$*),*

$$N = \operatorname{span}\{u \in \mathbb{R}^N : (A - \lambda I)^k(u) = 0 \textit{ for some } k \geqslant 1, \lambda < 0\}$$

and $\dim N = m$*, then* $d(A, \Omega, 0) = (-1)^m$.

Remark 3.8.9 So, for every quadratic function $\varphi(u) = \frac{1}{2}(A(u), u)_{\mathbb{R}^N}$ for all $u \in \mathbb{R}^N$ with A symmetric $N \times N$ matrix such that $\det A \neq 0$, we have $\varphi'(u) = A(u)$ and so $i(\varphi, 0, 0) = \operatorname{sgn} \det A = (-1)^m$ with m being the number of negative eigenvalues of A, counting multiplicities.

Proposition 3.8.10 *If* $f \in C^1(\mathbb{R}^N)$*,* $f'(u) \neq 0$ *for all* $||u|| \geqslant R$ *and* $f(u) \to +\infty$ *as* $||u|| \to \infty$*, then* $i(\varphi = f'(0), 0, 0) = 1$.

Proof Without any loss of generality we may assume that $f \in C^\infty(\mathbb{R}^N)$ by replacing f with its mollification f_ϵ for $\epsilon > 0$ small. We consider the Cauchy problem

$$y'(t) = -\varphi(y(t)), \ y(0) = x \in \mathbb{R}^N. \tag{3.120}$$

We know that (3.120) admits a unique local flow $y(t) = y(t; x)$. Let $g(t) = f(y(t))$. Then

$$\begin{aligned}&g'(t) = -||\varphi(u(t))||^2 \leqslant 0\\ \Rightarrow\ &f(y(t)) \leqslant f(x) \text{ on the interval of existence of } y(\cdot).\end{aligned}$$

The coercivity hypothesis on $f(\cdot)$ implies that $y(\cdot)$ remains bounded and so it is a global flow (that is, it holds for all $t \geqslant 0$). Without any loss of generality we may assume that $f(x) \geqslant 0$ (addition of a constant keeps $\varphi = f'$ the same). Let

$$M_R = \max\{f(u) : ||u|| \leqslant R\}.$$

Choose $r > R$ large such that

$$\varphi(u) \geqslant M_R + 1 \text{ for all } ||u|| \geqslant r.$$

Set $\hat{M} = \max[f(u) : ||u|| = r]$. We know that

$$f(y(t)) \leqslant f(x) - \int_0^t ||\varphi(y(s))||^2 ds \text{ for all } t \geqslant 0. \tag{3.121}$$

So, the solutions starting at $x \in \partial B_r$ satisfy

$$f(y(t)) \leqslant f(x) \leqslant \hat{M} \text{ for all } t \geqslant 0.$$

Let $\hat{m} = \min\{||\varphi(u)|| : ||u|| \geqslant r,\ \varphi(u) \leqslant \hat{M}\}$. Then from (3.121) we have

$$0 \leqslant \varphi(y(t)) \leqslant \varphi(x) - \hat{m}^2 t \leqslant \hat{M} - \hat{m}^2 t \text{ provided } ||y(t)|| \geqslant r.$$

Therefore $||y(t_0)|| \leqslant R$ for some $t_0 = t_0(x) < \vartheta = \frac{\hat{M}}{\hat{m}^2}$, hence $\varphi(y(t_0)) \leqslant M_R$. Let $P_t(x) = y(t; x)$ be the Poincaré map. Then $P_\vartheta(\partial B_r) \subseteq B_r$ since

$$f(P_\vartheta(x)) \leqslant M_R \leqslant M_R + 1 \leqslant f(x) \text{ for all } x \in \partial B_r.$$

Finally note that

$$\begin{aligned}&i(\varphi, 0, 0) = d(\varphi, B_r, 0) = d(i - P_\vartheta, B_r, 0) \text{ (see Krasnoselskii [250])}\\ \Rightarrow\ &i(\varphi, 0, 0) = 1\end{aligned}$$

since $d(i - P_\vartheta, B_r, 0) = 1$ via the homotopy invariance property, using the homotopy $h_t(u) = u - tP_\vartheta(u)$. $\square$

3.9 Remarks

3.1: The topological degree was introduced by Brouwer [74–76]. His approach as well as that of Hopf [213–215] used methods and techniques from combinatorial and algebraic topology. This approach can also be found in the books of Cronin [131], Granas and Dugundji [197], and Hatcher [203]. An alternative approach based on analysis was introduced by Nagumo [313]. This approach uses smooth approximations of the original vector field and the result of Sard [370] (see Proposition 3.1.17) concerning the measure of the critical values of a differentiable map. This approach can be found in the books of Istratescu [222], Krawcewicz and Wu [251] and Nirenberg [319]. A third approach, also analytic in nature and closely related to the second one, was proposed by Heinz [205]. This is the approach that we follow here. It is based on the so-called Kronecker integral (see Definition 3.1.1) and again on Sard's theorem. It appears that this is the most popular way of introducing Brouwer's degree and can be found in the books of Cioranescu [125], Deimling [142], Denkowski et al. [143], Lloyd [283], Rabinowitz [345], Schwartz [376] and Zeidler [426].

One natural question that arises about Brouwer's degree concerns its uniqueness. If we can establish uniqueness, then we know that it is useless to seek another similar tool with the same properties. Also, if we have different expressions for the Brouwer topological degree, then we can guarantee that they are all equivalent and we can always use the one that is more suitable for our needs. It turns out that the properties of normalization, domain additivity and homotopy invariance, define Brouwer's degree uniquely.

Theorem 3.9.1 *If* $\tau = \{(\varphi, \Omega, \xi) : \Omega \subseteq \mathbb{R}^N$ *bounded open,* $\varphi \in C(\overline{\Omega}, \mathbb{R}^N)$, $\xi \notin \varphi(\partial\Omega)\}$, *then there exists at most one function* $d : \tau \to \mathbb{Z}$ *satisfying the normalization, domain additivity and homotopy invariance properties. Moreover, these properties imply that* $d(A, \Omega, 0) = \operatorname{sgn} \det A$ *for all* $A \in GL(\mathbb{R}^N)$ *and* $0 \in \Omega$.

Remark 3.9.2 The solution property can be obtained as a consequence of the domain additivity property.

The uniqueness of Brouwer's degree was first proved independently by Führer [176] and Amann and Weiss [15]. In fact, Amann and Weiss [15] established the uniqueness of the Leray–Schauder degree and as a consequence obtained the uniqueness of Brouwer's degree in a normed space. For the uniqueness of the degree, see also the books of Deimling [142, pp. 10–12] and Lloyd [283, pp. 86–88]. We should also mention that Nagumo [313] asserted that the uniqueness of Brouwer's degree can be proved using simplicial approximations.

Another result that we should mention is the so-called product formula, which relates the degree of the composition $\psi \circ \varphi$ to those of ψ and φ separately. For a proof, we refer to Deimling [142, p. 24] and Lloyd [283, p. 29]. The result is attributed to Leray [263].

Proposition 3.9.3 *If* $\Omega \subseteq \mathbb{R}^N$ *is bounded open,* $\varphi \in C(\overline{\Omega}, \mathbb{R}^N)$, $\psi \in C(\mathbb{R}^N, \mathbb{R}^N)$, $\{C_k\}_{k \geqslant 1}$ *are the bounded connected components of* $\mathbb{R}^N \setminus \varphi(\partial\Omega)$ *and* $\xi \notin (\psi \circ \varphi)(\partial\Omega)$, *then* $d(\psi \circ \varphi, \Omega, \xi) = \sum_{k \geqslant 1} d(\varphi, \Omega, C_k) d(\psi, C_k, \xi)$.

Remark 3.9.4 Recall that $d(\varphi, \Omega, \cdot)$ is constant on every connected component of $\mathbb{R}^N \setminus \varphi(\partial\Omega)$ (see Theorem 3.1.25(e)). So, for convenience we denote this common value by $d(\varphi, \Omega, C_k)$.

Theorem 3.1.34 is the celebrated "Brouwer's fixed point theorem" (see Brouwer [75]) and has found many applications. We mention one such application, which is the so-called Perron–Frobenius theorem (see Bellman [41]).

Theorem 3.9.5 *If $A = (a_{ij})_{i,j=1}^N$ is an $N \times N$-matrix such that $a_{ij} \geqslant 0$ for all i, j, then A has a nonnegative eigenvector $\hat{u} = (u_k)_{k=1}^N$, $u_k \geqslant 0$, corresponding to a nonnegative eigenvalue.*

Remark 3.9.6 This result has an infinite-dimensional counterpart known as the Krein–Rutman theorem. It is useful in the spectral theory of differential operators (see Volume 2). Matrices like those in the above theorem are usually called stochastic matrices.

Theorem 3.9.7 (Krein–Rutman) *If X is an ordered Banach space with a solid order cone K (that is, $\operatorname{int} K \neq \emptyset$) which is total (that is, $X = \overline{K - K}$), $\lambda \in \mathscr{L}_c(X)$, A is positive (that is, $A(K) \subseteq K$) and the spectral radius $r(A) = \lim_{n\to\infty} \|A^n\|_{\mathscr{L}}^{1/n} > 0$, then $r(A) > 0$ is an eigenvalue of A with positive eigenvector.*

Theorem 3.1.37 is due to Hartman and Stampacchia [202] (see also Kinderlehrer and Stampacchia [234]). Proposition 3.1.43 is a key tool in the definition of the Leray–Schauder degree and is also known as the Leray–Schauder Lemma (see Leray and Schauder [266]). Theorem 3.1.44 is due to Borsuk [58], while Theorem 3.1.45 was conjectured by Ulam and proved by Borsuk [58]. Further discussion of Borsuk's theorem and its equivalent formulations can be found in Granas and Dugundji [197] (Sect. I.5).

3.2: Since the finite-dimensional topological degree of Brouwer turned out to be a very valuable tool and many problems of interest are infinite-dimensional in nature, soon mathematicians started looking for infinite-dimensional extensions of Brouwer's theory. This was achieved by Leray and Schauder [266], who considered compact perturbations of the identity. The two key tools in their constructions were Theorem 2.1.7, due to Schauder [374], which permits the approximation of compact maps by finite rank maps, and Proposition 3.1.43 (the Leray–Schauder lemma), which allows one to shift between finite-dimensional spaces without changing the value of the degree map. The question of whether the Brouwer degree can be extended in infinite dimensions to all continuous functions was answered in the negative by Leray [265], who produced an example on the space $C[0, 1]$ (see also Cronin [131] and Fonseca and Gangbo [172]). The counterexample presented in Example 3.2.1 is due to Kakutani [226]. As we already mentioned the uniqueness of the Leray–Schauder degree was proved by Amann and Weiss [15]. In this respect, we should also mention the earlier work of O'Neil [323], who proved that the fixed point index is uniquely determined by certain basic properties (axioms). His approach was based on algebraic topology (cohomology theory). An alternative method of proving the uniqueness of the fixed point index can be found in Brown [95].

From Definition 3.2.3 it is clear that $\Omega \subseteq X$ need not be bounded. It is enough to assume that Ω intersects every finite-dimensional subspace of X in a bounded subset. Such sets are called "finitely bounded". So, the whole theory remains valid if throughout, we replace "bounded" by "finitely bounded". The whole theory can also be extended to locally convex spaces (see Leray [265]).

Theorem 3.2.20 is due to Schauder [374] and is the first infinite-dimensional generalization of the Brouwer fixed point theorem (see Theorem 3.1.34). Proposition 3.2.22 is due to Schaefer [371]. An alternative formulation of Theorem 3.2.16 is the following:

Theorem 3.9.8 *If X is a Banach space, $\Omega \subseteq X$ is bounded open symmetric with $0 \in \Omega$, $\varphi \in K(\overline{\Omega}, X)$ and*

$$\frac{\varphi(u)}{||\varphi(u)||} \neq \frac{\varphi(-u)}{||\varphi(-u)||} \text{ for all } u \in \partial\Omega,$$

then $d_{LS}(\varphi, \Omega, 0)$ is odd.

Another result from the Leray–Schauder degree theory worth mentioning is the so-called "mod p-theorem" which is useful in asymptotic fixed point theory. The result is due to Steinlein [391].

Theorem 3.9.9 *If X is a Banach space, $\varphi : \Omega_0 \subseteq X \to X$, $n = p^m$ with p a prime number and $m \in \mathbb{Z}$, there is a bounded open $\Omega \subseteq X$ such that $\overline{\Omega}$ is contained in the domain of φ^n, $\varphi|_{\overline{\Omega}}$, $\varphi^n|_{\overline{\Omega}}$ are compact maps and $\varphi(\{u \in \overline{\Omega} : \varphi^n(u) = u\}) \subseteq \Omega$, then $d_{LS}(i - \varphi^n, \Omega, 0) = d_{LS}(i - \varphi, \Omega, 0)$ $(mod\ p)$.*

3.3: Proposition 3.3.2 is due to Cellina [111] and it was used to prove a multivalued version of the Kakutani fixed point theorem. This proposition was used by Cellina and Lasota [114] to provide an analytic approach to the definition of the Leray–Schauder degree for multifunctions. Earlier definitions were provided by Granas [195] (based on tools from algebraic topology) and by Hukuhara [220] (an analytic approach based on the approximation of F by Hausdorff continuous multifunctions). Extensions to locally convex spaces can be found in the work of Ma [285]. We should also mention the work of Lasry and Robert [259], who defined in a finite-dimensional context a degree for multifunctions $F(\cdot)$ with unbounded values. Their work was extended to multifunctions with values in a uniformly convex Banach space by De Blasi and Myjak [139].

3.4: The degree for $(S)_+$-maps is useful because it can be used in the definition of the degree for other operators of monotone type. The first degree for $(S)_+$-operators was produced by Skrypnik [382] (see also Skrypnik [386]). Browder [88] (see also [87, 89, 90]) produced a degree theory for $(S)_+$-maps. Both definitions were based on Brouwer's degree, but their constructions differ. In particular, Skrypnik [386] considers a separable reflexive Banach space X and $\{e_k\}_{k \geqslant 1}$ a fundamental set (that is, if $x^* \in X^*$ and $\langle x^*, e_k \rangle = 0$, then $x^* = 0$, see Lindenstrauss and Tzafriri [272, p. 43]) with linear independent elements. Let

$$X_n = \text{span}\,\{e_k\}_{k=1}^n$$

and let $\Omega \subseteq X$ be bounded open and $\varphi : \overline{\Omega} \to X^*$ a bounded, demicontinuous $(S)_+$-map. Let $\varphi_n : \overline{\Omega} \cap X_n \to X_n$ be the finite-dimensional approximation of φ defined by

$$\varphi_n(u) = \sum_{k=1}^{m} \langle \varphi(u), e_k \rangle \, e_k \text{ for all } u \in \overline{\Omega} \cap X_n.$$

Skrypnik [382] establishes that if $0 \notin \varphi(\partial\Omega)$, then $d(\varphi_n, \Omega, \cap X_n, 0)$ stabilizes for large $n \geqslant 1$ and so it can be used to define a degree for the triple $(\varphi, \Omega, 0)$ depending on the fundamental set $\{e_k\}_{k\geqslant 1}$. In fact, exploiting the uniqueness of the degree for $(S)_+$-maps due to Browder [87] and Berkovits and Mustonen [47], we can show that the definition of Skrypnik is in fact independent of the choice of the fundamental set $\{e_k\}_{k\geqslant 1}$. The earlier justification provided by Skrypnik in this direction appears to have a gap.

Browder [88] uses different approximations of φ. His approximations are the usual Galerkin approximations of φ. So, let $F \in \mathscr{F}$ and let $j : F \to X$ be the injection map and $j^* : X^* \to F^*$ its adjoint. We define

$$\varphi_F(u) = j^*(\varphi(u)) \text{ for all } u \in \overline{\Omega} \cap F.$$

So, $\varphi_F : \overline{\Omega} \cap F \to F^*$, that is, the approximations of the Browder map F into its dual F^*, and for this reason the Leray–Schauder lemma is not available. The whole construction then depends on the existence and uniqueness of a degree (satisfying the usual properties and with the duality map being the normalizing map) for the Galerkin approximations of bounded demicontinuous $(S)_+$-maps from a finite-dimensional Banach space to its dual.

Still a third distinct construction was provided by Berkovits and Mustonen [46, 47] and it is based on the Leray–Schauder degree and the Browder–Ton embedding theorem. This embedding theorem established by Browder and Ton [94] (see also Berkovits [45]) says that for every separable reflexive Banach space X, there is a separable Hilbert space H and a compact injective linear map $\vartheta : H \to X$ such that $\vartheta(H)$ is dense in X. Identifying H with its dual, we can consider the adjoint $\vartheta^* : X^* \to H$ of ϑ. Then Berkovits and Mustonen [47] consider the following approximation of φ

$$\varphi_\epsilon = i + \frac{1}{\epsilon}(\vartheta \circ \vartheta^* \circ \varphi).$$

The map $\varphi_\epsilon : \overline{\Omega} \subseteq X \to X$ has the form of a compact perturbation of the identity. So, we use the Leray–Schauder degree. It can be proved that if $0 \notin \varphi(\partial\Omega)$, then $d_{LS}(\varphi_\epsilon, \Omega, 0)$ stabilizes for $\epsilon > 0$ small and so it can be used to define a degree for the triple $(\varphi, \Omega, 0)$. This degree is unique.

Here the construction of the degree that we present is due to Oinas [322]. Note that this construction does not require boundedness of φ and the separability of X

provided we fix a basis on F_0. If X is also separable, then the degree is independent of the basis we fix on F_0. Propositions 3.4.18, 3.4.19 and 3.4.20 are due to Motreanu et al. [309] and extend earlier finite-dimensional results of Amann [14].

3.5: The resolvent and Yosida approximation for a general maximal monotone map $A : X \to 2^{X^*}$ (see Definition 2.5.1) were introduced by Brezis et al. [67]. The degree d_M on triples $(\varphi + A, \Omega, 0)$, with $\varphi : \overline{\Omega} \to X^*$ an $(S)_+$-map and $A : X \to 2^{X^*}$ maximal monotone (see Definition 3.5.6), was defined by Browder [87]. Extensions to maps of the form $\varphi + A + G$ with G multivalued can be found in Hu and Papageorgiou [217, 218]. We mention also the work of Kobayashi and Otani [243] involving subdifferential operators. Some other works in this direction are those by Aizicovici et al. [5, 6], Kien et al. [232], Kobayashi and Otani [244], and Wang and Huang [412, 413].

3.6: This section is based on the work of Kobayashi and Otani [243]. The degree defined here is useful in the study of problems with unilateral constraints (see also Kobayashi and Otani [241, 242, 244]).

3.7: The Kuratowski measure of noncompactness $\alpha(\cdot)$ (see Definition 3.7.1(a)) was introduced by Kuratowski [253] for topological purposes. More precisely, he established the following generalization of Cantor's intersection theorem (see, for example Denkowski et al. [143, p. 46]). Note that the notion of measure of noncompactness $\alpha(\cdot)$ makes perfect sense in a metric space.

Proposition 3.9.10 *A metric space X is complete if and only if for every decreasing sequence $\{C_n\}_{n\geqslant 1}$ of closed sets with $\alpha(C_n) \to 0$ as $n \to \infty$, we have $\bigcap_{n\geqslant 1} C_n$ is nonempty and compact.*

The Hausdorff or ball measure of noncompactness was in fact introduced by Goldenstein et al. [192].

For $\alpha(\cdot)$ the properties in Proposition 3.7.6 were proved by Darbo [136]. Maps which are γ-contractions were first considered by Darbo [136] and Goldenstein et al. [192], while γ-condensing maps were first considered by Furi and Vignoli [177], Nussbaum [320, 321] and Sadovskii [367, 368]. For the corresponding notions for multifunctions we refer to Kamenskii et al. [227]. The degree d_C in Definition 3.7.13 is due to Nussbaum [320]. A nice application of this degree is the following invariance of domain result.

Proposition 3.9.11 *If X is a Banach space, $\Omega \subseteq X$ is open, $\varphi : \Omega \to X$ is locally γ-condensing and $i - \varphi$ is locally one-to-one, then $i - \varphi$ is open.*

The A-proper maps (see Definition 3.7.22) were introduced by Browder and Petryshyn [92]. These operators and the theory built on them require projection schemes. However, most separable Banach spaces that arise in applications exhibit such projection schemes (see Proposition 3.7.17). More about projection schemes can be found in Browder [85, 86] and Petryshyn [337, 340]. Theorem 3.7.29 is taken from Petryshyn [337]. In general most of the properties of A-proper maps can be found in the works of Petryshyn [337–340]. Definition 3.7.35 is due to Browder and Petryshyn

[92, 93]. The degree theory of A-proper maps is also discussed in the books of Cioranescu [125], Deimling [142] and Lloyd [283].

3.8: The index at an isolated ξ-point is as old as the degree theory and it is discussed in the books of Deimling [142], Lloyd [283], Rothe [364] and Zeidler [426]. In Chap. 4, we will discuss an infinite-dimensional analog based on the Leray–Schauder degree, the fixed point index.

Chapter 4
Partial Order, Fixed Point Theory, Variational Principles

To know that you do not know is the best.
To pretend to know when you do not know is a disease.
Lao Tzu

Many problems arising in applications impose nonnegativity requirements on the solutions that we obtain. What we understand by nonnegativity can be described using the concept of a cone. A cone is a special closed and convex subset of the underlying Banach space (the state space). Using the cone we can define a relation "$\leqslant$" on the space which allows the comparison of different elements, which is more precise than the crude estimates produced using the norm. The ordering $\leqslant$ induced by the cone leads to an extension of the fundamental concept of monotonicity of a map (increasing or decreasing map). So, studying nonlinear problems in terms of partial orders induced by means of order cones is an important part of nonlinear analysis and leads to important results. In the first section of this chapter, we deal with cones and the partial order they induce on the space. We investigate how this order structure interacts with the metric, topological and linear structure of the space. One part of Nonlinear Analysis where the order structure leads to remarkable results is fixed point theory. Fixed point theorems are one of the basic mathematical tools used in showing the existence of solution concepts in various problems from partial differential equations all the way to mathematical economics and game theory. There is a rough classification of fixed point theorems into three basic classes:

(a) Metric fixed point theorems.
(b) Topological fixed point theorems.
(c) Order fixed point theorems.

In this first class, we include all those results in which geometric conditions on the underlying spaces and/or the maps, making use of metric structures, are involved.

N. S. Papageorgiou et al., *Nonlinear Analysis–Theory and Methods*,
Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-030-03430-6_4

In the second class belong all those results that depend in a more fundamental way on the topological structure of the space. The prototype results are the Brouwer and Schauder fixed point theorems, which we encountered in Chap. 3 as by-products of degree theory. Finally, in the third class belong all those fixed point results which exploit the order structure induced by a cone. Of course this classification is not strict and there are no clear boundaries separating the three classes. From fixed point theory, we pass to the study of the minimization method. We pay attention to the existence of minimizers and not their regularity, which will be discussed later in the framework of nonlinear boundary value problems (see Volume 2). So, we formulate variational principles, starting with the well-known Lax–Milgram theorem, a basic tool in the study of semilinear boundary value problems. In parallel we also discuss Galerkin approximations, which we already used in degree theory. Then we formulate the Ekeland variational principle and examine some of its many applications. The Ekeland variational principle turned out to be equivalent to some other important results of nonlinear analysis, which we discuss in this chapter (Caristi's fixed point theorem, the Takahashi variational principle, the drop theorem and the Brezis–Browder order principle). Finally we discuss Young measures, which are a basic tool in the relaxation of minimization problems.

4.1 Cones and Partial Order

We start with a definition that extends to general Banach spaces the notion of a positive cone in $\mathbb{R}^N$, that is, of $\mathbb{R}^N_+ = \{\hat{u} = (u_k)_{k=1}^N \in \mathbb{R}^N : u_k \geqslant 0 \text{ for all } k \in \{1, \cdots, N\}\}$.

Definition 4.1.1 Let X be a Banach space. A nonempty, closed and convex set $K \subseteq X$ is said to be a cone if it satisfies the following conditions:

(a) If $u \in K$ and $\lambda \geqslant 0$, then $\lambda u \in K$ (that is, $\lambda K \subseteq K$ for all $\lambda \geqslant 0$).
(b) If $u, -u \in K$, then $u = 0$ (that is, $K \cap (-K) = \{0\}$).

A cone induces a partial order $\leqslant$ on X as follows:

"$u \leqslant v$ if and only if $v - u \in K$" (so the elements of K are called positive).

For concepts related to this partial order, we use the usual terminology. So, a sequence $\{u_n\}_{n\geqslant 1} \subseteq X$ is said to be monotonic if

$$u_n \leqslant u_{n+1} \text{ (increasing) or } u_n \geqslant u_{n+1} \text{ (decreasing) for all } n \geqslant 1.$$

Also, a set $C \subseteq X$ is said to be bounded above (or below) with respect to the partial order $\leqslant$ if there exists a $y \in X$ such that

$$u \leqslant y \text{ for all } u \in C \text{ (or } y \leqslant u \text{ for all } u \in C).$$

By $\sup C$ (resp. $\inf C$) we denote the least upper bound of C (resp. the greatest lower bound of C), if it exists.

There are many geometric and topological properties of cones which in finite-dimensional spaces are evident, but need not hold in infinite-dimensional spaces. This leads to the following concepts:

Definition 4.1.2 Let X be a Banach space, $K \subseteq X$ a cone and denote by $\leqslant$ the partial order on X induced by the cone K.

(a) We say that K is solid if $\operatorname{int} K \neq \emptyset$.
(b) We say that K is generating (or reproducing) if $X = K - K$ and total if $X = \overline{K - K}$.
(c) We say that K is normal if there exists a $\delta > 0$ such that

$$u, y \in K, \ ||u|| = ||y|| = 1 \Longrightarrow ||u + y|| \geqslant \delta.$$

(d) We say that the norm on X is monotonic if $0 \leqslant u \leqslant y$ implies $||u|| \leqslant ||y||$ and semimonotonic if $0 \leqslant u \leqslant y$ implies $||u|| \leqslant \xi||y||$ for some $\xi > 0$.
(e) We say that K is regular if every increasing and bounded above sequence converges (that is, if $\{u_n\}_{n\geqslant 1} \subseteq X$ and $u_1 \leqslant u_2 \leqslant \cdots \leqslant u_n \leqslant \cdots \leqslant y$ for all $n \geqslant 1$, then $u_n \to u$ in X).
(f) We say that K is fully regular if every increasing and norm bounded sequence converges (that is, if $\{u_n\}_{n\geqslant 1} \subseteq X$ and $u_1 \leqslant u_2 \leqslant \cdots \leqslant u_n \leqslant \cdots$ and $\sup_{n\geqslant 1} ||u_n|| < \infty$, then $u_n \to u$ in X).
(g) We say that K is minihedral if for any $u, y \in X$, $\sup\{u, y\}$ exists (recall that by $\sup C$ we denote the least upper bound of $C \subseteq X$).
(h) We say that K is strongly minihedral if $\sup C$ exists for any set $C \subseteq X$ which is bounded from above.

Remark 4.1.3 Geometrically, the notion of normality means that the angle between two positive unit vectors has to be bounded away from π. So, a normal cone cannot be too large. The cone K is regular if and only if every decreasing and bounded from below sequence converges in X. Similarly, the cone K is fully regular if and only if every decreasing and norm bounded sequence converges in X. Finally, K is minihedral (resp. strongly minihedral) if and only if $\inf\{u, y\}$ exists for any $u, y \in X$ (resp. $\inf C$ exists for any $C \subseteq X$ bounded from below).

Proposition 4.1.4 *If X is a Banach space and $K \subseteq X$ is a solid cone, then K is generating.*

Proof Let $e \in \operatorname{int} K$ and $\delta > 0$ such that $\overline{B}_\delta(e) \subseteq K$. For $u \in X\backslash\{0\}$, we have $e + \delta\frac{u}{||u||} \in K$ and

$$u = ||u||\frac{1}{\delta}\left(e + \delta\frac{u}{||u||}\right) - \frac{||u||}{\delta}e \in K - K.$$

The proof is now complete. $\square$

Remark 4.1.5 The converse in not true in general. Consider $L^p[0, 1]$ $(1 \leqslant p < \infty)$ and let $K = \{u \in L^p[0, 1] : u(t) \geqslant 0 \text{ a.e. on } [0, 1]\}$. Evidently, K is generating since $u = u^+ - u^-$ (with $u^+ = \max\{u, 0\}$ and $u^- = \max\{-u, 0\}$), but $\operatorname{int} K = \emptyset$.

Proposition 4.1.6 *If X is a Banach space and $K \subseteq X$ is a cone, then K is normal if and only if $||\cdot||$ is semimonotonic.*

Proof $\underline{\Rightarrow}$: Suppose that $||\cdot||$ is not semimonotonic. Then we can find $\{u_n, y_n\}_{n\geqslant 1} \subseteq X$ such that

$$0 \leqslant u_n \leqslant y_n \text{ and } 0 < n||y_n|| < ||u_n|| \text{ for all } n \geqslant 1.$$

Let $v_n = \frac{u_n}{||u_n||}$, $w_n = \frac{y_n}{||y_n||}$ and $z_n = \left(\frac{1}{n}w_n - v_n\right)/||\frac{1}{n}w_n - v_n||$. Then $v_n, z_n \in K \cap \partial B_1$ and $||v_n + z_n|| \to 0$ as $n \to \infty$, which means that K is not normal.

$\underline{\Leftarrow}$: Let $u, y \in K \cap \partial B_1$, then $1 = ||u|| \leqslant \xi||u + y||$ (by virtue of the semimonotonicity of $||\cdot||$) and so K is normal (see Definition 4.1.2(c)). ☐

Proposition 4.1.7 *If X is a Banach space and $K \subseteq X$ is a cone, then K is normal if and only if $(\overline{B}_1 + K) \cap (\overline{B}_1 - K)$ is bounded ($\overline{B}_1 = \{x \in X : ||x|| \leqslant 1\}$).*

Proof $\underline{\Rightarrow}$: Let $u \in (\overline{B}_1 + K) \cap (\overline{B}_1 - K)$. Then $v \leqslant u \leqslant w$ with $v, w \in \overline{B}_1$. We have

$$\begin{aligned}
&0 \leqslant u - v \leqslant w - v \\
\Rightarrow\ & ||u - v|| \leqslant \xi||w - v|| \text{ for some } \xi > 0 \text{ (see Proposition 4.1.6)} \\
\Rightarrow\ & ||u|| \leqslant \xi||w|| + 2||v|| = \xi + 2 \\
\Rightarrow\ & (\overline{B}_1 + K) \cap (\overline{B}_1 - K) \text{ is bounded.}
\end{aligned}$$

$\underline{\Leftarrow}$: We can find $r > 0$ such that $(\overline{B}_1 + K) \cap (\overline{B}_1 - K) \subseteq r\overline{B}_1$. So, $0 \leqslant u \leqslant y$ implies

$$\begin{aligned}
&\frac{y - u}{||y||} \in (\overline{B}_1 - K) \cap K \\
\Rightarrow\ & ||u|| \leqslant (r + 1)||y|| \\
\Rightarrow\ & ||\cdot|| \text{ is semimonotonic with } \xi = r + 1
\end{aligned}$$

(see Proposition 4.1.6). The proof is now complete. ☐

Proposition 4.1.8 *If X is a Banach space and $K \subseteq X$ is a cone, then K is normal if and only if every order interval $[u, y] = \{v \in X : u \leqslant v \leqslant y\}$ is bounded.*

Proof $\underline{\Rightarrow}$: From Proposition 4.1.7 we know that there exists an $r > 0$ such that $(\overline{B}_1 + K) \cap (\overline{B}_1 - K) \subseteq r\overline{B}_1$. Let $\rho = \max\{||u||, ||y||\}$ and $v \in [u, y]$. Then $\frac{v}{\rho} \in (\overline{B}_1 + K) \cap (\overline{B}_1 - K)$ and so $||v|| \leqslant r\rho$.

$\underline{\Leftarrow}$: Arguing by contradiction, suppose that K is not normal. Then we can find $\{u_n\}_{n\geqslant 1}$ $\{y_n\}_{n\geqslant 1} \subseteq K$ such that $||u_n|| = ||y_n|| = 1$ and $||u_n + y_n|| < \frac{1}{4^n}$ for all $n \geqslant 1$. Let

$$v_n = \frac{u_n}{||u_n + y_n||^{1/2}} \text{ and } z_n = \frac{u_n + y_n}{||u_n + y_n||^{1/2}} \text{ for all } n \geqslant 1.$$

We have $0 \leqslant v_n \leqslant z_n$ and $\sum_{n \geqslant 1} ||z_n|| < \sum_{n \geqslant 1} \frac{1}{2^n} < \infty$. So, $z = \sum_{n \geqslant 1} z_n \in X, 0 \leqslant v_n \leqslant z_n \leqslant z$ and

$$||v_n|| = \frac{1}{||u_n + y_n||^{1/2}} > 2^n \text{ for all } n \geqslant 1$$
$$\Rightarrow [0, z] \text{ is unbounded, a contradiction.}$$

The proof is now complete. □

Proposition 4.1.9 *If X is a Banach space and $K \subseteq X$ is a cone, then K is normal if and only if $u_n \leqslant v_n \leqslant y_n$ for all $n \geqslant 1$ and $||u_n - x||,\ ||y_n - x|| \to 0$ imply $||v_n - x|| \to 0$.*

Proof $\underline{\Rightarrow}$: We have

$$\begin{aligned} &0 \leqslant v_n - u_n \leqslant y_n - u_n \text{ for all } n \geqslant 1 \\ \Rightarrow\ &||v_n - u_n|| \leqslant \xi ||y_n - u_n|| \text{ for all } n \geqslant 1 \text{ and some } \xi > 0 \\ &\qquad \text{(see Proposition 4.1.6)}. \end{aligned} \tag{4.1}$$

Then

$$\begin{aligned} ||v_n - x|| \leqslant ||v_n - u_n|| + ||u_n - x|| &\leqslant \xi ||y_n - u_n|| + ||u_n - x|| \text{ (see (4.1))} \\ &\leqslant \xi ||y_n - x|| + (\xi + 1)||u_n - x|| \\ &\qquad \text{for all } n \geqslant 1 \end{aligned}$$
$$\Rightarrow ||v_n - x|| \to 0 \text{ as } n \to \infty.$$

$\underline{\Leftarrow}$: According to Proposition 4.1.7 it suffices to show that $(\overline{B}_1 + K) \cap (\overline{B}_1 - K)$ is bounded. Proceeding indirectly, suppose that the set is unbounded. So, we can find $\{v_n\}_{n \geqslant 1} \subseteq (\overline{B}_1 + K) \cap (\overline{B}_1 - K)$ such that $||v_n|| \to \infty$. We have

$$u_n \leqslant v_n \leqslant y_n \text{ for all } n \geqslant 1 \text{ with } u_n, y_n \in \overline{B}_1,\ n \geqslant 1.$$

We set

$$x_n = \frac{u_n}{||v_n||},\ e_n = \frac{v_n}{||v_n||},\ w_n = \frac{y_n}{||v_n||} \text{ for all } n \geqslant 1.$$

Then $x_n \leqslant e_n \leqslant w_n$ for all $n \geqslant 1$ and $||x_n||,\ ||w_n|| \to 0$. On the other hand, $||e_n|| = 1$ for all $n \geqslant 1$, contradicting our hypothesis. □

Proposition 4.1.10 *If X is a Banach space and $K \subseteq X$ is a cone, then we have the following implications*

K is fully regular $\Rightarrow$ *K is regular* $\Rightarrow$ *K is normal.*

Proof Regular $\Rightarrow$ Normal.

Arguing by contradiction, suppose that K is not normal. Then by virtue of Proposition 4.1.6, we can find $\{u_n\}_{n\geqslant 1},\ \{y_n\}_{n\geqslant 1} \subseteq K$ such that

$$0 \leqslant u_n \leqslant y_n \text{ and } 2^n||y_n|| < ||u_n|| \text{ for all } n \geqslant 1. \tag{4.2}$$

Let $v_n = \frac{u_n}{||u_n||}$ and $z_n = \frac{y_n}{2^n||y_n||}$ for all $n \geqslant 1$. Then from (4.2) we have

$$0 \leqslant v_n \leqslant z_n \text{ for all } n \geqslant 1 \text{ and } \sum_{n\geqslant 1} ||z_n|| = \sum_{n\geqslant 1} \frac{1}{2^n} = 1. \tag{4.3}$$

Therefore $z = \sum_{n\geqslant 1} z_n \in X$. We set

$$w_n = \begin{cases} \sum_{k=1}^{2m} z_k & \text{if } n = 2m,\ m \geqslant 1 \\ \sum_{k=1}^{2m} z_k + v_{2m+1} & \text{if } n = 2m+1,\ m \geqslant 1. \end{cases}$$

From (4.3) we have $\{w_n\}_{n\geqslant 1} \subseteq K$ is increasing and $w_n \leqslant z$ for all $n \geqslant 1$. Therefore by regularity $w_n \to w$ in X. On the other hand

$$||w_{2m+1} - w_{2m}|| = ||v_{2m+1}|| = 1 \text{ for all } m \geqslant 1,$$

a contradiction. This proves that K is normal.

Fully Regular $\Rightarrow$ Regular.

We show that K is normal. Indeed, if K is not normal, then we can find $\{u_n\}_{n\geqslant 1},\ \{y_n\}_{n\geqslant 1} \subseteq K$ such that

$$||u_n|| = ||y_n|| = 1 \text{ and } ||u_n + y_n|| \leqslant \frac{1}{2^n} \text{ for all } n \geqslant 1. \tag{4.4}$$

As before, let

$$z_{2m} = \sum_{k=1}^{2m} (u_k + y_k) \text{ and } z_{2m+1} = z_{2m} + u_{2m+1}.$$

Evidently, $\{z_n\}_{n\geqslant 1} \subseteq K$ is increasing and bounded, hence convergent. On the other hand, we have $||z_{2m+1} - z_{2m}|| = 1$ for all $m \geqslant 1$, a contradiction. So, K is normal. Hence, if $\{u_n\}_{n\geqslant 1} \subseteq K$ is increasing and order bounded from above by $y \in K$, then

$$
\begin{aligned}
& 0 \leqslant y - u_n \leqslant y - u_1 \text{ for all } n \geqslant 1 \\
\Rightarrow\ & \|y - u_n\| \leqslant \xi \|y - u_1\| \text{ for all } n \geqslant 1 \text{ (since } K \text{ is normal, see Corollary 1.4.6)} \\
\Rightarrow\ & \{u_n\}_{n\geqslant 1} \text{ is norm bounded, thus convergent (by the full regularity of } K).
\end{aligned}
$$

This proves that K is regular. The proof is now complete. □

Example 4.1.11 None of the reverse implications are true in general.

(a) Normal $\not\Rightarrow$ Regular: Let $X = C[0, 1]$ and let

$$K = \{u \in C[0, 1] : u(t) \geqslant 0 \text{ for all } t \in [0, 1]\}.$$

Since $\|\cdot\|_{C[0,1]}$ is monotonic, K is normal. However, K is not regular. To see this let $\{u_n(t) = 1 - t^n : t \in [0, 1]\}_{n\geqslant 1} \subseteq K$. Then $\{u_n\}_{n\geqslant 1}$ is increasing and order bounded from above by 1, but it does not converge in X.

(b) Regular $\not\Rightarrow$ Fully Regular: Let $X = c_0 = \{\hat{u} = (u_n)_{n\geqslant 1} : u_n \to 0\}$ furnished with the norm $\|\hat{u}\| = \sup\limits_{n\geqslant 1} |u_n|$. Let $K = \{\hat{u} = (u_n)_{n\geqslant 1} : u_n \geqslant 0 \text{ for all } n \geqslant 1\}$. It is easy to see that K is regular, since increasing and bounded from above sequences in $\mathbb{R}$ converge. On the other hand, if $\hat{u}_n = (u_{n,k})_{k\geqslant 1}$ is defined by

$$u_{n,k} = \begin{cases} 1 \text{ if } k \leqslant n \\ 0 \text{ if } k > n \end{cases}$$

then $\{\hat{u}_n\}_{n\geqslant 1}$ is increasing and norm bounded (since $\|\hat{u}_n\| = 1$ for all $n \geqslant 1$). But $\{\hat{u}_n\}_{n\geqslant 1}$ does not converge and so K is not fully regular.
However, with additional structure on X, we can have the reverse implications.

Proposition 4.1.12 *If X is a reflexive Banach space and $K \subseteq X$ is a cone, then K is normal $\Leftrightarrow$ K is regular $\Leftrightarrow$ K is fully regular.*

Proof By virtue of Proposition 4.1.10, it suffices to prove that normality of K implies full regularity of K.

So, suppose that $\{u_n\}_{n\geqslant 1} \subseteq X$ is an increasing and bounded sequence. The reflexivity of X implies that we can find a subsequence $\{u_{n_k}\}_{k\geqslant 1}$ of $\{u_n\}_{n\geqslant 1}$ such that

$$u_{n_k} \xrightarrow{w} u \text{ in } X. \tag{4.5}$$

We must have $u_{n_k} \leqslant u$ for all $k \geqslant 1$. Otherwise, there exists a $k_0 \geqslant 1$ such that $u - u_{n_{k_0}} \notin K$. So, by the strong separation theorem, we can find $u^* \in X^*\backslash\{0\}$ such that

$$
\begin{aligned}
& \langle u^*, u - u_{n_{k_0}} \rangle < \eta < \langle u^*, v \rangle \text{ for all } v \in K \text{ and some } \eta \in \mathbb{R} \qquad (4.6) \\
\Rightarrow\ & \langle u^*, u \rangle < \langle u^*, u_{n_{k_0}} \rangle + \eta \\
\Rightarrow\ & \langle u^*, u_{n_k} \rangle < \langle u^*, u_{n_{k_0}} \rangle + \eta \text{ for } k > k_0 \text{ large} \\
\Rightarrow\ & \eta < \langle u^*, u_{n_k} - u_{n_{k_0}} \rangle < \eta \text{ (see (4.6)), a contradiction.}
\end{aligned}
$$

Therefore we have

$$u_{n_k} \leqslant u \text{ for all } k \geqslant 1. \tag{4.7}$$

Next we show that $\{u_{n_k}\}_{k\geqslant 1} \subseteq X$ contains a subsequence which converges strongly to u. Again we argue by contradiction. So, if no such strongly convergent subsequence can be found, then there exist $\epsilon > 0$ and $m \geqslant 1$ such that

$$||u_{n_k} - u|| \geqslant \epsilon \text{ for all } k \geqslant m. \tag{4.8}$$

Let $C_k = \{v \in X : v \leqslant u_{n_k}\}$, $k \geqslant 1$. Each C_k is convex and since $\{u_{n_k}\}_{k\geqslant 1}$ is increasing, we have $C_k \subseteq C_{k+1}$. We set $C = \bigcup_{k\geqslant 1} C_k$. Clearly C is convex. Let $v \in C$. Then $v \in C_k$ for some $k \geqslant 1$. From (4.7) we have

$$\begin{aligned}
&0 \leqslant u - u_{n_k} \leqslant u - v \\
\Rightarrow\ &\epsilon \leqslant ||u_{n_k} - u|| \leqslant \xi ||u - v|| \text{ for } k \geqslant m \text{ (see Proposition 4.1.6 and (4.8))} \\
\Rightarrow\ &\frac{\epsilon}{\xi} \leqslant ||u - v|| \text{ for every } v \in \overline{C} \\
\Rightarrow\ &u \notin \overline{C}.
\end{aligned}$$

Again the strong separation theorem implies that we can find $v^* \in X^* \setminus \{0\}$ such that

$$\begin{aligned}
&\langle v^*, u\rangle < \hat{\eta} < \langle v^*, v\rangle \text{ for all } v \in \overline{C} \\
\Rightarrow\ &\langle v^*, u\rangle < \hat{\eta} < \langle v^*, u_{n_k}\rangle \text{ for all } k \geqslant m
\end{aligned}$$

which contradicts (4.5). Therefore we have proved that $\{u_{n_k}\}_{k\geqslant 1} \subseteq X$ contains a strongly convergent subsequence. For notational economy, we still denote it by $\{u_{n_k}\}_{k\geqslant 1}$. Then for any $m \geqslant 1$, we have

$$\begin{aligned}
&u_m \leqslant u_{n_k} \text{ for } k \geqslant 1 \text{ large} \\
\Rightarrow\ &u_n \leqslant u \text{ for all } n \geqslant 1 \\
\Rightarrow\ &0 \leqslant u - u_n \leqslant u - u_{n_k} \text{ for all } n \geqslant n_k \\
\Rightarrow\ &||u - u_n|| \leqslant \xi ||u - u_{n_k}|| \text{ (due to the normality of } K, \text{ see Proposition 4.1.6)} \\
\Rightarrow\ &u_n \to u \text{ in } X \\
\Rightarrow\ &K \text{ is fully regular (see Definition 4.1.2(f)).}
\end{aligned}$$

The proof is now complete. □

From Definitions 4.1.2(g) and (h), a strongly minihedral cone is minihedral. The converse is not in general true.

Example 4.1.13 Let $X = C[0, 1]$ and $K = \{u \in X : u(t) \geqslant 0 \text{ for all } t \in [0, 1]\}$. Evidently, K is minihedral since for $u, v \in K$, $\sup\{u, v\} = y$ ($y(t) = \max\{u(t), v(t)\}$ for all $t \in [0, 1]$) is still in $C[0, 1]$. On the other hand K is not strongly minihedral. Indeed, let $C = \left\{u \in C[0, 1] : u(t) < \frac{1}{2} \text{ for } t \in [0, \frac{1}{2}] \text{ and } u(t) < 1 \text{ for } t \in [\frac{1}{2}, 1]\right\}$. Clearly $\sup C$ does not exist in X.

With more structure on X and K, we can have the equivalence of minihedral and strongly minihedral.

Proposition 4.1.14 *If X is a separable Banach space and $K \subseteq X$ is a regular minihedral cone, then K is strongly minihedral.*

Proof Let $C \subseteq X$ and suppose that there exists a $v \in X$ such that $u \leqslant v$ for all $u \in C$. The separability of X implies the separability of C. So, we can find $\{u_n\}_{n \geqslant 1} \subseteq C$ which is dense in C. Let $y_n = \sup\{u_k\}_{k=1}^n$ (it exists since by hypothesis K is minihedral). We have

$$\{y_n\}_{n \geqslant 1} \text{ is increasing and order bounded from above by } v. \tag{4.9}$$

Since K is regular, (4.9) implies that $y_n \to y$ in X. We claim $y = \sup C$. Indeed, first note that

$$u_n \leqslant y \text{ for all } n \geqslant 1. \tag{4.10}$$

For any $z \in C$, we can find a subsequence of $\{u_n\}_{n \geqslant 1}$ which converges to z in X. Because of (4.10), we see that

$$z \leqslant y, \text{ that is, } y \text{ is an upper bound of } C.$$

Let $w \in X$ be any upper bound of C. We have

$$\begin{aligned} & y_n \leqslant w \text{ for all } n \geqslant 1 \\ \Rightarrow\ & y \leqslant w, \text{ hence } y = \sup C. \end{aligned}$$

The proof is now complete. □

From Propositions 4.1.12 and 4.1.14, we obtain:

Corollary 4.1.15 *If X is a separable reflexive Banach space and $K \subseteq X$ is a normal, minihedral cone, then K is strongly minihedral.*

Using a cone $K \subseteq X$, we can define positive elements for the dual space X^*.

Definition 4.1.16 Let X be a Banach space and $K \subseteq X$ a cone. The "dual cone" of K is the set

$$K^* = \{u^* \in X^* : \langle u^*, u \rangle \geqslant 0 \text{ for all } u \in K\}.$$

The elements of K^* are called "positive linear functionals" (or simply "positive").

Remark 4.1.17 It is easy to see that K^* satisfies all the requirements of a cone (see Definition 4.1.1), except that $K^* \cap (-K^*) = \{0\}$. Still we call K^* the dual cone. However, if K is generating (see Definition 4.1.2(b)), then $K^* \cap (-K^*) = \{0\}$.

Proposition 4.1.18 *If X is a Banach space, $K \subseteq X$ is a cone and $K^* \subseteq X^*$ the dual cone, then*

(a) *$u \in K$ if and only if $\langle u^*, u\rangle \geqslant 0$ for all $u^* \in K$;*
(b) *for any $u \in K\backslash\{0\}$, there exists a $u^* \in K^*$ such that $\langle u^*, u\rangle > 0$;*
(c) *for any $v \notin K$, there exists a $u^* \in K^*$ such that $\langle u^*, v\rangle < 0$;*
(d) *if K is solid, then $u \in \operatorname{int} K$ if and only if $\langle u^*, u\rangle > 0$ for all $u^* \in K^*\backslash\{0\}$;*
(e) *if X is separable, then there exists a $u^* \in K^*$ such that $\langle u^*, u\rangle > 0$ for all $u \in K\backslash\{0\}$.*

Proof (a) Suppose that $v \notin K$. By the strong separation theorem, we can find $u^* \in X^*\backslash\{0\}$ and $\epsilon > 0$ such that

$$\left\langle u^*, v\right\rangle + \epsilon \leqslant \left\langle u^*, u\right\rangle \text{ for all } u \in K.$$

Since K is a cone, $\langle u^*, u\rangle \geqslant 0$ for all $u \in K$ and so $u^* \in K^*$. This proves that

$$K = \{u \in X : \left\langle u^*, u\right\rangle \geqslant 0 \text{ for all } u^* \in K^*\}.$$

(b) We have $-u \notin K$. By the strong separation theorem there exists a $u^* \in X^*$ and $\epsilon > 0$ such that

$$\begin{aligned} &\left\langle u^*, -u\right\rangle + \epsilon \leqslant \left\langle u^*, y\right\rangle \text{ for all } y \in K \\ \Rightarrow\ & u^* \in K^* \text{ and } \frac{\epsilon}{2} \leqslant \left\langle u^*, u\right\rangle. \end{aligned}$$

(c) See the proof of part (a).

(d) Let $e \in \operatorname{int} K$. Then we can find $\delta > 0$ such that $\overline{B}_\delta(e) \subseteq K$. This means that

$$\begin{aligned} &e \pm \delta h \geqslant 0 \text{ for all } h \in X,\ ||h|| \leqslant 1 \\ \Rightarrow\ &\left\langle u^*, e \pm \delta h\right\rangle \geqslant 0 \text{ for all } u^* \in K^* \\ \Rightarrow\ &\left\langle u^*, e\right\rangle \geqslant \delta||u^*|| > 0. \end{aligned}$$

Next, suppose that $v \notin \operatorname{int} K$ and let $K_1 = \{\hat{v} = \lambda v : \lambda \geqslant 0\}$. Evidently, K_1 is a cone and $K_1 \cap \operatorname{int} K = \emptyset$. So, by the weak separation theorem, we can find $u^* \in X^*\backslash\{0\}$ such that

$$\begin{aligned} &\left\langle u^*, \hat{v}\right\rangle \leqslant \left\langle u^*, u\right\rangle \text{ for all } \hat{v} \in K_1 \text{ and all } u \in K \\ \Rightarrow\ &\left\langle u^*, v\right\rangle \leqslant 0. \end{aligned}$$

(e) The separability of X implies that $\overline{B}_1^* = \{u^* \in X^* : ||u^*||_* \leqslant 1\}$ furnished with the w^*-topology is compact, metrizable, hence separable too. Let $\{u_n^*\}_{n\geqslant 1}$ be w^*-dense in $K^* \cap \overline{B}_1^*$ and let $\hat{u}^* = \sum_{n\geqslant 1} \frac{1}{n^2} u_n^*$. Then $\hat{u}^* \in K^*$ and $\langle \hat{u}^*, u\rangle = 0$ for some $u \in K$ implies $\langle u^*, u\rangle = 0$ for all $u^* \in K^*$, hence $u = 0$ (see (b)). $\square$

The next result presents a duality in the properties of K and K^* and is known in the literature as "Krein's theorem".

Theorem 4.1.19 *If X is a Banach space, $K \subseteq X$ is a cone and K^* is its dual cone, then*

(a) *K is generating if and only if K^* is normal;*
(b) *K is normal if and only if K^* is generating.*

Proof *(a)* K generating $\Rightarrow K^*$ normal

Recall that K^* is a cone (see Remark 4.1.17). By hypothesis, $X = K - K$. So, if $C = K \cap \overline{B}_1$ ($\overline{B}_1 = \{u \in X : ||u|| \leqslant 1\}$), then

$$X = \bigcup_{\text{n}\geqslant 1} n(C - C)$$
$$\Rightarrow \text{int}(\overline{C - C}) \neq \emptyset \text{ (by Baire's theorem).}$$

The set $\overline{C - C}$ is closed, convex and symmetric. So, we can find $\epsilon > 0$ such that

$$\epsilon\overline{B}_1 \subseteq \overline{C - C}$$
$$\Rightarrow \frac{\epsilon}{2}\overline{B}_1 \subseteq C - C. \tag{4.11}$$

Let $u^*, y^* \in X^*$ with $0 \leqslant u^* \leqslant y^*$. For any $x \in X$, we have $x = u - y$ with $u, y \in K$ and $||u|| \leqslant \frac{2}{\epsilon}||x||$, $||y|| \leqslant \frac{2}{\epsilon}||x||$ (see (4.11)). Then

$$\langle u^*, x\rangle = \langle u^*, u - y\rangle \leqslant \langle u^*, u\rangle \leqslant \langle y^*, u\rangle \leqslant ||y^*||_*||u|| \leqslant \frac{2}{\epsilon}||y^*||_*||x|| \tag{4.12}$$

$$\text{and } -\langle u^*, x\rangle \leqslant \langle u^*, y - x\rangle \leqslant \langle u^*, v\rangle \leqslant \langle y^*, v\rangle \leqslant ||y^*||_*||v|| \leqslant \frac{2}{\epsilon}||y^*||_*||x||. \tag{4.13}$$

From (4.12) and (4.13) it follows that

$$|\langle u^*, x\rangle| \leqslant \frac{2}{\epsilon}||y^*||_*||x||$$
$$\Rightarrow ||u^*||_* \leqslant \frac{2}{\epsilon}||y^*||_*$$
$$\Rightarrow K^* \text{is normal (see Proposition 4.1.6).}$$

$\underline{K^* \text{ normal } \Rightarrow K \text{ generating}}$

Arguing by contradiction, suppose that K is not generating. Then $E = \overline{C - C}$ is not a neighborhood of the origin (see the proof of the first part). So, for every $n \geqslant 1$, we can find $u \notin E$ with $||u|| \leqslant \frac{1}{n}$ and there exists a $u^* \in X^*$ satisfying $\langle u^*, u\rangle > 1$, $\langle u^*, y\rangle < 1$ for all $y \in \overline{C - C}$ (by the strong separation theorem). It follows that $||u^*||_* > n$. On the other hand we have $u^* \in (\overline{B}_1^* + K^*) \cap (\overline{B}_1^* - K^*)(\overline{B}_1^* = \{v^* \in X^* : ||v^*||_* \leqslant 1\})$. Proposition 4.1.7 implies that K^* is not normal, a contradiction.

(b) $\underline{K \text{ normal} \Rightarrow K^* \text{ generating}}$

Proceeding indirectly, suppose that K^* is not generating. Let $C^* = K^* \cap \overline{B}_1^*$. As in the proof of part *(a)*, we have that $C^* - C^*$ is not a neighborhood of the origin. We can find $u \in X$ of arbitrarily large norm such that $\langle y^*, u\rangle < 1$ for all $y^* \in C^* - C^*$. Moreover, as in the proof of the second implication in part *(a)*, we can show that $u \in (\overline{B}_1 + K) \cap (\overline{B}_1 - K)$, contradicting the normality of K (see Proposition 4.1.7).

$\underline{K^* \text{ is generating} \Rightarrow K \text{ normal}}$

From *(a)* we know that K^{**} is normal. Since $K \subseteq K^{**}$, we conclude that K must be normal, too. The proof is now complete. □

Combining Proposition 4.1.4 and Theorem 4.1.19, we have the following result.

Corollary 4.1.20 *If X is a Banach space and $K \subseteq X$ is a solid cone, then K^* is normal.*

Also, a by-product of the proof of Theorem 4.1.19 worth mentioning is the following result.

Corollary 4.1.21 *If X is a Banach space and $K \subseteq X$ is a generating cone, then there exists a $\delta > 0$ such that*

$$\delta \overline{B}_1 \subseteq C - C$$

where $\overline{B}_1 = \{u \in X : ||u|| \leqslant 1\}$ and $C = K \cap \overline{B}_1$.

We conclude with a simple observation concerning solid cones which is useful in the study of positive solutions for boundary value problems.

Proposition 4.1.22 *If X is a Banach space, $K \subseteq X$ is a solid cone and $e \in \operatorname{int} K$, then for every $u \in X$ we can find $t_u > 0$ such that $t_u e - u \in K$.*

Proof Let $r > 0$ be such that $B_r(e) \subseteq K$ and let $t_u > \frac{||u||}{r}$. Then

$$\begin{aligned} & e - \frac{1}{t_u} u \in B_r(e) \subseteq K \\ \Rightarrow\ & t_u e - u \in K \text{ (since } K \text{ is a cone).} \end{aligned}$$

The proof is now complete. □

4.2 Metric Fixed Points

As we already mentioned in the introduction of this chapter metric fixed point theory, refers to those fixed point results which depend on the metric structure of the underlying space and/or the corresponding metric properties of the maps involved.

The starting point of metric fixed point theory is the well-known "Banach contraction principle", one of the most versatile elementary results of fixed point theory.

Definition 4.2.1 Let (X, d) be a metric space and $\varphi : X \to X$.

(a) We say that φ is Lipschitz if there exists a $k \geqslant 0$ such that

$$d(\varphi(u), \varphi(v)) \leqslant kd(u, v) \text{ for all } u, v \in X.$$

The smallest $k \geqslant 0$ for which the above inequality holds is said to be the Lipschitz constant of φ and is denoted by $l(\varphi)$.

(b) We say that φ is a contraction if it is Lipschitz with Lipschitz constant $l(\varphi) < 1$.

(c) We say that φ is nonexpansive if it is Lipschitz with Lipschitz constant $l(\varphi) = 1$, that is,

$$d(\varphi(u), \varphi(v)) \leqslant d(u, v) \text{ for all } u, v \in X,\ u \neq v.$$

Remark 4.2.2 In the sequel we will use compositions of φ with itself. So, for every $u \in X$ we define $\varphi^{(n)}(u)$ inductively by $\varphi^{(0)}(u) = u$ and $\varphi^{(n+1)}(u) = \varphi(\varphi^{(n)}(u))$ for all $n \geqslant 0$. If $\varphi, \psi : X \to X$ are two Lipschitz functions, we have

$$l(\varphi \circ \psi) \leqslant l(\varphi)l(\psi) \text{ and } l(\varphi^{(n)}) \leqslant l(\varphi)^n \text{ for all } n \geqslant 1.$$

Moreover, if X is a normed space, then $l(\varphi + \psi) \leqslant l(\varphi) + l(\psi)$ and for every $\lambda \geqslant 0$, $l(\lambda\varphi) = \lambda l(\varphi)$. Finally, for every $u \in X$, the set of iterates $\{\varphi^{(n)}(u)\}_{n\geqslant 0}$ is called the orbit of u under φ and it is denoted by $\mathrm{orb}(u)$.

The next theorem is the celebrated "Banach contraction principle".

Theorem 4.2.3 (Banach)*: If (X, d) is a complete metric space and $\varphi : X \to X$ is a contraction, then φ has a unique fixed point $\hat{u} \in X$ and for every $u \in X$*

$$\varphi^{(n)}(u) \to \hat{u} \text{ in } X.$$

Proof First we show that there is at most one fixed point. So, suppose $\hat{u}, \hat{u}_0$ are distinct fixed points of φ. Then we have

$$d(\hat{u}, \hat{u}_0) = d(\varphi(\hat{u}), \varphi(\hat{u}_0)) \leqslant l(\varphi)d(\hat{u}, \hat{u}_0) < d(\hat{u}, \hat{u}_0),$$

a contradiction. So, we have at most one fixed point.

Next we establish the existence of a fixed point. Let $v \in X$ and consider the orbit of v under φ, that is, the sequence $\{u_n = \varphi^{(n)}(v)\}_{n \geqslant 0} \subseteq X$. For every $n, k \in \mathbb{N}$, we have

$$\begin{aligned} d(u_n, u_{n+k}) &= d(\varphi^{(u)}(v), \varphi^{(n+k)}(v)) \\ &= d(\varphi^{(n)}(v), \varphi^{(n)} \circ \varphi^{(k)}(v)) \\ &\leqslant l(\varphi^{(n)}) d(v, \varphi^{(k)}(v)) \\ &\leqslant l(\varphi)^n \left[d(v, \varphi(v)) + d(\varphi(v), \varphi^{(2)}(v)) + \cdots \right. \\ &\qquad \left. + d(\varphi^{(k-1)}(v), \varphi^{(k)}(v)) \right] \\ &\leqslant l(\varphi)^n \left[1 + l(\varphi) + \cdots + l(\varphi)^{k-1} \right] d(v, \varphi(v)) \\ &= l(\varphi)^n \left(\frac{1 - l(\varphi)^k}{1 - l(\varphi)} \right) d(v, \varphi(v)), \\ &\Rightarrow \{u_n\}_{n \geqslant 1} \subseteq X \text{ is Cauchy.} \end{aligned}$$

The completeness of X implies that there exists a $\hat{u} \in X$ such that $u_n \to \hat{u}$ in X. We have

$$\hat{u} = \lim_{n \to \infty} u_n = \lim_{n \to \infty} u_{n+1} = \lim_{n \to \infty} \varphi(u_n) = \varphi(\hat{u}).$$

The proof is now complete. $\square$

Remark 4.2.4 A careful reading of the above proof reveals that we do not need to assume that φ is a contraction. It suffices to assume that $l(\varphi^{(n)}) < 1$ for at least one fixed $n \in \mathbb{N}$. Also, from the above proof, we have

$$d(\varphi^{(n)}(v), u) = \lim_{k \to \infty} d(\varphi^{(n)}(v), \varphi^{(n+k)}(v)) \leqslant \frac{l(\varphi)^n}{1 - l(\varphi)} d(v, \varphi(v)),$$

which implies that the error at the $n \stackrel{th}{=}$-iteration is completely determined by the contraction constant $l(\varphi)$ and the initial displacement $d(v, \varphi(v))$. Moreover, the rate of convergence is determined by

$$d(\varphi^{(n+1)}(v), u) \leqslant l(\varphi) d(\varphi^{(n)}(v), u) \text{ for all } n \geqslant 0.$$

The Banach fixed point theorem has a local version.

Proposition 4.2.5 *If (X, d) is a complete metric space, $x_0 \in X$, $r > 0$, $B_r(x_0) = \{u \in X : d(u, x_0) < r\}$ and $\varphi : B_r(x_0) \to X$ is a contraction such that*

$$d(\varphi(x_0), x_0) < (1 - l(\varphi))r,$$

then φ admits a fixed point.

Proof Let $\rho \in (0, r)$ such that $d(\varphi(x_0), x_0) \leqslant (1 - l(\varphi))\rho < (1 - l(\varphi))r$ and let $\overline{B}_\rho(x_0) = \{u \in X : d(u, x_0) \leqslant \rho\}$. For $u \in \overline{B}\rho(x_0)$, we have

$$\begin{aligned} d(\varphi(u), x_0) &\leqslant d(\varphi(u), \varphi(x_0)) + d(\varphi(x_0), x_0) \\ &\leqslant l(\varphi)d(u, x_0) + (1 - l(\varphi))\rho \leqslant \rho. \end{aligned}$$

So, $\varphi : \overline{B}_\rho(x_0) \to \overline{B}_\rho(x_0)$ and by Theorem 4.2.3 it admits a fixed point. □

In many applications the contractive map φ also depends on an additional parameter λ. So, the fixed point equation is now the following:

$$u_\lambda = \varphi_\lambda(u_\lambda) \text{ for all } \lambda \in P \ (P \text{ is the parameter space}). \tag{4.14}$$

Proposition 4.2.6 *If (X, d) is a complete metric space, P (the parameter space) is a metric space, for every $\lambda \in P$, $\varphi_\lambda : X \to X$ is a contraction with contraction constant $l(\varphi_\lambda) = l_0$ independent of $\lambda \in P$ and if $\lambda \to \lambda_0$ in P, then*

$$\varphi_\lambda(u) \to \varphi_{\lambda_0}(u) \text{ for all } u \in X,$$

then problem (4.14) has exactly one solution $u_\lambda \in X$ and $\lambda \to u_\lambda$ is continuous from P into X.

Proof Let u_λ be the unique fixed point of φ_λ (see Theorem 4.2.3) and suppose $\lambda \to \lambda_0 \in P$. Then

$$\begin{aligned} d(u_\lambda, u_{\lambda_0}) &= d(\varphi_\lambda(u_\lambda), \varphi_{\lambda_0}(u_{\lambda_0})) \\ &\leqslant d(\varphi_\lambda(u_\lambda), \varphi_\lambda(u_{\lambda_0})) + d(\varphi_\lambda(u_{\lambda_0}), \varphi_{\lambda_0}(u_{\lambda_0})) \\ &\leqslant l_0 d(u_\lambda, u_{\lambda_0}) + d(\varphi_\lambda(u_{\lambda_0}), \varphi_{\lambda_0}(u_{\lambda_0})) \\ \Rightarrow d(u_\lambda, u_{\lambda_0}) &\leqslant \frac{1}{1 - l_0} d(\varphi_\lambda(u_{\lambda_0}), \varphi_{\lambda_0}(u_{\lambda_0})) \to 0 \text{ as } \lambda \to \lambda_0. \end{aligned}$$

The proof is now complete. □

Remark 4.2.7 This proposition establishes the stability of the approximation method introduced by the proof of the Banach principle.

An interesting descendant of the Banach fixed point theorem is the following result.

Theorem 4.2.8 *If (X, d) is a compact metric space and $\varphi : X \to X$ satisfies*

$$d(\varphi(u), \varphi(v)) < d(u, v) \text{ for all } u, v \in X, \ u \neq v,$$

then φ has a unique fixed point $\hat{u} \in X$ and for any $v \in X$ the sequence of iterates $\{\varphi^{(n)}(v)\}_{n \geqslant 0}$ converges to $\hat{u}$.

Proof Let $\tau : X \to \mathbb{R}_+$ be defined by $\tau(u) = d(u, \varphi(u))$. This function is continuous and so we can find $\hat{u} \in X$ such that

$$\tau(\hat{u}) = \inf\{\tau(u) : u \in X\}.$$

We claim that $\hat{u} = \varphi(\hat{u})$. If this is not true, then

$$\tau(\varphi(\hat{u})) = d(\varphi(\hat{u}), \varphi(\varphi(\hat{u}))) < d(\hat{u}, \varphi(\hat{u})) = \tau(\hat{u}),$$

which contradicts the fact that $\hat{u} \in X$ is a global minimizer of τ. Therefore $\hat{u} \in X$ is a fixed point of φ and clearly it is unique.

Next let $v \in X$ and let $\xi_n = d(\varphi^{(n)}(v), \hat{u})\ n \geqslant 0$. Then

$$\begin{aligned}&\xi_{n+1} = d(\varphi^{(n+1)}(v), \hat{u}) = d(\varphi^{(n+1)}(v), \varphi(\hat{u})) \leqslant d(\varphi^{(n)}(v), \hat{u}) = \xi_n \\ \Rightarrow\ &\{\xi_n\}_{n\geqslant 0} \subseteq \mathbb{R}_+ \text{ is decreasing.}\end{aligned}$$

Hence $\xi_n \to \xi \geqslant 0$. The compactness of X implies that $\{\varphi^{(n)}(v)\}_{n\geqslant 0}$ admits a convergent subsequence $\{\varphi^{(n_k)}(v)\}_{k\geqslant 1}$. Assume that $\varphi^{(n_k)}(v) \to y$ in X. Then

$$d(y, \hat{u}) = \xi. \tag{4.15}$$

If $\xi > 0$, then

$$\begin{aligned}&\xi = \lim_{k\to\infty} d(\varphi^{(n_k+1)}(v), \hat{u}) = d(\varphi(y), \hat{u}) = d(\varphi(y), \varphi(\hat{u})) < d(y, \hat{u}) = \xi \\ &\text{(see (4.15) and recall } \xi > 0),\end{aligned}$$

a contradiction. So $y = \hat{u}$ and then by the compactness of X and the Urysohn criterion of the convergence of sequences for the original sequence, we have

$$\varphi^{(n)}(v) \to \hat{u} \text{ in } X.$$

The proof is now complete. □

Example 4.2.9 The above theorem fails if we drop the compactness requirement on X. Let

$$X = \{u \in C[0, 1] :\ 0 = u(0) \leqslant u(t) \leqslant u(1) = 1 \text{ for all } t \in [0, 1]\}.$$

This set is bounded, closed and convex. Then $(X, ||\cdot||_\infty)$ is a complete metric space but not compact. Let $\varphi : X \to X$ be defined by

$$\varphi(u)(t) = tu(t) \text{ for all } t \in [0, 1].$$

Evidently, $||\varphi(u) - \varphi(v)||_\infty < ||u - v||_\infty$ but it does not have a fixed point.

We present some more extensions of Theorem 4.2.3.

Proposition 4.2.10 *If (X, d) is a complete metric space and $\varphi : X \to X$ satisfies*

(i) $d(\varphi^{(n)}(u), \varphi^{(n)}(v)) \leqslant \xi d(u, v)$ *for all $u, v \in X$, all $n \geqslant 1$ and some $\xi > 0$;*

(ii) *there exists an* $x_0 \in X$ *such that* $\operatorname{orb}(x_0) = \{\varphi^{(n)}(x_0)\}_{n \geqslant 0}$ *is bounded and contains a convergent subsequence;*

(iii) *if* $0 < \operatorname{diam}(\operatorname{orb}(u)) < \infty$, *then* $\operatorname{diam}(\operatorname{orb}(\varphi^{(n)}(u))) < \operatorname{diam}(\operatorname{orb}(u))$ *for some* $n = n(u) \geqslant 1$,

then φ *admits a fixed point.*

Proof Without any loss of generality, we may assume that $\xi \geqslant 1$ (otherwise we fall within the realm of Theorem 4.2.3). Let $\tau(u) = \operatorname{diam}(\operatorname{orb}(u))$. By hypothesis *(ii)* $\tau(x_0) < \infty$. Then for all $u \in X$, we have

$$\tau(u) \leqslant \tau(x_0) + 2\xi d(u, x_0) < \infty. \tag{4.16}$$

From (4.16) it follows that

$$|\tau(u) - \tau(v)| \leqslant 2\xi d(u, v) \text{ for all } u, v \in X. \tag{4.17}$$

By hypothesis *(ii)* for some subsequence $\{\varphi^{(n_k)}(x_0)\}_{k \geqslant 1}$ of $\{\varphi^{(n)}(x_0)\}_{n \geqslant 0}$, we have $\varphi^{(n_k)}(x_0) \to \hat{u} \in X$. From the continuity of φ and τ and since $\{\tau(\varphi^{(n)}(x_0))\}_{n \geqslant 0}$ is decreasing, we have

$$\begin{aligned}
&\tau(\hat{u}) = \lim_{k \to \infty} \tau(\varphi^{(n_k)}(x_0)) = \lim_{k \to \infty} \tau(\varphi^{(n_k+m)}(x_0)) = \tau(\varphi^{(m)}(\hat{u})) \text{ for all } m \geqslant 1 \\
&\Rightarrow \tau(\hat{u}) = 0 \text{ (see hypothesis } (iii)) \\
&\Rightarrow \hat{u} = \varphi(\hat{u}).
\end{aligned}$$

The proof is now complete. □

Corollary 4.2.11 *If* (X, d) *is a complete metric space and* $\varphi : X \to X$ *satisfies*

$$d(\varphi^{(n)}(u), \varphi^{(n)}(v)) \leqslant k_n d(u, v)$$

for all $u, v \in X$ *with* $\sum_{n \geqslant 1} k_n < \infty$, *then* φ *admits a unique fixed point.*

When the complete metric space X is a Banach space, the richer structure leads to the following result, which is useful in applications.

Proposition 4.2.12 *If* X *is a Banach space,* $U \subseteq X$ *is an open set and* $\varphi : U \to X$ *is a contraction then* $(i - \varphi)(U) = V \subseteq X$ *is open and* $i - \varphi : U \to V$ *is a homeomorphism (here* $i : X \to X$ *is the identity operator).*

Proof Let $g = i - \varphi$, $u \in U$ and $r > 0$ such that $\overline{B}_r(u) \subseteq U$. Choose $y \in X$ such that $||g(u) - y|| < (1 - l(\varphi))r$. If we can show that $y \in g(\overline{B}_r(u))$, then $g = i - \varphi$ is open. Consider the map $e(v) = v - (g(v) - y)$, $v \in \overline{B}_r(u)$. Then for all $v \in \overline{B}_r(u)$, we have

$$\begin{aligned}||e(v)-u|| &= ||\varphi(v)+y-u|| \\ &\leqslant ||\varphi(v)-\varphi(u)||+||\varphi(u)+y-u|| \\ &\leqslant l(\varphi)||v-u||+||y-g(u)|| \\ &\leqslant l(\varphi)r+(1-l(\varphi))r=r.\end{aligned}$$

Therefore $e : \overline{B}_r(u) \to \overline{B}_r(u)$ and it is a contraction with constant $l(e) = l(\varphi) < 1$. So, Theorem 4.2.3 implies that there exists a unique $\hat{u} \in \overline{B}_r(u)$ such that $e(\hat{u}) = \hat{u}$. Hence $g(\hat{u}) = y$, which establishes the openness of $g = i - \varphi$.

If $u, y \in U$, then

$$||g(u)-g(y)|| \geqslant ||u-y|| - ||\varphi(u)-\varphi(y)|| \geqslant (1-l(\varphi))||u-y||$$
$$\Rightarrow g \text{ is } i-1$$
$$\Rightarrow g : u \to g(u) = (i-\varphi)(u) \text{ is a continuous}$$
$$\text{open bijection, hence a homeomorphism.}$$

The proof is now complete. □

Corollary 4.2.13 *If X is a Banach space and $\varphi : X \to X$ is a contraction, then $g = i - \varphi : X \to X$ is a homeomorphism.*

Proof We claim that g is surjective. Let $u_0 \in X$ and consider the map $h : X \to X$ defined by $h(u) = u_0 + \varphi(u)$. Evidently, h is a contraction. So, by Theorem 4.2.3 we can find $\hat{u} \in X$ such that $\hat{u} = h(\hat{u}) = u_0 + \varphi(\hat{u})$, hence $g(\hat{u}) = u_0$. Since $u_0 \in X$ is arbitrary, we conclude that $g(\cdot)$ is surjective. Then Proposition 4.2.12 implies that g is a homeomorphism. □

We have already seen a parametric version of the Banach fixed point theorem (see Proposition 4.2.6). Next we will examine this issue further and develop a continuation method for contractions. So, we will approach the fixed point problem $u = \varphi(u)$ by embedding the contraction φ in a parametric family $\{h_\lambda\}_{\lambda \in P}$ (P being the parameter metric space) which connects φ with a simpler function ψ and reduces the original fixed point problem to the simpler one $u = \psi(u)$.

So, let (X, d) be a complete metric space and let (P, ρ) be the parameter metric space. Also, let $C \subseteq X$ be a closed subset with nonempty interior $U = \operatorname{int} C$ and let $A = \partial C$ =the boundary of C. By Con (C, X) we denote the set of all contractions from C into X. By $\mathrm{Con}_A(C, X)$ we denote the subset of Con (C, X) consisting of all maps $\varphi \in$ Con (C, X) such that $\varphi\big|_A : A \to X$ is fixed point free. The parametric families $\{h_\lambda\}_{\lambda \in P}$ we will consider will consist of functions in Con (C, X). More precisely, they will have the following properties:

Definition 4.2.14 A parametric family $\{h_\lambda\}_{\lambda \in P} \subseteq \mathrm{Con}\,(C, X)$ is said to be l_0-contractive with $l_0 \in [0, 1)$ if we can find $M > 0$ and $\vartheta \in (0, 1]$ such that

(a) $d(h_\lambda(u), h_\lambda(y)) \leqslant l_0 d(u, y)$ for all $\lambda \in P$ and all $u, y \in C$;
(b) $d(h_\lambda(u), h_\mu(u)) \leqslant M\rho(\lambda, \mu)^\vartheta$ for all $u \in C$ and all $\lambda, \mu \in P$.

Remark 4.2.15 Evidently, the map $(\lambda, u) \to H_\lambda(u)$ is continuous from $P \times C$ into X. Also, for every $\lambda \in P$, the fixed point set of h_λ is either empty or a singleton. Moreover, if $u_\lambda = h_\lambda(u_\lambda)$ and $u_\mu = h_\mu(u_\mu)$, then

$$\begin{aligned} d(u_\lambda, u_\mu) &\leqslant d(h_\lambda(u_\lambda), h_\mu(u_\lambda)) + d(h_\mu(u_\lambda), h_\mu(u_\mu)) \\ &\leqslant M\rho(\lambda, \mu)^\vartheta + l_0 d(u_\lambda, u_\mu) \\ \Rightarrow \quad d(u_\lambda, u_\mu) &\leqslant \frac{M}{1 - l_0}\rho(\lambda, \mu)^\vartheta. \end{aligned} \tag{4.18}$$

Proposition 4.2.16 *If P is connected and $\{h_\lambda\}_{\lambda \in P}$ is an l_0-contractive family in* $\mathrm{Con}_A(C, X)$ *then*

(a) if for some $\lambda \in P$ the fixed point problem $h_\lambda(u) = u$ has a solution, then for every $\lambda \in P$ the fixed point problem has a unique solution $u_\lambda \in C$;
(b) if $u_\lambda = h_\lambda(u_\lambda)$ for all $\lambda \in P$, then $\lambda \to u_\lambda$ from P into $U = \mathrm{int}\, C$ is Hölder continuous.

Proof (a) Let $P_0 = \{\lambda \in P : u_\lambda = h_\lambda(u_\lambda) \text{ for some } u_\lambda \in U\}$. By hypothesis $P_0 \neq \emptyset$.

Claim 1. *$P_0 \subseteq P$ is closed.*

Let $\{\lambda_n\}_{n \geqslant 1} \subseteq P_0$ and assume that $\lambda_n \to \lambda$. Then we can find $u_{\lambda_n} \in U$ such that $u_{\lambda_n} = h_{\lambda_n}(u_{\lambda_n})$ for all $n \geqslant 1$. From (4.18) we have

$$\begin{aligned} & d(u_{\lambda_n}, u_{\lambda_m}) \leqslant \frac{M}{1 - l_0}\rho(\lambda_n, \lambda_m)^\vartheta \text{ for all } n, m \in \mathbb{N} \\ \Rightarrow\ & \{u_{\lambda_n}\}_{n \geqslant 1} \subseteq C \text{ is Cauchy} \\ \Rightarrow\ & u_{\lambda_n} \to u \in C \text{ (recall that } C \text{ is complete being closed in } X). \end{aligned}$$

Definition 4.2.14 *(b)* implies

$$\begin{aligned} & h_{\lambda_n}(u_{\lambda_n}) \to h_\lambda(u) \\ \Rightarrow\ & u = h_\lambda(u) \text{ and } u \in U \text{ (since } h_\lambda \in \mathrm{Con}_A(C, X)) \\ \Rightarrow\ & \lambda \in P_0 \text{ and so we have proved that } P_0 \subseteq P \text{ is closed.} \end{aligned}$$

Claim 2. *$P_0 \subseteq P$ is open.*

Let $\lambda \in P_0$ with $u_\lambda = h_\lambda(u_\lambda),\ u_\lambda \in U$. We can find $r > 0$ such that

$$B_r(u_\lambda) = \{u \in C : d(u, u_\lambda) < r\} \subseteq U.$$

Also, let $\epsilon > 0$ such that $\epsilon^\vartheta \leqslant \frac{(1-l_0)r}{M}$. If $\mu \in B_\epsilon(\lambda) = \{s \in P : \rho(s, \lambda) < \epsilon\}$, then

$$d(h_\mu(u_\lambda), u_\lambda) = d(h_\mu(u_\lambda), h_\lambda(u_\lambda)) \leqslant M\rho(\mu, \lambda)^\vartheta < M\epsilon^\vartheta \leqslant (1 - l_0)r.$$

Proposition 4.2.5 implies that $h_\mu(\cdot)$ has a fixed point $u_\mu \in U$ and so

$$B_\epsilon(\lambda) \subseteq P_0$$
$$\Rightarrow P_0 \subseteq P \text{ is open.}$$

Claims 1, 2, and the connectedness of P imply that $P_0 = P$.

(b) This part is an immediate consequence of *(a)* and (4.18). The proof is now complete. □

This proposition leads to the following alternative result for contractions. So, assume that X is a Banach space and $C \subseteq X$ is a closed convex set. Also $P = [0, 1]$.

Proposition 4.2.17 *If V is a relatively open subset of C with $0 \in V$, then any bounded contraction $\varphi : \overline{V} \to C$ has at least one of the following two properties:*

(a) *φ has a unique fixed point;*
(b) *there exist $u_0 \in \partial V$ and $\lambda \in (0, 1)$ such that $u_0 = \lambda\varphi(u_0)$.*

(Here, bounded φ means that $\varphi(\overline{V}) \subset X$ is bounded.)

Proof For $(\lambda, u) \in [0, 1] \times \overline{V}$, let $h_\lambda(u) = \lambda\varphi(u)$. Evidently, $\{h_\lambda\}_{\lambda\in[0,1]} \subseteq \mathrm{Con}\,(\overline{V}, C)$ is an $l_0 = l(\varphi)$-contractive family.

First suppose that $h_\lambda \in \mathrm{Con}_{\partial V}(\overline{V}, C)$ for all $\lambda \in [0, 1]$. Since $h_0(0) = 0$ from Proposition 4.2.16 we have that $h_1(\cdot) = \varphi(\cdot)$ admits a fixed point in V. If $h_\lambda \notin \mathrm{Con}_{\partial V}(\overline{V}, C)$ for some $\lambda \in [0, 1]$, h_λ must have a fixed point on ∂V and since $0 \in V$, we infer that $\lambda \neq 0$. Therefore either φ has a fixed point on ∂U or statement *(b)* holds. The proof is now complete. □

This alternative result leads to fixed point theorems by imposing conditions on φ which prevent the second alternative from occurring (see also Sect. 3.2).

Corollary 4.2.18 *If V is a relatively open subset of C with $0 \in V$ and $\varphi : \overline{V} \to C$ is a bounded contraction which for all $u \in \partial V$ satisfies one of the following conditions:*

(i) $||\varphi(u)|| \leqslant ||u||$;
(ii) $||\varphi(u)|| \leqslant ||u - \varphi(u)||$;
(iii) $||\varphi(u)||^2 \leqslant ||u||^2 + ||u - \varphi(u)||^2$;
(iv) *$X = H =$ a Hilbert space with inner product $(\cdot, \cdot)_H$ and*

$$(\varphi(u), u)_H \leqslant ||u||^2,$$

then φ has a unique fixed point.

Also, we can have a “contractive” version of Borsuk’s antipodal theorem.

Corollary 4.2.19 *If X is a Banach space, $V \subseteq X$ is open, symmetric and $0 \in V$ and $\varphi : \overline{V} \to X$ is a bounded contraction such that $\varphi(u) = -\varphi(-u)$ for all $u \in \partial V$, then φ has a unique fixed point.*

Proof Since $\varphi\Big|_{\partial V}$ is odd and V is symmetric, we have $||\varphi(u)|| \leqslant ||u||$ for all $u \in \partial V$. So, we can apply Corollary 4.2.18(i) and conclude that φ admits a unique fixed point. The proof is now complete. $\square$

Now we turn our attention to nonexpansive maps (see Definition 4.2.1(c)). Such maps may be fixed point free.

Example 4.2.20 $X = c_0 = \{\hat{u} = (u_n)_{n\geqslant 1} : u_n \to 0\}$ with the norm $||\hat{u}||_\infty = \sup\limits_{n\geqslant 1} |u_n|$. Let $\varphi : X \to X$ be defined by

$$(\varphi(\hat{u}))_1 = \frac{1}{2}(1 + ||\hat{u}||) \text{ and } (\varphi(\hat{u}))_{n+1} = u_n \text{ for all } n \geqslant 1.$$

Evidently, φ is continuous and in fact $||\varphi(\hat{u}) - \varphi(\hat{v})|| = ||\hat{u} - \hat{v}||$ for all $\hat{u}, \hat{v} \in X$. It maps $\overline{B}_1$ into itself. But φ has no fixed point, since $\hat{u} = \varphi(\hat{u})$ implies

$$\begin{aligned} u_n &= \frac{1}{2}[1 + ||\hat{u}||] \text{ for all } n \geqslant 1 \\ \Rightarrow \hat{u} &= (u_n)_{n\geqslant 1} \notin c_0. \end{aligned}$$

So, we want to find out what additional assumptions must be added on the structure of X and/or of the nonexpansive map φ in order to guarantee the existence of at least one fixed point.

We start with the following simple observation.

Lemma 4.2.21 *If X is a Banach space, $C \subseteq X$ is nonempty, closed, convex and bounded and $\varphi : C \to C$ is nonexpansive, then* $\inf\{||u - \varphi(u)|| : u \in C\} = 0$.

Proof We fix $u_0 \in C$ and $t \in (0, 1)$ and consider the map $\varphi_t : C \to C$ defined by

$$\varphi_t(u) = tu_0 + (1 - t)\varphi(u).$$

For all $u, v \in C$, we have

$$\begin{aligned} &||\varphi_t(u) - \varphi_t(v)|| \leqslant (1 - t)||\varphi(u) - \varphi(v)|| \leqslant (1 - t)||u - v|| \\ &\Rightarrow \varphi_t \text{ is a contraction.} \end{aligned}$$

Theorem 4.2.3 implies that there exists a $u_t \in C$ such that $u_t = \varphi_t(u_t)$. Then

$$\begin{aligned} ||u_t - \varphi(u_t)|| &= ||tu_0 + (1 - t)\varphi(u_t) - \varphi(u_t)|| \\ &\leqslant t||u_0 - \varphi(u_t)|| \\ &\leqslant t \text{ diam } C \text{ (recall that } C \text{ is bounded).} \end{aligned}$$

Letting $t \to 0^+$, we conclude that $\inf\{||u - \varphi(u)|| : u \in C\} = 0$. $\square$

Using this lemma, we can state the first elementary fixed point theorem for nonexpansive maps, which is a special case of Schauder's fixed point theorem (see Theorem 3.2.20).

Theorem 4.2.22 *If X is a Banach space, $C \subseteq X$ is nonempty compact convex and $\varphi : C \to C$ is nonexpansive, then φ has a fixed point.*

Proof Let $\{u_n\}_{n\geqslant 1} \subseteq C$ such that $||u_n - \varphi(u_n)|| \downarrow 0$ (see Lemma 4.2.21). The compactness of C implies that we may assume that $u_n \to u \in C$. Then

$$||u - \varphi(u)|| = 0 \Rightarrow u = \varphi(u).$$

The proof is now complete. □

Another straightforward situation is also a consequence of Lemma 4.2.21.

Proposition 4.2.23 *If X is a Banach space, $C \subseteq X$ is nonempty, closed, convex and bounded, $\varphi : C \to C$ is nonexpansive and $(i - \varphi)(C)$ is closed, then φ has a fixed point.*

Proof From Lemma 4.2.21, we have $0 \in \overline{(i - \varphi)(C)} = (i - \varphi)(C)$. So, there exists a $\hat{u} \in C$ such that $\hat{u} = \varphi(\hat{u})$. The proof is now complete. □

To produce fixed points in a more general setting, we need some new ideas. So, let

$$\mathscr{F} = \{D \subseteq C : D \text{ is nonempty, closed, convex and } \varphi(D) \subseteq D \text{ (i.e., } \varphi\text{-invariant)}\}.$$

We partially order $\mathscr{F}$ by setting

$$D_1 \leqslant D_2 \text{ if and only if } D_2 \subseteq D_1.$$

Let $\hat{D} = \bigcap_{D\in\mathscr{F}} D$. Then $\hat{D}$ is closed, convex and φ-invariant, but it may be empty. However, if C is weakly compact (or even stronger compact), then $\hat{D} \neq \emptyset$ (recall that closed convex sets are weakly closed). Then $\hat{D}$ is an upper bound for every chain in $\mathscr{F}$ and so by Zorn's lemma there is a maximal element D_0 of $\mathscr{F}$ (so D_0 is a nonempty, closed, convex, φ-invariant set which is minimal with respect to inclusion). Then $\overline{\text{conv}}\, \varphi(D_0) = D_0$ and so if D_0 is a singleton, we have a fixed point for φ. However, in general if C is weakly compact and convex, the set D_0 need not be a singleton (see Alspach [10]). So, we introduce the following notions:

Definition 4.2.24 Let X be a Banach space and $C \subseteq X$ nonempty.

(a) A point $u \in C$ is a diametral point of C if

$$\sup\{||y - u|| : y \in C\} = \text{diam}\, C.$$

(b) A convex set $C \subseteq X$ is said to have normal structure if each bounded convex subset K of C with diam $K > 0$ contains a nondiametral point.

(c) We set

$$\begin{aligned} r_y(C) &= \sup\{||y-u|| : u \in C\}, \\ r(C) &= \inf\{r_u(C) : u \in X\}, \\ C_0 &= \{u \in X : r_u(C) = r(C)\}. \end{aligned}$$

Remark 4.2.25 $r(C)$ is called the Chebyshev radius of C and C_0 is the Chebyshev center of C.

Proposition 4.2.26 *If X is a Banach space and $C \subseteq X$ is compact convex, then C has normal structure.*

Proof Without any loss of generality we assume that diam $C > 0$. Suppose that C does not have normal structure. Then given $u_1 \in C$ we can find $u_2 \in C$ such that diam $C = ||u_1 - u_2||$. The convexity of C implies that $\frac{1}{2}[u_1 + u_2] \in C$. Then we can find $u_3 \in C$ such that

$$\operatorname{diam} C = \left\| u_3 - \frac{1}{2}[u_1 + u_2] \right\|.$$

Continuing in this way, inductively we can generate a sequence $\{u_n\}_{n\geqslant 1} \subseteq C$ such that

$$\begin{aligned} \operatorname{diam} C &= \left\| u_{n+1} - \frac{1}{n}\sum_{k=1}^{n} u_k \right\| \text{ for all } n \geqslant 2 \\ \Rightarrow \operatorname{diam} C &\leqslant \frac{1}{n}\sum_{k=1}^{n} ||u_{n+1} - u_k|| \leqslant \operatorname{diam} C \\ \Rightarrow \operatorname{diam} C &= ||u_{n+1} - u_k|| \text{ for all } k \in \{1, \cdots, n\}. \end{aligned}$$

This means that $\{u_n\}_{n\geqslant 1} \subseteq C$ has no convergent subsequence, contradicting the compactness of C. The proof is now complete. ☐

Lemma 4.2.27 *If X is a reflexive Banach space and $C \subseteq X$ is nonempty, closed, convex and bounded, then C_0 is nonempty and convex.*

Proof Let $u \in C$ and let $C_n(u) = \{v \in C : ||v - u|| \leqslant r(C) + \frac{1}{n}\}$. We set $\hat{C}_n = \bigcap_{u\in C} C_n(u)$. Then each $\hat{C}_n$ is nonempty, closed, convex and $\{\hat{C}_n\}_{n\geqslant 1}$ is decreasing. Therefore by virtue of the reflexivity of X we have

$$C_0 = \bigcap_{n\geqslant 1} \hat{C}_n \text{ is nonempty, closed and convex.}$$

The proof is now complete. ☐

Lemma 4.2.28 *If X is a Banach space and $C \subseteq X$ is nonempty, closed, convex and bounded, and has normal structure, then* $\operatorname{diam} C_0 < \operatorname{diam} C$.

Proof Since C has normal structure, we can find $u \in C$ such that $r_u(C) < \operatorname{diam} C$ (see Definition 4.2.24). If $v, y \in C_0$, then

$$\begin{aligned} &||y - v|| \leqslant r_v(C) = r(C) \leqslant r_u(C) < \operatorname{diam} C \\ \Rightarrow\ &\operatorname{diam} C_0 < \operatorname{diam} C. \end{aligned}$$

The proof is now complete. □

Now we can state the main fixed point theorem for nonexpansive maps.

Theorem 4.2.29 *If X is a reflexive Banach space, $C \subseteq X$ is nonempty, closed, convex, bounded and has normal structure and $\varphi : C \to C$ is nonexpansive, then φ has a fixed point.*

Proof Let $\mathscr{S} = \{D \subseteq C :$ nonempty, closed, convex and $\varphi(D) \subseteq D\}$. The reflexivity of X and Zorn's lemma imply that $\mathscr{S}$ has a minimal element C^*. Suppose that $\operatorname{diam} C^* > 0$ and let $u \in C_0^*$. Then

$$\begin{aligned} &||\varphi(u) - \varphi(v)|| \leqslant ||u - v|| \leqslant r(C^*) \text{ for all } v \in C^* \\ \Rightarrow\ &\varphi(C^*) \subseteq \overline{B}_{r(C^*)}(\varphi(u)) = \overline{B}^*. \end{aligned}$$

Then $\varphi(C^* \cap \overline{B}^*) \subseteq C^* \cap \overline{B}^*$ and so the minimality of C^* implies $C^* \subseteq \overline{B}^*$. Therefore $\varphi(u) \in C_0^*$ and so $\varphi(C_0^*) \subseteq C_0^*$. From Lemma 4.2.27 we have that $C_0^* \in \mathscr{S}$. Since $\operatorname{diam} C^* > 0$, from Lemma 4.2.28 it follows that C_0^* is a proper subset of C^*, which contradicts the minimality of C^*. Therefore C^* is a singleton and it is a fixed point of φ. The proof is now complete. □

A careful reading of the proofs of Lemma 4.2.27 and Theorem 4.2.29 reveals that we can drop the reflexivity requirement of X and instead assume that $C \subseteq X$ is nonempty convex and weakly compact. Recall that in a reflexive Banach space a closed, convex and bounded set is weakly compact.

Theorem 4.2.30 *If X is a Banach space, $C \subseteq X$ is nonempty, weakly compact, convex and has normal structure and $\varphi : C \to C$ is nonexpansive, then φ has a fixed point.*

The normal structure hypothesis on the set C is satisfied in the following case.

Proposition 4.2.31 *If X is a uniformly convex Banach space and $C \subseteq X$ is nonempty, closed, convex and bounded, then C has normal structure.*

Proof Since X is uniformly convex (see Definition 2.7.29), given $\epsilon \in (0, 2]$, we can find $\delta(\epsilon) > 0$ such that

$$||u||, ||v|| \leqslant 1 \text{ and } ||u - v|| \geqslant \epsilon \Rightarrow \left\| \frac{1}{2}[u + v] \right\| \leqslant 1 - \delta(\epsilon).$$

Because C is bounded, without any loss of generality, we may assume that

$$C \subseteq \overline{B}_1 = \{u \in X : ||u|| \leqslant 1\}.$$

Let $D \subseteq C$ be nonempty, closed, convex, $u_1 \in D$ and $\epsilon = \frac{1}{2}$. Choose $u_2 \in D$ such that

$$||u_2 - u_1|| \geqslant \frac{1}{2} \text{diam } D.$$

Then for any $u \in D$ we have

$$\begin{aligned} \left\| u - \frac{1}{2}[u_1 + u_2] \right\| &= \left\| \frac{1}{2}[u - u_1] + \frac{1}{2}[u - u_2] \right\| \\ &\leqslant \text{diam } D \left[1 - \delta\left(\frac{1}{2}\right) \right]. \end{aligned}$$

Since $u \in D$ is arbitrary and $\delta\left(\frac{1}{2}\right) > 0$, we conclude that D has normal structure. The proof is now complete. $\square$

Remark 4.2.32 Recall that a uniformly convex Banach space is reflexive (Milman–Pettis theorem). If X is uniformly convex, then in Theorem 4.2.29 the hypothesis that C has normal structure is redundant.

In the case of Hilbert spaces, we can say more. More precisely, let H be a Hilbert space and $\overline{B}_\rho = \{u \in H : ||u|| \leqslant \rho\}$. We consider nonexpansive maps $\varphi : \overline{B}_\rho \to H$. In our search for fixed points of φ, we will use the radial retraction map, $r : H \to \overline{B}_\rho$, defined by

$$r(u) = \begin{cases} u & \text{if } ||u|| \leqslant \rho \\ \rho \dfrac{u}{||u||} & \text{if } ||u|| > \rho. \end{cases} \tag{4.19}$$

Evidently, $r(u) = \text{proj}_{\overline{B}_\rho}(u)$ (=the metric projection onto $\overline{B}_\rho$).

Proposition 4.2.33 *The radial retraction $r : H \to \overline{B}_\rho$ defined by (4.19) is nonexpansive.*

Proof For all $u, v \in H$, we have

$$(u - r(u), r(v) - r(u))_H \leqslant 0 \tag{4.20}$$

(by $(\cdot, \cdot)_H$ we denote the inner product of H). We write

$$u - v = r(u) - r(v) + u - r(u) + r(v) - v = r(u) - r(v) + h$$

with $h = u - r(u) + r(v) - v$. Then

$$||u - v||^2 = ||r(u) - r(v)||^2 + ||h||^2 + 2(h, r(u) - r(v))_H. \tag{4.21}$$

We have

$$(h, r(u) - r(v))_H = -(u - r(u), r(v) - r(u))_H - (v - r(v), r(u) - r(v))_H \geqslant 0$$
(see (4.7))

$\Rightarrow ||r(u) - r(v)||^2 \leqslant ||u - v||^2$ (see (4.20))

$\Rightarrow r$ is nonexpansive.

The proof is now complete. □

This observation leads to an alternative theorem for nonexpansive maps in a Hilbert space.

Theorem 4.2.34 *If H is a Hilbert space, $\overline{B}_\rho = \{u \in H : ||u|| \leqslant \rho\}$ and $\varphi : \overline{B}_\rho \to H$ is nonexpansive, then φ satisfies at least one of the following statements:*

(a) φ has a fixed point;
(b) there exist $u \in \partial B_\rho$ and $\lambda \in (0, 1)$ such that $u = \lambda\varphi(u)$.

Proof We consider the map $r \circ \varphi : \overline{B}_\rho \to \overline{B}_\rho$, where r is the radial retraction defined by (4.19). Proposition 4.2.33 implies that $r \circ \varphi$ is nonexpansive. Recalling that a Hilbert space is uniformly convex, from Proposition 4.2.31 we infer that $\overline{B}_r$ has normal structure. So, we can apply Theorem 4.2.29 and find $\hat{u} \in \overline{B}_\rho$ such that $\hat{u} = r \circ \varphi(\hat{u})$. If $\varphi(\hat{u}) \in \overline{B}_\rho$, then $\hat{u} = r(\varphi(\hat{u})) = \varphi(\hat{u})$ (see (4.19)) and so φ has a fixed point. If $\varphi(\hat{u}) \in H \setminus \overline{B}_\rho$, then $\hat{u} = \rho \frac{\varphi(\hat{u})}{||\varphi(\hat{u})||}$ (see (4.19)) and so $\hat{u} \in \partial B_\rho$. Thus taking $\lambda = \frac{\rho}{||\varphi(\hat{u})||} < 1$, we see that statement *(b)* holds. □

As before (see Corollary 4.2.18), by imposing conditions on φ which prohibit the occurrence of the second possibility, we obtain several fixed point results.

Corollary 4.2.35 *If H is a Hilbert space, $\overline{B}_\rho = \{u \in H : ||u|| \leqslant \rho\}$, $\varphi : \overline{B}_\rho \to H$ is nonexpansive and for all $u \in \partial B_\rho$ one of the following conditions holds*

(i) $||\varphi(u)|| \leqslant ||u||$;
(ii) $||\varphi(u)|| \leqslant ||u - \varphi(u)||$;
(iii) $||\varphi(u)||^2 \leqslant ||u||^2 + ||u - \varphi(u)||^2$;
(iv) $(\varphi(u), u)_H \leqslant ||u||^2$;
(v) $\varphi\big|_{\partial\Omega}$ *is odd;*

then φ has a fixed point.

As an application of these fixed point theorems, we prove the following surjectivity result.

Corollary 4.2.36 *If H is a Hilbert space, $\varphi : H \to H$ is nonexpansive and*

$$(u, u - \varphi(u))_H \geqslant \vartheta(||u||)||u|| \text{ for all } u \in H$$

with $\vartheta(t) \to \infty$ as $t \to \infty$, then $u \to g(u) = u - \varphi(u)$ is surjective on H.

Proof Let $u_0 \in H$ and consider the map $h(u) = g(u) - u_0$ for all $u \in H$. We have

$$\frac{(u, h(u))_H}{||u||} = \frac{(u, g(u))_H}{||u||} - \frac{(u, u_0)_H}{||u||} \geqslant \vartheta(||u||) - ||u_0||.$$

So, for large $\rho > 0$ we have

$$(u, h(u))_H \geqslant 0 \text{ for all } ||u|| = \rho.$$

Invoking Corollary 4.2.35(d), we know that we can find a $\hat{u} \in H$ such that

$$\begin{aligned} \hat{u} &= \varphi(\hat{u}) + u_0 \\ \Rightarrow u_0 &= \hat{u} - \varphi(\hat{u}) = g(\hat{u}). \end{aligned}$$

Since $u_0 \in H$ is arbitrary, we conclude that g is surjective. The proof is now complete. □

4.3 Topological Fixed Points

Topological fixed point theorems use more fundamental results from topology in order to establish the existence of fixed points and so go outside the framework of metric spaces. We have already seen two prototype such results in Chap. 3 in conjunction with the study of degree theories. The first was Brouwer's fixed point theorem (finite-dimensional case, see Theorem 3.1.34) and the second was Schauder's fixed point theorem (infinite-dimensional case, see Theorem 3.2.20). In this section, we continue in this direction and, using mainly topological tools, we prove some new fixed point theorems, which are of interest in applications.

We start by extending the Brouwer fixed point theorem (see Theorem 3.1.34).

Proposition 4.3.1 *If $C \subseteq \mathbb{R}^N$ is a compact set homeomorphic to the closed unit ball of $\mathbb{R}^N$ and $\varphi : C \to C$ is continuous, then φ has a fixed point.*

Proof Let $h : C \to \overline{B}_1$ be the homeomorphism between C and the unit ball $\overline{B}_1 = \{u \in \mathbb{R}^N : ||u|| \leqslant 1\}$. Then $\psi = h \circ \varphi \circ h^{-1} : \overline{B}_1 \to \overline{B}_1$ is a continuous map and so we can apply Theorem 3.1.34 and find $\hat{y} \in \overline{B}_1$ such that $\psi(\hat{y}) = \hat{y}$. Then $\hat{u} = h^{-1}(\hat{y}) \in C$ is a fixed point of φ. □

Proposition 4.3.2 *If $C \subseteq \mathbb{R}^N$ is compact and convex, then C is homeomorphic to the closed unit ball of $\mathbb{R}^m$ for some $m \leqslant N$.*

Proof First assume that $\operatorname{int} C \neq \emptyset$. Translating things if necessary (recall that the translation is a homeomorphism), we may assume that $0 \in \operatorname{int} C$. So, we can find $r > 0$ such that $B_r \subseteq C$. We introduce the Minkowski functional corresponding to C, defined by

$$j_C(u) = \inf\left\{\lambda > 0 : \frac{1}{\lambda}u \in C\right\}.$$

Then $j_C(u) \leqslant \frac{1}{r}||u||$ for all $u \in \mathbb{R}^N$ and so $j_C : \mathbb{R}^N \to \mathbb{R}_+$. Since C is compact, we can find $R > 0$ such that $C \subseteq \overline{B}_R$ and so $\frac{||u||}{R} \leqslant j_C(u)$. Also we have

$$\cdot \;\; u \in C \Rightarrow j_C(u) \leqslant 1. \tag{4.22}$$

$$\cdot \;\; j_C(tu) = t j_C(u) \text{ for all } t \geqslant 0, \text{ all } u \in \mathbb{R}^N. \tag{4.23}$$

$$\cdot \;\; j_C(u+v) \leqslant j_C(u) + j_C(v) \text{ for all } u, v \in \mathbb{R}^N. \tag{4.24}$$

Properties (4.22) and (4.23) are evident from the definition of the Minkowski functional j_C. We prove (4.24). Let $\lambda_1 > 0$ and $\lambda_2 > 0$ be such that $\frac{u}{\lambda_1}, \frac{v}{\lambda_2} \in C$. Then

$$\begin{aligned} &\frac{u+v}{\lambda_1+\lambda_2} = \frac{\lambda_1}{\lambda_1+\lambda_2}\frac{u}{\lambda_1} + \frac{\lambda_2}{\lambda_1+\lambda_2}\frac{v}{\lambda_2} \in C \text{ (due to the convexity of } C) \\ &\Rightarrow j_C(u+v) \leqslant \lambda_1 + \lambda_2 \\ &\Rightarrow j_C(u+v) \leqslant j_C(u) + j_C(v). \end{aligned}$$

Moreover, the converse of (4.22) holds, namely "$j_C(u) \leqslant 1 \Rightarrow u \in C$". To see this let $\lambda_n \to j_C(y)$ and $y_n \in C$ such that $u = \lambda_n y_n + (1-\lambda_n)0$. If $\lambda_n \in [0,1]$ for all $n \geqslant n_0$, then $u \in C$, being the convex combination of elements in C. If $\lambda_n > 1$ for all $n \geqslant 1$, then $\lambda_n \to j_C(u) = 1$ and so $y_n \to u$. Since C is closed, we conclude that $u \in C$.

For all $u, v \in \mathbb{R}^N$ we have

$$-j_C(-v) \leqslant j_C(u+v) - j_C(u) \leqslant j_C(v) \text{ (see (4.24))}$$

and $\max\{|j_C(-v)|, |j_C(v)|\} \leqslant \frac{||v||}{r}$. Therefore j_C is continuous.

We introduce the following two functions, $f, g : \mathbb{R}^N \to \mathbb{R}^N$:

$$f(u) = \begin{cases} \frac{j_C(u)}{||u||}u & \text{if } u \neq 0 \\ 0 & \text{if } u = 0 \end{cases} \text{ and } g(y) = \begin{cases} \frac{||y||}{j_C(y)}y & \text{if } y \neq 0 \\ 0 & \text{if } y = 0. \end{cases} \tag{4.25}$$

It is easy to see that $f \circ g = g \circ f = i_{\mathbb{R}^N}$. Moreover, from the continuity of $j_C(\cdot)$ we see that both are continuous on $\mathbb{R}^N \setminus \{0\}$. In addition, we have

$$
\begin{aligned}
&||f(u)|| \leqslant j_C(u) \leqslant \frac{||u||}{r}\\
&\Rightarrow f \text{ is continuous at } u = 0, \text{ hence so is } g\\
&\Rightarrow f, g \text{ are homeomorphisms of } \mathbb{R}^N.
\end{aligned}
$$

Finally, we show that $f(C) = \overline{B}_1 = \{u \in \mathbb{R}^N : ||u|| \leqslant 1\}$. If $u \in C$, then $j_C(u) \leqslant 1$ (see (4.22)) and so $||f(u)|| \leqslant 1$ (see (4.25)), hence $f(C) \subseteq \overline{B}_1$. On the other hand, if $y \in \overline{B}_1$, then

$$
\begin{aligned}
&j_C(g(u)) = \frac{||y||}{j_C(y)} j_C(y) \leqslant 1 \text{ (see (4.23), (4.25))}\\
&\Rightarrow g(\overline{B}_1) \subseteq C\\
&\Rightarrow \overline{B}_1 \subseteq f(C) \text{ (recall } f = g^{-1})\\
&\Rightarrow f(C) = \overline{B}_1.
\end{aligned}
$$

If $\operatorname{int} C = \emptyset$, then due to the convexity of C we know that its relative interior is nonempty (recall that $\operatorname{rint} C$ is the interior of C in the affine hull of C, which is $\mathbb{R}^m$ for some $m < N$). So, we repeat the above argument in $\mathbb{R}^m$. The proof is now complete. □

Combining Propositions 4.3.1 and 4.3.2, we obtain.

Corollary 4.3.3 *If $C \subseteq \mathbb{R}^N$ is compact convex and $\varphi : C \to C$ is continuous, then φ has a fixed point.*

Remark 4.3.4 In general, in contrast to the Banach fixed point theorem (see Theorem 4.2.3), the fixed point obtained in the above corollary need not be unique.

Proposition 4.3.5 *If $\varphi : \mathbb{R}^N \to \mathbb{R}^N$ is continuous and there exists an $r > 0$ such that*

$$(\varphi(u), u)_{\mathbb{R}^N} \geqslant 0 \text{ for all } u \in \overline{B}_r,$$

then there exists a $\hat{u} \in \overline{B}_r$ such that $\varphi(\hat{u}) = 0$.

Proof Arguing by contradiction suppose that $||\varphi(u)|| > 0$ for all $||u|| \leqslant r$. Then we can define the map $\psi : \overline{B}_r \to \partial B_r$ by setting

$$\psi(u) = -r \frac{\varphi(u)}{||\varphi(u)||}.$$

Evidently, ψ is continuous. So, Corollary 4.3.3 implies that there exists a $u_0 \in \overline{B}_r$ such that

$$
\begin{aligned}
&\psi(u_0) = u_0 = -r \frac{\varphi(u_0)}{||\varphi(u_0)||}\\
&\Rightarrow ||u_0|| = r.
\end{aligned}
$$

Then

$$r^2 = ||u_0||^2 = (\psi(u_0), u_0)_{\mathbb{R}^N} = -\frac{r}{||\varphi(u_0)||}(\varphi(u_0), u_0)_{\mathbb{R}^N} \leqslant 0,$$

a contradiction. So, there exists a $\hat{u} \in \overline{B}_r$ such that $\varphi(\hat{u}) = 0$. $\square$

Remark 4.3.6 The hypothesis $(\varphi(u), u)_{\mathbb{R}^N} \geqslant 0$ for $u \in \partial B_r$ means that the map φ points to the exterior of $\overline{B}_r$ on ∂B_r. Evidently, Proposition 4.3.5 implies the lack of retraction of $\overline{B}_r$ onto ∂B_r (see Proposition 3.1.32). Indeed, if a retraction existed, then we could use it as φ in the above proposition and reach a contradiction. So, Proposition 4.3.5 is in fact equivalent to Brouwer's fixed point theorem. Next we present some more interesting topological results which are equivalent to the Brouwer fixed point theorem. First a definition.

Definition 4.3.7 Let X, Y be Hausdorff topological spaces.

(a) Two continuous maps $\varphi, \psi : X \to Y$ are said to be homotopic if there exists a continuous map $h : [0, 1] \times X \to Y$ such that

$$h(0, \cdot) = \varphi(\cdot) \text{ and } h(1, \cdot) = \psi(\cdot).$$

The map h is called a homotopy (or deformation) of φ to ψ. If φ, ψ are homotopic we write $\varphi \simeq \psi$.

(b) A map $\varphi : X \to Y$ homotopic to a constant map is said to be nullhomotopic and we write $\varphi \simeq 0$.

(c) A space X is said to be contractible if the identity map $i_X : X \to X$ is nullhomotopic.

Theorem 4.3.8 *The following statements are equivalent:*

(a) $\partial B_1 = \{u \in \mathbb{R}^N : ||u|| = 1\}$ *is not contractible in itself.*

(b) Every continuous map $\varphi : \overline{B}_1 = \{u \in \mathbb{R}^N : ||u|| \leqslant 1\} \to \mathbb{R}^N$ *has at least one of the following properties:*

(b_1) φ *has a fixed point;*

(b_2) there exists $\hat{u} \in \partial B_1$ *and* $\lambda \in (0, 1)$ *such that* $\hat{u} = \lambda\varphi(\hat{u})$.

(c) Every continuous map $\varphi : \overline{B}_1 \to \overline{B}_1$ *has a fixed point.*

(d) ∂B_1 *is not a retract of* $\overline{B}_1$.

Proof <u>(a) ⇒ (b)</u>

Arguing by contradiction, suppose that

$$\varphi(u) \neq u \text{ for all } u \in \overline{B}_1 \text{ and } y \neq \lambda\varphi(y) \text{ for all } \lambda \in (0, 1) \text{ and all } y \in \partial B_1.$$

In fact, $y \neq \lambda\varphi(y)$ also for $\lambda = 0$ and for $\lambda = 1$. Let $r : \mathbb{R}^N \setminus \{0\} \to \partial B_1$ be the map defined by $r(u) = \frac{u}{||u||}$. We consider the homotopy $h : [0, 1] \times \partial B_1 \to \partial B_1$ defined by

$$h(\lambda, y) = \begin{cases} r(y - 2\lambda\varphi(u)) & \text{if } \lambda \in \left[0, \frac{1}{2}\right] \\ r((2-2\lambda)y - \varphi((2-2\lambda)y)) & \text{if } \lambda \in \left[\frac{1}{2}, 1\right]. \end{cases}$$

This shows that ∂B_1 is contractible in itself, a contradiction.

$\underline{(b) \Rightarrow (c)}$

Property (b_2) cannot occur since $\varphi(\partial B_1) \subseteq \overline{B}_1$. So, φ has a fixed point.

$\underline{(c) \Rightarrow (d)}$

See Proposition 3.1.36.

$\underline{(d) \Rightarrow (a)}$

Arguing by contradiction, suppose that $i_{\mathbb{R}^N} \simeq 0$. Let h be a homotopy such that $h(0, \partial B_1) = u_0 \in \partial B_1$. We introduce the map $r : \overline{B}_1 \to \partial B_1$ defined by

$$r(u) = \begin{cases} u_0 & \text{if } ||u|| \leqslant \frac{1}{2} \\ h\left(2||u|| - 1, \frac{u}{||u||}\right) & \text{if } \frac{1}{2} \leqslant ||u|| \leqslant 1. \end{cases}$$

Then r is a retraction of $\overline{B}_1$ onto ∂B_1, a contradiction. The proof is now complete. $\square$

A purely topological version of this theorem has the following form:

Theorem 4.3.9 *The following statements are equivalent.*

(a) The Ljusternik–Schnirelmann–Borsuk theorem holds (see Proposition 3.1.51).
(b) There is no continuous odd map $\varphi : \partial B_1^{N+1} = \{u \in \mathbb{R}^{N+1} : ||u|| = 1\} \to \partial B_1^N = \{u \in \mathbb{R}^N : ||u|| = 1\}$.
(c) A continuous odd map $\varphi : \partial B_1^N \to \partial B_1^N$ *is not nullhomotopic.*
(d) For every continuous map $\varphi : \partial B_1^N \to \mathbb{R}^{N-1}$ *we can find* $\hat{u} \in \partial B_1^N$ *such that*

$$\varphi(\hat{u}) = \varphi(-\hat{u}).$$

Proof $\underline{(a) \Rightarrow (b)}$

Arguing by contradiction, suppose that there exists a continuous odd map $\varphi : \partial B_1^{N+1} \to \partial B_1^N$. Consider the N-simplex centered at the origin. Its boundary is homeomorphic to ∂B_1^N. Let $\{C_k\}_{k=1}^{N+1}$ be the images of the $(N-1)$-faces. So, each C_k is a closed set and contains no antipodal points. Let $D_k = \varphi^{-1}(C_k)$, $k = 1, \dots, N+1$. These are closed sets and cover ∂B_1^{N+1}. So, by the Ljusternik–Schnirelmann–Borsuk theorem, there exists a $u \in D_k \cap (-D_k)$ for some $k = 1, \dots, N+1$. Since φ is odd, we have $\varphi(-u), -\varphi(u) \in C_k$, a contradiction.

$\underline{(b) \Rightarrow (c)}$

Again we proceed indirectly. So, suppose that there exists a continuous odd map $\varphi : \partial B_1^N \to \partial B_1^N$ which is nullhomotopic. This means that there exists a homotopy h deforming φ to a constant map. Let

$$\hat{\varphi}(u) = \begin{cases} h(0, \partial B_1^N) & \text{if } 0 \leqslant ||u|| \leqslant \frac{1}{2} \\ h\left(2||u|| - 1, \frac{u}{||u||}\right) & \text{if } \frac{1}{2} \leqslant ||u|| \leqslant 1. \end{cases}$$

Then $\hat{\varphi} : \overline{B}_1^N \to \partial B_1^N$ is continuous and $\hat{\varphi}\Big|_{\partial B_1^N} = \varphi$. Let

$$\partial B_{1,+}^{N+1} = \{u \in \partial B_1^{N+1} : u_{N+1} \geqslant 0\} \text{ and}$$
$$\partial B_{1,-}^{N+1} = \{u \in \partial B_1^{N+1} : u_{N+1} \leqslant 0\} \ (u = (u_k)_{k=1}^{N+1} \in \partial B_1^{N+1}).$$

We know that $\overline{B}^N$ is homeomorphic to each of the above hemispheres. So, the following map is well-defined

$$g(u) = \begin{cases} \hat{\varphi}(u) & \text{if } u \in \partial B_{1,+}^{N+1} \\ -\hat{\varphi}(u) & \text{if } u \in \partial B_{1,-}^{N+1}. \end{cases}$$

Then $g : \partial B_1^{N+1} \to \partial B_1^N$ is continuous and odd, a contradiction.

(c) ⇒ (d)

Proceeding by contradiction, suppose that we can find a continuous map $\varphi : \partial B_1^N \to \mathbb{R}^{N+1}$ such that $\varphi(u) \neq \varphi(-u)$ for all $u \in \partial B_1^N$. Let $\psi : \partial B_1^N \to \partial B_1^{N-1}$ be defined by

$$\psi(u) = \frac{\varphi(u) - \varphi(-u)}{||\varphi(u) - \varphi(-u)||}.$$

Then $\psi\Big|_{\partial B_1^{N-1}} : \partial B_1^{N-1} \to \partial B_1^{N-1}$ is odd and since $\psi\Big|_{\partial B_{1,+}^{N+1}}$ is an extension over $\overline{B}_1^N$, $\psi\Big|_{\partial B_1^{N-1}}$ is nullhomotopic (just take the deformation $h(t,u) = \psi\Big|_{\partial B_{1,+}^{N+1}}(tu)$), a contradiction.

(d) ⇒ (a)

Suppose that $\{C_k\}_{k=1}^{N+1}$ is a closed cover of ∂B_1^{N+1} and no C_k contains a pair of antipodal points, that is, $C_k \cap (-C_k) = \emptyset$ for every $k = 1, \cdots, N+1$. Using Urysohn's theorem for each $k = 1, \cdots, N$, we can find $f_k : \partial B_1^{N+1} \to [0,1]$ such that

$$f_k \mid_{C_k} = 0 \text{ and } f_k \mid_{-C_k} = 1.$$

Let $\hat{f} : \partial B_1^{N+1} \to \mathbb{R}^N$ be defined by

$$\hat{f}(u) = (f_k(u))_{k=1}^N.$$

By virtue of *(d)*, there exists a $\hat{u} \in \partial B_1^{N+1}$ such that $\hat{f}(\hat{u}) = \hat{f}(-\hat{u})$. Hence $f_k(\hat{u}) = f_k(-\hat{u})$ for all $k = 1, \cdots, N$ and so $\hat{u} \in \partial B_1^{N+1} \setminus \left(\bigcup_{k=1}^N C_k \cup \bigcup_{k=1}^N (-C_k) \right)$. The families $\{C_k\}_{k=1}^{N+1}$ and $\{-C_k\}_{k=1}^{N+1}$ cover ∂B_1^{N+1}, so we must have $\hat{u} \in C_{N+1} \cap (-C_{N+1})$, a contradiction. □

We present two interesting consequences of this theorem.

The first relaxes the oddness hypothesis in part *(c)* of the theorem.

Proposition 4.3.10 *If* $\varphi : \partial B_1^N \to \partial B_1^N$ *is continuous and* $\varphi(u) \neq \varphi(-u)$ *for all* $u \in \partial B_1^N$*, then* φ *is not nullhomotopic.*

Proof Let $\psi : \partial B_1^N \to \partial B_1^N$ be defined by

$$\psi(u) = \frac{\varphi(u) - \varphi(-u)}{||\varphi(u) - \varphi(-u)||} \text{ for all } u \in \partial B_1^N.$$

Then ψ is continuous and odd. Suppose that for some $y \in \partial B_1^N$, we have $\psi(y) = -\varphi(y)$. Then

$$\begin{aligned} & [1 + ||\varphi(y) - \varphi(-y)||]\varphi(y) = \varphi(-y) \\ \Rightarrow\ & 1 + ||\varphi(y) - \varphi(-y)|| = 1 \text{ (recall that } ||\varphi(y)|| = ||\varphi(-y)|| = 1) \\ \Rightarrow\ & \varphi(y) = \varphi(-y), \text{ a contradiction to our hypothesis.} \end{aligned}$$

So, the maps φ and ψ are never antipodal. We consider the homotopy

$$h_t(u) = \frac{(1-t)\varphi(u) + t\psi(u)}{||(1-t)\varphi(u) + t\psi(u)||} \text{ for all } (t, u) \in [0, 1] \times \partial B_1^N$$
$$\Rightarrow \varphi \simeq \psi.$$

But from Theorem 4.3.9(c) we know that ψ is not nullhomotopic. Since $\simeq$ is an equivalence relation, we conclude that φ is not nullhomotopic too. □

The second consequence of Theorem 4.3.9 is the so-called "Borsuk Fixed Point Theorem" (see also Theorem 3.1.45, the Borsuk–Ulam theorem).

Proposition 4.3.11 *If* U *is a bounded, open, convex and symmetric neighborhood of the origin,* $\varphi : \overline{U} \to \mathbb{R}^N$ *is continuous and* $\varphi\big|_{\partial U}$ *is odd, then* φ *has a fixed point.*

Proof Let $j_U : \mathbb{R}^N \to \mathbb{R}_+$ be the Minkowski functional for the set U (see the proof of Proposition 4.3.2). Let E^N be the space $\mathbb{R}^N$ furnished with the norm $|\cdot| = j_U(\cdot)$. Then the identity map $h : \mathbb{R}^N \to E^N$ is a homeomorphism which maps $\overline{U}$ onto $\overline{B}_1^{|\cdot|} = \{u \in E^N : |u| \leqslant 1\}$. Let $\psi = h \circ \varphi \circ h^{-1} : \overline{B}_1^{|\cdot|} \to \mathbb{R}^N$. Then ψ is continuous and odd on $\partial B_1^{|\cdot|}$. Suppose that $\psi(u) \neq u$ for all $u \in \overline{B}_1^{|\cdot|}$. We introduce the map $\tau : \overline{B}_1^{|\cdot|} \to \partial B_1^{|\cdot|}$ defined by

$$\tau(u) = \frac{\psi(u) - u}{|\psi(u) - u|} \text{ for all } u \in \overline{B}_1^{|\cdot|}.$$

Then τ is continuous and $\tau\big|_{\partial B_1^{|\cdot|}}$ is nullhomotopic (since E^N is finite-dimensional). But note that $\tau\big|_{\partial B_1^{|\cdot|}}$ is odd. This then contradicts Theorem 4.3.9(c). So, we can find $\hat{u} \in \overline{B}_1^{|\cdot|}$ such that

$$\begin{aligned}
&\psi(\hat{u}) = \hat{u} \\
\Rightarrow\ &(h \circ \varphi \circ h^{-1})(\hat{u}) = \hat{u} \\
\Rightarrow\ &\varphi(h^{-1}(\hat{u})) = h^{-1}(\hat{u}).
\end{aligned}$$

The proof is now complete. □

Now we turn our attention to infinite-dimensional Banach spaces (generalizations of the Schauder fixed point theorem, see Theorem 3.2.20).

In the next result, the set D need not be convex and it is only the image of the boundary ∂D under the map φ that matters.

Theorem 4.3.12 *If X is a Banach space, $D \subseteq X$ is bounded and closed with* $\operatorname{int} D \neq \emptyset$, *$\varphi \in K(D, X)$ and there exists a $u_0 \in \operatorname{int} D$ such that*

$$\varphi(u) - u_0 \neq \lambda(u - u_0) \text{ for all } u \in \partial D \text{ and all } \lambda > 1, \tag{4.26}$$

then φ has a fixed point.

Proof Consider the homotopy $\{h_t\}_{t\in[0,1]}$ of compact maps defined by

$$h_t(u) = t(\varphi(u) - u_0) - u_0 \text{ for all } (t, u) \in [0, 1] \times (\overline{\operatorname{int} D}).$$

We may always assume that $0 \notin (i - \varphi)(\partial D)$ (or otherwise we already have a fixed point for φ). Suppose that

$$u - h_t(u) = 0 \text{ for some } t \in (0, 1) \text{ and some } u \in \partial(\operatorname{int} D).$$

Then hypothesis (4.26) is violated for $\lambda = \frac{1}{t} > 1$. Therefore

$$0 \notin (i - h_t)(\partial(\operatorname{int} D)) \text{ for all } t \in [0, 1].$$

Hence the homotopy invariance property of the Leray–Schauder degree (see Theorem 3.2.15 (c)) implies

$$d_{LS}(i - \varphi, \operatorname{int} D, 0) = d_{LS}(i - u_0, \operatorname{int} D, 0). \tag{4.27}$$

Since $u_0 \in \operatorname{int} D$, we have $d_{LS}(i - u_0, \operatorname{int} D, 0) = d_{LS}(i, \operatorname{int} D, u_0) = 1$ (normalization property, see Theorem 3.2.15 (a)). Therefore

$$d_{LS}(i - \varphi, \operatorname{int} D, 0) = 1 \text{ see (4.27)}.$$

So, by the solution property (see Theorem 3.2.15 (d)), we can find $\hat{u} \in \operatorname{int} D$ such that

$$\hat{u} = \varphi(\hat{u}).$$

This completes the proof. □

Remark 4.3.13 Condition (4.26) relates to $\partial(\text{int } D)$ and we know that $\partial(\text{int } D) \subseteq \partial D$. It is satisfied if D is convex and $\varphi(\partial D) \subseteq D$.

Theorem 4.3.12 can be extended to γ-condensing maps using this time the degree map d_C established in Definition 3.7.13. So, we have the following property.

Theorem 4.3.14 *If X is a Banach space, $\Omega \subseteq X$ is bounded open, $\varphi : \overline{\Omega} \to X$ is γ-condensing, $0 \notin (i - \varphi)(\partial\Omega)$ and there exists a $u_0 \in \Omega$ such that*

$$\varphi(u) - u_0 \neq \lambda(u - u_0) \text{ for all } u \in \partial\Omega \text{ and all } \lambda > 1, \tag{4.28}$$

then φ has a fixed point.

Condition (4.28) is quite general and incorporates as special cases several concrete conditions existing in the literature of fixed point theory.

Corollary 4.3.15 *If X is a Banach space, $\Omega \subseteq X$ is bounded open with $0 \in \Omega$, then any of the following conditions implies the existence of a fixed point for φ:*

(a) *Ω is convex and $\varphi(\partial\Omega) \subseteq \overline{\Omega}$ (Rothe–Potter).*
(b) *$||\varphi(u) - u||^2 + ||u||^2 \geqslant ||\varphi(u)||^2$ for all $u \in \partial\Omega$ (Altman).*
(c) *If $X = H$ is a Hilbert space with inner product $(\cdot, \cdot)_H$ and $(\varphi(u), u)_H \leqslant ||u||^2$ for all $u \in \partial\Omega$ (Browder).*

Remark 4.3.16 All the above conditions are special cases of (4.28) with $u_0 = 0$.

Next we will present another generalization of the Schauder fixed point theorem in which the set C need not be convex (see Theorem 3.2.20) and int C may be empty (see Theorem 4.3.12). First we introduce the following fundamental topological notion.

Definition 4.3.17 A set X is an absolute retract (AR for short), if

(a) X is metrizable;
(b) for every metrizable space Y and $A \subseteq Y$ closed, for every continuous $\varphi : A \to X$, we can find a continuous extension on all Y (that is, there exists a continuous map $\hat{\varphi} : Y \to X$ such that $\hat{\varphi}\big|_A = \varphi$).

Remark 4.3.18 So, an AR is that metrizable space which can replace $[0, 1]$ in the classical Tietze extension theorem. By virtue of Proposition 2.1.9, every convex set in a normed space is an AR. Moreover, it is clear that if X is an AR, then every space homeomorphic to X is also an AR.

Proposition 4.3.19 *If X is an AR and D is a retract of X (see Definition 3.1.30), then D is an AR too.*

Proof Let Y be a metrizable space, $A \subseteq Y$ a closed set and $\varphi : A \to D$ a continuous map. Let $r : X \to D$ be a retraction for D. Since X is an AR, we can find a continuous extension $\hat{\varphi} : Y \to X$ of φ. Then $r \circ \hat{\varphi} : Y \to D$ is a continuous extension into D and this proves that D is an AR. $\square$

Proposition 4.3.20 *If X is an AR, then X is contractible (see Definition 4.3.7(c)).*

Proof Fix $u_0 \in X$ and let $h : \{0, 1\} \times X \to X$ be the continuous map defined by

$$h(t, u) = \begin{cases} u & \text{if } t = 0 \\ u_0 & \text{if } t = 1. \end{cases}$$

Since X is an AR, there exists a continuous extension $\hat{h} : [0, 1] \times X \to X$ of h. So

$$\begin{aligned} &\hat{h}(0, u) = u \text{ and } \hat{h}(1, u) = u_0 \\ \Rightarrow\ &X \text{ is contractible.} \end{aligned}$$

The proof is now complete. □

Now we can prove the generalized version of the Schauder fixed point theorem.

Theorem 4.3.21 *If X is an AR and $\varphi : X \to X$ is a continuous map such that $\varphi(X)$ is relatively compact, then φ has a fixed point.*

Proof The set $K = \overline{\varphi(X)}$ is compact, metric and so it is homeomorphic to a closed subset $\hat{K}$ of the Hilbert cube $I^\infty = [0, 1]^{\mathbb{N}}$. Let $h : K \to \hat{K}$ be the homeomorphism. We consider the maps

$$X \xrightarrow{\varphi} K \xrightarrow{h} \hat{K} \xrightarrow{h^{-1}} K \xrightarrow{i} X \ (i = \text{the identity (inclusion) map}).$$

The set $\hat{K} \subseteq I^\infty$ is closed and I^∞ is metrizable. Since X is an AR, the map $i \circ h^{-1} : \hat{K} \to X$ admits a continuous extension $\psi : I^\infty \to X$. Let $\tau = \hat{i} \circ h \circ \varphi$, where $\hat{i} : \hat{K} \to I^\infty$ is the identity (inclusion) map. Then $\tau : X \to I^\infty$ is continuous. The map $\tau \circ \psi : I^\infty \to I^\infty$ is continuous. The Hilbert cube I^∞ is a fixed point space (that is, every continuous map $f : I^\infty \to I^\infty$ admits a fixed point). This follows from the fact that an infinite product of nonempty fixed point spaces is also a fixed point space, if every finite product of those spaces is a fixed point space (see Dyer [152]). So, using Corollary 4.3.3, we infer that I^∞ is a fixed point space. Hence we can find $\hat{y} \in I^\infty$ such that $(\tau \circ \psi)(\hat{y}) = \hat{y}$. Then for $\hat{u} = \psi(\hat{y})$, we have

$$\begin{aligned} &\varphi(\hat{u}) = (\psi \circ \tau)(\hat{u}) = \psi(\tau(\psi(\hat{y}))) = \psi(\hat{y}) = \hat{u} \\ \Rightarrow\ &\varphi \text{ has a fixed point.} \end{aligned}$$

The proof is now complete. □

Remark 4.3.22 As we already mentioned every convex set of a normed space is an AR (see Proposition 2.1.9). So Theorem 4.3.21 is a generalization of the Schauder fixed point theorem. The product of two fixed point spaces need not be a fixed point space (even if they are compact, see Bredon [61]). If X is an infinite-dimensional

Banach space, then we know that $\partial B_1 = \{u \in X : ||u|| = 1\}$ is a retract of $\overline{B}_1 = \{u \in X : ||u|| \leqslant 1\}$. So, we can apply Proposition 4.3.19 to conclude that ∂B_1 is an AR.

Now we present a generalization of Theorem 4.3.14 for maps satisfying condition (4.28) (sometimes known in the literature as the Leray–Schauder boundary condition).

Theorem 4.3.23 *If X is a Banach space, $\Omega \subseteq X$ is open, $\varphi : \overline{\Omega} \to X$ is continuous and*

(i) for some $u_0 \in \Omega$ we have

$$\varphi(u) - u_0 \neq \lambda(u - u_0) \text{ for all } u \in \partial\Omega, \text{ all } \lambda > 1$$

(the Leray–Schauder boundary condition);

(ii) if $C \subseteq \overline{\Omega}$ is countable and $C \subseteq \overline{\text{conv}}\,[\{u_0\} \cap \varphi(C)]$, then $\overline{C}$ is compact,

then φ has a fixed point in Ω.

Proof By translating things if necessary, without any loss of generality, we may assume that $u_0 = 0 \in \Omega$ (indeed, if this is not the case, we replace Ω by $\Omega - u_0$ and $\varphi(\cdot)$ by $\varphi(\cdot + u_0) - u_0$ defined on $\Omega - u_0$).

Let $\Omega_0 = \{0\}$ and inductively define $\Omega_{n+1} = \text{conv}[\{0\} \cup \varphi(\Omega_n \cap \Omega)]$ for all $n \geqslant 0$. Evidently, $\Omega_n \subseteq \Omega_{n+1}$ and $\overline{\Omega}_n$ is compact for every $n \geqslant 0$. So, we can find a countable set C_n such that $\overline{C}_n = \overline{\Omega_n \cap \Omega}$. Let $V = \bigcup\limits_{n\geqslant 0} \Omega_n$. Since $\{\Omega_n\}_{n\geqslant 0}$ is increasing, we have

$$V = \text{conv}\,[\{0\} \cup \varphi(V \cap \Omega)]. \tag{4.29}$$

For $C = \bigcup\limits_{n\geqslant 0} C_n$, we obtain

$$\begin{aligned} C \subseteq \overline{\bigcup_{n\geqslant 0} C_n} = \overline{\bigcup_{n\geqslant 0}(\Omega_n \cap \Omega)} \subseteq \overline{V} &= \overline{\text{conv}}\ [\{0\} \cup \varphi(V \cap \Omega)] \ \text{ (see (4.29))} \\ &= \overline{\text{conv}} \left[\{0\} \cup \varphi\left(\bigcup_{n\geqslant 0}(\Omega_n \cap \Omega) \right) \right] \\ &= \overline{\text{conv}}\ [\{0\} \cup \varphi(C)] \end{aligned} \tag{4.30}$$

$\Rightarrow \overline{C}$ is compact see hypothesis *(ii)*)

$\Rightarrow D = \overline{V}$ is compact too (see (4.30)).

If $\varphi\big|_{\partial\Omega}$ has a fixed point, then we are done. Otherwise, $\lambda\varphi(u) \neq u$ for all $(\lambda, u) \in [0, 1] \times \partial\Omega$ and so $F = \bigcup\limits_{\lambda\in[0,1]} \text{Fix}(\lambda\varphi)$ is compact and satisfies $F \cap \partial\Omega = \emptyset$ (here

$\mathrm{Fix}(\lambda\varphi) = \{u \in \Omega : \lambda\varphi(u) = u\}$). The sets $F \cap D$ and $\partial_D(\Omega \cap D)$ (=the boundary of $\Omega \cap D$ in D) are closed in D and disjoint since $D \subseteq \Omega$ and $\partial_D(\Omega \cap D) \subseteq \partial\Omega \cap D$. So, by Urysohn's theorem we can find a continuous function $\vartheta : D \to [0, 1]$ such that

$$\vartheta\Big|_{\partial_D(\Omega\cap D)} = 0 \text{ and } \vartheta\Big|_{F\cap D} = 1. \tag{4.31}$$

Let $\psi : D \to X$ be the map defined by

$$\psi(u) = \begin{cases} \vartheta(u)\varphi(u) & \text{if } u \in \overline{\Omega \cap D} \\ 0 & \text{otherwise.} \end{cases} \tag{4.32}$$

Since $\partial_D(\overline{\Omega \cap D}) = \partial_D(\Omega \cap D)$ and $\vartheta\Big|_{\partial_D(\Omega\cap D)} = 0$ (see (4.31)), from (4.32) we infer that ψ is continuous. In addition we have

$$\psi(D) \subseteq \mathrm{conv}\,[\{0\} \cup \varphi(\overline{\Omega} \cap \overline{V})] \subseteq \overline{\mathrm{conv}}\,[\{0\} \cup \varphi(\Omega \cap V)] = \overline{V} = D \text{ (see (4.29))}.$$

Invoking the Schauder fixed point theorem (see Theorem 3.2.20), we can find $\hat{u} \in D$ such that $\psi(\hat{u}) = \hat{u}$. Hence $\hat{u} \in \overline{\Omega \cap D}$ and $\vartheta(\hat{u})\varphi(\hat{u}) = \hat{u}$ (see (4.32)). So, $\hat{u} \in \mathrm{Fix}\,(\vartheta(\hat{u})\varphi) \subseteq F$ and so $\vartheta(\hat{u}) = 1$ (see (4.31)). Therefore finally we have $\varphi(\hat{u}) = \hat{u}$. The proof is now complete. $\square$

For self-maps, Theorem 4.3.23 takes the following form.

Theorem 4.3.24 *If X is a Banach space, $C \subseteq X$ is closed and convex, $\varphi : C \to C$ is continuous and for some $u_0 \in C$ the following condition holds*

$$\text{“}\textit{if } E \subseteq C \textit{ is countable and } \overline{E} = \overline{\mathrm{conv}}\,[\{u_0\} \cup \varphi(E)], \\ \textit{then } \overline{E} \textit{ is compact”},$$

then φ has a fixed point.

Proof As in the proof of Theorem 4.3.23, we obtain $V = \mathrm{conv}[\{u_0\} \cup \varphi(V)] \subseteq C$ and $D = \overline{V}$ is compact. Then $\varphi : D \to D$ and so by the Schauder fixed point theorem, we have a fixed point. $\square$

We will conclude this section with one more fixed point theorem for nonself-maps. It concerns the so-called weakly inward maps. So, first we introduce this class of maps.

Definition 4.3.25 (a) X is a Banach space, $C \subseteq X$ is nonempty and $\{u_n\}_{n\geqslant 1}$ a bounded sequence in X. The asymptotic center of $\{u_n\}_{n\geqslant 1}$ relative to C is defined by

$$A(C, \{u_n\}) = \left\{ y \in C : \limsup_{n\to\infty} ||y - u_n|| = \inf_{v\in C} \left[\limsup_{n\to\infty} ||v - u_n|| \right] \right\};$$

(b) X is a Banach space, $C \subseteq X$ is nonempty, closed, convex and for $u \in C$ we define

$$I_C(u) = \{(1-\lambda)u + \lambda y : \lambda \geqslant 0, y \in C\},$$

the inward set of $u \in C$ with respect to C. A map $\varphi : C \to X$ is said to be weakly inward if $\varphi(u) \in \overline{I_C(u)}$ for all $u \in C$.

Remark 4.3.26 If we set $r(y) = \limsup_{n\to\infty} ||y - u_n||$, then $r(\cdot)$ is convex and nonexpansive. Also, if the space X is uniformly convex and $C \subseteq X$ is nonempty, bounded, closed and convex, then the set $A(C, \{u_n\})$ is a singleton. Finally, it is easy to check that $\varphi : C \to X$ is weakly inward if and only if $\lim_{\lambda\to 0^+} \frac{1}{\lambda} d(u + \lambda(\varphi(u) - u), C) = 0$ for all $u \in C$. This condition is closely connected with the existence of solutions of initial-value problems on closed convex sets. So, consider the initial value problem $u'(t) = \varphi(u(t)) - u(t)$, $t \in T$ and $u(0) = u_0 \in C$. If $u(\cdot)$ is a local solution of this problem, then

$$\begin{aligned} & u_0 + \lambda(\varphi(u_0) - u_0) + o(\delta) = u(t) \in C \text{ as } t \to 0^+ \\ \Rightarrow\ & d(u_0 + \lambda(\varphi(u_0) - u_0), C) = d(u(t) + o(t), C) = o(t) \text{ as } t \to 0^+, \end{aligned}$$

which is the inwardness condition for u_0. In fact this boundary condition is also sufficient if, for example, φ is Lipschitz (see Lakshmikantham and Leela [257] and Martin [291]). Weak inwardness can also be characterized in terms of linear functionals. Namely, $\varphi : C \to X$ is weakly inward if and only if

$$\begin{gathered} \text{“}u \in \partial C,\ u^* \in X^* \text{ and } \langle u^*, u\rangle = \sigma_C(u^*) = \sup\{\langle u^*, v\rangle : v \in C\} \\ \text{imply} \\ \langle u^*, \varphi(u) - u\rangle \leqslant 0\text{”}. \end{gathered}$$

Proposition 4.3.27 *If X is a uniformly convex Banach space, $C \subseteq X$ is nonempty, closed and convex, $\{u_n\}_{n\geqslant 1} \subseteq C$ is bounded and u is its asymptotic center with respect to C (see Definition 4.3.25 and Remark 4.3.26), then u is also its asymptotic center with respect to $\overline{I_C(u)}$.*

Proof Let v be the asymptotic center of $\{u_n\}_{n\geqslant 1}$ with respect to $\overline{I_C(u)}$. Assume that $v \neq u$. Since $C \subseteq \overline{I_C(u)}$, we have $v \in \overline{I_C(u)} - C$ and $r(v) < r(u)$. Since $r(\cdot)$ is continuous, we can find $z \in I_C(u) - C$ such that $r(z) < r(u)$. Therefore $z = (1 - \lambda)u + \lambda x$ for some $x \in C$ and $\lambda > 1$. The convexity of $r(\cdot)$ implies

$$\begin{aligned} r(x) = r\left[\frac{1}{\lambda}z + \left(1 - \frac{1}{\lambda}\right)u\right] &\leqslant \frac{1}{\lambda}r(z) + \left(1 - \frac{1}{\lambda}\right)r(u) \\ &\leqslant \frac{1}{\lambda}r(u) + \left(1 - \frac{1}{\lambda}\right)r(u) = r(u), \end{aligned}$$

contradicting Definition 4.3.25(a). This proves that $v = u$. The proof is now complete. $\square$

We will need the following fixed point theorem of Caristi [101].

Proposition 4.3.28 *If X is a Banach space, $C \subseteq X$ is nonempty, closed and convex and $\varphi : C \to X$ is a contraction and weakly inward, then φ has a unique fixed point.*

Using this proposition, we can prove the following fixed point theorem for weakly inward maps.

Theorem 4.3.29 *If X is a uniformly convex Banach space, $C \subseteq X$ is nonempty, bounded closed and convex and $\varphi : C \to X$ is nonexpansive and weakly inward, then φ has a fixed point.*

Proof Let $u_0 \in C$ and for every $n \geqslant 1$ let

$$\varphi_n(u) = (1 - \lambda_n)u_0 + \lambda_n \varphi(u) \text{ for all } u \in C \text{ with } \lambda_n \in (0, 1).$$

Assume that $\lambda_n \to 1^-$. Evidently, φ_n is a contraction with constant $l(\varphi_n) = \lambda_n < 1$. By Theorem 4.2.3, φ_n has a unique fixed point $u_n \in C$. We have

$$\begin{aligned} ||u_n - \varphi(u_n)|| &= \left\| u_n - \frac{1}{\lambda_n} u_n - \left(\frac{1}{\lambda_n} - 1 \right) u_0 \right\| \\ &= \left(\frac{1}{\lambda_n} - 1 \right) ||u_n - u_0|| \to 0 \text{ as } n \to \infty, \end{aligned}$$

since C is bounded and $\lambda_n \to 1^-$. Let u be the asymptotic center of $\{u_n\}_{n \geqslant 1}$ with respect to C. Then

$$\begin{aligned} r(\varphi(u)) &= \limsup_{n \to \infty} ||u_n - \varphi(u)|| \\ &\leqslant \limsup_{n \to \infty} ||\varphi(u_n) - \varphi(u)|| \\ &\leqslant \limsup_{n \to \infty} ||u_n - u|| \text{ (since } \varphi \text{ is nonexpansive)} \\ &= r(u). \end{aligned} \tag{4.33}$$

Since φ is weakly inward, $\varphi(u) \in \overline{I_C(u)}$. Also from Proposition 4.3.27 we know that u is also the asymptotic center of $\{u_n\}_{n \geqslant 1}$ with respect to $\overline{I_C(u)}$. So, from (4.33), we conclude that $\varphi(u) = u$. $\square$

In fact, a similar result, due to Deimling [142, p. 211], is also true for γ-condensing maps.

Theorem 4.3.30 *If X is a Banach space, $C \subseteq X$ is nonempty, bounded, closed and convex and $\varphi : C \to X$ is γ-condensing and weakly inward, then φ has a fixed point.*

4.4 Order Fixed Points and the Fixed Point Index

In this section we discuss fixed point theorems resulting from the order structure of the ambient Banach space.

So, let X be an ordered Banach space (OBS for short), that is, a Banach space together with a cone K inducing a partial order on X (by setting $u \leqslant v$ if and only if $v - u \in K$, see Definition 4.1.1). We call K the order cone of X.

Definition 4.4.1 Let X be an OBS with order cone K.

(a) For every $u, v \in X$, the set

$$[u, v] = \{y \in X : u \leqslant y \leqslant v\}$$

is the order interval determined by u, v and it is nonempty if and only if $u \leqslant v$ (that is, $v - u \in K$). A set $C \subseteq X$ is said to be order bounded if it is contained in some order interval. Also, we say that C is order convex if $u, v \in C$ imply $[u, v] \subseteq C$. By $[C]$ we denote the order convex hull of C, which is the smallest order convex subset of X which contains C, that is, $[C] = \bigcup\{[u, v] : u, v \in C\}$.

(b) A map $\varphi : X \to X$ is said to be

- increasing if $u \leqslant v \Rightarrow \varphi(u) \leqslant \varphi(v)$;
- strictly increasing if $v - u \in K\backslash\{0\} \Rightarrow \varphi(v) - \varphi(u) \in K\backslash\{0\}$;
- strongly increasing (provided $\operatorname{int} K \neq \emptyset$) if $v - u \in K\backslash\{0\} \Rightarrow \varphi(v) - \varphi(u) \in \operatorname{int} K$.

Also, we say that $\varphi : X \to X$ is decreasing (resp. strictly, strongly) decreasing if $-\varphi$ is increasing (resp. strictly, strongly) increasing.

(c) A linear operator $A : X \to X$ is said to be

- positive if $A(K) \subseteq K$;
- strictly positive if $A(K\backslash\{0\}) \subseteq K\backslash\{0\}$;
- strongly positive (provided $\operatorname{int} K \neq \emptyset$) if $\varphi(K\backslash\{0\}) \subseteq \operatorname{int} K$.

Remark 4.4.2 Evidently, a positive linear operator $A : X \to X$ is increasing.

Proposition 4.4.3 *If X is an OBS with order cone K, $u_0, v_0 \in K$ with $v_0 - u_0 \in K\backslash\{0\}$, $\varphi : [u_0, v_0] \to X$ is an increasing operator such that*

$$u_0 \leqslant \varphi(u_0) \text{ and } \varphi(v_0) \leqslant v_0 \tag{4.34}$$

and one of the following conditions holds:

(i) φ is γ-condensing and K is normal; or

(ii) φ is demicontinuous (that is, $u_n \to u$ in $X \Rightarrow \varphi(u_n) \xrightarrow{w} \varphi(u)$ in X) and K is regular,

then φ has a maximal fixed point $\hat{v}$ and a minimal fixed point $\hat{u}$ in $[u_0, v_0]$.

Proof Let $u_n = \varphi(u_{n-1})$ and $v_n = \varphi(v_{n-1})$ for all $n \geqslant 1$. Since φ is increasing and (4.34) holds, we have

$$u_0 \leqslant u_1 \leqslant \cdots \leqslant u_n \leqslant \cdots \leqslant v_n \leqslant \cdots \leqslant v_1 \leqslant v_0. \tag{4.35}$$

First assume that hypothesis *(i)* is in effect. Let $C = \{u_n\}_{n \geqslant 0}$. The set C is bounded and $C = \varphi(C) \cup \{u_0\}$. So, $\gamma(C) = \gamma(\varphi(C))$. But according to hypothesis *(i)* φ is γ-condensing. It follows that $\gamma(C) = 0$ and so $\overline{C}$ is compact in X. Hence there exists a subsequence $\{u_{n_k}\}_{k \geqslant 1}$ of $\{u_n\}_{n \geqslant 1}$ such that $u_{n_k} \to \hat{u}$ in X. We have

$$\begin{aligned}
& u_n \leqslant \hat{u} \leqslant v_n \text{ for all } n \geqslant 1 \\
\Rightarrow\ & 0 \leqslant \hat{u} - u_m \leqslant \hat{u} - u_{n_k} \text{ for all } m \geqslant n_k \\
\Rightarrow\ & ||\hat{u} - u_m|| \leqslant \xi ||\hat{u} - u_{n_k}|| \text{ for all } m \geqslant n_k \\
& \qquad\qquad\qquad \text{(due to the normality of } K, \text{ see Proposition 4.1.6)} \\
\Rightarrow\ & u_m \to \hat{u} \text{ in } X.
\end{aligned} \tag{4.36}$$

Since $u_n = \varphi(u_{n-1})$ for all $n \geqslant 1$, passing to the limit as $n \to \infty$ and using (4.36) and the continuity of φ, we obtain

$$\hat{u} = \varphi(\hat{u}).$$

Reasoning in a similar fashion, we show that

$$\begin{aligned}
& v_n \to \hat{v} \text{ in } X \\
\Rightarrow\ & \hat{v} = \varphi(\hat{v}).
\end{aligned}$$

Next, assume that hypothesis *(ii)* holds. The regularity of K and (4.35) imply

$$u_n \to \hat{u} \text{ in } X \text{ (see Definition 4.1.2 (e))}. \tag{4.37}$$

The demicontinuity of φ implies

$$\varphi(u_n) \xrightarrow{w} \varphi(\hat{u}) \text{ in } X. \tag{4.38}$$

Since $u_n = \varphi(u_{n-1})$ for all $n \geqslant 1$, from (4.37) and (4.38) we obtain

$$\hat{u} = \varphi(\hat{u}).$$

Similarly we obtain $u_n \to \hat{v}$ in X and $\hat{u} = \varphi(\hat{v})$.

Finally we show that $\hat{u}, \hat{v} \in [u_0, v_0]$ are the extremal fixed points of φ in the order interval $[u_0, v_0]$. To this end, let $\tilde{u} \in [u_0, v_0]$ and assume that $\tilde{u} = \varphi(\tilde{u})$. Since φ is increasing, we have

$$\varphi(u_0) \leqslant \varphi(\tilde{u}) \leqslant \varphi(v_0)$$
$$\Rightarrow u_1 \leqslant \varphi(\tilde{u}) = \tilde{u} \leqslant v_1.$$

In this fashion we obtain inductively

$$u_n \leqslant \tilde{u} \leqslant v_n \text{ for all } n \geqslant 1$$
$$\Rightarrow \hat{u} \leqslant \tilde{u} \leqslant \hat{v}.$$

The proof is now complete. □

Remark 4.4.4 If φ has only one fixed point in $[u_0, v_0]$, then starting from any $y_0 \in [u_0, v_0]$ the sequence of successive iterates $u_n = \varphi(u_{n-1})$ for all $n \geqslant 1$ converges to this unique fixed point $\hat{u} \in [u_0, v_0]$.

In the next three theorems, we do not require φ to be continuous.

Theorem 4.4.5 *If X is an OBS, $u_0, v_0 \in X$ with $v_0 - u_0 \in K \backslash \{0\}$, $\varphi : [u_0, v_0] \to X$ is an increasing map which satisfies (4.34) and the set $\varphi([u_0, v_0]) \subseteq X$ is relatively compact, then φ has a fixed point.*

Proof Let $E = \{u \in \varphi([u_0, v_0]) : u \leqslant \varphi(u)\}$. Note that $E \neq \emptyset$ since $\varphi(u_0) \in E$ (recall that φ is increasing). Suppose that $M \subseteq E$ is a totally ordered subset. We have

$$M \subseteq E \subseteq \varphi([u_0, v_0])$$
$$\Rightarrow M \text{ is relatively compact in } X.$$

So, we can find $\{u_n\}_{n\geqslant 1} \subseteq M$ dense in M. Let $y_n = \sup\{u_k\}_{k=1}^n \in M$ (recall that M is totally ordered and so y_n equals one of the $u_k's$, $k \in \{1, \cdots, n\}$). So, we can find a subsequence $\{y_{n_k}\}_{k\geqslant 1}$ of $\{y_n\}_{n\geqslant 1}$ such that

$$y_{n_k} \to \hat{y} \text{ in } X.$$

Evidently, $\{y_n\}_{n\geqslant 1}$ is increasing and so

$$u_n \leqslant y_n \leqslant \hat{y} \text{ for all } n \geqslant 1. \tag{4.39}$$

We have

$$\hat{y} \in \overline{M} \subseteq \overline{E} \subseteq \overline{\varphi([u_0, v_0])} \subseteq [u_0, v_0] \text{ (see (4.34))}$$
$$\Rightarrow u \leqslant \hat{y} \text{ for all } u \in M$$
$$\Rightarrow u \leqslant \varphi(u) \leqslant \varphi(\hat{y}) \text{ for all } u \in M.$$

Hence $\varphi(\hat{u})$ is an upper bound of M and $y_n \leqslant \varphi(\hat{y})$ for all $n \geqslant 1$. So, $\hat{y} \leqslant \varphi(\hat{y})$, therefore $\varphi(\hat{y}) \leqslant \varphi(\varphi(\hat{y}))$ and this implies that $\varphi(\hat{y}) \in E$. We have proved that every totally ordered subset of E has an upper bound in E. Applying Zorn's lemma, we infer that E has a maximal element $\hat{v}$ (for the partial order on E induced by (X, K)). From the definition of E we have $\hat{v} \leqslant \varphi(\hat{v})$ and so the maximality of $\hat{v}$ implies $\hat{v} = \varphi(\hat{v})$. □

We can produce extremal fixed points provided we strengthen the condition on the order cone K.

Theorem 4.4.6 *If X is an OBS with a minihedral order cone $K \subseteq X$, $u_0, v_0 \in X$ such that $v_0 - u_0 \in K \setminus \{0\}$, $\varphi : [u_0, v_0] \to X$ is an increasing map which satisfies (4.34) and $\varphi([u_0, v_0]) \subseteq X$ is relatively compact, then φ has a maximal fixed point $\hat{v} \in [u_0, v_0]$ and a minimal fixed point $\hat{u} \in [u_0, v_0]$.*

Proof As in the proof of Proposition 4.2.5, we set $E = \{u \in \varphi([u_0, v_0]) : u \leqslant \varphi(u)\}$. In that proof, using Zorn's lemma, we obtained a maximal element $\hat{v} \in E$ such that $\varphi(\hat{v}) = \hat{v}$. We show that this is the maximal fixed point of φ in $[u_0, v_0]$. To this end, let $\tilde{u} \in [u_0, v_0]$ be a fixed point of φ. Since K is minihedral, $y = \sup\{\tilde{u}, \tilde{v}\} \in X$. We have

$$
\begin{aligned}
& \tilde{u} \leqslant y \text{ and } \hat{v} \leqslant y \\
\Rightarrow\ & \varphi(\tilde{u}) = \tilde{u} \leqslant \varphi(y) \text{ and } \varphi(\hat{v}) = \hat{v} \leqslant \varphi(y) \text{ (since } \varphi \text{ is increasing)} \\
\Rightarrow\ & y \leqslant \varphi(y) \text{ and so } \varphi(y) \leqslant \varphi(\varphi(y)).
\end{aligned}
$$

This means that $\varphi(y) \in E$ and then the maximality of $\hat{v}$ implies that $\hat{v} = \varphi(y)$. Hence $\tilde{u} \leqslant \hat{v}$.

Similarly, considering the set $G = \{u \in \varphi([u_0, v_0]) : \varphi(u) \leqslant u\}$ and reasoning as in the proof of Theorem 4.4.5, via Zorn's lemma, we produce a minimal element $\hat{u} \in G$ which, as above, we show is the minimal fixed point of φ on $[u_0, v_0]$. □

We can drop the topological condition on $\varphi([u_0, v_0])$ at the expense of strengthening the hypothesis on the order cone K.

Theorem 4.4.7 *If X is an OBS with a strongly minihedral order cone $K \subseteq X$, $u_0, v_0 \in X$ such that $v_0 - u_0 \in K \setminus \{0\}$ and $\varphi : [u_0, v_0] \to X$ is an increasing map which satisfies (4.34), then φ has a maximal fixed point $\hat{v} \in [u_0, v_0]$ and a minimal fixed point $\hat{u} \in [u_0, v_0]$.*

Proof As before, let $E = \{u \in [u_0, v_0] : u \leqslant \varphi(u)\}$. Since by hypothesis K is strongly minihedral, $\sup E = \hat{v} \in [u_0, v_0]$ exists (see Definition 4.1.2(h)). We claim that $\hat{v}$ is the maximal fixed point of φ in $[u_0, v_0]$. Since φ is increasing, we have

$$
\begin{aligned}
& u \leqslant \varphi(\hat{u}) \text{ for all } u \in E \\
\Rightarrow\ & \hat{v} \leqslant \varphi(\hat{u}) \\
\Rightarrow\ & \varphi(\hat{v}) \leqslant \varphi(\varphi(\hat{v})), \text{ hence } \varphi(\hat{v}) \in E \\
\Rightarrow\ & \varphi(\hat{v}) \leqslant \hat{v} = \sup E, \text{ hence } \varphi(\hat{v}) = \hat{v}.
\end{aligned}
$$

If $\tilde{u} \in [u_0, v_0]$ is a fixed point of φ, then $\tilde{u} \in E$ and so $\tilde{u} \leqslant \hat{v}$.

Similarly, if we set $G = \{u \in [u_0, v_0] : \varphi(u) \leqslant u\}$, then as above via Zorn's lemma we show that G has a minimal element $\hat{u} \in G$ and $\varphi(\hat{u}) = \hat{u}$. Moreover, for every fixed point $\tilde{u} \in [u_0, v_0]$ of φ, we have $\hat{u} \leqslant \tilde{u}$. $\square$

So far we have imposed a monotonicity condition (with respect to K) on φ. Next we will drop the monotonicity hypothesis and instead we are going to use degree theory in order to prove fixed point results. To do this we will introduce the notion of a fixed point index for the map.

Let X be an OBS with order cone K. Recall that $C \subseteq X$ is a retract of K if there exists a continuous map $r : X \to C$ such that $r\big|_C = i\big|_C$ (see Definition 3.1.30). We know that every closed and convex set in X is a retract (this is a consequence of Proposition 2.1.9). A retract is always closed and if $\partial B_1 = \{u \in X : ||u|| = 1\}$, then we know that it is a retract of X provided X is infinite-dimensional (see Remark 3.1.33). This is no longer true if X is finite-dimensional (see Proposition 3.1.32).

Suppose that C is a retract of X and $\Omega \subseteq C$ a bounded relatively open set. Assume that $\varphi : \overline{\Omega} \to C$ is compact and $0 \notin (i - \varphi)(\partial\Omega)$. Let $r : X \to C$ be a retraction and let $R > 0$ be such that $\Omega \subseteq B_R = \{u \in X : ||u|| < R\}$. Note that $B_R \cap r^{-1}(\Omega)$ is a bounded open set in X. For the triple (φ, Ω, C) we can make the following definition.

Definition 4.4.8 $i_F(\varphi, \Omega, C) = d_{LS}(i - \varphi \circ r, B_R \cap r^{-1}(\Omega), 0)$. This quantity is called the fixed point index of φ on Ω with respect to C.

Remark 4.4.9 Note that

$$\overline{B_R \cap r^{-1}(\Omega)} \subseteq \overline{r^{-1}(\Omega)} \subseteq r^{-1}(\overline{\Omega})$$

and $u \in r^{-1}(\overline{\Omega})$, $\varphi(r(u)) = u$ imply $u \in \Omega$ and $\varphi(u) = u$. In order for Definition 4.4.8 to be meaningful, we need to show that it is independent of the choice of R and r. So, let $R_1 > R$. Then

$$\Omega \subseteq B_R \cap r^{-1}(\Omega) \subseteq B_{R_1} \cap r^{-1}(\Omega).$$

From the initial observations, we know that $\varphi \circ r$ has no fixed points in the set $\overline{B_{R_1} \cap r^{-1}(\Omega)} \backslash (B_R \cap r^{-1}(\Omega))$. So, the excision property of the Leray–Schauder degree (see Theorem 3.2.15 (g)), we have

$$\begin{aligned} &d_{LS}(i - \varphi \circ r, B_{R_1} \cap r^{-1}(\Omega), 0) = d_{LS}(i - \varphi \circ r, B_R \cap r^{-1}(\Omega), 0) \\ &\Rightarrow i(\varphi, \Omega, C) \text{ is independent of } R > 0 \text{ (see Definition4.4.8).} \end{aligned}$$

Next, let $r_1 : X \to C$ be another retraction of X onto C. Let

$$\hat{\Omega} = B_R \cap r^{-1}(\Omega) \cap r_1^{-1}(\Omega).$$

Then $\hat{\Omega} \subseteq X$ is bounded open and $\Omega \subseteq \hat{\Omega}$. As above, we see that $\varphi \circ r$ has no fixed points on $\overline{B_R \cap r^{-1}(\Omega)} \backslash \hat{\Omega}$ and $\varphi \circ r_1$ has no fixed points on $\overline{B_R \cap r_1^{-}(\Omega)} \backslash \hat{\Omega}$. So, invoking the excision property of the Leray–Schauder degree, we have

$$d_{LS}(i - \varphi \circ r, B_R \cap r^{-1}(\Omega), 0) = d_{LS}(i - \varphi \circ r, \hat{\Omega}, 0), \tag{4.40}$$
$$d_{LS}(i - \varphi \circ r_1, B_R \cap r_1^{-1}(\Omega), 0) = d_{LS}(i - \varphi \circ r_1, \hat{\Omega}, 0). \tag{4.41}$$

We consider the homotopy $h(t, u)$ defined by

$$h(t, u) = t(\varphi \circ r)(u) + (1 - t)(\varphi \circ r_1)(u) \text{ for all } (t, u) \in [0, 1] \times \overline{\hat{\Omega}}.$$

This is a compact homotopy. We claim that

$$u - h(t, u) \neq 0 \text{ for all } t \in [0, 1], \text{ all } u \in \partial\hat{\Omega}. \tag{4.42}$$

Arguing by contradiction, suppose that we can find $(t_0, u_0) \in [0, 1] \times \partial\hat{\Omega}$ such that

$$\begin{aligned}
& u_0 - h(t_0, u_0) = 0 \\
\Rightarrow\ & u_0 = t_0\varphi(r(u_0)) + (1 - t)\varphi(r_1(u_0)) \\
\Rightarrow\ & u_0 \in C \text{ and so } r(u_0) = u_0, r_1(u_0) = u_0, u_0 = \varphi(u_0) \\
\Rightarrow\ & u_0 \in \Omega \subseteq \hat{\Omega}, \text{ a contradiction since } u_0 \in \partial\hat{\Omega}.
\end{aligned}$$

Therefore (4.42) holds. Then by the homotopy invariance of the Leray–Schauder degree (see Theorem 3.2.15 (c)), we have

$$\begin{aligned}
& d_{LS}(i - \varphi \circ r, \hat{\Omega}, 0) = d_{LS}(i - \varphi \circ r_1, \hat{\Omega}, 0) \\
\Rightarrow\ & d_{LS}(i - \varphi \circ r, B_R \cap r^{-1}(\Omega), 0) = d_{LS}(i - \varphi \circ r_1, B_R \cap r_1^{-1}(\Omega), 0) \\
& \qquad \text{(see (4.40), (4.41))} \\
\Rightarrow\ & i_F(\varphi, \Omega, C) \text{ is independent of the retraction } r \text{ (see Definition 4.4.8).}
\end{aligned}$$

So, Definition 4.4.8 makes perfect sense.

Using Theorem 3.2.15 (the main properties of the Leray–Schauder degree), we can state the following theorem summarizing the main properties of the fixed point index.

Theorem 4.4.10 *If* $\mathscr{F} = \{(\varphi, \Omega, C) : C \subseteq X$ *is a retract,* $\Omega \subseteq X$ *is bounded, and relatively open,* $\varphi : \overline{\Omega} \to C$ *is compact and* $0 \notin (i - \varphi)(\partial\Omega)\}$, *then there exists a map* $i_F : \mathscr{F} \to \mathbb{Z}$, *known as the fixed point index such that*

(a) Normalization: $i_F(\varphi, \Omega, C) = 1$ *if* $\varphi(u) = u_0 \in \Omega$ *for all* $u \in \overline{\Omega}$.
(b) Domain Additivity: $i_F(\varphi, \Omega, C) = i_F(\varphi, \Omega_1, C) + i_F(\varphi, \Omega_2, C)$ *with* Ω_1, Ω_2 *disjoint open subsets of* Ω *such that* φ *has no fixed points on* $\overline{\Omega} \backslash (\Omega_1 \cup \Omega_2)$.

(c) *Homotopy Invariance:* $i_F(h_t, \Omega, C)$ *is independent of* $t \in [0, 1]$*, when* $\{h_t\}_{t\in[0,1]}$ *is a homotopy of compact maps and*

$$u \neq h_t(u) \text{ for all } t \in [0, 1] \text{ and all } u \in \partial\Omega.$$

(d) *Solution Property:* $i_F(\varphi, \Omega, C) \neq 0$ *implies that* φ *has at least one fixed point in* Ω*.*
(e) *Excision Property:* $i_F(\varphi, \Omega, C) = i_F(\varphi, \Omega_1, C)$ *for every open set* $\Omega_1 \subseteq \Omega$ *such that* φ *has no fixed points in* $\overline{\Omega}\backslash\Omega_1$*.*
(f) *Stability with respect to* C*:* $i_F(\varphi, \Omega, C) = i_F(\varphi, \Omega, C_1)$ *if* C_1 *is a retract of* C *and* $\varphi(\overline{\Omega}) \subseteq C_1$*.*

Remark 4.4.11 In fact the homotopy invariance property has the following more general formulation:

"Let X be a Banach space, $C \subseteq X$ a nonempty retract, $\Omega \subseteq [0, 1] \times C$ a relatively open set, and $h : \overline{\Omega} \to C$ a compact map such that $h(t, u) \neq u$ for all $(t, u) \in \partial\Omega$, then $i_F(h(t, \cdot), \Omega_t, C)$ is independent of $t \in [0, 1]$ (here we set $\Omega_t = \{u \in C : (t, u) \in \Omega\}$)."

Next we do some computations of fixed point indices which we will use later to prove the existence of fixed points.

So, let X be an OBS with order cone $K \subseteq X$ and $\Omega \subseteq X$ a bounded open set. Then $\Omega \cap K$ is bounded and relatively open in K (which is a retract, being a closed and convex subset of the Banach space X) and $\partial(\Omega \cap K) = \partial\Omega \cap K$, $\overline{\Omega \cap K} = \overline{\Omega} \cap K$.

Proposition 4.4.12 *If* $0 \in \Omega$*,* $\varphi : \overline{\Omega} \cap K \to K$ *is compact and* $\varphi(u) \neq \lambda u$ *for all* $u \in \partial\Omega \cap K$ *and* $\lambda \geqslant 1$*, then* $i_F(\varphi, \Omega \cap K, K) = 1$*.*

Proof We consider the compact homotopy $h_t(u) = t\varphi(u)$. Then $h_t(u) \neq u$ for all $(t, u) \in [0, 1] \times (\partial\Omega \cap K)$. We have

$$\begin{aligned} &i_F(\varphi, \Omega \cap K, K) = i_F(i, \Omega \cap K, K) \text{ (by the homotopy invariance property)}\\ \Rightarrow\ &i_F(\varphi, \Omega \cap K, K) = 1 \text{ (by the normalization property).}\end{aligned}$$

The proof is now complete. □

Proposition 4.4.13 *If* $\varphi : \overline{\Omega} \cap K \to K$ *and* $g : \partial\Omega \cap K \to K$ *are both compact maps and*

(i) $\inf\{||g(u)|| : u \in \partial\Omega \cap K\} > 0$*;*
(ii) $u - \varphi(u) \neq \lambda g(u)$ *for all* $u \in \partial\Omega \cap K$ *and all* $\lambda \geqslant 0$*,*

then $i_F(\varphi, \Omega \cap K, K) = 0$*.*

Proof By virtue of Proposition 2.1.9 (Dugundji's extension theorem), there exists a compact map $\hat{g} : \overline{\Omega} \cap K \to K$ such that $\hat{g}\Big|_{\partial\Omega\cap K} = g$ and $\hat{g}(\overline{\Omega} \cap K) \subseteq \overline{\text{conv}}\, g(\partial\Omega \cap K)$. Let $D = g(\partial\Omega \cap k)$. We claim that

$$\inf\{||v|| : v \in \overline{\text{conv}}\, D\} = \vartheta > 0. \tag{4.43}$$

Let $Y = \overline{\text{span}}\, D$. Since g is compact, the set $D \subseteq X$ is relatively compact and so Y is a separable Banach subspace of X. Let $\hat{K} = K \cap Y$. Then $\hat{K}$ is a cone in Y and $\overline{\text{conv}}\, D \subseteq \hat{K}$. According to Proposition 4.1.18(e) we can find $\hat{y}^* \in \hat{K}^*$ such that $\langle \hat{y}^*, y\rangle > 0$ for all $y \in \hat{K}\backslash\{0\}$. We claim that

$$\inf\left\{\langle \hat{y}^*, u\rangle : u \in D\right\} = \hat{m} > 0. \tag{4.44}$$

Suppose that $\hat{m} = 0$, then we can find $\{v_n\}_{n\geqslant 1} \subseteq D$ such that $\langle \hat{y}^*, v_n\rangle \to 0$. Recall that D is relatively compact. So, by passing to a suitable subsequence if necessary, we may assume that $v_n \to \hat{v}$ in X, $\hat{v} \in \hat{K}$. Then $\langle \hat{y}^*, v_n\rangle \to \langle \hat{y}^*, \hat{v}\rangle = 0$ and so $\hat{v} = 0$ (see Proposition 4.1.18(e)). Hence $v_n \to 0$ in X, which contradicts hypothesis *(i)*. This proves that (4.44) holds. Let $y \in \text{conv}\, D$. Then $y = \sum_{\text{k}=1}^{n} \lambda_k v_k$, $v_k \in D$, $\lambda_k \geqslant 0$, $\sum_{\text{k}=1}^{n} \lambda_k = 1$, $n \geqslant 1$. We have

$$\begin{aligned} &\langle \hat{y}^*, y\rangle = \sum_{\text{k}=1}^{n} \lambda_k \langle \hat{y}^*, v_k\rangle \geqslant \sum_{\text{k}=1}^{n} \lambda_k \hat{m} = \hat{m} \text{ (see (4.44))} \\ \Rightarrow &\langle \hat{y}^*, v\rangle \geqslant \hat{m} \text{ for all } v \in \overline{\text{conv}}\, D. \end{aligned} \tag{4.45}$$

The set $\overline{\text{conv}}\, D$ is compact, since D is. So, we can find $\hat{v} \in \overline{\text{conv}}\, D$ such that

$$\inf\{||u|| : v \in \overline{\text{conv}}\, D\} = ||\hat{v}||. \tag{4.46}$$

From (4.45) we see that $\hat{v} \neq 0$ and so from (4.46) we see that (4.43) holds.

Suppose that $i_F(\varphi, \Omega \cap K, K) \neq 0$. By virtue of hypothesis *(ii)* and the homotopy invariance property of the fixed point index, we have

$$i_F(\varphi + t\hat{g}, \Omega \cap K, K) = i_F(\varphi, \Omega \cap K, K) \neq 0 \text{ for all } t > 0. \tag{4.47}$$

Let $\eta_1 = \sup\{||u|| : u \in \overline{\Omega} \cap K\}$ and $\eta_2 = \sup\{||\varphi(u)|| : u \in \overline{\Omega} \cap K\}$ and choose $\hat{t} > \frac{1}{\vartheta}[\eta_1 + \eta_2]$. From (4.47) we have

$$i_F(\varphi + \hat{t}\hat{g}, \Omega \cap K, K) \neq 0.$$

By the solution property of the fixed point index (see Theorem 4.4.10(d)), we have that there exists a $\hat{u} \in \Omega \cap K$ such that

$$\varphi(\hat{u}) + \hat{t}\hat{g}(\hat{u}) = \hat{u}$$
$$\Rightarrow \hat{t} = \frac{||\hat{u} - \varphi(\hat{u})||}{||\hat{g}(\hat{u})||} \leqslant \frac{1}{\vartheta}[\eta_1, \eta_2], \text{ a contradiction.}$$

The proof is now complete. □

If $g : \partial\Omega \cap K \to K$ is the constant map $g(u) = u_0 \in K\backslash\{0\}$, then from the previous proposition, we deduce the following corollary.

Corollary 4.4.14 *If $\varphi : \overline{\Omega} \cap K \to K$ is a compact map, $u_0 \in K\backslash\{0\}$ and*

$$u - \varphi(u) \neq \lambda u_0 \textit{ for all } u \in \partial\Omega \cap K \textit{ and all } \lambda \geqslant 0,$$

then $i_F(\varphi, \Omega \cap K, K) = 0$.

Next we present some fixed point theorems in conical shells, that is, sets of the form $(\Omega_2\backslash\overline{\Omega_1}) \cap K$ where Ω_1, Ω_2 are two bounded open sets and $\overline{\Omega_1} \subseteq \Omega_2$. Of special interest is the case where $\Omega_1 = B_\rho = \{u \in X : ||u|| < \rho\}$ and $\Omega_2 = B_R = \{u \in X : ||u|| \leqslant R\}$ with $\rho < R$. Such results are also known as fixed point theorems of cone expansion and compression.

Theorem 4.4.15 *If $\Omega_1, \Omega_2 \subseteq X$ are bounded open sets with $0 \in \Omega_1$, $\overline{\Omega}_1 \subseteq \Omega_2$, $\varphi : \overline{\Omega}_2 \cap K \to K$ is a compact map and the following conditions hold*

(i) $\varphi(u) \neq \lambda u$ for all $u \in \partial\Omega_2$, all $\lambda > 1$;
(ii) there exists a $u_0 \in K\backslash\{0\}$ such that

$$u - \varphi(u) \neq \lambda u_0 \textit{ for all } u \in \partial\Omega \textit{ and all } \lambda > 0,$$

then φ has a fixed point in $(\overline{\Omega}_2\backslash\Omega_1) \cap K$.

Proof From Proposition 4.4.12, we have

$$i_F(\varphi, \Omega_2 \cap K, K) = 1. \tag{4.48}$$

Moreover, from Corollary 4.4.14, we have

$$i_F(\varphi, \Omega_1 \cap K, K) = 0. \tag{4.49}$$

If there are no fixed points on $\partial\Omega_2 \cap K$ or on $\partial\Omega_1 \cap K$ (otherwise we are done) then from the domain additivity property of the fixed point index (see Theorem 4.4.10(b)) we have

$$\begin{aligned} i_F(\varphi, (\Omega_2\backslash\overline{\Omega_1}) \cap K, K) &= i_F(\varphi, \Omega_2 \cap K, K) - i_F(\varphi, \Omega_1 \cap K, K) \\ &= 1 - 0 = 1 \text{ (see (4.48) and (4.49)).} \end{aligned}$$

Then the solution property of the fixed point index (see Theorem 4.4.10(d)) implies that there exists a $\hat{u} \in (\Omega_1\backslash\overline{\Omega}_2) \cap K$ such that $\varphi(\hat{u}) = \hat{u}$. □

Corollary 4.4.16 *If* $\Omega_1, \Omega_2 \subseteq X$ *are bounded open sets with* $0 \in \Omega_1, \overline{\Omega}_1 \subseteq \Omega_2$, $\varphi : \overline{\Omega}_2 \cap K \to K$ *is compact and the following conditions hold*

(i) $\varphi(u) - u \notin K$ *for all* $u \in \partial\Omega_2$;
(ii) $u - \varphi(u) \notin K$ *for all* $u \in \partial\Omega_1$,

then φ *has a fixed point in* $(\Omega_2 \backslash \overline{\Omega}_1) \cap K$.

Remark 4.4.17 In fact the above fixed point theorem remains valid if we reverse both conditions in *(i)* and *(ii)*, namely if we assume that $u - \varphi(u) \notin K$ for all $u \in \partial\Omega_2$ and $\varphi(u) - u \notin K$ for all $u \in \partial\Omega_1$. In this case, we have $i_F(\varphi, \Omega_1 \cap K, K) = 0$ and $i_F(\varphi, \Omega_2 \cap K, K) = 1$.

Theorem 4.4.18 *If* $\Omega_1, \Omega_2 \subseteq X$ *are bounded open sets with* $0 \in \Omega_1, \overline{\Omega}_1 \subseteq \Omega_2$, $\varphi : \overline{\Omega}_2 \cap K \to K$ *is compact and one of the following conditions holds*

(i) $||\varphi(u)|| \geqslant ||u||$ *for all* $u \in \partial\Omega_2 \cap K$ *and* $||\varphi(u)|| \leqslant ||u||$ *for all* $u \in \partial\Omega_1 \cap K$; *or*
(ii) $||\varphi(u)|| \leqslant ||u||$ *for all* $u \in \partial\Omega_2 \cap K$ *and* $||\varphi(u)|| \geqslant ||u||$ *for all* $u \in \partial\Omega_1 \cap K$,

then φ *has at fixed point in* $(\overline{\Omega}_2 \backslash \Omega_1) \cap K$.

Proof We focus on the case when *(i)* holds, the proof is similar if *(ii)* holds.

If condition *(i)* is in effect, then we have

$$\varphi(u) \neq \lambda u \text{ for all } u \in \partial\Omega_2 \cap K \text{ and all } \lambda \in (0, 1]. \tag{4.50}$$

Indeed, if (4.50) is not true, then there exist $u_0 \in \partial\Omega_1 \cap K$ and $\lambda_0 \in (0, 1)$ such that $\varphi(u_0) = \lambda_0 u_0$, hence $||\varphi(u_0)|| = \lambda_0 ||u_0||$ and so $||\varphi(u_0)|| < ||u_0||$, a contradiction to our hypothesis.

In a similar fashion we show that

$$\varphi(u) \neq \vartheta u \text{ for all } u \in \partial\Omega_1 \cap K \text{ and all } \vartheta > 1. \tag{4.51}$$

Finally, note that

$$\inf\{\varphi(u) : u \in \partial\Omega_2 \cap K\} \geqslant \inf\{||u|| : u \in \partial\Omega_2 \cap K\} > 0. \tag{4.52}$$

From (4.50) and (4.52), we see that we can use Proposition 4.4.13 and have

$$i_F(\varphi, \Omega_2 \cap K, K) = 0. \tag{4.53}$$

Assume that $\varphi(u) \neq u$ for $\partial\Omega_1$ (or otherwise we already have the desired fixed point), from (4.51) and Proposition 4.4.12, we have

$$i_F(\varphi, \Omega_1 \cap K, K) = 1. \tag{4.54}$$

From (4.53), (4.54) and the domain additivity of the fixed point index (see Theorem 4.4.10(b)), we obtain

$$\begin{aligned} & i_F(\varphi, (\Omega_2 \backslash \overline{\Omega}_1) \cap K, K) = 1 \\ \Rightarrow\ & \varphi \text{ has a fixed point in } (\overline{\Omega}_2 \backslash \Omega_1) \cap K \text{ (by the solution property).} \end{aligned}$$

The proof is now complete. □

The next lemma is a straightforward consequence of the definitions.

Lemma 4.4.19 *If X is an OBS with order cone K, $C \subseteq K$ is compact and $0 \notin C$, then $0 \notin \operatorname{conv} C$.*

Using this lemma, we can prove the following fixed point theorem.

Theorem 4.4.20 *If $\Omega_1, \Omega_2 \subseteq X$ are bounded open sets with $0 \in \Omega_1$, $\overline{\Omega}_1 \subseteq \Omega_2$ and $\varphi : \overline{\Omega}_2 \cap K \to K$ is a compact map which satisfies the following conditions:*

(i) $\varphi(u) \neq \lambda u$ *for all* $u \in \partial\Omega_2 \cap K$ *and all* $\lambda > 1$*;*
(ii) $\varphi(u) \neq \vartheta u$ *for all* $u \in \partial\Omega_1 \cap K$ *and all* $\vartheta \in (0, 1)$*;*
(iii) $\inf\{||\varphi(u)|| : u \in \partial\Omega_1 \cap K\} > 0$*,*

then φ has a fixed point in $(\overline{\Omega}_2 \backslash \Omega_1) \cap K$.

Proof We may assume that the map φ has no fixed points on $\partial\Omega_2 \cap K$ and $\partial\Omega_1 \cap K$. Using the domain additivity property (see Theorem 4.4.10(b)) and Proposition 4.4.12, we have

$$i_F(\varphi, \Omega_2 \cap K \backslash \overline{\Omega}_1 \cap K, K) = 1 - i_F(\varphi, \Omega_1 \cap K, K). \tag{4.55}$$

Using Proposition 2.1.9 (Dugundji's extension theorem) and Lemma 4.4.19, we can find a compact map $\hat{\varphi} : X \to \operatorname{conv}\{\varphi(u) : u \in \partial\Omega_1 \cap K\}$ such that $\hat{\varphi}\Big|_{\partial\Omega_1 \cap K} = \varphi\Big|_{\partial\Omega_1 \cap K}$ and $\inf\{||\hat{\varphi}(u)|| : u \in X\} = \xi > 0$. We consider the homotopy

$$h_t(u) = (1 - t)\varphi(u) + tk\hat{\varphi}(u) \text{ for all } (t, u) \in [0, 1] \times \overline{\Omega}_2, \text{ with } k > 1.$$

Since $k > 1$, we see that $h_t(u) \neq 0$ for all $(t, u) \in [0, 1] \times (\partial\Omega_1 \cap K)$. Hence from the homotopy invariance property (see Theorem 4.4.10(c)), we have

$$i_F(\varphi, \Omega_1 \cap K, K) = i_F(k\hat{\varphi}, \Omega_1 \cap K, K). \tag{4.56}$$

Note that $u = k\hat{\varphi}(u)$ implies that $k \leqslant \xi_0$ for some $\xi_0 > 0$ large enough. Therefore for $k > \max\{1, \xi_0\}$ large we have

$$i_F(k\hat{\varphi}, \Omega_1 \cap K, K) = 0. \tag{4.57}$$

Then from (4.55), (4.56), (4.57) and the solution property (see Theorem 4.4.10(d)), we conclude that φ has a fixed point in $(\overline{\Omega}_2 \backslash \Omega_1) \cap K$. $\square$

The following multiple fixed point theorems are direct consequences of Corollary 4.4.16 and Theorem 4.4.18 respectively.

Theorem 4.4.21 *If $\Omega_1, \Omega_2, \Omega_3$ are bounded open sets with $0 \in \Omega_1 \subseteq \overline{\Omega}_1 \subseteq \Omega_2 \subseteq \overline{\Omega}_2 \subseteq \Omega_3$ and $\varphi : (\overline{\Omega}_3 \backslash \Omega_1) \cap K \to K$ is a compact map which satisfies the following conditions:*

(i) $u - \varphi(u) \notin K$ for all $u \in \partial\Omega_1 \cap K$;
(ii) $\varphi(u) - u \notin K$ for all $u \in \partial\Omega_2 \cap K$;
(iii) $u - \varphi(u) \notin K$ for all $u \in \partial\Omega_3 \cap K$,

then φ has at least two fixed points $u_1^*, u_2^* \in (\overline{\Omega}_3 \backslash \Omega_1) \cap K$ such that

$$u_1^* \in (\Omega_3 \backslash \overline{\Omega}_2) \cap K \text{ and } u_2^* \in (\Omega_2 \backslash \overline{\Omega}_1) \cap K.$$

Theorem 4.4.22 *If $\Omega_1, \Omega_2, \Omega_3$ are bounded open sets such that $0 \in \Omega_1 \subseteq \overline{\Omega}_1 \subseteq \Omega_2 \subseteq \overline{\Omega}_2 \subseteq \Omega_3$ and $\varphi : (\overline{\Omega}_3 \backslash \Omega_1) \cap K \to K$ is a compact map which satisfies the following conditions:*

(i) $||\varphi(u)|| \geqslant ||u||$ for all $u \in \partial\Omega_1 \cap K$;
(ii) $||\varphi(u)|| \leqslant ||u||$ and $\varphi(u) \neq u$ for all $u \in \partial\Omega_2 \cap K$;
(iii) $||\varphi(u)|| \geqslant ||u||$ for all $u \in \partial\Omega_3 \cap K$,

then φ has at least two fixed points $u_1^*, u_2^* \in (\overline{\Omega}_3 \backslash \Omega_1) \cap K$ such that

$$u_1^* \in (\Omega_2 \backslash \Omega_1) \cap K \text{ and } u_2^* \in (\overline{\Omega}_3 \backslash \overline{\Omega}_2) \cap K.$$

4.5 Fixed Points for Multifunctions

In this section we extend some of the fixed point theorems from the previous sections to multifunctions. We start with the Banach fixed point theorem (see Theorem 4.2.3). To formulate its multivalued counterpart, we use the Hausdorff metric defined on the hyperspace $P_{bf}(X) = \{B \subseteq X : \text{nonempty, bounded and closed}\}$, with (X, d) being a metric space (see Definition 3.7.4). Recall that $(P_{bf}(X), h)$ is a complete metric space if and only if (X, d) is a complete metric space.

Theorem 4.5.1 *If (X, d) is a complete metric space and $F : X \to P_{bf}(X)$ is a multifunction which satisfies*

$$h(F(u), F(y)) \leqslant kd(u, y) \text{ for all } u, y \in X \text{ with } k \in [0, 1)\,, \tag{4.58}$$

then F has a fixed point, that is, there exists a $\hat{u} \in X$ such that $\hat{u} \in F(\hat{u})$.

Proof Let $\hat{k} \in (k, 1)$ and $u_0 \in X$. We choose $u_1 \in F(u_0)$ such that $d(u_1, u_0) > 0$. If no such point exists then u_0 is the desired fixed point of F. We have

$$d(u_1, F(u_1)) \leqslant h(F(u_0), F(u_1)) < \hat{k}d(u_0, u_1) \text{ (see Definition 3.7.4).}$$

So, we can find $u_2 \in F(u_1)$ such that $d(u_1, u_2) < \hat{k}d(u_0, u_1)$. Inductively, we generate a sequence $\{u_n\}_{n \geqslant 1} \subseteq X$ such that

$$u_{n+1} \in F(u_n) \text{ and } d(u_n, u_{n+1}) < \hat{k}^n d(u_0, u_1) \text{ for all } n \geqslant 1. \tag{4.59}$$

From (4.46) and the completeness of (X, d), we have that $u_n \to \hat{u}$ in (X, d). Also, we have

$$\begin{aligned} &d(u_{n+1}, F(\hat{u})) \leqslant h(F(u_n), F(\hat{u})) \leqslant kd(u_n, \hat{u}) \to 0 \\ \Rightarrow\ &\hat{u} \in F(\hat{u}) \text{ (since } F \text{ has a closed values).} \end{aligned}$$

The proof is now complete. □

Remark 4.5.2 In contrast to the single-valued case (see Theorem 4.2.3), the fixed point $\hat{u} \in X$ obtained in the above theorem need not be unique. To see this let $X = [0, 1]$ with the usual metric and consider the function $\vartheta : X \to X$ defined by

$$\vartheta(u) = \begin{cases} \frac{1}{2}(u+1) & \text{if } u \in [0, \frac{1}{2}] \\ -\frac{1}{2}(u-2) & \text{if } u \in [\frac{1}{2}, 1]. \end{cases}$$

Evidently, this is a continuous function. Let $F : [0, 1] \to P_f([0, 1])$ be defined by

$$F(u) = \{0\} \cup \{\vartheta(u)\}.$$

Then $F(\cdot)$ satisfies (4.58) and $u = 0$, $u = \frac{1}{2}$ are both fixed points of $F(\cdot)$.

Next we present multivalued counterparts of the Schauder fixed point theorem (see Theorem 3.2.20). We start with the easy case of a lower semicontinuous (lsc) multifunction (see Definition 2.5.2 (b)).

Theorem 4.5.3 *If X is a Banach space, $C \subseteq X$ is nonempty bounded closed convex and $F : C \to P_{f_c}(C)$ is an lsc multifunction with $F(C) \subseteq X$ relatively compact, then $F(\cdot)$ has a fixed point.*

Proof Invoking Theorem 2.5.14 (the Michael selection theorem), we can find a continuous map $f : C \to C$ such that $f(u) \in F(u)$ for all $u \in C$. Applying Schauder's fixed point theorem (see Theorem 3.2.20), we can find $\hat{u} \in C$ such that $\hat{u} = f(\hat{u}) \in F(\hat{u})$. □

The continuous selection approach can also be used for upper semicontinuous (usc) multifunctions (see Definition 2.5.2 (a)), but not directly since, as we already saw in Example 2.5.13, usc multifunctions need not have a continuous selection. To overcome this difficulty, we need the following lemma.

Lemma 4.5.4 *If (M, d) is a metric space, X is a Banach space, $F : M \to 2^X \setminus \{\emptyset\}$ and for every $s \in M$ and $r > 0$, we define*

$$\hat{F}_r(s) = F(B_r(s)) = \bigcup_{t \in B_r(s)} F(t) \; (B_r(s) = \{t \in M : d(t, s) < r\}),$$

then the multifunction $s \to \hat{F}_r(s)$ from M into $2^X \setminus \{\emptyset\}$ is lsc.

Proof Let $\{s_\alpha\}_{\alpha \in J}$ be a net in M such that $d(s_\alpha, s) \to 0$ and let $u \in \hat{F}_r(s)$. From the definition of $\hat{F}_r(\cdot)$, we know that there exists a $t \in B_r(s)$ such that $u \in F(t)$. Since $s_\alpha \to s$ in (M, d) for all $\alpha \in J$ with $\alpha \geqslant \alpha^*$, we have $s_\alpha \in B_r(s)$. Hence $u \in \bigcup_{t \in B_{r(s_\alpha)}} F(t) = \hat{F}_r(s_\alpha)$. So, for $\alpha \geqslant \alpha^*$ we take $u_\alpha = u$ and by virtue of Proposition 2.5.4 we conclude that $\hat{F}_r(\cdot)$ is lsc. ☐

With this lemma, we have the usc multivalued counterpart of the Schauder fixed point theorem.

Theorem 4.5.5 *If X is a Banach space, $C \subseteq X$ is nonempty, bounded, closed and convex and $F : C \to P_{f_c}(C)$ is compact (see Definition 3.3.1), then F has a fixed point.*

Proof For every $n \geqslant 1$, let

$$\tilde{F}_n(u) = \overline{\text{conv}}\, \hat{F}_n(u) = \overline{\text{conv}}\, F(B_{\frac{1}{n}}(u)).$$

From Lemma 4.5.4 we have that $\tilde{F}_n(\cdot)$ is lsc and so we can apply Proposition 3.5.3 and find a $\hat{u}_n \in C$ such that $\hat{u}_n \in \tilde{F}_n(\hat{u}_n)$. Since F is compact and C is bounded, by passing to a subsequence if necessary, we may assume that $\hat{u}_n \to \hat{u}$ in X.

For every $u^* \in X^*$ we have

$$\begin{aligned} \langle u^*, \hat{u}_n \rangle &\leqslant \sigma(u^*, \tilde{F}_n(\hat{u}_n)) = \sup \left\{ \langle u^*, u \rangle : u \in \tilde{F}_n(\hat{u}_n) \right\} \\ &\qquad \text{(the support function of the set } \tilde{F}_n(u_n)\text{)} \\ &= \sup(u^*, F(B_{\frac{1}{n}}(\hat{u}_n))) \leqslant \sup \left\{ \sigma(u^*, F(u)) : u \in B_{\frac{1}{n}}(\hat{u}_n) \right\}. \end{aligned}$$

So, for every $n \geqslant 1$, we can find $u_n \in B_{\frac{1}{n}}(\hat{u}_n)$ such that

$$\langle u^*, u_n \rangle - \frac{1}{n} \leqslant \sigma(u^*, F(u_n)).$$

Note that $u_n \to \hat{u}_0$ in X and so

$$\langle u^*, \sigma u \rangle \leqslant \limsup_{n\to\infty} \sigma(u^*, F(u_n)) \leqslant \tau(u^*, F(\hat{u})) \tag{4.60}$$
$$\text{(see } F(\cdot) \text{ is usc)}.$$

Since $u^* \in X^*$ is arbitrary and $F(\hat{u})$ is closed convex, from (4.60) it follows that

$$\hat{u} \in F(\hat{u}).$$

The proof is now complete. □

Remark 4.5.6 A careful reading of the above proof reveals that the result remains true for any locally convex topological vector space and in particular for a Banach space endowed with its weak topology. Moreover, we need not assume that F is a self-map, that is, that maps C into itself. To do this, we introduce the following notion.

Definition 4.5.7 Let X be a locally convex vector space, $C \subseteq X$ a nonempty convex set and $u \in \overline{C}$. The tangent cone $T_C(u)$ to the convex set C at u is the closed cone spanned by $C - u$, that is,

$$T_C(u) = \overline{\bigcup_{\lambda>0} \frac{1}{\lambda}(C - u)}.$$

Remark 4.5.8 It is easy to see that $T_C(u)$ is convex. If X is a Banach space, then

$$T_C(u) = \{h \in X : \liminf_{\lambda\to 0^+} \frac{1}{\lambda} d(u + \lambda h, C) = 0\}.$$

We will need the following result, known in the literature as the Ky Fan inequality. For its proof, we refer to Aubin and Ekeland [21, p. 327].

Theorem 4.5.9 *If X is a locally convex space, $C \subseteq X$ is nonempty, compact and convex and $\xi : C \times C \to \mathbb{R}$ is a function satisfying*

(i) for every $y \in C$, $u \to \xi(u, y)$ is lower semicontinuous;
(ii) for every $u \in C$, $y \to \xi(u, y)$ is concave,

then there exists a $\overline{u} \in C$ such that

$$\sup\{\xi(\overline{u}, y) : y \in C\} \leqslant \sup\{\xi(u, y) : y \in C\}.$$

Using this theorem, we can establish the existence of equilibrium points for usc multifunctions.

Proposition 4.5.10 *If X is a locally convex space, $F : X \to P_{f_c}(X)$ is usc, $C \subseteq X$ is nonempty, compact and convex and*

$$F(u) \cap T_C(u) \neq \emptyset \text{ for all } u \in C, \tag{4.61}$$

then there exists a $\hat{u} \in C$ *such that* $0 \in F(\hat{u})$.

Proof We argue indirectly. So, suppose that the result of the proposition is not true. Then for every $u \in C$, we have $0 \notin F(u)$. Then by the strong separation theorem, we can find $u_u^* \in X^*$ such that $\sigma(u_u^*, F(u)) = \sup\left\{\left\langle u_u^*, y\right\rangle : y \in F(u)\right\}$. We set

$$U(u^*) = \{u \in C : \sigma(u^*, F(u)) < 0\}.$$

Since $F(\cdot)$ is usc, the mapping $u \to \sigma(u^*, F(u))$ is an upper semicontinuous $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$-valued function and so $U(u^*)$ is relatively open. The sets $\{U(u^*)\}_{u^* \in X^*}$ form an open cover of C and so by compactness, we can find $\{u_k^*\}_{k=1}^m \subseteq X^*$ such that $\{U(u_k^*)\}_{k=1}^m$ still covers C. The compactness of C implies that it is a paracompact space. So, we can find a continuous partition of unity $\{\vartheta_k\}_{k=1}^m$ corresponding to the finite open cover $\{U(u_k^*)\}_{k=1}^m$. Let $\xi : C \times C \to \mathbb{R}$ be the function defined by

$$\xi(u, y) = \sum_{\mathrm{k}=1}^{m} \vartheta_k(u) \left\langle u_k^*, u - y\right\rangle.$$

Evidently, the mapping $u \mapsto \xi(u, y)$ is continuous, while $y \mapsto \xi(u, y)$ is affine. Hence we can apply Theorem 4.5.9 and find $\overline{u} \in C$ such that for $\overline{u}^* = \sum_{\mathrm{k}=1}^{m} \vartheta_k(\overline{u}) u_k^*$, we have

$$\begin{aligned} &\xi(\overline{u}, y) = \left\langle \overline{u}^*, \overline{u} - y\right\rangle \leqslant 0 \text{ for all } y \in C \\ \Rightarrow\ &\overline{u}^* \in T_C(\overline{u})^* \text{ (see Definition 4.1.16).} \end{aligned}$$

Condition (4.61) implies that there exists a $v \in F(\overline{u}) \cap T_C(\overline{u})$ such that

$$0 \leqslant \left\langle \overline{u}^*, v\right\rangle \leqslant \sigma(\overline{u}^*, F(\overline{u})). \tag{4.62}$$

Let $I(\overline{u}) = \{k \in \{1, \ldots, m\} : \vartheta_k(\overline{u}) > 0\} \neq \emptyset$. We have

$$\begin{aligned} &\sigma(\overline{u}^*, F(\overline{u})) \leqslant \sum_{\mathrm{k} \in \mathrm{I}(\overline{\mathrm{u}})} \vartheta_k(\overline{u}) \sigma(u_k^*, F(\overline{u})) < 0 \\ \Rightarrow\ &\sigma(u_k^*, F(\overline{u})) < 0 \text{ for all } k \in I(\overline{u}) \\ \Rightarrow\ &\sigma(\overline{u}^*, F(\overline{u})) < 0, \text{ which contradicts (4.62).} \end{aligned}$$

The proof is now complete. □

This leads to an alternative proof of the multivalued Schauder fixed point theorem (usc case), known as the Kakutani–Ky Fan fixed point theorem (see Theorem 4.5.5).

Theorem 4.5.11 *If* X *is a locally convex space,* $C \subseteq X$ *is nonempty, compact and convex and* $G : X \to P_{f_c}(K)$ *is usc, then* $G(\cdot)$ *has a fixed point* $\overline{u} \in K$.

Proof Let $F(u) = G(u) - u$ for all $u \in X$. Evidently, $F(\cdot)$ is still usc with values in $P_{f_c}(X)$. Note that $C - u \subseteq T_C(u)$ (see Definition 4.5.7) and since $G(C) \subseteq C$, we have

$$F(u) = G(u) - u \subseteq C - u \subseteq T_C(u).$$

So, we can apply Proposition 4.5.10 and find $\overline{u} \in C$ such that $0 \in F(\overline{u}) = G(\overline{u}) - \overline{u}$. Hence $\hat{u} \in G(\hat{u})$, $\hat{u} \in C$. $\square$

As we already mentioned earlier, $G(\cdot)$ need not map C to itself. It suffices for a condition of the nature of (4.60) to hold. We formalize this in the following definition.

Definition 4.5.12 Let X be locally convex space, $C \subseteq X$ be nonempty and $G : C \to 2^X \backslash \{\emptyset\}$ be a multifunction. We say that

(a) $G(\cdot)$ is inward if and only if $G(u) \cap (u + T_C(u)) \neq \emptyset$ for all $u \in C$;
(b) $G(\cdot)$ is outward if and only if $G(u) \cap (u - T_C(u)) \neq \emptyset$ for all $u \in C$.

Remark 4.5.13 If $G(\cdot)$ is inward, then $F = G - id$ satisfies condition (4.61). Similarly, if $G(\cdot)$ is outward, then $F = id - G$ satisfies condition (4.61). These observations, combined with Proposition 4.5.10, lead to the following fixed point theorem.

Theorem 4.5.14 *If X is a locally convex space, $C \subseteq X$ is nonempty, compact and convex and $G : C \to P_{f_c}(X)$ is a usc multifunction which is inward or outward, then $G(\cdot)$ has a fixed point $\overline{u} \in C$.*

4.6 Abstract Variational Principles

In this section we present some basic abstract variational principles of nonlinear analysis and investigate their connections.

We start with the Lax–Milgram theorem, which is an important tool in the study of linear boundary value problems, and revisit the Galerkin approximation method (see also Sect. 3.7).

Theorem 4.6.1 *If H is a Hilbert space with inner product $(\cdot, \cdot)_H$ and $a : H \times H \to \mathbb{R}$ is a bilinear form such that*

(i) a is continuous, that is, there exists an $M > 0$ such that

$$|a(u, y)| \leqslant M||u|| \, ||y|| \text{ for all } u, y \in H;$$

(ii) a is coercive, that is, there exists a $c > 0$ such that

$$a(u, u) \geqslant c||u||^2 \text{ for all } u \in H,$$

then for any $h^* \in H^*$, there exists a unique $u_0 \in H$ such that

$$a(u_0, y) = \langle h^*, y \rangle \text{ for all } y \in H$$

($\langle \cdot, \cdot \rangle$ being the duality brackets for the pair (H^*, H)).

Moreover, if $a(\cdot, \cdot)$ is symmetric, that is, $a(u, y) = a(y, u)$ for all $u, y \in H$, then $u_0 \in H$ is the unique minimizer of the functional

$$\varphi(u) = \frac{1}{2} a(y, y) - \langle h^*, y \rangle \text{ for all } y \in H.$$

Proof We are going to present two proofs. The first is based on the Banach fixed point theorem (see Theorem 4.2.3) and so it has the advantage of being a constructive proof, while the second is based on the surjectivity properties of continuous linear operators.

First proof: By virtue of the Riesz representation theorem for Hilbert spaces we can find $h \in H$ such that $\langle h^*, y \rangle = (h, y)_H$ for all $y \in H$. By hypothesis *(i)*, for every fixed $u \in H$, $y \to a(u, y)$ is continuous and linear on H. So, a new application of the Riesz representation theorem implies that there exists a unique element $A(u) \in H$ such that

$$a(u, y) = (A(u), y)_H \text{ for all } y \in H. \tag{4.63}$$

For any $u_1, u_2, y \in H$, we have

$$\begin{aligned} (A(u_1 + u_2), y)_H = a(u_1 + u_2, y) = a(u_1, y) + a(u_2, y) = \\ \langle A(u_1), y \rangle + \langle A(u_2), y \rangle \\ \Rightarrow A(u_1 + u_2) = A(u_1) + A(u_2). \end{aligned} \tag{4.64}$$

Also, for every $\lambda \in \mathbb{R}$ and $u, y \in H$, we have

$$\begin{aligned} (A(\lambda u), y)_H = a(\lambda u, y) = \lambda a(u, y) = \lambda (A(u), y)_H \\ \Rightarrow A(\lambda u) = \lambda A(u). \end{aligned} \tag{4.65}$$

From (4.63), (4.64), (4.65) and hypothesis *(i)* we conclude that $A \in \mathscr{L}(H)$. Our problem takes the following form:

"Find $u_0 \in H$ such that $(A(u_0), y)_H = (h, y)_H$ for all $y \in H$".

This of course means that we must look for $u_0 \in H$ such that

$$A(u_0) = h. \tag{4.66}$$

We know that $A \in \mathscr{L}(H)$. Moreover, using hypothesis *(ii)*, we have

$$(A(u), u)_H \geqslant c||u||^2 \text{ for all } u \in H. \tag{4.67}$$

Let $\lambda > 0$. To solve (4.66) is equivalent to finding $u_0 \in H$ such that

$$u_0 - \lambda(A(u_0) - h) = u_0. \tag{4.68}$$

Problem (4.68) is a fixed point problem for the map $\psi_\lambda : H \to H$ defined by

$$\psi_\lambda(u) = u - \lambda(A(u) - h).$$

We examine this map. So, let $u_1, u_2 \in H$. Then

$$\begin{aligned}
&\psi_\lambda(u_1) - \psi_\lambda(u_2) = u_1 - u_2 - \lambda A(u_1 - u_2) \\
\Rightarrow\ &||\psi_\lambda(u_1) - \psi_\lambda(u_2)|| = ||u - \lambda A(u)||^2 \text{ with } u = u_1 - u_2 \\
&\qquad\qquad\qquad\qquad = ||u||^2 - 2\lambda(A(u), u)_H + \lambda^2||A(u)||^2 \\
\Rightarrow\ &||\psi_\lambda(u_1) - \psi_\lambda(u_2)||^2 \leqslant (1 - 2\lambda c + \lambda^2 M^2)||u||^2 \text{ (see (4.67) and hypothesis } (i)).
\end{aligned}$$

To be able to apply the Banach fixed point theorem (see Theorem 4.2.3), we have to find $\lambda > 0$ such that $1 - 2\lambda c + \lambda^2 M^2 < 1$. Let $\hat{\lambda}$ be the minimizer of $1 - 2\lambda c + \lambda^2 M^2$, that is, $\hat{\lambda} = \frac{C}{M^2}$. Then $1 - 2\hat{\lambda}c + \hat{\lambda}^2 M^2 = 1 - \frac{C^2}{M^2} < 1$. Hence $\psi_{\hat{\lambda}}$ is a contradiction and so by Theorem 4.2.3 it has a unique fixed point $u_0 \in H$. This is the solution of our problem.

Now suppose that $a(\cdot, \cdot)$ is symmetric. For all $y \in H$ we have

$$\begin{aligned}
\varphi(y) = \varphi(u_0 + y - u_0) &= \frac{1}{2}a(u_0 + y - u_0, u_0 + y - u_0) - (h, u_0 + y - u_0)_H \\
&= \frac{1}{2}a(u_0, u_0) - (h, u_0)_H + a(u_0, y - u_0) - (h, y - u_0)_H \\
&\quad + \frac{1}{2}a(y - u_0, y - u_0) \ \text{ (since } a(\cdot, \cdot) \text{ is symmetric)} \\
&\geqslant \varphi(u_0) + \frac{c}{2}||y - u_0||^2 \\
\Rightarrow \quad & u_0 \in H \text{ is the unique minimizer of } \varphi.
\end{aligned}$$

Second proof: Again we reduce the problem to equation (4.66). We are going to show that $R(A) = H$. First we show that $R(A)$ is closed. Recall that

$$\begin{aligned}
&||A(u)||\ ||u|| \geqslant (A(u), u)_H \geqslant c||u||^2 \text{ for all } u \in H \\
\Rightarrow\ &||A(u)|| \geqslant c||u|| \text{ for all } u \in H.
\end{aligned} \tag{4.69}$$

Suppose that $A(u_n) \to v$ in H. Then for all $n, m \geqslant 1$ we have

$$\begin{aligned} c||u_n - u_m|| &\leqslant ||A(u_n - u_m)|| \text{ (see (4.69))} \\ &\leqslant ||A(u_n) - A(u_m)|| \\ &\Rightarrow \{u_n\}_{n\geqslant 1} \subseteq H \text{ is Cauchy} \\ &\Rightarrow u_n \to u \text{ in } H \\ &\Rightarrow A(u_n) \to A(u) = v \text{ (since } A \in \mathscr{L}(H)). \end{aligned}$$

This proves the closedness of $R(A)$.

Next, we show that $R(A)$ is dense. So, let $v \in R(A)^\perp$. Then

$$\begin{aligned} &(A(u), v)_H = 0 \text{ for all } u \in H \\ \Rightarrow\ & c||v||^2 \leqslant (A(v), v)_H = 0, \text{ hence } v = 0 \\ \Rightarrow\ & R(A) \text{ is dense in } H. \end{aligned}$$

We conclude that $R(A) = H$ and so (4.66) has a solution u_0, which is clearly unique. When $a(\cdot, \cdot)$ is symmetric, as before we show that u_0 is the unique minimizer of the functional φ. □

The result can be extended to variational inequalities.

Theorem 4.6.2 *If H is a Hilbert space with inner product $(\cdot, \cdot)_H$, $a : H \times H \to \mathbb{R}$ is a bilinear form which is continuous and coercive, $C \subseteq H$ is nonempty, closed and convex and $h \in H$, then there exists a unique $u_0 \in C$ such that*

$$a(u_0, y - u_0) \geqslant (h, y - u_0)_H \text{ for all } y \in C. \tag{4.70}$$

Moreover, if $a(\cdot, \cdot)$ is symmetric, then $u_0 \in H$ is the unique minimizer of the functional

$$\varphi(y) = \frac{1}{2}a(y, y) - (h, y)_H \text{ for all } y \in H.$$

Proof As before (see the proof of Theorem 4.6.1), we can find $A \in \mathscr{L}(H)$ such that

$$\|A(u)\| \geqslant c\|u\| \text{ for all } u \in H \tag{4.71}$$

$$\text{and } (A(u), y)_H = a(u, y) \text{ for all } u, y \in H. \tag{4.72}$$

Let i_C be the indicator function of the set C, that is

$$i_C(y) = \begin{cases} 0 & \text{if } y \in C \\ +\infty & \text{if } y \notin C. \end{cases}$$

Since C is nonempty, closed and convex, i_C is a lower semicontinuous, convex function which is not identically $+\infty$. Identifying H^* with H, for all $u \in C$, we have

$$\partial i_C(u) = N_C(u) = \{v \in H : (v, y-u)_H \leqslant 0 \text{ for all } y \in C\} = -\,T_C(u)^*$$
$$\text{(the normal cone of } C \text{ at } u).$$

We see that $u_0 \in C$ solves (4.70) if and only if $0 \in A(u_0) + \partial i_C(u_0) - h$. Theorem 2.8.5 guarantees the existence of such a solution of (4.70) (see (4.72)). Also, by virtue of (4.71) this solution is unique.

Finally, when $a(\cdot,\cdot)$ is symmetric, as in the proof of Theorem 4.6.1 we show that $u_0 \in H$ is the unique minimizer of the functional φ. □

Remark 4.6.3 If $C = H$, then (4.70) becomes $a(u, y) \geqslant (h, y)_H$ for all $y \in H$, hence $a(u, y) = (h, y)_H$ and we recover the Lax–Milgram theorem (see Theorem 4.6.1).

Suppose that H is a separable Hilbert space. From Proposition 3.7.17 we know that H admits a projection scheme (Galerkin scheme) $\{P_n, H_n\}_{n\geqslant 1}$. Let $a : H \times H \to \mathbb{R}$ be a continuous and coercive bilinear form and $h^* \in H^*$. We consider the following problems:

$$\text{“Find } u \in H \text{ such that } a(u, y) = \langle h^*, y\rangle \text{ for all } y \in H\text{”.} \tag{4.73}$$

$$\text{“Find } u_n \in H_n \text{ such that } a(u_n, y) = \langle h^*, y\rangle \text{ for all } y \in H_n\text{”.} \tag{4.74}$$

From Theorem 4.6.1 we know that both problems have unique solutions $u \in H$ and $u_n \in H_n$, $n \geqslant 1$, respectively.

Proposition 4.6.4 *In the above setting,* $\|u_n - u\| \leqslant \frac{M}{c} d(u, H_n)$ *for all* $n \geqslant 1$.

Proof From (4.73) and (4.74) we have

$$a(u - u_n, y) = 0 \text{ for all } y \in H_n \text{ and } n \geqslant 1 \tag{4.75}$$

$$\Rightarrow a(u - u_n, u_n) = 0 \text{ for all } n \geqslant 1. \tag{4.76}$$

Then for every $y \in H_n$ we have

$$\begin{aligned} a(u-u_n, u-u_n) &= a(u-u_n, u-y) + a(u-u_n, y-u_n) \\ &= a(u-u_n, u-y) \text{ (see (4.75), (4.76))} \end{aligned}$$

$$\begin{aligned} &\Rightarrow c\|u-u_n\|^2 \leqslant M\|u-u_n\|\|u-y\| \\ &\Rightarrow c\|u-u_n\| \leqslant M\|u-y\| \text{ for all } y \in H_n \\ &\Rightarrow \|u-u_n\| \leqslant \frac{M}{c} d(y, H_n) \text{ for all } n \geqslant 1. \end{aligned}$$

The proof is now complete. □

Remark 4.6.5 The above proposition not only tells us that $u_n \to u$ in H, but also provides an explicit bound for $\|u - u_n\|$.

The Galerkin method also works for variational inequalities. In fact, we will be more general. So, let X be a separable, reflexive Banach space and X^* its topological dual. By $\langle \cdot, \cdot \rangle$ we denote the duality brackets for the pair (X^*, X). Let $A : X \to X^*$ be a bounded, hemicontinuous (see Definition 2.6.10), monotone map, $C \subseteq X$ a nonempty, closed, convex set and $h^* \in X^*$. We study the following variational inequality problem:

"Find $u \in C$ such that

$$\langle A(u) - h^*, y - u \rangle \geqslant 0 \text{ for all } y \in C \text{''}. \tag{4.77}$$

We will use the Galerkin method to deal with this problem. First we produce finite-dimensional convex sets approximating the constraint set C.

Lemma 4.6.6 *There exists an increasing sequence $\{C_n\}_{n\geqslant 1}$ of finite-dimensional convex sets such that $\overline{\bigcup_{n\geqslant 1} C_n} = C$.*

Proof Let $\{u_k\}_{k\geqslant 1} \subseteq C$ be dense in C. We set

$$C_n = \text{conv}\{u_k\}_{k=1}^n.$$

Then C_n is closed, convex and $\dim C_n \leqslant n$. Moreover, $C_n \subseteq C_{n+1} \subseteq C$ for all $n \geqslant 1$. Since $\{u_n\}_{n\geqslant 1} \subseteq \bigcup_{n\geqslant 1} C_n$, it follows that $\overline{\bigcup_{n\geqslant 1} C_n} = C$. $\square$

Then we consider the following finite-dimensional approximation of problem (4.77).

"Find $u_n \in C$ such that

$$\langle A(u_n) - h^*, y - u_n \rangle \geqslant 0 \text{ for all } y \in C_n \text{''}. \tag{4.78}$$

Lemma 4.6.7 *For every $n \geqslant 1$, problem (4.78) has at least one solution $u_n \in C_n$.*

Proof Let $X_n = \text{span}\, C_n$, $n \geqslant 1$. Evidently, X_n is finite-dimensional. We endow X_n with a Euclidean structure and denote the corresponding inner product by $(\cdot, \cdot)_n$. Then we can find a continuous linear map $L_n : X^*_w \to X_n$ such that

$$\langle u^*, y \rangle = (L_n(u^*), y)_n \text{ for all } n \geqslant 1, \text{ all } u^* \in X^* \text{ and all } y \in X_n. \tag{4.79}$$

Let $\xi_n : X_n \to C_n$, $n \geqslant 1$, be the metric projection map. We know that

$$(y - \xi_n(y), \xi_n(y) - v)_n \geqslant 0 \text{ for all } v \in C_n. \tag{4.80}$$

We introduce the map $\vartheta_n : C_n \to C_n$ defined by

$$\vartheta_n(v) = \xi_n(v - L_n(A(v)) + L_n(h^*)),\ n \geqslant 1. \tag{4.81}$$

Evidently, ϑ_n is continuous and so we can apply Corollary 4.3.3 and find $u_n \in C_n$ such that $\vartheta_n(u_n) = u_n$ for all $n \geqslant 1$. From (4.80) and (4.81) it follows that

$$\begin{aligned}
&(u_n - L_n(A(u_n)) + L_n(h^*) - \vartheta_n(u_n), \vartheta_n(u_n) - v)_n \geqslant 0 \text{ for all } v \in C_n \\
\Rightarrow\ &(L_n(h^*) - L_n(A(u_n)), u_n - v)_n \geqslant 0 \text{ for all } v \in C_n \\
\Rightarrow\ &\langle h^* - A(u_n), u_n - v\rangle \geqslant 0 \text{ for all } v \in C_n \text{ (see (4.79))}.
\end{aligned}$$

The proof is now complete. □

Remark 4.6.8 Note that so far the only place where we have used the monotonicity of A is to deduce that A is demicontinuous and so infer the continuity of ϑ_n. Hence if from the beginning we assume the demicontinuity of A and we drop the monotonicity, the result of Lemma 4.6.7 remains true.

Lemma 4.6.9 *By passing to a subsequence if necessary, we may assume that*

$$u_n \xrightarrow{w} u \text{ in } X \text{ and } A(u_n) \xrightarrow{w} u^* \text{ in } X^*.$$

Proof We have $u_n \in C_n \subseteq C$ for all $n \geqslant 1$ and so $\{u_n\}_{n\geqslant 1} \subseteq X$ is bounded. Since A is bounded, $\{A(u_n)\}_{n\geqslant 1} \subseteq X^*$ is bounded too. Then the reflexivity of X (hence of X^* too) implies that by passing to a suitable subsequence if necessary, we may assume that

$$u_n \xrightarrow{w} u \text{ in } X \text{ and } A(u_n) \xrightarrow{w} u^* \text{ in } X^*.$$

The proof is now complete. □

Lemma 4.6.10 *We have $\langle A(u_n), u_n\rangle \to \langle u^*, u\rangle$ and $u^* = A(u)$.*

Proof For every $n \geqslant k$, we have $C_k \subseteq C_n$ and

$$\begin{aligned}
&\langle A(u_n) - h^*, v - u_n\rangle \geqslant 0 \text{ for all } v \in C_k \\
\Rightarrow\ &\limsup_{n\to\infty}\langle A(u_n), u_n\rangle \leqslant \langle h^*, u - v\rangle + \langle u^*, v\rangle \\
\Rightarrow\ &\limsup_{n\to\infty}\langle A(u_n), u_n - u\rangle \leqslant \langle h^* - u^*, u - v\rangle \text{ for all } v \in C_k,\ k \geqslant 1.
\end{aligned} \tag{4.82}$$

Since $\bigcup\limits_{k\geqslant 1} C_k$ is dense in C, we can find $v_k \in C_k$ such that $v_k \to u$. So, from (4.82) it follows that

$$\limsup_{n\to\infty}\langle A(u_n), u_n - u\rangle \leqslant 0. \tag{4.83}$$

From Proposition 2.6.12 we know that A is maximal monotone. Then Proposition 2.10.4 implies that A is generalized pseudomonotone. Hence by virtue of (4.83) we have

$$\langle A(u_n), u_n\rangle \to \langle u^*, u\rangle \text{ and } u^* = A(u) \text{ (see Definition 2.10.1 } (b)).$$

The proof is now complete. □

This last lemma leads to the following existence theorem for problem (4.77).

Theorem 4.6.11 *If X is a separable reflexive Banach space, $A : X \to X^*$ is bounded, hemicontinuous and monotone, $C \subseteq X$ is nonempty, bounded, closed and convex and $h^* \in X^*$, then problem (4.77) admits at least one solution $u_0 \in C$, that is*

$$\langle A(u_0) - h^*, y - u_0\rangle \geqslant 0 \textit{ for all } y \in C.$$

Remark 4.6.12 It is easy to see that this solution is unique if A is strictly monotone.

If C is unbounded, then we can still produce a solution provided we introduce an additional hypothesis on $A(\cdot)$.

Theorem 4.6.13 *If X is a separable reflexive Banach space, $A : X \to X^*$ is bounded, hemicontinuous and monotone, $C \subseteq X$ is nonempty, closed and convex, $h^* \in X^*$, and there exists a $v_0 \in C$ such that*

$$\lim_{\|v\|\to\infty} \frac{\langle A(v), v - v_0\rangle}{\|v\|} = +\infty, \tag{4.84}$$

then problem (4.77) admits at least one solution $u_0 \in C$, that is,

$$\langle A(u_0) - h^*, y - u_0\rangle \geqslant 0 \textit{ for all } y \in C.$$

Proof For every $r > 0$, we set $C_r = \{v \in C : \|v\| \leqslant r\}$. For $r > 0$ big enough (say, $r \geqslant r^*$) we will have that C_r is nonempty, bounded, closed and convex and $v_0 \in C_r$ (see (4.84)). So, we can apply Theorem 4.6.11 and find $u_r \in C_r$ such that

$$\langle A(u_r) - h^*, y - u_r\rangle \geqslant 0 \text{ for all } y \in C_r.$$

Let $y = v_0$. Then we have

$$\langle A(u_r), u_r - v_0\rangle \leqslant \|h^*\|_*(\|v_0\| + \|u_r\|).$$

We will show that the family $\{u_r\}_{r\geqslant r^*} \subseteq X$ is bounded. We may assume that all $u_r \neq 0$, $r \geqslant r^*$. We have

$$\frac{\langle A(u_r), u_r - v_0 \rangle}{\|u_r\|} \leqslant \|h^*\|_* \left(\frac{\|v_0\|}{\|u_r\|} + 1 \right). \tag{4.85}$$

Suppose we can find $r_n \to +\infty$ such that $\|u_{r_n}\| \to +\infty$. Then using (4.84), from (4.85) we have a contradiction. Therefore we can find $M > 0$ such that

$$\|u_r\| \leqslant M \text{ for all } r \geqslant r^*.$$

Let $r = M + 1$. For every $y \in C$ and $t \in [0, 1]$, $ty + (1 - t)u_r \in C$. Moreover,

$$\|ty + (1 - t)u_r\| \leqslant t\|y\| + (1 - t)\|u_r\| \leqslant t\|y\| + M.$$

Let $t \in \left(0, \frac{1}{\|y\|}\right]$. Then $ty + (1 - t)u_r \in C_r$ (recall $r = M + 1$). Hence we have

$$\begin{aligned} &\langle A(u_r) - h^*, ty + (1 - t)u_r - u_r \rangle \geqslant 0 \\ \Rightarrow\ & t\langle A(u_r) - h^*, y - u_r \rangle \geqslant 0 \\ \Rightarrow\ & \langle A(u_r) - h^*, y - u_r \rangle \geqslant 0 \text{ for all } y \in C. \end{aligned}$$

The proof is now complete. □

The direct method of the calculus of variations is based on the so-called Weierstrass–Tonelli theorem, which asserts that a lower semicontinuous functional on a compact space attains its infimum. In infinite-dimensional spaces, it is more difficult to satisfy the compactness and lower semicontinuity conditions of the theorem and so the existence of a minimizer may fail. We can find only approximate minimizers. Nevertheless, some interesting things can be said even in this case. More precisely, if $\varphi(u_0)$ is an approximate minimum value of the lower semicontinuous functional φ, then a small Lipschitz perturbation of φ attains a strict minimum at a point $\hat{u}$ relatively close to u_0 (that is, there is a Lipschitz function ψ with small Lipschitz constant such that $\varphi + \psi$ has a strict minimum at $\hat{u}$). This is the essence of the so-called "Ekeland variational principle" which since its appearance (see Ekeland [157]) has found numerous applications. Also, it turns out that it is equivalent to some other interesting results of Nonlinear Analysis.

Theorem 4.6.14 *(Ekeland): If (X, d) is a complete metric space, $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is a lower semicontinuous function which is bounded below, $\varepsilon > 0$ and $u_0 \in X$ satisfies*

$$\varphi(u_0) \leqslant \inf_X \varphi + \varepsilon,$$

then given $\lambda > 0$, we can find $u_\lambda \in X$ such that

(a) $\varphi(u_\lambda) \leqslant \varphi(u_0)$ *and* $d(u_\lambda, u_0) \leqslant \lambda$*;*
(b) $\varphi(u_\lambda) < \varphi(u) + \frac{\varepsilon}{\lambda} d(u_\lambda, u)$ *for all* $u \neq u_\lambda$.

Proof By replacing φ with $\frac{1}{\varepsilon}\varphi$ and the metric d by $\frac{1}{\lambda}d$, without any loss of generality, we may assume that $\varepsilon = \lambda = 1$.

We introduce the following relation on X:

$$\text{“}v \preccurlyeq u \text{ if and only if } \varphi(v) \leqslant \varphi(u) - d(v, u).\text{”} \tag{4.86}$$

Evidently, $\preccurlyeq$ is reflexive (that is, $u \preccurlyeq u$). Also, if $v \preccurlyeq u$ and $u \preccurlyeq y$, we have

$$\varphi(v) \leqslant \varphi(u) - d(v, u) \text{ and } \varphi(u) \leqslant \varphi(y) - d(u, y) \text{ (see (4.86))}. \tag{4.87}$$

Then

$$\begin{aligned}
\varphi(v) &\leqslant \varphi(y) - (d(v, u) + d(u, y)) \text{ (see (4.87))} \\
&\leqslant \varphi(y) - d(v, y) \text{ (by the triangle inequality)} \\
\Rightarrow v &\preccurlyeq y \text{ (see (4.86))}.
\end{aligned}$$

This proves that the relation $\preccurlyeq$ is transitive. Finally suppose that $v \preccurlyeq u$ and $u \preccurlyeq v$. Then

$$\begin{aligned}
&\varphi(v) \leqslant \varphi(u) - d(v, u) \text{ and } \varphi(u) \leqslant \varphi(v) - d(u, v) \text{ (see (4.86))} \\
\Rightarrow\ & d(v, u) = 0 \text{ and so } v = u.
\end{aligned}$$

Therefore $\preccurlyeq$ is antisymmetric. Hence we have established that $\preccurlyeq$ is a partial order (that is, it is reflexive, transitive and antisymmetric).

Now let $u_1 = u_0$ and define

$$\begin{aligned}
&S_1 = \{u \in X : u \preccurlyeq u_1\} \\
\text{and} \quad & u_2 \in S_1 \text{ such that } \varphi(u_2) \leqslant \inf_{S_1} \varphi + \frac{1}{2^2}.
\end{aligned}$$

Then inductively we define

$$\begin{aligned}
&S_n = \{u \in X : u \preccurlyeq u_n\} \\
\text{and} \quad & u_{n+1} \in S_n \text{ such that } \varphi(u_{n+1}) \leqslant \inf_{S_n} \varphi + \frac{1}{2^{n+1}}.
\end{aligned} \tag{4.88}$$

Since $u_{n+1} \preccurlyeq u_n$, we see that $S_{n+1} \subseteq S_n$ for all $n \geqslant 1$. Also, the lower semicontinuity of φ and (4.86) imply that for each $n \geqslant 1$, S_n is closed. Let $u \in S_{n+1}$. Then $u \preccurlyeq u_{n+1} \preccurlyeq u_n$ and so

$$\begin{aligned}
d(u, u_n) &\leqslant \varphi(u_{n+1}) - \varphi(u) \text{ (see (4.86))} \\
&\leqslant \inf_{S_n} \varphi + \frac{1}{2^{n+1}} - \varphi(u) \\
&\leqslant \varphi(u) + \frac{1}{2^{n+1}} - \varphi(u) = \frac{1}{2^{n+1}} \\
\Rightarrow \operatorname{diam} S_{n+1} &\leqslant \frac{1}{2^{n+1}} \text{ for all } n \geqslant 0 \\
\Rightarrow \operatorname{diam} S_n &\to 0 \text{ as } n \to \infty.
\end{aligned}$$

Since (X, d) is complete, by Cantor's theorem, we have that

$$\bigcap_{n \geqslant 1} S_n = \{u_\lambda\}. \tag{4.89}$$

Then $u_\lambda \in S_1$ and so $u_\lambda \preccurlyeq u_1 = u_0$. Hence

$$\varphi(u_\lambda) \leqslant \varphi(u_0) - d(u_\lambda, u_0) \leqslant \varphi(u_0).$$

Also, we have

$$\begin{aligned}
d(u_\lambda, u_0) &\leqslant \varphi(u_0) - \varphi(u_\lambda) \\
&\leqslant \inf_X \varphi + 1 - \varphi(u_\lambda) \text{ (recall } \varepsilon = 1) \\
&\leqslant 1 = \lambda.
\end{aligned}$$

Therefore, we have proved statement (a).

To prove statement (b), we need to show that $v \preccurlyeq u_\lambda$ implies $v = u_\lambda$. Indeed, in this case we have $v \preccurlyeq u_n$ for all $n \geqslant 1$. It follows that $v \in \bigcap_{n \geqslant 1} S_n$ and so $v = u_\lambda$ (see (4.89)). ☐

Remark 4.6.15 The conclusions $d(u_\lambda, u_0) \leqslant \lambda$ and $\varphi(u_\lambda) < \varphi(u) + \frac{\varepsilon}{\lambda} d(u, u_\lambda)$ for $u \neq u_\lambda$ are complementary and the value of $\lambda > 0$ determines which one contains substantial information. Indeed, if $\lambda > 0$ is small, then the condition $d(u_\lambda, u_0) \leqslant \lambda$ contains important information since it tells us that u_λ is close to u_0, while the other condition provides little information since the perturbation of φ is big. On the other hand, if $\lambda > 0$ is big, then the conclusion $\varphi(u_\lambda) < \varphi(u) + \frac{\varepsilon}{\lambda} d(u, u_\lambda)$ for all $u \neq u_\lambda$ says that u_λ is close to being a global minimizer of φ (the perturbation term $\frac{\varepsilon}{\lambda} d(u, u_\lambda)$ is small), while the conclusion $d(u_\lambda, u_0) \leqslant \lambda$ provides little information about the whereabouts of u_λ. Often we try to strike a balance between the two conclusions by taking $\lambda = \sqrt{\varepsilon}$, $\varepsilon > 0$. Another interesting case is when $\lambda = 1$ and $\varepsilon > 0$. This case implies that we are not interested in how u_λ is located with respect to u_0 and we want only to have u_λ very close to being a global minimizer of φ. We present both cases as corollaries of Theorem 4.6.14.

Corollary 4.6.16 *If (X, d) is a complete metric space and $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is a lower semicontinuous function which is bounded below, $\varepsilon > 0$ and $u_0 \in X$ satisfies*

$$\varphi(u_0) \leqslant \inf_X \varphi + \varepsilon,$$

then we can find $\overline{u} \in X$ such that

(a) $\varphi(\overline{u}) \leqslant \varphi(u_0)$ *and* $d(\overline{u}, u_0) \leqslant \sqrt{\varepsilon}$*;*
(b) $\varphi(\overline{u}) < \varphi(u) + \sqrt{\varepsilon} d(u, \overline{u})$ *for all* $u \neq \overline{u}$*.*

Corollary 4.6.17 *If (X, d) is a complete metric space, $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is a lower semicontinuous function which is bounded below, then for every $\varepsilon > 0$ we can find $u_\varepsilon \in X$ such that*

(a) $\varphi(u_\varepsilon) \leqslant \inf\limits_X \varphi + \varepsilon$*;*
(b) $\varphi(u_\varepsilon) < \varphi(u) + \varepsilon d(u, u_\varepsilon)$ *for all* $u \neq u_\varepsilon$*.*

By introducing more structure on X and φ, we obtain improved versions of Theorem 4.6.14.

Theorem 4.6.18 *If X is a Banach space and $\varphi : X \to \mathbb{R}$ is a lower semicontinuous function which is bounded below and Gâteaux differentiable, then for every $\varepsilon > 0$ we can find $u_\varepsilon \in X$ such that*

(a) $\varphi(u_\varepsilon) \leqslant \inf\limits_X \varphi + \varepsilon$*;*
(b) $\|\varphi'(u_\varepsilon)\|_* \leqslant \varepsilon$*.*

Proof From Corollary 4.6.17, we know that there exists a $u_\varepsilon \in X$ such that $\varphi(u_\varepsilon) \leqslant \inf\limits_X \varphi + \varepsilon$ and

$$\begin{aligned}
&\varphi(u_\varepsilon) - \varphi(u) \leqslant \varepsilon \|u - u_\varepsilon\| \text{ for all } u \in X \\
\Rightarrow\ &\frac{1}{\lambda} [\varphi(u_\varepsilon) - \varphi(u_\varepsilon + \lambda h)] \leqslant \varepsilon \|h\| \text{ for all } \lambda > 0 \text{ and all } h \in X \\
\Rightarrow\ &-\langle \varphi'(u_\varepsilon), h \rangle \leqslant \varepsilon \|h\| \text{ for all } h \in X \\
\Rightarrow\ &|\langle \varphi'(u_\varepsilon), h \rangle| \leqslant \varepsilon \|h\| \text{ for all } h \in X \\
\Rightarrow\ &\|\varphi'(u_\varepsilon)\|_* \leqslant \varepsilon.
\end{aligned}$$

The proof is now complete. □

Remark 4.6.19 In general, the Gâteaux differentiability of φ is not enough to conclude that it is lower semicontinuous.

Next, we present some interesting applications of the Ekeland variational principle. We start with the Caristi fixed point theorem.

Theorem 4.6.20 *If (X, d) is a complete metric space, $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is a lower semicontinuous function which is bounded below and $F : X \to 2^X \setminus \{\emptyset\}$ is a multifunction such that*

$$\varphi(y) \leqslant \varphi(u) - d(u, y) \text{ for all } u \in X \text{ and some } y \in F(u), \tag{4.90}$$

then $F(\cdot)$ has a fixed point $\hat{u} \in X$.

Proof Corollary 4.6.17 with $\varepsilon = 1$ implies that there exists a $\hat{u} \in X$ such that

$$\varphi(\hat{u}) < \varphi(u) + d(u, \hat{u}) \text{ for all } u \neq \hat{u}. \tag{4.91}$$

Suppose that $\hat{u} \notin F(\hat{u})$. Then for all $y \in F(\hat{u})$ we have that $y \neq \hat{u}$. Let $y \in F(\hat{u})$ be as postulated by hypothesis (4.90). We have

$$\begin{aligned} &\varphi(y) \leqslant \varphi(\hat{u}) - d(\hat{u}, y) \text{ and } \varphi(\hat{u}) < \varphi(y) + d(y, \hat{u}) \text{ (see (4.91))} \\ \Rightarrow\ & d(\hat{u}, y) < d(\hat{u}, y), \text{ a contradiction.} \end{aligned}$$

So, $F(\cdot)$ has a fixed point $\hat{u} \in X$ (that is, $\hat{u} \in F(\hat{u})$). □

Remark 4.6.21 We stress that in the above theorem, we did not assume anything about the multifunction $F(\cdot)$. If $F : X \to X$ is a single-valued contraction (see Definition 4.2.1(b)) with contraction constant $k \in (0, 1)$ and $\varphi(u) = \frac{1}{1-k} d(u, F(u))$ for all $u \in X$, then for all $u \in X$ we have

$$\begin{aligned} \varphi(u) - \varphi(F(u)) &= \frac{1}{1-k}\left[d(u, F(u)) - d(F(u), F(F(u)))\right] \\ &\geqslant \frac{1}{1-k}\left[d(u, F(u)) - kd(u, F(u))\right] \\ &= d(u, F(u)). \end{aligned}$$

So, we have satisfied condition (4.90) and we can apply Theorem 4.6.20 to conclude that there exists a $\hat{u} \in X$ such that $\hat{u} = F(\hat{u})$. Hence we have derived the Banach fixed point theorem (see Theorem 4.2.3). Nevertheless, the fixed point principle of Banach contains much more information (see Sect. 4.2).

In the proof of Theorem 4.6.20 (Caristi's fixed point theorem), we have used the Ekeland variational principle (in particular, Corollary 4.6.17). Conversely we can show that Caristi's fixed point theorem implies conclusion (b) of Corollary 4.6.17 (that is, the existence of a strict minimizer for a small perturbation of φ).

Proposition 4.6.22 *Conclusion (b) of Corollary 4.6.17 can be derived from Theorem 4.6.20.*

Proof Let $d_1 = \varepsilon d$. This is an equivalent metric on X. Arguing by contradiction, suppose that there is no $u_\varepsilon \in X$ satisfying conclusion (b) of Corollary 4.6.17. For each $u \in X$, the set $F(u) = \{y \in X : \varphi(u) \geqslant \varphi(y) + d_1(u, y), y \neq u\}$ is nonempty and condition (4.90) is satisfied. Invoking Theorem 4.6.20, we find $\hat{u} \in X$ such that $\hat{u} \in F(\hat{u})$, a contradiction to the definition of $F(\cdot)$. □

Another geometrical result of nonlinear analysis, which is closely related to the Ekeland variational principle, is the so-called "Drop theorem". First a definition.

Definition 4.6.23 Let X be a normed space, $C \subseteq X$ a nonempty convex set and $u \in X$. The "drop" associated to the pair (C, u), denoted by $D(C, u)$, is the convex hull of $C \cup \{u\}$, that is,

$$D(C, u) = \{u + \lambda(y - u) : y \in C \text{ and } \lambda \in [0, 1]\}.$$

Remark 4.6.24 The set $D(C, u)$ is called a "drop", in view of its evocative geometry.

Theorem 4.6.25 *If X is a normed space, $A \subseteq X$ is a complete subset, $y \in X \setminus A$, $R = d(y, A)$ and $0 < r < R < \rho$, then there exists a $u_0 \in A$ such that*

$$\begin{aligned} &u_0 \in \overline{B}_\rho(y) = \{v \in X : \|v - y\| \leqslant \rho\}, \\ &D(\overline{B}_r(y), u_0) \cap A = \{u_0\}. \end{aligned}$$

Proof By translating things if necessary, we may assume without any loss of generality that $y = 0$. Let $S = \overline{B}_\rho(0) \cap A$. This is a closed subset of A, hence a complete metric space with the metric induced by the norm of X. Let $\varphi : S \to \mathbb{R}_+$ be the continuous function defined by

$$\varphi(u) = \frac{\rho + r}{R - r}\|u\|.$$

We apply Corollary 4.6.17 with $\varepsilon = 1$ and obtain $u_0 \in S$ such that

$$\varphi(u_0) < \varphi(u) + \|u - u_0\| \text{ for all } u \in S, u \neq u_0. \tag{4.92}$$

We will show that $D(\overline{B}_r(0), u_0) \cap A = \{u_0\}$ and this will finish the proof. Arguing by contradiction, suppose that $v \in D(\overline{B}_r(0), u_0) \cap A,\ v \neq u_0$. Then

$$v \in A \text{ and } v = (1 - \lambda)u_0 + \lambda w \text{ with } w \in \overline{B}_r(0),\ \lambda \in [0, 1].$$

Since $v \neq u_0$ and $r < R$, we see that $\lambda \in (0, 1)$. We have

$$\|v\| \leqslant (1 - \lambda)\|u_0\| + \lambda\|w\|. \tag{4.93}$$

Because $u_0 \in A$, we have $\|u_0\| \geqslant R$ and so it follows that

$$\lambda(R - r) \leqslant \lambda\left(\|u_0\| - \|w\|\right) \leqslant \|u_0\| - \|v\| \text{ (see (4.93))}. \tag{4.94}$$

In (4.92) we choose $u = v$ and have

$$\begin{aligned}
\frac{\rho+r}{R-r}\|u_0\| &< \frac{\rho+r}{R-r}\|v\| + \|v-u_0\| \\
&= \frac{\rho+r}{R-r}\|v\| + \lambda\|u_0-w\| \\
&\Rightarrow \frac{\rho+r}{R-r}\,(\|u_0\|-\|v\|) < \lambda\|u_0-w\| \\
&\Rightarrow \rho+r < \|u_0-w\| \text{ (see (4.94))}.
\end{aligned} \tag{4.95}$$

But we know that $\|u_0\| \leqslant \rho$ (recall $u_0 \in S$) and $w \in \overline{B}_r(0)$. Therefore

$$\|u_0-w\| \leqslant \rho+r. \tag{4.96}$$

Comparing (4.95) and (4.96) we reach a contradiction. □

Remark 4.6.26 In fact it can be shown that Theorem 4.6.25 (the drop theorem) is equivalent to the Ekeland variational principle in the form of Corollary 4.6.17. For the details, we refer to Penot [333].

Continuing with the applications of the Ekeland variational principle, we have the following result for lower semicontinuous, Gâteaux differentiable functionals.

Proposition 4.6.27 *If X is a Banach space, $\varphi : X \to \mathbb{R}$ is a lower semicontinuous and Gâteaux differentiable functional and there exist $a, c > 0$ such that*

$$\varphi(u) \geqslant a\|u\| - c \textit{ for all } u \in X, \tag{4.97}$$

then $\varphi'(X)$ is dense in $a\overline{B}_1^*$, where $\overline{B}_1^* = \{u^* \in X^* : \|u^*\|_* \leqslant 1\}$.

Proof Let $u^* \in a\overline{B}_1^*$ and consider the functional $\psi : X \to \mathbb{R}$ defined by

$$\psi(u) = \varphi(u) - \langle u^*, u\rangle.$$

As always by $\langle\cdot,\cdot\rangle$ we denote the duality brackets for the pair (X^*, X). Evidently, ψ is lower semicontinuous and Gâteaux differentiable. Also using (4.97) we have

$$\psi(u) \geqslant a\|u\| - c - \|u^*\|_*\|u\| \geqslant -c.$$

So, ψ is bounded below and we can apply Theorem 4.6.18 and find $u_\varepsilon \in X$ such that

$$\begin{aligned}
&\|\psi'(u_\varepsilon)\|_* \leqslant \varepsilon \\
\Rightarrow\ &\|\varphi'(u_\varepsilon) - u^*\|_* \leqslant \varepsilon \\
\Rightarrow\ &\varphi'(X) \text{ is dense in } a\overline{B}_1^*.
\end{aligned}$$

The proof is now complete. □

Corollary 4.6.28 *If X is a Banach space, $\varphi : X \to \mathbb{R}$ is a lower semicontinuous and Gâteaux differentiable functional and there exists a continuous function $\vartheta : \mathbb{R}_+ = [0, \infty) \to \mathbb{R}$ such that $\frac{\vartheta(t)}{t} \to +\infty$ as $t \to +\infty$ and*

$$\varphi(u) \geqslant \vartheta(\|u\|) \text{ for all } u \in X,$$

then $\varphi'(X)$ is dense in X^*.

Proof From the superlinearity of ϑ, we see that given $a > 0$, we can find $t_0 > 0$ such that $\vartheta(t) \geqslant at$ for all $t \geqslant t_0$, hence $\varphi(u) \geqslant a\|u\|$ for all $\|u\| \geqslant t_0$. Also let

$$m = \inf\{\vartheta(t) : t \in [0, t_0]\} \geqslant 0.$$

Then $\varphi(u) \geqslant m$ for all $\|u\| \leqslant t_0$. So, if $c > m$, we have

$$\varphi(u) \geqslant a\|u\| - c \text{ for all } u \in X$$

and applying Proposition 4.6.27 we infer that $\varphi'(X)$ is dense in $a\overline{B}_1^*$. Since $a > 0$ is arbitrary, we conclude that $\varphi'(X)$ is dense in X^*. □

Anticipating a notion that plays a central role in Chap. 5, we introduce the following definition.

Definition 4.6.29 Let X be a Banach space and $\varphi \in C^1(X)$. We say that φ satisfies the Palais–Smale condition (PS-condition for short) if the following holds: "every sequence $\{u_n\}_{n\geqslant 1} \subseteq X$ such that

$$\{\varphi(u_n)\}_{n\geqslant 1} \subseteq \mathbb{R} \text{ is bounded and } \varphi'(u_n) \to 0 \text{ in } X^* \text{ as } n \to \infty$$

admits a strongly convergent subsequence."

Proposition 4.6.30 *If X is a Banach space and $\varphi \in C^1(X)$ satisfies the PS-condition and it is bounded below, then there exists a $u_0 \in X$ such that $\varphi(u_0) = \inf_X \varphi$.*

Proof By virtue of Theorem 4.6.18 we can find $\{u_n\}_{n\geqslant 1} \subseteq X$ such that

$$\varphi(u_n) \leqslant \inf_X \varphi + \frac{1}{n} \text{ and } \|\varphi'(u_n)\|_* \leqslant \frac{1}{n}.$$

Since φ satisfies the PS-condition, by passing to a suitable subsequence if necessary, we may assume that $u_n \to u_0$ in X. Then $\varphi(u_0) = \inf_X \varphi$. □

Remark 4.6.31 In fact we do not need the full strength of the hypothesis $\varphi \in C^1(X)$. It is enough to assume that φ is Fréchet differentiable. Note that the PS-condition still makes sense. In this case, we apply Corollary 4.6.17 and obtain $\{u_n\}_{n\geqslant 1} \subseteq X$ such that

$$\varphi(u_n) \leqslant \inf_X \varphi + \frac{1}{n} \text{ and } \varphi(u_n) \leqslant \varphi(u) + \frac{1}{n}\|u - u_n\| \text{ for all } u \in X.$$

Let $u = u_n + \lambda h$ with $\lambda > 0$, $h \in X$. Then

$$\begin{aligned} &\frac{1}{\lambda}\left[\varphi(u_n) - \varphi(u_n + \lambda h)\right] \leqslant \frac{1}{n}\|h\| \\ &\Rightarrow -\langle \varphi'(u_n), h\rangle \leqslant \frac{1}{n}\|h\| \\ &\Rightarrow \|\varphi'(u_n)\|_* \leqslant \frac{1}{n} \text{ (since } h \in X \text{ is arbitrary).} \end{aligned}$$

Again we apply the PS-condition to reach the desired conclusion.

Using the Ekeland variational principle, we can also show that for a lower semicontinuous, convex functional $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ (as always not identically $+\infty$), the domain of the subdifferential $\partial\varphi$, $\operatorname{dom}\partial\varphi = \{u \in X : \partial\varphi(u) \neq \emptyset\}$ is dense in the effective domain of φ, $\operatorname{dom}\varphi = \{u \in X : \varphi(u) < +\infty\}$ (see also Corollary 2.7.12). More precisely we have:

Proposition 4.6.32 *If X is a Banach space and $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is a lower semicontinuous convex functional not identically $+\infty$, then for any $\hat{u} \in \operatorname{dom}\varphi$, we can find a sequence $\{u_n\}_{n\geqslant 1} \subseteq X$ such that*

$$\|u_n - \hat{u}\| \leqslant \frac{1}{n},\ \varphi(u_n) \to \varphi(\hat{u}) \text{ and } \partial\varphi(u_n) \neq \emptyset \text{ for all } n \geqslant 1.$$

Proof From Proposition 2.7.5, we know that φ admits a continuous affine minorant, that is, there exist $u^* \in X^*$ and $c \in \mathbb{R}$ such that $\varphi(u) \geqslant \langle u^*, u\rangle - c$ for all $u \in X$. We introduce the function $\psi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ defined by

$$\psi(u) = \varphi(u) - \langle u^*, u\rangle + c \geqslant 0 \text{ for all } u \in X.$$

Evidently, ψ is lower semicontinuous, bounded below and of course not identically $+\infty$. We apply Theorem 4.6.14 with $\varepsilon = \psi(\hat{u}) - \inf_X \psi$ and $\lambda = \frac{1}{n}$, $n \in \mathbb{N}$. We obtain a sequence $\{u_n\}_{n\geqslant 1} \subseteq X$ such that

$$\psi(u_n) \leqslant \psi(\hat{u}) \text{ and } \|u_n - \hat{u}\| \leqslant \frac{1}{n} \text{ for all } n \geqslant 1, \tag{4.98}$$

$$\psi(u_n) < \psi(u) + \varepsilon n\|u - u_n\| \text{ for all } u \neq u_n \text{ and all } n \geqslant 1. \tag{4.99}$$

Consider the functionals $\xi_n : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ defined by

$$\xi_n(u) = \psi(u) + \varepsilon n\|u - u_n\|,\ n \geqslant 1.$$

From (4.99) we see that u_n is the unique global minimizer of ξ_n. Hence

$$\begin{aligned} & 0 \in \partial\xi_n(u_n) \\ \Rightarrow\ & 0 = v_n^* + \varepsilon n y_n^* \text{ where } v_n^* \in \partial\psi_n(u_n),\ \ y_n^* \in \partial\mu_n(u_n), \end{aligned}$$

where $\mu_n(u) = \varepsilon n\|u - u_n\|$, $n \geqslant 1$ (see Proposition 2.7.20). Also, from the definition of ψ and Proposition 2.7.20, we have

$$v_n^* = u_n^* - u^* \text{ with } u_n^* \in \partial\varphi(u_n),\ \ n \geqslant 1.$$

Hence $\partial\varphi(u_n) \neq 0$ for all $n \geqslant 1$. From (4.98) we have

$$\varphi(u_n) \leqslant \varphi(\hat{u}) + \langle u^*, u_n - \hat{u}\rangle \text{ for all } n \geqslant 1. \tag{4.100}$$

Since $u_n \to \hat{u}$ in X (see (4.98)), passing to the limit as $n \to \infty$ in (4.100), we obtain

$$\limsup_{n\to\infty} \varphi(u_n) \leqslant \varphi(\hat{u}). \tag{4.101}$$

But φ is lower semicontinuous and $u_n \to \hat{u}$. Hence

$$\varphi(\hat{u}) \leqslant \liminf_{n\to\infty} \varphi(u_n). \tag{4.102}$$

From (4.101) and (4.102) we conclude that $\varphi(u_n) \to \varphi(\hat{u})$. $\square$

From linear operator theory, we recall the following basic result.

Proposition 4.6.33 *If X, Y are Banach spaces and $L \in \mathscr{L}(X, Y)$, then the following statements are equivalent:*

(a) L is surjective;
(b) there exists a $c > 0$ such that $\|y^*\|_{Y^*} \leqslant c\|L^*(y^*)\|_{X^*}$ for all $y^* \in Y^*$;
(c) $N(L^*) = \ker L^* = \{0\}$ and $R(L^*)$ is closed.

Remark 4.6.34 So according to this theorem, for $L \in \mathscr{L}(X, Y)$ we have

$$L \text{ is surjective } \Rightarrow\ L^* \text{ is injective.}$$

If one of the spaces X or Y is finite-dimensional, then the converse is also true. Also, note that a corresponding result is also true with the roles of L and L^* interchanged.

Next we will prove a nonlinear analog of Proposition 4.6.33.

Proposition 4.6.35 *If X, Y are Banach spaces, $\varphi : X \to Y$ is Gâteaux differentiable map, $\varphi(X)$ is closed in Y, $y \in Y$ and there exist $\rho_0 > 0$ and $k \in [0, 1)$ such that*

$$\varphi^{-1}(\overline{B}_{\rho_0}(y)) \neq \emptyset, \tag{4.103}$$

$$\inf\left\{\|y - \varphi(u) - v\|_Y : v \in \overline{R(\varphi'(u))}\right\} \leqslant k\|y - \varphi(u)\|_Y, \tag{4.104}$$

for all $u \in \varphi^{-1}(\overline{B}_{\rho_0}(y))$ *(see (4.103)), then* $y \in \varphi(X)$.

Proof Let $S = \varphi(X)$. By hypothesis S is closed. Arguing by contradiction, suppose that $y \notin \varphi(X) = S$ and let $R = d_Y(y, S)$. We choose $\rho, r > 0$ such that

$$r < R < \rho \text{ and } k\rho < r.$$

Note that (4.103), (4.104) still hold for any $\rho \in (R, \rho_0]$. Let $C = \overline{B}_r(y)$. Then invoking Theorem 4.6.25 (the Drop theorem), we can find $y_0 \in \overline{B}_r(y)$ such that

$$D(C, y_0) \cap S = \{y_0\}. \tag{4.105}$$

Let $u_0 \in X$ such that $\varphi(u_0) = y_0$. From (4.104) and since $k\rho < r$, we have

$$\inf\left[\|y - \varphi(u_0) - v\|_Y : v \in \overline{R(\varphi'(u_0))}\right] \leqslant k\|y - \varphi(u_0)\|_Y < r.$$

So, we can find $h \in X$ such that

$$\|y - \varphi(u_0) - \varphi'(u_0)h\|_Y < r. \tag{4.106}$$

Then for small $\lambda > 0$ we have

$$\|y - \varphi(u_0) - \frac{\varphi(u_0 + \lambda h) - \varphi(u_0)}{\lambda}\|_Y < r.$$

Let $z_\lambda = y - \varphi(u_0) - \frac{\varphi(u_0+\lambda h)-\varphi(u_0)}{\lambda} \in Y$. Then

$$\begin{aligned}
& y - z_\lambda \in D(C, y_0) \text{ (see Definition 4.6.23)} \\
\Rightarrow\ & (1-\lambda)y_0 + \lambda(y - z_\lambda) \in D(C, u_0) \text{ for } \lambda > 0 \text{ small} \\
\Rightarrow\ & \varphi(u_0 + \lambda h) \in D(C, y_0) \text{ for } \lambda > 0 \text{ small} \\
\Rightarrow\ & \varphi(u_0 + \lambda h) = y_0 \text{ for } \lambda > 0 \text{ small (see (4.105))} \\
\Rightarrow\ & \varphi'(u_0) = 0.
\end{aligned}$$

Then from (4.106), we have

$$\|y - \varphi(u_0)\|_Y < r < R,$$

a contradiction. ∎

Remark 4.6.36 If conditions (4.103), (4.104) hold for all $y \in Y$, we can conclude that φ is surjective. As in the case of the Ekeland variational principle (see Remark 4.6.15), conditions (4.103), (4.104) are complementary. That is, the larger $\rho > 0$, the more difficult it is to verify (4.104).

Corollary 4.6.37 *If X, Y are Banach spaces, $\varphi : X \to Y$ is Gâteaux differentiable, $\varphi(X)$ is closed in Y and $N(\varphi'_G(u)^*) = \{0\}$ for all $u \in X$, then φ is surjective.*

Proof From linear operator theory (see, for example, Denkowski, Migorski and Papageorgiou [143, p. 320]), we know that

$$\overline{R(\varphi'(u))} = N(\varphi'(u)^*)^{\perp}$$

(recall that if $V^* \subseteq Y^*$ is a subspace, $(V^*)^{\perp} = \{y \in Y : \langle v^*, y\rangle_{Y^*,Y} = 0$ for all $v^* \in V^*\}$). So, $\overline{R(\varphi'(u))} = Y$ and (4.104) holds for $k = 0$. Then for given $y \in Y$, let $\rho > 0$ be such that $d(y, \varphi(X)) < \rho$. Applying Remark 4.6.36 with this $\rho > 0$ and with $k = 0$, we obtain $y \in \varphi(X)$, hence φ is surjective. The proof is now complete. □

Remark 4.6.38 The above proof suggests that the condition $N(\varphi'_G(u)^*) = \{0\}$ for all $u \in X$ can be replaced by the hypothesis that $R(\varphi'(u))$ is dense in Y for all $u \in X$.

We will conclude this section with a result which is motivated by the proof of the Ekeland variational principle (see the proof of Theorem 4.6.14). So, we will formulate a variational principle for partially ordered spaces, from which we can deduce the Ekeland variational principle. Recall that a set X is partially ordered if there is a relation $\preccurlyeq$ which is reflexive (that is, $u \preccurlyeq u$ for all $u \in X$), antisymmetric (that is, $u \preccurlyeq v$ and $v \preccurlyeq u$ imply $u = v$) and transitive (that is, $u \preccurlyeq v$ and $v \preccurlyeq w$ imply $u \preccurlyeq w$). A partially ordered set is denoted by $(X, \preccurlyeq)$. A sequence $\{u_n\}_{n\geqslant 1}$ in $(X, \preccurlyeq)$ is said to be increasing if $u_n \preccurlyeq u_{n+1}$ for all $n \geqslant 1$ and bounded above if there exists a $\hat{u} \in X$ such that $u_n \preccurlyeq \hat{u}$ for all $n \geqslant 1$. A function $\varphi : (X, \preccurlyeq) \to \mathbb{R}$ is increasing if $u \preccurlyeq v$ implies $\varphi(u) \leqslant \varphi(v)$.

Theorem 4.6.39 *If $(X, \preccurlyeq)$ is a partially ordered set in which every increasing sequence is bounded above and $\varphi : (X, \preccurlyeq) \to \mathbb{R}$ is an increasing function which is bounded above, then there exists a $\hat{u} \in X$ such that $\hat{u} \preccurlyeq u$ implies $\varphi(\hat{u}) = \varphi(u)$.*

Proof Let $u_1 \in X$. Using induction we will generate an increasing sequence $\{u_n\}_{n\geqslant 1} \subseteq X$. Suppose we have produced $u_n \in X$. We set

$$C_n = \{u \in X : u_n \preccurlyeq u\} \text{ and } M_n = \sup_{C_n} \varphi.$$

If for u_n, we have that $u_n \preccurlyeq u$ implies $\varphi(u_n) = \varphi(u)$, then we are done. Otherwise $\varphi(u_n) < M_n$ and so we can find $u_{n+1} \in C_n$ such that

$$M_n \leqslant \varphi(u_n) + \frac{1}{2}\left[M_n - \varphi(u_n)\right]. \tag{4.107}$$

Then by induction we have an increasing sequence $\{u_n\}_{n\geqslant 1} \subseteq X$ and it satisfies (4.107). By hypothesis we can find $\hat{u} \in X$ such that $u_n \preccurlyeq \hat{u}$ for all $n \geqslant 1$. We claim that $\hat{u} \in X$ is the desired solution. If this is not true, then we can find $u \in X$ with $\hat{u} \preccurlyeq u$ and $\varphi(\hat{u}) < \varphi(u)$. The sequence $\{\varphi(u_n)\}_{n\geqslant 1} \subseteq \mathbb{R}$ is increasing and bounded by $\varphi(u)$. So, it converges and we have

$$\lim_{n\to\infty} \varphi(u_n) \leqslant \varphi(\hat{u}). \tag{4.108}$$

Since $u_n \preccurlyeq \hat{u}$ and $\hat{u} \preccurlyeq y$, by transitivity we have $u_n \preccurlyeq y$ for all $n \geqslant 1$ and so $y \in C_n$ for all $n \geqslant 1$. Then from (4.107) it follows that

$$\begin{aligned} &\varphi(y) \leqslant M_n \leqslant 2\varphi(u_{n+1}) - \varphi(u_n) \text{ for all } n \geqslant 1 \\ \Rightarrow\ &\varphi(y) \leqslant \varphi(\hat{u}), \text{ a contradiction.} \end{aligned}$$

So, indeed $\hat{u} \in X$ is the desired solution. The proof is now complete. ☐

Remark 4.6.40 We can give Theorem 4.6.39 a physical interpretation. Think of φ as measuring the entropy of a system. Then the theorem guarantees the existence of a state of maximal entropy. To these states correspond stable equilibria of the system.

Corollary 4.6.41 *If X is a Hausdorff topological space equipped with a partial order $\preccurlyeq$ and $\psi : X \to \mathbb{R}$ is a function which is bounded below and*

(i) for every $u \in X$, the set $\{v \in X : u \preccurlyeq v\}$ is closed;
(ii) $u \preccurlyeq v$, $u \neq v$ imply $\psi(v) < \psi(u)$ (ψ is strictly decreasing);
(iii) any increasing sequence in $(X, \preccurlyeq)$ is relatively compact,

then for each $u \in X$ we can find $\hat{u} \in X$ such that $u \preccurlyeq \hat{u}$ and $\hat{u}$ is maximal.

Proof Let $\{u_n\}_{n\geqslant 1} \subseteq (X, \preccurlyeq)$ be an increasing sequence. By virtue of hypothesis (iii), $\{u_n\}_{n\geqslant 1} \subseteq X$ is compact and so we can find a subsequence $\{u_{n_k}\}_{k\geqslant 1}$ of $\{u_n\}_{n\geqslant 1}$ such that $u_{n_k} \to v$ in X. We claim that $u_n \preccurlyeq v$ for all $n \geqslant 1$. Indeed, given $n \geqslant 1$ we have $n \leqslant n_k$ for all $k \geqslant k_n$. So, $u_n \leqslant u_{n_k}$ for all $k \geqslant k_n$. Invoking hypothesis (i) we have $u_n \preccurlyeq v$ for all $n \geqslant 1$. Taking $\varphi = -\psi$, we see that because of hypothesis (ii) we can apply Theorem 4.6.39 starting from $u_1 = u$ and obtain $\hat{u} \in X$ such that $u \preccurlyeq \hat{u}$ and $\hat{u}$ is maximal. ☐

Finally, we show that conclusion (b) in Corollary 4.6.17 can be derived from Theorem 4.6.39.

Proposition 4.6.42 *Theorem 4.6.39 implies conclusion (b) of Corollary 4.6.17.*

Proof Without any loss of generality, we may take $\varepsilon = 1$. Consider the following partial order $\preccurlyeq$ on the complete metric space (X, d)

$$u \preccurlyeq v \text{ if and only if } \varphi(v) - \varphi(u) \leqslant -d(v, u). \tag{4.109}$$

For any increasing sequence $\{u_n\}_{n\geqslant 1} \subseteq (X, \preccurlyeq)$, $\{\varphi(u_n)\}_{n\geqslant 1} \subseteq \mathbb{R}$ is decreasing and bounded below. So, it converges. Also, again from (4.109), we infer that $\{u_n\}_{n\geqslant 1} \subseteq X$ is Cauchy. The completeness of X implies that $u_n \to u$ in (X, d). Therefore, we can apply Corollary 4.6.41 to obtain conclusion (b) of Corollary 4.6.17. $\square$

4.7 Young Measures

From the theory of L^p-spaces, we know that if (Ω, Σ, μ) is a finite measure space and $\{u_n\}_{n\geqslant 1} \subseteq L^1(\Omega)$, then the following is true:

$$\text{“}u_n \xrightarrow{w} u \text{ in } L^1(\Omega) \text{ and } u(z) \leqslant \liminf_{n\to\infty} u_n(z)\ \mu\text{-a.e. in } \Omega \Rightarrow u_n \to u \text{ in } L^1(\Omega).\text{”}$$

This result illustrates the difference between weak and strong convergence in $L^1(\Omega)$. A sequence $\{u_n\}_{n\geqslant 1} \subseteq L^1(\Omega)$ which converges weakly but not strongly oscillates rapidly around its weak limit. However, in the limit all this information about the oscillation is lost and only a mean value is recorded. This is not helpful, because if we consider a Nemitsky (superposition) operator N_f (that is, if $N_f(y)(\cdot) = f(y(\cdot))$), we cannot say that $N_f(u_n) \xrightarrow{w} N_f(u)$ in $L^1(\Omega)$, unless f is an affine function. To recover this lost information, we embed the sequence $\{u_n\}_{n\geqslant 1}$ into a larger space and consider the limit there. This larger space is that of parametrized measures (probability-valued functions). These are the Young measures. This idea is present in many applications (such as stochastic analysis, optimal control, game theory, etc.) and for this reason Young measures appear in the literature under different names, such as Markov kernels, relaxed controls, mixed strategies, etc.

Our setting is the following: (Ω, Σ, μ) is a complete finite measure space with Σ countably generated and (X, d) is a locally compact, σ-compact complete metric space (for example, think of $\mathbb{R}^N$). Let us remark that much of what is done in this section extends without any difficulty to a σ-finite measure space. In what follows by $B(X)$ we denote the Borel σ-field of X and by $\Sigma \times B(X)$ the product σ-field of Σ and $B(X)$. By M_b we will denote the vector space of real bounded measures on some space (for example, $M_b(\Omega \times X)$ is the space of bounded real measures on $\Omega \times X$). Also by M_1^+ we denote the probability measures on a space and by SM_1^+ the subprobability measures on a space (for example, $SM_+^1(X) = \{\mu \in M_b(X) : \mu \geqslant 0, \mu(X) \leqslant 1\}$). Recall that a Radon measure on X is a Borel measure μ (that is, a measure defined on $B(X)$) such that for every $A \in B(X)$ and $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subseteq X$ such that $|\mu|(A \setminus K_\varepsilon) < \varepsilon$. Here by $|\mu|$ we denote the total variation of μ. From topological measure theory, we know that every Radon measure on $B(X)$ is regular, that for every $A \in B(X)$ and every $\varepsilon > 0$, we can find $U \subseteq X$ open and $C \subseteq X$ closed such that $C \subseteq A \subseteq U$ and $|\mu|(U \setminus C) < \varepsilon$. We denote the space of Radon measures on X by $M_r(X)$ and we have $M_r(X) = M_b(X)$. By $C_c(X)$ we denote the space of continuous functions on X with compact support, by $C_0(X)$ the space of continuous functions on X which vanish at infinity, that is, $\varphi \in C_0(X)$, if for every $\varepsilon > 0$, there exists a compact $K_\varepsilon \subseteq X$ such that $|\varphi(u)| \leqslant \varepsilon$ for all $u \in X \setminus K_\varepsilon$ and finally by $C_b(X)$ we denote the space of bounded continuous functions on X. We have the following inclusions

$$C_c(X) \subseteq C_0(X) \subseteq C_b(X).$$

If X is compact, then the three spaces coincide. If X is not compact, then each of the above inclusions is strict. We endow $C_b(X)$ with the supremum norm

$$\|\varphi\|_\infty = \sup\,[|\varphi(u)| : u \in X] \text{ for all } \varphi \in C_b(X).$$

This norm is inherited by the subspaces $C_0(X)$ and $C_c(X)$. We have

- $(C_b(X), \|\cdot\|_\infty)$ is a Banach space;
- $C_0(X)$ is a $\|\cdot\|_\infty$- closed subspace of $C_b(X)$ and so $(C_0(X), \|\cdot\|_\infty)$ is a Banach space too;
- $C_c(X)$ is $\|\cdot\|_\infty$- dense in $C_0(X)$.

From the Riesz representation theorem we have that

$$C_0(X)^* = M_r(X). \tag{4.110}$$

We consider the linear functionals $l_\varphi : M_r(X) \to \mathbb{R}$ defined by

$$l_\varphi(\mu) = \int_X \varphi(u)\mu(du) \text{ for all } \mu \in M_r(X),$$

with $\varphi \in C_0(X)$ or $\varphi \in C_b(X)$. Then

- $w(M_r(X), C_0(X))$ is the usual w^*-topology on the Banach space $M_r(X)$;
- $w(M_r(X), C_b(X))$ is the so-called narrow topology on the Banach space $M_r(X)$.

Remark 4.7.1 Probabilists prefer to call convergence in the w^*-topology $w(M_r(X), C_0(X))$ weak convergence and not w^*-convergence.

Definition 4.7.2 A transition probability (resp. transition subprobability) on Ω is a map $\lambda : \Omega \to M_1^+(X)$ (resp. $\lambda : \Omega \to SM_1^+(X)$) which is measurable in the following sense: for every $A \in B(X)$, $z \to \lambda(z)(A)$ from Ω into $\mathbb{R}$ is Σ-measurable. By $R(\Omega, X)$ (resp. $SR(\Omega, X)$) we denote the space of all transition probabilities (resp. transition subprobabilities) on Ω.

In the sequel we will restrict ourselves to transition probabilities, although the results are also valid for transition subprobabilities.

Proposition 4.7.3 *$\lambda \in R(\Omega, X)$ if and only if $\lambda : \Omega \to M_1^+(X)_{w^*}$ is $(\Sigma, B(M_1^+(X)_{w^*}))$-measurable where $M_1^+(X)_{w^*}$ denotes the space $M_1^+(X)$ furnished with the weak* topology.*

Proof $\Rightarrow$: Evidently, $z \mapsto \int_X \varphi(u)\lambda(z)(du)$ is measurable for every characteristic function φ of an element in $B(X)$. Hence it is also measurable for every simple function φ. Then by density it is true for every bounded Borel function φ, in particular for every $\varphi \in C_0(X)$.

$\Leftarrow$: Let $\hat{X}$ be the Alexandroff one point compactification of X. Since X is separable (being σ-compact), $\hat{X}$ is metrizable. Let $\hat{d}$ be a compatible metric. Let $C \subseteq X$ be compact and set $\varphi_n(u) = \left(1 - n\hat{d}(u, C)\right)^+$. Then $\varphi_n \in C(\hat{X})$ and $\varphi_n(\infty) = 0$ for all $n \geqslant n_0$. Hence $\varphi_n\Big|_X \in C_0(X)$. Moreover, $\varphi_n \downarrow \chi_C$ and so by the monotone convergence theorem $\int_X \varphi_n(u)\lambda(z)(du) \downarrow \lambda(z)(C)$. By hypothesis $z \to \int_X \varphi_n(u)\lambda(z)(du)$ is Σ-measurable for all $n \geqslant n_0$, hence $z \to \lambda(z)(C)$ is Σ-measurable. Let

$$\mathscr{D} = \{A \in B(X) : z \to \lambda(z)(A) \text{ is } \Sigma\text{-measurable}\}.$$

Clearly, $\mathscr{D}$ is a Dynkin class and it contains all compact sets in X. So, by the Dynkin class theorem (see, for example, Denkowski et al. [143, p. 220]), $\mathscr{D} = B(X)$. Therefore $\lambda \in R(\Omega, X)$. $\square$

To introduce Young measures we will need the notion of the image of a measure.

Definition 4.7.4 Let (Ω_k, Σ_k), $k = 1, 2$, be two measurable spaces, $\xi : \Omega_1 \to \Omega_2$ a measurable map and λ a measure on (Ω_1, Σ_1), then the image of λ by ξ is defined by

$$\hat{\lambda} = \lambda \circ \xi^{-1},$$

that is, for all $A \in \Sigma_2$, $\hat{\lambda}(A) = \lambda(\xi^{-1}(A))$.

Remark 4.7.5 If $\varphi : \Omega_2 \to \mathbb{R}$ is $\hat{\lambda}$-integrable, then

$$\int_{\Omega_2} \varphi d\hat{\lambda} = \int_{\Omega_1} (\varphi \circ \xi) d\lambda_1.$$

Also, if (Ω_3, Σ_3) is a third measurable space and $\vartheta : \Omega_2 \to \Omega_3$ is measurable, then the image of λ by $\vartheta \circ \xi$ equals the image of $\hat{\lambda}$ by ϑ.

Now we can define Young measures. By $\Pi_\Omega : \Omega \times X \to \Omega$ we denote the projection $\Pi_\Omega(z, u) = z$ and by $\Pi_X : \Omega \times X \to X$ the projection $\Pi_X(z, u) = u$.

Definition 4.7.6 (a) The space of Young measures with respect to μ, denoted by $\mathscr{Y}(\Omega, \mu; X)$ is the set of all positive measures λ on $\Omega \times X$ such that $\mu = \lambda \circ \Pi_\Omega^{-1}$ (that is, μ is the image of λ under the projection map Π_Ω and so for every $A \in \Sigma$, $\lambda(A \times X) = \mu(A)$). The space of sub-Young measures with respect to μ, denoted by $S\mathscr{Y}(\Omega, \mu; X)$, is the set of all positive measures (including the zero measure), whose projection (marginal) on Ω is $\leqslant \mu$ (that is, $\lambda \circ \Pi_\Omega^{-1} \leqslant \mu$ and so for all $A \in \Sigma$, $\lambda(A \times X) \leqslant \mu(A)$).

(b) Let $u : \Omega \to X$ be a Σ-measurable map. The Young measure associated to u is the image λ of μ under the map $z \to (z, u(z))$.

Remark 4.7.7 If $\varphi : \Omega \times \mathbb{R} \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is either positive or λ-integrable, then

$$\int_{\Omega \times X} \varphi(z, u)\lambda(dz, du) = \int_{\Omega} \varphi(z, u(z))d\mu.$$

Moreover, λ is the unique measure on $\Omega \times X$ supported by graph u and such that for all $A \in \Sigma$, $\lambda(A \times X) = \mu(A)$. If $u_1, u_2 : \Omega \to X$ are both Σ-measurable maps, then

$$\lambda_1 = \lambda_2 \text{ if and only if } u_1(z) = u_2(z) \ \mu\text{-a.e. in } \Omega.$$

The link connecting transition subprobabilities and Young submeasures is the disintegration theorem (see Valadier [408]).

Theorem 4.7.8 *If λ is a positive measure on $\Omega \times X$ such that $\lambda \circ \Pi_{\Omega}^{-1} \leqslant \mu$, then there exists a unique (up to equality μ-a.e.) transition subprobability $\hat{\lambda} \in SR(\Omega, X)$ such that*

$$\lambda(A) = \int_{\Omega} \left[\int_X \chi_A(z, u)\hat{\lambda}(z)(du) \right] \mu(dz) \text{ for all } A \in \Sigma \times B(X).$$

Remark 4.7.9 In general we do not distinguish between $\lambda \in S\mathscr{Y}(\Omega, \mu; X)$ and its disintegration $\hat{\lambda} \in SR(\Omega, X)$. Also, if $u : \Omega \to X$ is Σ-measurable, the disintegration of the Young measure associated to u (see Definition 4.7.6) is given by $\hat{\lambda}(z) = \delta_{u(z)}$.

This identification of sub-Young measures with transition subprobabilities via the disintegration theorem leads to the identification of $S\mathscr{Y}(\Omega, \mu; X)$ with a subset of $L^\infty(\Omega, M_r(X)_{w^*})$. Recall that the Dinculeanu and Foias theorem (see, for example, Gasinski and Papageorgiou [182, p. 131]) says that $L^\infty(\Omega, M_r(X)_{w^*}) = L^1(\Omega, C_0(X))^*$. Moreover, since Σ is countably generated and $C_0(X)$ is a separable Banach space, then the Lebesgue–Bochner space $L^1(\Omega, C_0(X))$ is separable too.

To get the desired identification, we will need the following abstract result about multifunctions.

Proposition 4.7.10 *If Y is a separable Banach space and $F : \Omega \to 2^{Y^*} \setminus \{\emptyset\}$ is a multifunction with w^*-closed and convex values in the unit ball of Y^* and for every $y \in Y$, $\omega \to \sigma(y, F(\omega)) = \sup\{\langle y^*, y\rangle : y \in F(\omega)\}$ is Σ-measurable, then $S_F^\infty = \{f \in L^\infty(\Omega, Y^*_{w^*}) = L^1(\Omega, Y)^* : f(\omega) \in F(\omega)\ \mu\text{-a.e. in } \Omega\}$ is convex and w^*-compact.*

Proof Evidently, S_F^∞ is relatively w^*-compact. Since $L^1(\Omega, X)$ (the predual) is separable, the w^*-topology on bounded sets of the dual $L^\infty(\Omega, X^*_{w^*})$ is metrizable. So, we can work with sequences. Let $\{f_n\}_{n \geqslant 1} \subseteq S_F^\infty$ and assume that $f_n \xrightarrow{w^*} f$. Arguing by contradiction, suppose $f \notin S_F^\infty$. We consider the set

$$C = \{z \in \Omega : f(z) \notin F(z)\} \in \Sigma.$$

Due to the separability of Y, we can find $\{y_n\}_{n\geqslant 1} \subseteq Y$ such that

$$C = \bigcup_{n\geqslant 1}\{z \in \Omega : \langle f(z), y_n\rangle_Y > \sigma(y_n, F(z)) = \sup\left\{\langle y^*, y_n\rangle_Y : y^* \in F(z)\right\}\}.$$

For some $k \in \mathbb{N}$, for $C_k = \{z \in \Omega : \langle f(z), y_k\rangle_Y > \sigma(y_k, F(z))\}$, we have $\mu(C_k) > 0$. Since $f_n \xrightarrow{w^*} f$, we have

$$\int_{C_k} \langle f(z), y_k\rangle_Y d\mu = \lim_{n\to\infty} \int_{C_k} \langle f_n(z), y_k\rangle_Y d\mu \leqslant \int_{C_k} \sigma(y_k, F(z))d\mu. \quad (4.111)$$

On the other hand, from the definition of C_k we have

$$\int_{C_k} \sigma(y_k, F(z))d\mu < \int_{C_k} \langle f(z), y_k\rangle_Y dz. \quad (4.112)$$

Comparing (4.111) and (4.112), we reach a contradiction. So, $f \in S^\infty_F$ and we conclude that S^∞_F is w^*-compact. □

Using this proposition and the identification of subYoung measures with transition subprobabilities, we obtain the following very useful identification.

Proposition 4.7.11 *The set $S\mathscr{Y}(\Omega, \mu; X)$ is homeomorphic to a closed subset of the unit ball of $L^\infty(\Omega, M_r(X)_{w^*}) = L^1(\Omega, C_0(X))^*$ furnished with the w^*-topology.*

This leads to the introduction of the following topology on $S\mathscr{Y}(\Omega, \mu; X)$.

Definition 4.7.12 The relative $w(L^\infty(\Omega, M_r(X)_{w^*}), L^1(\Omega, C_0(X)))$-topology on $S\mathscr{Y}(\Omega, \mu; X)$ is called the w^*-topology (or the $w(S\mathscr{Y}(\Omega, \mu; X), L^1(\Omega, C_0(X)))$-topology).

Corollary 4.7.13 *Every sequence $\{\lambda_n\}_{n\geqslant 1} \subseteq S\mathscr{Y}(\Omega, \mu; X)$ admits a w^*-convergent subsequence.*

Proposition 4.7.14 *If X is compact, then $\mathscr{Y}(\Omega, \mu; X)$ is w^*-compact.*

Proof The result follows from Proposition 4.7.10 by considering the multifunction $F(z) = M_1^+(X)$ for all $z \in \Omega$. Recall that since X is compact, then $M_1^+(X)$ furnished with the weak* topology (that is, the $w(M_r(X), C(X))$-topology) is compact (see, for example, Denkowski et al. [143, p. 198]). □

Proposition 4.7.15 *If Ω is also a locally compact, σ-compact metric space then the w^*-topology on $S\mathscr{Y}(\Omega, \mu; X)$ (see Definition 4.7.12) coincides with the $w(M_r(\Omega \times X), C_0(\Omega \times X))$-topology.*

Proof Note that the $w(M_r(\Omega \times X), C_0(\Omega \times X))$-topology is a Hausdorff topology which is weaker that the w^*-topology on $S\mathscr{Y}(\Omega, \mu; X)$, which is compact. So, the two topologies must coincide (see, for example, Denkowski et al. [143, p. 31]). $\square$

Definition 4.7.16 (a) A C_0-Carathéodory integrand is a function $\varphi : \Omega \times X \to \mathbb{R}$ such that

(i) for all $x \in X$, $z \to \varphi(z, x)$ is Σ-measurable;
(ii) for all $z \in \Omega$, $\varphi(z, \cdot) \in C_0(X)$;
(iii) $z \to \|\varphi(z, \cdot)\|_\infty$ is μ-integrable.

We denote the vector space of C_0-Carathéodory integrands by $V_{C_0}(\Omega, \mu; X)$.
(b) A normal integrand is a function $\varphi : \Omega \times X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ such that

(i) $(z, x) \to \varphi(z, x)$ is $\Sigma \times B(X)$-measurable;
(ii) for all $z \in \Omega$, $x \to \varphi(z, x)$ is lower semicontinuous.

We denote the cone of normal integrands by $V(\Omega, \Sigma; X)$ and by $V_+(\Omega, \Sigma; X)$ the subset of normal integrands with values in $\mathbb{R}_+ = [0, \infty)$.

Remark 4.7.17 Recall that a Carathéodory integrand is $\Sigma \times B(X)$-measurable. Also, a $\Sigma \times B(X)$-measurable function $\varphi : \Omega \times X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is superpositionally measurable, that is, if $u : \Omega \to X$ is Σ-measurable, then so is the function $z \to \varphi(z, u(z))$. In what follows we consider the equivalence classes of $V_{C_0}(\Omega, \mu; X)$ for the relation

$$\varphi_1 \sim \varphi_2 \text{ if and only if } \mu\{z \in \Omega : \varphi_1(z, \cdot) \neq \varphi_2(z, \cdot)\} = 0.$$

Proposition 4.7.18 *The map* $V_{C_0}(\Omega, \mu; X) \ni \varphi \longrightarrow (z \to \varphi(z, \cdot)) \in L^1(\Omega, C_0(X))$ *(see Definition 4.7.16(a)(iii)) is a bijection.*

Proof Let $\varphi \in V_{C_0}(\Omega, \mu; X)$ and set $\xi(z) = \varphi(z, \cdot)$. We have

- $\xi(z) \in C_0(X)$ for all $z \in \Omega$ (see Definition 4.7.16(a)(i));
- ξ is measurable; indeed if $\nu \in M_r(X) = C_0(X)^*$ (see (4.110)), then

$$z \to \int_X \varphi(z, u)\nu(du) \text{ is } \Sigma\text{-measurable.}$$

So, $\xi(\cdot)$ is weakly measurable and since $C_0(X)$ is a separable Banach space it is also Σ-measurable.

- $z \to \|\xi(z)\|_\infty$ is μ-integrable (see Definition 4.7.16(a)(iii)).

Thus we have a one-to-one map from $V_{C_0}(\Omega, \mu; X)$ into $L^1(\Omega, C_0(X))$. In fact we claim that this map is onto. To see this, let $\vartheta \in L^1(\Omega, C_0(X))$. We set $\varphi(z, x) = \vartheta(z)(x)$. Clearly $\varphi \in V_{C_0}(\Omega, \mu; X)$ (recall that the map $\eta \to \eta(u)$ is continuous on $C_0(X)$, thus measurable). $\square$

Proposition 4.7.19 *If* $\varphi \in V_+(\Omega, \Sigma; X)$*, then there exists an increasing sequence* $\{\varphi_n\}_{n\geqslant 1} \subseteq V_{C_0}(\Omega, \mu; X)$ *such that*

$$\varphi_n \uparrow \varphi \text{ (that is, } \varphi = \sup_{n\geqslant 1} \varphi_n).$$

Proof Let $\hat{\varphi}_n(z, u) = \inf\{\varphi(z, y) + nd(y, u) : y \in X\}$. Clearly, $\{\hat{\varphi}_n\}_{n\geqslant 1}$ is increasing.

First we show that for all $u \in X$, $z \to \varphi_n(z, u)$ is Σ-measurable. To this end, for any $u \in X$ and $\eta \in \mathbb{R}$ we have

$$\{z \in \Omega : \hat{\varphi}_n(z, u) < \eta\} = \Pi_\Omega\{(z, y) \in \Omega \times X : \varphi(z, y) + nd(y, u) < \eta\}.$$

The measurability of φ and the continuity of the distance function imply that

$$D = \{(z, y) \in \Omega \times X : \varphi(z, y) + nd(y, u) < \eta\} \in \Sigma \times B(X).$$

Since Σ is complete, from the Yankov–von Neumann–Aumann projection theorem we have $\Pi_\Omega(D) \in \Sigma$ and so we conclude that $z \to \hat{\varphi}_n(z, u)$ is Σ-measurable.

Next, we show that $u \to \hat{\varphi}_n(z, u)$ is Lipschitz continuous. For $v \in X$, we have

$$\begin{aligned}\hat{\varphi}_n(z, u) &\leqslant \varphi(z, y) + nd(y, u)\\ &\leqslant \varphi(z, y) + nd(y, v) + nd(v, u) \text{ (by the triangle inequality).}\end{aligned}$$

Since $y \in X$ is arbitrary, we obtain

$$\hat{\varphi}_n(z, u) \leqslant \hat{\varphi}_n(z, v) + nd(v, u). \tag{4.113}$$

Reversing the roles of $u, v \in X$ in the above argument, we also have

$$\begin{aligned}&\hat{\varphi}_n(z, v) \leqslant \hat{\varphi}_n(z, u) + nd(v, u)\\ \Rightarrow\ &|\hat{\varphi}_n(z, u) - \hat{\varphi}_n(z, v)| \leqslant nd(v, u) \text{ (see (4.113)).}\end{aligned}$$

Note that $\varphi_n(z, u) \leqslant \varphi(z, u)$ for all $(z, u) \in \Omega \times X$ and all $n \geqslant 1$. Hence

$$\lim_{n\to\infty} \varphi_n(z, u) \leqslant \varphi(z, u) \text{ (recall } \{\varphi_n\}_{n\geqslant 1} \text{ is increasing).} \tag{4.114}$$

On the other hand, for fixed $z \in \Omega$ and for every $n, k \in \mathbb{N}$, we can find $y_k \in X$ such that

$$\varphi(z, y_k) + nd(y_k, u) - \frac{1}{k} \leqslant \varphi_n(z, u). \tag{4.115}$$

Clearly, $d(y_k, u) \to 0$ as $n \to \infty$. So, if in (4.115) we pass to the limit as $k \to \infty$ and use the lower semicontinuity of $\varphi(z, \cdot)$, we obtain

$$\begin{aligned} &\varphi(z, u) \leqslant \varphi_n(z, u) \text{ for all } n \geqslant 1 \\ \Rightarrow\ &\varphi(z, u) \leqslant \lim_{n\to\infty} \varphi_n(z, u) \\ \Rightarrow\ &\varphi_n \uparrow \varphi \text{ (see (113))}. \end{aligned}$$

Finally, let $\{\beta_n\}_{n\geqslant 1} \subseteq C_0(X)$ be an increasing sequence such that $\beta_n \to 1$. If we set

$$\varphi_n(z, u) = \inf\{\hat{\varphi}_n(z, u); n\beta_n(u)\},$$

then $\varphi_n \in V_{C_0}(\Omega, \mu; X)$, $\varphi_n \geqslant 0$ for all $n \geqslant 1$ and $\varphi_n \uparrow \varphi$. $\square$

Using the last two propositions, we can prove the following theorem for Young measures.

Theorem 4.7.20 *If $\varphi \in V_+(\Omega, \Sigma; X)$, then*

$$S\mathscr{Y}(\Omega, \mu; X) \ni \lambda \to \int_{\Omega\times X} \varphi(z, u)\lambda(dz, du)$$

is lower semicontinuous.

Proof By virtue of Proposition 4.7.19, we can find an increasing sequence $\{\varphi_n\}_{n\geqslant 1} \subseteq V_{C_0}(\Omega, \mu; X)$ such that $\varphi_n \uparrow \varphi$. Then by the monotone convergence theorem, we have

$$\int_{\Omega\times X} \varphi_n d\lambda \uparrow \int_{\Omega\times X} \varphi d\lambda.$$

Using Proposition 4.7.18, we see that for each $n \geqslant 1$ the mapping $\lambda \mapsto \int_{\Omega\times X} \varphi_n d\lambda$ is continuous. Therefore, $\lambda \mapsto \int_{\Omega\times X} \varphi d\lambda$, being the supremum of continuous functions, is itself lower semicontinuous. The proof is now complete. $\square$

Example 4.7.21 The result fails if φ is not positive (or more generally bounded below). To see this let $X = \mathbb{R}$, $\varphi(z, u) = \varphi(u) = -1$ and $u_n = n$. Then

$$\delta_{u_n} \to \lambda = 0 \text{ in } S\mathscr{Y}(\Omega, \mu; \mathbb{R}).$$

But $\displaystyle\int_{\Omega\times X} \varphi d\lambda = \int_X \varphi d\lambda = 0$ while $\displaystyle\liminf_{n\to\infty} \int_{\Omega\times X} \varphi d\delta_{u_n} = \liminf_{n\to\infty} \int_X \varphi d\delta_{u_n} = -1$.

Now we introduce a second topology on $\mathscr{Y}(\Omega, \mu; X)$. To do this, we need to expand the notion of a Carathéodory integrand.

Definition 4.7.22 (a) A C_b-Carathéodory integrand is a function $\varphi : \Omega \times X \to \mathbb{R}$ such that

(i) for all $u \in X$, $z \to \varphi(z, u)$ is Σ-measurable;
(ii) for all $z \in \Omega$, $\varphi(z, \cdot) \in C_b(X)$;
(iii) $z \to \|\varphi(z, \cdot)\|_\infty$ is μ-integrable.

We denote the vector space of C_b-Carathéodory integrands by $V_{C_b}(\Omega, \mu; X)$.

(b) The narrow topology on $\mathscr{Y}(\Omega, \mu; X)$ is the weakest topology which makes the maps $\lambda \to \int_{\Omega\times X} \varphi d\lambda$ continuous, where φ varies over $V_{C_b}(\Omega, \mu; X)$.

Remark 4.7.23 This topology is Hausdorff and it coincides with the w^*-topology (see Definition 4.7.12) on narrow relatively compact sets.

Proposition 4.7.24 *If $u_n, u : \Omega \to X$, $n \geqslant 1$, are Σ-measurable functions and $\{\lambda_n, \lambda\}_{n\geqslant 1}$ are the corresponding Young measures (see Definition 4.7.6(b)) then $u_n \xrightarrow{\mu} u$ if and only if $\lambda_n \to \lambda$ narrowly.*

Proof $\underline{\Rightarrow}$: For every $\varphi \in V_{C_b}(\Omega, \mu; X)$ we have

$$\int_{\Omega\times X} \varphi d\lambda_n = \int_\Omega \varphi(z, u_n(z))d\mu \text{ for all } n \geqslant 1 \text{ and} \tag{4.116}$$
$$\int_{\Omega\times X} \varphi d\lambda = \int_\Omega \varphi(z, u(z))d\mu.$$

By passing to a suitable subsequence, we may assume that

$$u_n(z) \to u(z) \ \mu\text{-a.e. in } \Omega$$
$$\Rightarrow \varphi(z, u_n(z)) \to \varphi(z, u(z)) \text{ a.e. in } \Omega.$$

So, by the Lebesgue dominated convergence theorem we have

$$\int_\Omega \varphi(z, u_n(z))d\mu \to \int_\Omega \varphi(z, u(z))d\mu.$$

Hence by (4.116) and the Urysohn criterion, for the original sequence we have

$$\int_\Omega \varphi d\lambda_n \to \int_\Omega \varphi d\lambda.$$

$\underline{\Leftarrow}$: Let
$$\varphi(z, v) = \min\{1, \|v - u(z)\|\}.$$

Evidently, $\varphi \in V_{C_b}(\Omega, \mu; X)$. By the Chebyshev inequality and (4.116), for every $\varepsilon > 0$ and $n \geqslant 1$ we have

$$\mu\{z \in \Omega : d(u_n(z), u(z)) \geqslant \varepsilon\} \leqslant \frac{1}{\varepsilon} \int_{\Omega \times X} \varphi d\lambda_n \text{ for all } n \geqslant 1$$
$$\Rightarrow u_n \xrightarrow{\mu} u \text{ as } n \to \infty.$$

The proof is now complete. ☐

Theorem 4.7.25 *If $\varphi \in V_+(\Omega, \Sigma; X)$, then $\mathscr{Y}(\Omega, \mu; X) \ni \lambda \to \int_{\Omega \times X} \varphi d\lambda$ is narrowly lower semicontinuous.*

Proof As in the proof of Proposition 4.7.19 we define

$$\hat{\varphi}_n(z, u) = \inf\{\varphi(z, y) + nd(y, u) : y \in X\}.$$

From the proof, we know that $\hat{\varphi}_n(z, u)$ is measurable in $z \in \Omega$ and Lipschitz continuous in $u \in X$. Setting $\varphi_n(z, u) = \min\{n, \hat{\varphi}_n(z, u)\}$ we see that $\varphi_n \in V_{C_b}(\Omega, \mu; X)$ and $\varphi_n \uparrow \varphi$. By the monotone convergence theorem we have

$$\int_{\Omega \times X} \varphi d\lambda = \sup_{n \geqslant 1} \int_{\Omega \times X} \varphi_n d\lambda$$
$$\Rightarrow \lambda \to \int_{\Omega \times X} \varphi d\lambda \text{ is narrowly lower semicontinuous.}$$

The proof is now complete. ☐

Let $\hat{X}$ denote the Alexandroff one point compactification of X (recall that X is locally compact).

Proposition 4.7.26 *The narrow topology on $\mathscr{Y}(\Omega, \mu; X)$ can also be defined as the weakest topology which makes the maps $\mathscr{Y}(\Omega, \mu; X) \ni \lambda \to \int_{\Omega \times X} \hat{\varphi}\Big|_{\Omega \times X} d\lambda$ continuous, where $\hat{\varphi}$ varies over $V_{C_b}(\Omega, \mu; \hat{X})$.*

Proof Let $\hat{\tau}$ denote the weak topology defined on $\mathscr{Y}(\Omega, \mu; X)$ by the statement of the proposition and let τ denote the narrow topology. Evidently, $\hat{\tau} \subseteq \tau$. Let $\varphi \in V_{C_b}(\Omega, \mu; X)$ and let $\beta(z) = \|\varphi(z, \cdot)\|_\infty$. Then we have

$$\int_{\Omega \times X} \varphi d\lambda = \int_{\Omega \times X} (\varphi + \beta) d\lambda - \int_\Omega \beta d\mu.$$

By virtue of Theorem 4.7.25, we infer that $\lambda \to \int_{\Omega \times X} \varphi d\lambda$ is $\hat{\tau}$-lower semicontinuous. But $V_{C_b}(\Omega, \mu; X)$ is a vector space. So, $\lambda \to \int_{\Omega \times X} \varphi d\lambda$ is $\hat{\tau}$-continuous, hence $\tau \subseteq \hat{\tau}$ and we conclude that $\tau = \hat{\tau}$. ☐

Remark 4.7.27 In fact the narrow topology is the weakest topology on $\mathscr{Y}(\Omega, \mu; X)$ which makes the maps $\mathscr{Y}(\Omega, \mu; X) \ni \lambda \to \int_{\Omega \times X} \varphi d\lambda$ continuous as φ varies over the vector space of integrands of the form

$$\varphi(z, u) = \sum_{k=1}^{n} \chi_{A_k}(z)\hat{\vartheta}_k\Big|_X(u),$$

where $\{A_k\}_{k=1}^n$ is a Σ-partition of Ω and $\hat{\vartheta}_k \in C(\hat{X})$.

Definition 4.7.28 A subset C of $\mathscr{Y}(\Omega, \mu; X)$ is said to be uniformly tight if given $\varepsilon > 0$ we can find $K_\varepsilon \subseteq X$ compact such that

$$\sup_{\lambda \in C} \lambda(\Omega \times (X \setminus K_\varepsilon)) < \varepsilon.$$

Remark 4.7.29 Note that using the disintegration of λ, we have

$$\lambda(\Omega \times (X \setminus K_\varepsilon)) = \int_\Omega \lambda(z)(X \setminus K_\varepsilon) d\mu.$$

Also, the uniform tightness is equivalent to saying that

$$D = \{\lambda \circ (\Pi_X)^{-1} : \lambda \in C\} \subseteq M_r^+(X)$$

is uniformly tight in the sense of Prokhorov. So, we have the following property.

Proposition 4.7.30 *If $C \subseteq \mathscr{Y}(\Omega, \mu; X)$, then C is uniformly tight if and only if there exists an inf-compact function $\xi : X \to \mathbb{R}_+ = [0, \infty)$ (that is, for every $\eta \geqslant 0$, the level set $\{u \in X : \xi(u) \leqslant \eta\}$ is compact) such that*

$$\sup_{\lambda \in C} \int_\Omega \left[\int_X \xi(u)\lambda(z)(du) \right] d\mu < \infty.$$

Remark 4.7.31 If $C \subseteq \mathscr{Y}(\Omega, \mu; X)$ is uniformly tight, then so is its narrow closure.

As in the case with the classical Prokhorov theorem (see, for example, Denkowski et al. [143, p. 201]), uniform tightness in the sense of Definition 4.7.28 is in fact equivalent to relative narrow compactness.

Theorem 4.7.32 *$C \subseteq \mathscr{Y}(\Omega, \mu; X)$ is uniformly tight if and only if it is relatively narrow compact.*

Proof $\underline{\Rightarrow}$: Let $\hat{X}$ be the Alexandroff one point compactification of X. Then $V_{C_b}(\Omega, \mu; \hat{X})$ can be identified with $L^1(\Omega, C(\hat{X}))$ and $\mathscr{Y}(\Omega, \mu; \hat{X})$ can be identified with a w^*-closed bounded subset of $L^\infty(\Omega, M_r(\hat{X})_{w^*}) = L^1(\Omega, C(\hat{X}))^*$. Therefore the narrow topology on C is metrizable and we can work with sequences. Using the

Alaoglu theorem and Proposition 4.7.26, we conclude that C is relatively narrow compact.

$\Leftarrow$: Consider the map $\lambda \to \nu = \lambda \circ (\Pi_X)^{-1}$ from $\mathscr{Y}(\Omega, \mu; X)$ into $M_r(X)$. We claim that this map is narrow continuous. To see this, note that for all $\varphi \in C_b(X)$, we have

$$\int_X \varphi(u)d\nu = \int_{\Omega \times X} \varphi(u)d\lambda$$

and the latter integral is continuous in λ. So, the image $\hat{C}$ of C is relatively compact in $M_r(X)$ furnished with the standard narrow topology. By the classical Prokhorov theorem (see, for example, Denkowski et al. [143, p. 201]), this implies that $\hat{C}$ is uniformly tight in the sense of Prokhorov. Hence C is uniformly tight (see Remark 4.7.29). ☐

Uniform tightness links w^*-compactness and narrow compactness.

Proposition 4.7.33 *If $C \subseteq \mathscr{Y}(\Omega, \mu; X)$ is uniformly tight, then the w^*-topology (see Definition 4.7.12) and the narrow topology (see Definition 4.7.22(b)) on C coincide.*

Proof Evidently, the w^*-topology is weaker than the narrow one (recall $C_0(X) \subseteq C_b(X)$) and both are Hausdorff. So, they coincide on compact sets, in particular then on C by virtue of Theorem 4.7.32. The proof is now complete. ☐

In the next result, $X = \mathbb{R}^N$.

Proposition 4.7.34 *If $\{u_n\}_{n\geqslant 1} \subseteq L^1(\Omega, \mathbb{R}^N)$ is norm bounded and $\{\lambda_n\}_{n\geqslant 1} \subseteq \mathscr{Y}(\Omega, \mu; \mathbb{R}^N)$ is the corresponding sequence of Young measures (see Definition 4.7.6(b)), then we can find a subsequence $\{\lambda_{n_k}\}_{k\geqslant 1}$ of $\{\lambda_n\}_{n\geqslant 1}$ and $\lambda \in \mathscr{Y}(\Omega, \mu; \mathbb{R}^N)$ such that*

$$\lambda_{n_k} \to \lambda \text{ narrowly.}$$

Proof By virtue of Proposition 4.7.30 with $h(u) = \|u\|$ for all $u \in \mathbb{R}^N$, we see that $\{\lambda_n\}_{n\geqslant 1}$ is uniformly tight. So, we can apply Theorem 4.7.32 to reach the desired conclusion. ☐

In the beginning of this section, we mentioned the difference between strongly and weakly convergent sequences in L^1. Now we will see what Young measures bring to the more complete understanding of this difference.

We start with a lower semicontinuity result, which can be viewed as a further elaboration of Theorem 4.7.25. In this result $X = \mathbb{R}^N$.

Theorem 4.7.35 *If $u_n : \Omega \to \mathbb{R}^N$, $n \geqslant 1$, is a sequence of Σ-measurable functions, $\{\lambda_n\}_{n\geqslant 1}$ is the corresponding sequence of Young measures, $\lambda_n \to \vartheta$ narrowly with ϑ a Young measure, $\varphi : \Omega \times \mathbb{R}^N \to \mathbb{R}$ is a measurable function such that $\varphi(z, \cdot)$ is lower semicontinuous for all $z \in \Omega$ and $\{\varphi(\cdot, u_n(\cdot))^-\}_{n\geqslant 1}$ is uniformly integrable, then $\int_{\Omega\times\mathbb{R}^N} \varphi d\vartheta \leqslant \liminf_{n\to\infty} \int_\Omega \varphi(z, u_n(z))d\mu$. Moreover*

$$\liminf_{n\to\infty}\int_{\Omega}\varphi(z,u_n(z))d\mu<+\infty \Rightarrow \int_{\Omega\times\mathbb{R}^N}\varphi^+ d\vartheta<+\infty.$$

Proof Let $r \geqslant 0$ and set $\varphi_r = \max\{-r,\varphi\}+r$. Then $\varphi_r \geqslant 0$ and we can apply Theorem 4.7.25 and have

$$\int_{\Omega\times\mathbb{R}^N}\varphi_r d\vartheta \leqslant \liminf_{n\to\infty}\int_{\Omega}\varphi_r(z,u_n(z))d\mu.$$

Subtracting $\int_{\Omega\times\mathbb{R}^N} r d\vartheta = \int_{\Omega} r d\mu$, we obtain

$$\int_{\Omega\times\mathbb{R}^N}\hat{\varphi}_r d\vartheta \leqslant \liminf_{n\to\infty}\int_{\Omega}\hat{\varphi}_r(z,u_n(z))d\mu,\ \text{ where } \hat{\varphi}_r=\max\{-r,\varphi\}. \tag{4.117}$$

Let $A_{nr}=\{z\in\Omega:\varphi(z,u_n(z))<-r\}$. Then

$$\int_{A_{nr}}\varphi(z,u_n(z))dz\leqslant 0$$

and by virtue of our hypothesis on the uniform integrability of the negative parts for $r>0$ large enough is greater than or equal to $-\varepsilon$. Then

$$\begin{aligned}\int_{\Omega}\varphi(z,u_n(z))d\mu &= \int_{A_{nr}}\varphi(z,u_n(z))d\mu+\int_{\Omega}\hat{\varphi}_r(z,u_n(z))d\mu-\int_{A_{nr}}\hat{\varphi}(z,u_n(z))d\mu\\ &\geqslant \int_{\Omega}\hat{\varphi}_r(z,u_n(z))d\mu-\varepsilon. \end{aligned}\tag{4.118}$$

If $r=0$, then $\hat{\varphi}_r=\varphi^+$ and so

$$\begin{aligned}&\int_{\Omega}\varphi(z,u_n(z))d\mu \geqslant \int_{\Omega}\varphi^+(z,u_n(z))d\mu - M\\ &\Rightarrow \int_{\Omega\times\mathbb{R}^N}\varphi^+ d\vartheta \leqslant \liminf_{n\to\infty}\int_{\Omega}\varphi^+(z,u_n(z))d\mu \text{ (see Theorem 4.7.25)}\\ &\qquad\leqslant \liminf_{n\to\infty}\int_{\Omega}\varphi(z,u_n(z))d\mu+M.\end{aligned}$$

This proves the last conclusion of the theorem.
We return to (4.118) to complete the proof. We have

$$\begin{aligned}\liminf_{n\to\infty}\int_{\Omega}\varphi(z,u_n(z))d\mu &\geqslant \liminf_{n\to\infty}\int_{\Omega}\hat{\varphi}(z,u_n(z))d\mu-\varepsilon\\ &\geqslant \int_{\Omega\times\mathbb{R}^N}\hat{\varphi}_r d\vartheta-\varepsilon \text{ (see (4.117))}.\end{aligned}\tag{4.119}$$

Note that $\varphi \leqslant \hat{\varphi}_r$. So, from (4.119) we have

$$\liminf_{n\to\infty} \int_\Omega \varphi(z, u_n(z))d\mu \geqslant \int_{\Omega\times\mathbb{R}^N} \varphi d\vartheta - \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary we let $\varepsilon \downarrow 0$ to finish the proof. $\square$

Remark 4.7.36 In applications, usually $\{u_n\}_{n\geqslant 1} \subseteq L^1(\Omega, \mathbb{R}^N)$ and it is weakly convergent. The uniform integrability requirement on $\{\varphi(\cdot, u_n(\cdot))^-\}_{n\geqslant 1}$ is satisfied if for example

$$\varphi(z, v) \geqslant a(z) - c\|v\| \text{ for } \mu\text{-a.a. } z \in \Omega,\ \text{all } v \in \mathbb{R}^N,\ \text{with } a \in L^1(\Omega),\ c > 0.$$

We will use this theorem to better understand the difference between strong and weak convergence in $L^1(\Omega, \mathbb{R}^N)$. First a definition.

Definition 4.7.37 We say that a Σ-measurable function $u : \Omega \to X$ is the barycenter of a transition probability $\lambda \in R(\Omega, X)$ ($u = \mathrm{bar}(\lambda)$) if

$$u(z) = \int_\Omega v\lambda(z)(dv) \text{ for all } z \in \Omega.$$

In the next result, $X = \mathbb{R}^N$.

Theorem 4.7.38 *If $u_n \xrightarrow{w} u$ in $L^1(\Omega, \mathbb{R}^N)$ and $\{\lambda_n, \lambda\}_{n\geqslant 1}$ are the Young measures corresponding to the functions $\{u_n, u\}_{n\geqslant 1}$, then*

(a) there exists a subsequence $\{\lambda_{n_k}\}_{k\geqslant 1}$ of $\{\lambda_n\}_{n\geqslant 1}$ and a Young measure ϑ such that

$$\lambda_{n_k} \to \vartheta \textit{ narrowly},\ u = \mathrm{bar}(\lambda)$$
$$\textit{and}\quad \|u_{n_k} - u\|_1 \to \int_{\Omega\times\mathbb{R}^N} \|v - u(z)\|d\vartheta;$$

moreover, if for μ-a.a. $z \in \Omega$, $\vartheta(z)$ is a Dirac measure, then $\vartheta = \lambda$ and $u_n \to u$ strongly in $L^1(\Omega, \mathbb{R}^N)$;

(b) if $\{u_n\}_{n\geqslant 1}$ does not converge strongly, then ϑ (as in part (a)) is not associated to a function;

(c) $u_n \to u$ strongly in $L^1(\Omega, \mathbb{R}^N)$ if and only if $\lambda_n \to \lambda$ narrowly.

Proof (a) Since $\{u_n\}_{n\geqslant 1} \subseteq L^1(\Omega, \mathbb{R}^N)$ is weakly convergent, it is norm bounded and so invoking Proposition 4.7.34, we can find a subsequence $\{\lambda_{n_k}\}_{k\geqslant 1}$ of $\{\lambda_n\}_{n\geqslant 1}$ and a Young measure ϑ such that

$$\lambda_n \to \vartheta \text{ narrowly}.$$

Using Theorem 4.7.25, we have

$$\int_\Omega \left[\int_{\mathbb{R}^N} \|v\|\vartheta(z)(dv)\right] d\mu = \int_{\Omega\times\mathbb{R}^N} \|v\| d\vartheta$$
$$\leqslant \liminf_{k\to\infty} \int_{\Omega\times\mathbb{R}^N} \|v\| d\lambda_{n_k}$$
$$= \liminf_{k\to\infty} \|u_{n_k}\|_1 \leqslant M < \infty.$$

So, for μ-a.a. $z \in \Omega$ bar(ϑ) exists. Let $h \in L^\infty(\Omega, \mathbb{R}^N)$ and set $\psi(z, v) = (v, h(z))_{\mathbb{R}^N}$. Evidently, ψ is a Carathéodory function (that is, measurable in $z \in \Omega$, continuous in $v \in \mathbb{R}^N$). Moreover, we have

$$\psi(z, u_n(z))^- = (u_n(z), h(z))^-_{\mathbb{R}^N} \leqslant \|h\|_\infty \|u_n(z)\| \text{ for } \mu\text{-a.e. in } \Omega.$$

But by the Dunford–Pettis theorem $\{u_n\}_{n\geqslant 1} \subseteq L^1(\Omega, \mathbb{R}^N)$ is uniformly integrable. So, $\{\psi(\cdot, u_n(\cdot))^-\}_{n\geqslant 1} \subseteq L^1(\Omega)$ is uniformly integrable. Therefore, we can apply Theorem 4.7.35 and obtain

$$\int_{\Omega\times\mathbb{R}^N} (v, h(z))_{\mathbb{R}^N} d\vartheta \leqslant \liminf_{n\to\infty} \int_{\Omega\times\mathbb{R}^N} (u_{n_k}(z), h(z))_{\mathbb{R}^N} d\mu.$$

Replacing h with $-h$ in the above argument, we conclude that

$$\int_{\Omega\times\mathbb{R}^N} (v, h(z))_{\mathbb{R}^N} d\vartheta = \lim_{n\to\infty} \int_{\Omega\times\mathbb{R}^N} (u_{n_k}(z), h(z))_{\mathbb{R}^N} d\mu. \tag{4.120}$$

We know that

$$\int_{\Omega\times\mathbb{R}^N} (v, h(z))_{\mathbb{R}^N} d\vartheta = \int_\Omega \left(\int_{\mathbb{R}^N} v\vartheta(z)(dv), h(z)\right)_{\mathbb{R}^N} d\mu \tag{4.121}$$
$$= \int_\Omega (\text{bar}(\vartheta), h(z))_{\mathbb{R}^N} d\mu.$$

Since $u_n \to u$ in $L^1(\Omega, \mathbb{R}^N)$, we have

$$\int_\Omega (u_n(z), h(z))_{\mathbb{R}^N} d\mu \to \int_\Omega (u(z), h(z)) d\mu. \tag{4.122}$$

From (4.120), (4.121) and (4.122), it follows that

$$\int_\Omega (u(z), h(z))_{\mathbb{R}^N} d\mu = \int_\Omega (\text{bar}(\vartheta), h(z))_{\mathbb{R}^N} d\mu.$$

Since this is true for all $h \in L^\infty(\Omega, \mathbb{R}^N)$, we conclude that $u = \text{bar}(\vartheta)$.

Let $\varphi(z, v) = \|v - u(z)\|$. This is a Carathéodory function and $\varphi(\cdot, u_n(\cdot))^- = 0$ for all $n \geqslant 1$. Also, if we consider $-\varphi$, then $(-\varphi(\cdot, u_n(\cdot)))^- = \|u_n(\cdot) - u(\cdot)\|$ for all $n \geqslant 1$ and by the Dunford–Pettis theorem $\{(-\varphi(\cdot, u_n(\cdot)))^-\}_{n\geqslant 1}$ is uniformly integrable. So, in both cases we can apply Theorem 4.7.35, to obtain

$$\|u_{n_k} - u\|_1 \to \int_{\Omega \times \mathbb{R}^N} \|v - u(z)\| d\vartheta. \tag{4.123}$$

Finally if for μ-a.a. $z \in \Omega$, $\vartheta(z)$ is a Dirac measure, then since $u = \mathrm{bar}(\vartheta)$, we have $\vartheta(z) = \delta_{u(z)}$ for μ-a.e. $z \in \Omega$ and so $\vartheta = \lambda$. Hence from (4.123) it follows that

$$u_n \to u \text{ strongly in } L^1(\Omega, \mathbb{R}^N).$$

(b) Arguing by contradiction suppose that ϑ is associated with a Σ-measurable function. Since u_n does not converge strongly, we can find $\varepsilon > 0$ and a subsequence of $\{u_n\}_{n \geqslant 1}$ (still denoted by index n) such that

$$\|u_n - u\|_1 \geqslant \varepsilon \text{ for all } n \geqslant 1.$$

From part (a) we know that there exists a subsequence $\{\lambda_{n_k}\}_{n \geqslant 1}$ of $\{\lambda_n\}_{n \geqslant 1}$ such that $\lambda_{n_k} \to \vartheta$ narrowly. Then from (a) we have $u_{n_k} \to u$ in $L^1(\Omega, \mathbb{R}^N)$, a contradiction.

(c) This follows from Proposition 4.7.24 and part (a). □

Definition 4.7.39 Let $(\Omega_k, \Sigma_k, \mu_k), k = 1, 2$, be two complete finite measure spaces, X_k, $k = 1, 2$, two locally compact, σ-compact, complete metric spaces and $\lambda_k \in \mathscr{Y}(\Omega_k, \mu_k; X_k)$, $k = 1, 2$, two Young measures. We define $\lambda_1 \otimes \lambda_2 \in \mathscr{Y}(\Omega_1 \times \Omega_2, \mu_1 \otimes \mu_2; X_1 \times X_2)$ by setting

$$(\lambda_1 \otimes \lambda_2)(z) = \lambda_1(z) \otimes \lambda_2(z) \text{ for all } z = (z_1, z_2) \in \Omega_1 \times \Omega_2.$$

For such "product" Young measures, we have the following continuity result. For its proof, we refer to Balder [28].

Lemma 4.7.40 *If $\lambda_1^n \to \lambda_1$ and $\lambda_2^n \to \lambda_2$ narrowly, then $\lambda_1^n \otimes \lambda_2^n \to \lambda_1 \otimes \lambda_2$ narrowly.*

Using this lemma, we can prove the following strong-weak lower semicontinuity result for integral functionals, which is important in the calculus of variations.

Theorem 4.7.41 *If $\varphi : \Omega \times \mathbb{R}^N \times \mathbb{R}^m \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is a measurable function which is lower semicontinuous in $(u, v) \in \mathbb{R}^N \times \mathbb{R}^m$, $u_n : \Omega \to \mathbb{R}^N$ $n \geqslant 1$ is a sequence of Σ-measurable functions such that $u_n \xrightarrow{\mu} u$, $\{v_n\}_{n \geqslant 1} \subseteq L^1(\Omega, \mathbb{R}^m)$ such that $v_n \xrightarrow{w} v$ in $L^1(\Omega, \mathbb{R}^m)$, for all $z \in \Omega$, $\varphi(z, u(z), \cdot)$ is convex and*

$$\{\varphi(\cdot, u_n(\cdot), v_n(\cdot))^-\}_{n \geqslant 1} \subseteq L^1(\Omega) \textit{ is uniformly integrable},$$

then $\displaystyle\int_\Omega \varphi(z, u(z), v(z)) d\mu \leqslant \liminf_{n \to \infty} \int_\Omega \varphi(z, u_n(z), v_n(z)) dz.$

Proof Let $\{\lambda_n\}_{n\geqslant 1}$ and $\{\beta_n\}_{n\geqslant 1}$ be the Young measures associated to the functions $\{u_n\}_{n\geqslant 1}$ and $\{v_n\}_{n\geqslant 1}$ respectively. Using Proposition 4.7.24, Theorem 4.7.38 and by passing to a suitable subsequence if necessary, we can find

$$\lambda_n \to \lambda \text{ and } \beta_n \to \beta \text{ narrowly, with } u = \mathrm{bar}(\beta). \tag{4.124}$$

Then by virtue of Lemma 4.7.40, we have

$$\eta_n = \lambda_n \otimes \beta_n \to \eta = \lambda \otimes \beta \text{ narrowly.}$$

So, we have

$$\begin{aligned}
\int_{\Omega\times\mathbb{R}^N\times\mathbb{R}^m} \varphi d\eta &= \int_\Omega \left[\int_{\mathbb{R}^N\times\mathbb{R}^m} \varphi(z,u,v)\eta(z)(d(u,v))\right] d\mu \\
&= \int_\Omega \left[\int_{\mathbb{R}^m} \varphi(z,u(z),v)\beta(z)(dv)\right] d\mu \\
&\geqslant \int_\Omega \varphi(z,u(z),v(z))d\mu \text{ (by Jensen's inequality and (4.124)).}
\end{aligned}$$

The proof is now complete. □

4.8 Remarks

4.1: Most of the results on the different types of cones are due to Krasnoselskii [250]. Various parts of this material can also be found in the books of Deimling [142], Guo and Lakshmikantham [199], Papageorgiou and Kyritsi [329] and Peressini [334]. Theorem 4.1.19 is due to Krein [252]. Proposition 4.1.22 can be found in Filippakis et al. [170].

4.2: Theorem 4.2.3 is due to Banach [30]. It is a versatile abstraction of the idea of successive approximations first used in the context of differential equations by Cauchy [110] and Picard [342]. Sometimes, in order to be able to apply the Banach fixed point theorem, we need to equivalently renorm the ambient Banach space. This renorming idea was first used by Bielecki [51] (see also Denkowski et al. [143, p. 217] and Goebel and Kirk [190, p. 17]). Theorem 4.2.8 is due to Edelstein [153]. The first fixed point theorem for nonexpansive maps was proved independently by Browder [83] and Göhde [191] and concerns such maps defined on a uniformly convex space. The more general version presented in Theorem 4.2.29 is due to Kirk [235], who realized that using the notion of normal structure (see Definition 4.2.24(b)), he could extend the Browder–Göhde fixed point result. The notion of normal structure was introduced by Brodskii and Milman [70], who used it to study fixed points of isometries. Concerning nonexpansive maps, we also have the following useful result due to Browder [83].

Theorem 4.8.1 *If X is a uniformly convex Banach space, $C \subseteq X$ is closed, convex and bounded, $\varphi : C \to X$ is nonexpansive and $h(u) = u - \varphi(u)$ for all $u \in C$, then h is demiclosed on C, that is, if $u_n \xrightarrow{w} u$ in C and $h(u_n) \to y$ in X, then we have $h(u) = y$.*

A very detailed presentation of the metric fixed point theory can be found in the book of Goebel and Kirk [190]. Several aspects of the theory can also be found in the books of Deimling [142], Granas and Dugundji [197], Papageorgiou and Kyritsi [329], and Zeidler [426].

4.3: Corollary 4.3.3 is one of the oldest and best known results of topology. It was proved by Brouwer [65] for $N = 3$. The proof for arbitrary $N \geqslant 1$ was provided by Hadamard [200]. We should mention that even before the above mentioned works, Bohl [53] had an equivalent result for differentiable maps. A combinatorial proof of Brouwer's fixed point theorem was given by Knaster et al. [240] who also noted that the condition $\varphi(\partial B_1^N) \subseteq \overline{B}_1^N$ suffices for a fixed point. Another proof was given by Milnor [299]. Theorems 4.3.8 and 4.3.9 are essentially due to Borsuk [56, 57], who also proved the following:

Proposition 4.8.2 *If $K \subseteq \mathbb{R}^N$ is compact and has nonempty interior, then ∂K is not a retract of K.*

The fixed point theorem in Proposition 4.3.11 is also due to Borsuk [57]. In Theorem 4.3.12, condition (3.26) is known as the Leray–Schauder boundary condition and guarantees that $d_{LS}(i - \varphi, \Omega, 0) \neq 0$. The theorem goes back to the seminal work of Leray and Schauder [266]. For the special cases mentioned in Corollary 4.3.15, we refer to Rothe [362] and Altman [11]. Theorems 4.3.23 and 4.3.24 are due to Mönch [302]. The notion of asymptotic center (see Definition 4.3.25(a)) was introduced by Edelstein [153], while the notion of the inward set of $u \in C$ with respect to C and of weakly inward maps (see Definition 4.3.25(b)) is due to Halpern and Bergman [201]. The alternative definition of weak inwardness given in Remark 4.3.26 is due to Caristi [101]. Theorem 4.3.29 is also due to Caristi [101].

Finally, we should mention the Tychonoff fixed point theorem (see Tychonoff [407]), which is the extension of the Schauder fixed point theorem to locally convex spaces.

Theorem 4.8.3 *If X is a locally convex space, $C \subseteq X$ is closed and convex, $\varphi : C \to C$ is continuous and $\overline{\varphi(C)}$ is compact, then φ has a fixed point.*

4.4: Fixed point theorems based on the order structure of the underlying spaces can be found in Amann [12, 13], Carl and Heikkila [102], Deimling [142], Granas and Dugundji [197], Guo and Lakshmikantham [199] and Papageorgiou and Kyritsi [329]. The fixed point index was studied in detail by Amann [14] who used it to solve various semilinear elliptic boundary value problems. Finally, we want to mention the classical Knaster–Tarski fixed point theorem (see Knaster [239] and Tarski [402]).

Theorem 4.8.4 *If $(X, \preccurlyeq)$ is a partially ordered space, $\varphi : X \to X$ is increasing (that is, $y \preccurlyeq u \Rightarrow \varphi(y) \preccurlyeq \varphi(u)$) and there exists a $u_0 \in X$ such that*

(i) $u_0 \preccurlyeq \varphi(u_0)$;
(ii) *every chain in* $\{v \in X : u_0 \preccurlyeq v\}$ *has a supremum,*

then the set of fixed points of φ *is nonempty and it admits a maximal element (that is, there exists a* $\hat{u} \in X$ *such that* $\varphi(\hat{u}) = \hat{u}$ *and there is no fixed point* $v \in X$ *with* $u_0 \preccurlyeq v$*).*

4.5: The multivalued extension of Banach's fixed point theorem presented in Theorem 4.5.1 was announced by Markin [289] and proved in a more general form (with $F(\cdot)$ having values in $P_f(X)$) by Nadler [312]. A continuation of that result is the following stability result due to Lim [271]. An earlier such result was proved by Markin [289], who assumed that X is a subset of a Hilbert space. In what follows, given a multifunction $F : X \to 2^X \setminus \{\emptyset\}$ by $\mathrm{Fix}(F)$ we denote the set of fixed points of F, that is, $\mathrm{Fix}(F) = \{u \in X : u \in F(u)\}$.

Proposition 4.8.5 *If* (X, d) *is a complete metric space and* $F_n, F : X \to P_f(X)$, $n \geqslant 1$, *are all multivalued contractions with the same contraction constant (that is, they all satisfy (4.58) with the same* $k \in [0, 1)$*) and* $\sup_{u \in X} h(F_n(u), F(u)) \to 0$ *as* $n \to \infty$, *then* $h(\mathrm{Fix}(F_n), \mathrm{Fix}(F)) \to 0$ *as* $n \to \infty$.

Theorem 4.5.5 (the usc multivalued version of the Schauder fixed point theorem) was first proved for $X = \mathbb{R}^N$ by Kakutani [226]. It was extended to Banach spaces by Bohenblust and Karlin [52] (see Theorem 4.5.5). Extensions to locally convex spaces were obtained by Fan [166] and Glicksberg [189]. More general fixed point theorems for inward and outward multifunctions (see Theorem 4.5.14), can be found in Deimling [142, Sect. 2.4.7].

In multivalued fixed point theory, an essential condition is the convexity of the values of $F(\cdot)$. By using tools from algebraic topology, we can replace the convexity hypothesis on the values of $F(\cdot)$ by the more general acyclicity condition. For results in this direction, we refer to Eilenberg and Montgomery [155] and Gorniewicz [194]. Another substitute for convexity is the notion of decomposability (see Definition 2.11.8). In this direction we have the following result of Cellina et al. [113].

Theorem 4.8.6 *If* Ω *is a compact Hausdorff topological space,* μ *is a finite nonatomic Borel measure,* X *is a separable Banach space,* $D \subseteq L^1(\Omega, X)$ *is closed and* $F : D \to P_f(D)$ *is a usc multifunction with decomposable values and* $\overline{F(D)}$ *is compact, then* F *has a fixed point.*

The books of Aubin and Frankowska [22], Border [54], Deimling [142], Hu and Papageorgiou [218, 219] and Klein and Thompson [238] contain results from the fixed point theory for multifunctions.

4.6: Theorem 4.6.1 was proved by Lax and Milgram [261]. Theorem 4.6.2 concerning variational inequalities is due to Stampacchia [385]. Theorem 4.6.13 is a classical existence result for nonlinear variational inequalities and can be found, for example, in Showalter [381, p. 84]. Theorem 4.6.14 was proved by Ekeland [157]. In Ekeland [158, 159] the reader can find comprehensive surveys of the many applications that this theorem has. Theorem 4.6.20 is due to Caristi [101], while Theorem 4.6.25 is due to Danes [135], who had a different proof, based on some results of

Krasnoselskii and Zabreiko. Proposition 4.6.32 is due to Brondsted and Rockafellar [72]. Theorem 4.6.39 was obtained by Brezis and Browder [66] and provides a very general variational principle from which all the other nonlinear results mentioned above can be derived.

The following generalization of Theorem 4.6.1 will be useful in critical point theory in conjunction with the Cerami compactness condition (the C-condition for short, see Sect. 5.1). The result is due to Zhang [428].

Proposition 4.8.7 *If $h : \mathbb{R}_+ \to \mathbb{R}_+$ is a continuous nondecreasing function such that $\int_0^\infty \frac{dr}{1+h(r)} = +\infty$, (X, d) is a complete metric space, $u_0 \in X$ is fixed, $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is a proper lower semicontinuous, bounded below function, $\varepsilon > 0$, $\varphi(y) \leqslant \inf_X \varphi + \varepsilon$ and $\lambda > 0$, then there exists a $u_\lambda \in X$ such that*

$$\begin{aligned}
&\varphi(u_\lambda) \leqslant \varphi(y), \ d(u_\lambda, u_0) \leqslant r_0 + \overline{r}, \\
&\varphi(u_\lambda) \leqslant \varphi(u) + \frac{\varepsilon}{\lambda} \frac{1}{1 + h(d(u_0, u_\lambda))} d(u, u_\lambda) \text{ for all } u \in X,
\end{aligned}$$

where $r_0 = d(u_0, y)$ and $\overline{r} > 0$ such that $\int_0^{r_0+r} \frac{dr}{1+h(r)} \geqslant \lambda$.

Remark 4.8.8 If $h = 0$ and $u_0 = y$, then Proposition 4.8.7 reduces to Theorem 4.6.1.

4.7: The theory of Young measures has its roots in the so-called "generalized curves" introduced by Young [420, 422, 423], in order to have a precise description of the limits of minimizing sequences in the calculus of variations and optimal control. Since then there has been an extensive development of the original ideas of Young, in order to meet the needs of the calculus of variations, optimal control, game theory, mathematical economics and more recently theoretical mechanics. This development can be traced in the books of Buttazzo [97], Castaing et al. [107], Gamkrelidze [179], Gasinski and Papageorgiou [182], Roubicek [365] and in the papers of Balder [27, 28], Ball and Zhang [29], Berliocchi and Lasry [48], DiPerna [145], DiPerna and Majda [146], Kinderlehrer and Pedregal [233], Sychev [395] and Valadier [409]. Theorem 4.7.41 can be extended to Banach spaces. More precisely, we have the following result due to Balder [28].

Proposition 4.8.9 *If (Ω, Σ, μ) is a finite measure space, X, Y are separable Banach spaces with Y also reflexive and $\varphi : \Omega \times X \times Y \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is a measurable function such that*

(i) for μ-a.a. $z \in \Omega$, $(u, y) \to \varphi(z, u, y)$ is lower semicontinuous;
(ii) for μ-a.a. $z \in \Omega$ and all $u \in X$, $\varphi(z, u, \cdot)$ is convex;
(iii) for μ-a.a. $z \in \Omega$ and all $(u, y) \in X \times Y$ we have

$$a(z) - c(\|u\|_X + \|y\|_Y) \leqslant \varphi(z, u, y)$$

with $a \in L^1(\Omega)$, $c > 0$,

then $(u, y) \to I_\varphi(u, y) = \int_\Omega \varphi(z, u(z), y(z))dz$ *is sequentially weakly lower semi-continuous from* $L^1(\Omega, X) \times L^1(\Omega, Y)_w$ *into* $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$.

Recall that, if $\{u_n\}_{n\geqslant 1} \subseteq L^1(\Omega)$, $u_n \xrightarrow{w} u$ in $L^1(\Omega)$ and $u_n(z) \leqslant \liminf_{n\to\infty} u(z)$ μ-a.e. in Ω, then $u_n \to u$. In the next theorem, we see how this results translates to vector-valued functions. The result is due to Visintin [411].

Theorem 4.8.10 *If* $\{u_n\}_{n\geqslant 1} \subseteq L^1(\Omega, \mathbb{R}^N)$, $u_n \xrightarrow{w} u$ *in* $L^1(\Omega, \mathbb{R}^N)$ *and* $u(z) \in \operatorname{ext}\left[\overline{\operatorname{conv}} \limsup_{n\to\infty}\{u_n(z)\}\right]$ *for* μ*-a.a.* $z \in \Omega$*, then* $u_n \to u$ *in* $L^1(\Omega, \mathbb{R}^N)$.

Chapter 5
Critical Point Theory

The study of mathematics, like the Nile, begins in minuteness but ends in magnificence

Charles Caleb Colton (1780–1832)

Critical point theory deals with variational problems and so it can be argued that it is as old as calculus. Nevertheless, in its modern form, critical point theory has its roots in the so-called "Dirichlet principle". The name was coined by Riemann in his thesis (1851) and the principle postulates that given a bounded open set $\Omega \subseteq \mathbb{R}^2$ and a continuous function $g : \partial\Omega \to \mathbb{R}$ ($\partial\Omega$ being the boundary of Ω), the boundary value problem (Dirichlet problem) $-\Delta u(z) = 0$ in Ω, $u|_{\partial\Omega} = g$, admits a smooth solution u which minimizes the energy functional

$$\varphi(u) = \int_\Omega \left[\left(\frac{\partial u}{\partial z_1} \right)^2 + \left(\frac{\partial u}{\partial z_2} \right)^2 \right] dz.$$

This principle was criticized by Weierstrass (1870) who produced an example showing that the existence of a minimum is not guaranteed even if the functional which is to be minimized is bounded below. In this example, Weierstrass considered the functional

$$\varphi(u) = \int_{-1}^{1} (tu'(t)^2)dt$$

to be minimized over the set

$$D = \{u \in C^1[-1, 1] : u(-1) = 0,\ u(1) = 1\}.$$

N. S. Papageorgiou et al., *Nonlinear Analysis–Theory and Methods*,
Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-030-03430-6_5

Then

$$u_n(t) = \frac{1}{2} + \frac{1}{2}\,\frac{\tan^{-1}\left(\frac{t}{n}\right)}{\tan^{-1}\left(\frac{1}{n}\right)} \qquad (n \geqslant 1)$$

is a minimizing sequence in D, that is, $\{u_n\}_{n\geqslant 1} \subseteq D$ and $\varphi(u_n) \downarrow 0$. However, there is no minimizer in D, since if $u \in D$ and $\varphi(u) = 0$, then $tu'(t) = 0$ for all $t \in [-1, 1]$, hence $u \equiv$constant, which contradicts the boundary conditions. Subsequently, analysts tried to correctly formulate the Dirichlet principle. This culminated in the so-called direct methods of the calculus of variations based on the concept of lower semicontinuity of variational integrals, introduced by Tonelli [404]. Until the early 20th century, mathematicians were looking for absolute minima of functionals bounded below. Birkhoff (1917) was the first to characterize critical points by a minimax principle. This was done more systematically in the 1930s by Ljusternik–Schnirelmann and Morse on finite-dimensional spaces. In the 1960s this theory was extended to infinite-dimensional spaces by Palais and Smale who, in order to compensate for the lack of local compactness of the ambient space, introduced a compactness condition on the functional. Their work opened the way for modern critical point theory and led to the mountain pass theorem of Ambrosetti and Rabinowitz [17] and the all encompassing linking principle of Benci and Rabinowitz [43]. In this chapter, we present some of the main aspects of modern critical point theory. In Sect. 5.1, we discuss compactness-type conditions on the functional (the Palais–Smale and Cerami conditions) and also establish the existence of a pseudogradient vector field, which is the main tool in obtaining deformation theorems for functionals defined on Banach spaces. In Sect. 5.2, we present some basic results from the direct method on the calculus of variations. In Sect. 5.3, we prove some deformation theorems. Such results are an effective tool in locating critical points of a C^1-functional. These results describe the deformations of the sublevel sets of the functional near a critical point where topologically interesting things may occur. In Sect. 5.4, we use the deformation theorems to prove a general linking principle and from it derive some well-known results of modern critical point theory, such as the mountain pass theorem and the saddle point theorem. Section 5.5 deals with critical points under constraints. So, we discuss the method of Lagrange multipliers, manifolds of codimension one and the so-called natural constraints. In Sect. 5.6, we investigate the effects of symmetries on the existence of critical points. So, there exists a group G acting in a continuous way on the space X and a functional φ that is invariant under this action. Under these conditions, it usually happens that the functional has many critical points. To study them we introduce the notions of Ljusternik–Schnirelmann category and of Krasnoselskii genus and we derive the Ljusternik–Schnirelmann multiplicity theorems. We also prove the so-called symmetric criticality principle due to Palais. In Sect. 5.7 we study the structure of the critical set of φ. Finally, in Sect. 5.8, we examine variational problems which exhibit a lack of some desirable compactness properties. In general, there are two ways to have such a lack of compactness. One is due to the action of a group which leaves the functional φ invariant (for example if we are in $\mathbb{R}^N$ and the functional is invariant under the group of translations $u(\cdot) \to u(\cdot + z_0)$ for some $z_0 \in \mathbb{R}^N$). The other occurs if the nonlinearity in the functional φ exhibits critical growth and so the compactness of the embedding in the Sobolev theorem fails. This leads us to the concentration-compactness theorem of P.-L. Lions.

5.1 Pseudogradients and Compactness Conditions

The deformation approach in modern critical point theory is based on deformation arguments along the gradient flow or of a suitable substitute of it when, due to the geometry of the space or to the lack of regularity of the functional, it is impossible to use the gradient flow. So, let $\varphi \in C^1(X)$ and consider the steepest descent flow

$$\dot{u}(t) = -\nabla\varphi(u(t)). \tag{5.1}$$

In many cases this differential system or just its trajectory are not defined, either due to the nature of X or due to the regularity of φ. For example, if X is a Banach space which is not a Hilbert space, then $\nabla\varphi(u) \in X^*$ for all $u \in X$ and so the differential system (5.1) cannot be defined. For this reason, we replace the gradient vector field by the so-called pseudogradient vector field with values in X. Then we study the deformation flow generated. To determine the basic properties of this flow, we need a compactness-type condition on the functional φ.

First let us introduce a substitute of the gradient vector field that will make a differential system like (5.1) valid under all circumstances.

Definition 5.1.1 Let X be a Banach space, X^* be its topological dual and denote by $\langle \cdot, \cdot \rangle$ the duality brackets for the pair (X, X^*). Let $\varphi \in C^1(X)$. A vector $v \in X$ is a "pseudogradient vector" of φ at u if

$$||v|| \leqslant 2||\varphi'(u)||_* \text{ and } ||\varphi'(u)||_* \leqslant \left\langle \varphi'(u), v \right\rangle.$$

We say that $V : \{u \in X : \varphi'(u) \neq 0\} \to X$ is a "pseudogradient vector field" for φ if V is locally Lipschitz and for every $u \in \{u \in X : \varphi'(u) \neq 0\}$ (this is the set of regular points of φ) $V(u)$ is a pseudogradient vector of φ at u.

Remark 5.1.2 Any convex combination of pseudogradient vectors (resp. of pseudogradient vector fields) is again a pseudogradient vector (resp. a pseudogradient vector field). So the pseudogradient vector field may exist, but is certainly not unique. If $X = H =$ Hilbert space, then $\nabla\varphi$ need not be a pseudogradient vector field due to the local Lipschitzness requirement. However, if $\varphi \in C^1(H)$ and has a locally Lipschitz derivative (for example, $\varphi \in C^2(H)$) $\varphi' : H \to H^*$, then the gradient of φ is a pseudogradient vector field. Recall that the gradient $\nabla\varphi : H \to H$ is defined using the Riesz–Fréchet representation theorem by $\nabla\varphi(u) \in H$ being the unique element such that

$$\left\langle \varphi'(u), h \right\rangle = (\nabla\varphi(u), h)_H \text{ for all } h \in H,$$

with $(\cdot, \cdot)_H$ denoting the inner product of H.

Lemma 5.1.3 *Assume that Y is a metric space, V is a normed space, for every $y \in Y$, $F(y)$ is a nonempty, convex subset of V and for every $y \in Y$ we can find a neighborhood U of y such that $\bigcap\limits_{y' \in U} F(y') \neq \emptyset$. Then there exists a locally Lipschitz map $h : Y \to V$ such that $h(y) \in F(y)$ for all $y \in Y$.*

Proof Let $y \in Y$. By hypothesis we can find $U(y)$, a neighborhood of y, such that

$$\bigcap_{y' \in U(y)} F(y') \neq \emptyset.$$

The family $\{U(y)\}_{y \in Y}$ is an open cover of Y. Note that Y, being a metric space, is paracompact. So, there is a locally finite refinement $\{V_i\}_{i \in I}$ of $\{U(y)\}_{y \in Y}$.

Suppose that $V_i \neq Y$ for all $i \in I$. We define

$$\sigma_i(y) = d_Y(y, Y \setminus V_i) \text{ and } \eta(y) = \sum_{i \in I} \sigma_i(y) \text{ for all } y \in Y.$$

Evidently, σ_i is Lipschitz continuous and since the cover $\{V_i\}_{i \in I}$ is locally finite, it follows that η is locally Lipschitz and $\eta(y) \neq 0$ for all $y \in Y$. We set

$$\vartheta_i(y) = \frac{\sigma_i(y)}{\eta(y)} \text{ for all } i \in I.$$

Then $\{\vartheta_i\}_{i \in I}$ is a locally Lipschitz partition of unity subordinate to the cover $\{V_i\}_{i \in I}$.

Now we assume that $V_{i_0} = Y$ for some $i_0 \in I$. We set $\vartheta_{i_0} \equiv 1$ and $\vartheta_i \equiv 0$ for $i \neq t_0$. Again we have that $\{\vartheta_i\}_{i \in I}$ is a locally Lipschitz partition of unity subordinate to $\{V_i\}_{i \in I}$.

Because the cover $\{V_i\}_{i \in I}$ refines the cover $\{U(y)\}_{y \in Y}$, we have that

$$\bigcap_{y' \in V_i} F(y') \neq \emptyset \text{ for all } i \in I.$$

Let $u_i \in \bigcap\limits_{y' \in V_i} F(y')$ for all $i \in I$ and introduce the function $h : Y \to V$ defined by

$$h(y) = \sum_{i \in I} \vartheta_i(y) u_i.$$

The function h is locally Lipschitz. Moreover, for every $y \in Y$, we can find a finite number of sets $V_{i_1}, \cdots, V_{i_m}$ of the cover $\{V_i\}_{i \in I}$ such that $y \in V_{i_k}$ for all $k \in \{1, \cdots, m\}$. We have

$$h(y) = \sum_{k=1}^{m} \vartheta_{i_k}(y) u_{i_k} \text{ and } \sum_{k=1}^{m} \vartheta_{i_k}(y) = 1.$$

Since $u_{i_k} \in F(y)$ for all $k \in \{1, \cdots, m\}$ and the latter is convex, it follows that $h(y) \in F(y)$ for all $y \in Y$. $\square$

This lemma leads to the existence of a pseudogradient vector field.

Theorem 5.1.4 *Assume that $\varphi \in C^1(X)$. Then there is a pseudogradient vector field for φ.*

Proof Let $N = \{u \in X : \varphi'(u) \neq 0\}$. For every $u \in N$, let $F(u)$ be the set of all pseudogradient vectors of φ at u (see Definition 5.1.1). From Remark 5.1.2, we know that $F(u)$ is a convex subset of X.

Given $u \in N$, we can find $v \in X$ such that

$$||v|| \leqslant 1 \text{ and } \frac{4}{5}||\varphi'(u)||_* \leqslant \langle \varphi'(u), v \rangle .$$

We set $y = \frac{5}{3}||\varphi'(u)||_* v$. Then

$$||y|| \leqslant \frac{5}{3}||\varphi'(u)||_* \text{ and } \langle \varphi'(u), y \rangle \geqslant \frac{4}{3}||\varphi'(u)||_* .$$

Since $\varphi \in C^1(X)$, we can find U, a neighborhood of u, such that

$$\begin{aligned} &||y|| < 2||\varphi'(z)||_* \text{ and } \langle \varphi'(z), y \rangle > ||\varphi'(z)||_* \text{ for all } z \in U \\ \Rightarrow\ & y \in \bigcap_{z \in U} F(z). \end{aligned}$$

Applying Lemma 5.1.3, we obtain a locally Lipschitz map $V : N \to X$ such that

$$\begin{aligned} & V(u) \in F(u) \text{ for all } u \in N \\ \Rightarrow\ & V \text{ is the desired pseudogradient vector field for } \varphi. \end{aligned}$$

The proof is now complete. $\square$

Remark 5.1.5 If φ is even, then it admits an odd pseudogradient vector field. Indeed, let V be any pseudogradient vector field and set

$$V_0(u) = \frac{1}{2}[V(u) - V(-u)] \text{ for all } u \in N.$$

Also, note that given positive constants $0 < \alpha < \beta$, we can find a locally Lipschitz map $V : N \to X$ such that

$$\alpha \leqslant \langle \varphi'(u), V(u) \rangle \leqslant ||\varphi'(u)||_* ||V(u)|| \leqslant \beta \text{ for all } u \in N.$$

Moreover, if $X = H$ is a Hilbert space and $\varphi \in C^2(H)$, then we can take

$$V(u) = \frac{\alpha + \beta}{2} \frac{\nabla\varphi(u)}{||\nabla\varphi(u)||^2},$$

with $\nabla\varphi$ being the gradient of φ (see Remark 5.1.2).

As we already mentioned in the Introduction, the deformation approach to the minimax theory of critical values of a C^1-functional φ is based on the deformation flow generated by the pseudogradient vector field. The topological properties of the sublevel sets $\varphi^\lambda = \{u \in X : \varphi(u) \leqslant \lambda\}$ change only when λ crosses a critical value. To be able to deform these sublevel sets of φ, we need a compactness-type condition, which compensates for the fact that the ambient space need not be locally compact (being infinite-dimensional).

Definition 5.1.6 Let X be a Banach space and X^* be its topological dual. Suppose that $\varphi \in C^1(X)$.

(a) We say that φ satisfies the Palais–Smale condition at the level $c \in \mathbb{R}$ (the PS_c-condition) if the following is true:

"every sequence $\{u_n\}_{n\geqslant 1} \subseteq X$ such that $\varphi(u_n) \to c$ and

$\varphi'(u_n) \to 0$ in X^* as $n \to \infty$

admits a strongly convergent subsequence".

If this condition holds at every level $c \in \mathbb{R}$, then we say that φ satisfies the Palais–Smale condition (the PS-condition for short).

(b) We say that φ satisfies the Cerami condition at the level $c \in \mathbb{R}$ (the C_c-condition for short) if the following is true:

"every sequence $\{u_n\}_{n\geqslant 1} \subseteq X$ such that $\varphi(u_n) \to c$ and

$(1 + ||u_n||)\varphi'(u_n) \to 0$ in X^* as $n \to \infty$

admits a strongly convergent subsequence".

If this condition holds at every level $c \in \mathbb{R}$, then we say that φ satisfies the Cerami condition (the C-condition for short).

Remark 5.1.7 Evidently, the C-condition is weaker than the PS-condition. Both conditions PS_c and C_c imply that the set $K_\varphi^c = \{u \in X : \varphi'(u) = 0,\ \varphi(u) = c\}$ is compact. However, neither the PS nor the C condition has any influence on the size of the critical set $K_\varphi = \{u \in X : \varphi'(u) = 0\}$. These compactness-type conditions are quite restrictive. For example, the constant functions and the trigonometric functions $\cos u$ and $\sin u$ do not satisfy them. Nevertheless in many applications, under suitable conditions on the data, the PS or C conditions are satisfied. In what follows, we examine these conditions in more detail and we see how they are related to the asymptotic behavior of the functional φ.

Proposition 5.1.8 *Let X be a Banach space, $\varphi \in C^1(X)$ be bounded from below, $c = \inf_X \varphi$ and φ satisfy the C_c-condition. Then there exists a $u_0 \in X$ such that $\varphi(u_0) = \inf_X \varphi = c$.*

Proof By virtue of Proposition 4.8.7 (with $h(r) = r$), we can produce a sequence $\{u_n\}_{n\geqslant 1} \subseteq X$ such that

$$\varphi(u_n) \to c \text{ and } (1 + ||u_n||)\varphi'(u_n) \to 0 \text{ in } X^*.$$

Since φ satisfies the C_c-condition (see Definition 5.1.6(b)), by passing to a suitable subsequence if necessary, we may assume that $u_n \to u_0$ in X as $n \to \infty$. Then

$$\varphi(u_n) \to \varphi(u_0) = c.$$

The proof is now complete. □

An analogous result is also valid in the case of the PS_c-condition. More precisely, we have the following proposition.

Proposition 5.1.9 *Let X be a Banach space, $\varphi \in C^1(X)$ be bounded from below, $c = \inf_X \varphi$ and assume that φ satisfies the PS_c-condition. Then every minimizing sequence $\{u_n\}_{n\geqslant 1}$ of φ admits a strongly convergent subsequence which converges to $u_0 \in X$, a global minimizer of φ.*

Proof By passing to a subsequence if necessary, we may assume that $\varphi(u_n) \leqslant c + \frac{1}{n^2}$ for all $n \geqslant 1$. Invoking Corollary 4.6.16, we can find a sequence $\{v_n\}_{n\geqslant 1} \subseteq X$ such that

$$\varphi(v_n) \downarrow c, \ \varphi'(v_n) \to 0 \text{ in } X^* \text{ and } ||v_n - u_n|| \leqslant \frac{1}{n} \text{ for all } n \geqslant 1. \tag{5.2}$$

Since φ satisfies the PS_c-condition, we may assume that

$$\begin{aligned} &v_n \to u_0 \text{ in } X \\ \Rightarrow\ &\varphi(v_n) \to \varphi(u_0) = c = \inf_X \varphi. \end{aligned}$$

The proof is now complete. □

Definition 5.1.10 Let X be a Banach space. We say that the functional $\varphi : X \to \mathbb{R}$ is coercive if $\varphi(u) \to +\infty$ as $||u|| \to \infty$.

Remark 5.1.11 Coercivity is equivalent to saying that for every $\lambda \in \mathbb{R}$, the sublevel set $\varphi^\lambda = \{u \in X : \varphi(u) \leqslant \lambda\}$ is bounded.

The next proposition relates the C-condition to the notion of coercivity.

Proposition 5.1.12 *Let X be a Banach space and $\varphi \in C^1(X)$ be a functional which is bounded from below and satisfies the C-condition. Then φ is coercive.*

Proof We argue by contradiction. So, suppose that φ is not coercive. Then we can find $c \in \mathbb{R}$ and a sequence $\{u_n\}_{n\geqslant 1} \subseteq X$ such that

$$\varphi(u_n) \leqslant c + \frac{1}{n} \text{ and } ||u_n|| \geqslant 2(e^n - 1) \text{ for all } n \geqslant 1. \tag{5.3}$$

We apply Proposition 4.8.7 with $h(r) = r$, $\epsilon = c + \frac{1}{n} - \inf\limits_X \varphi$, $\lambda = n$ and $\bar{r} = e^n - 1$ and we produce a sequence $\{y_n\}_{n\geqslant 1} \subseteq X$ such that

$$\varphi(y_n) \leqslant \varphi(u_n),\ ||y_n - u_n|| \leqslant e^n - 1 \text{ and } ||\varphi'(y_n)||_* \leqslant \frac{c + \frac{1}{n} - \inf\limits_X \varphi}{n(1 + ||y_n||)} \text{ for all } n \geqslant 1. \tag{5.4}$$

From (5.3) and (5.4) it follows that

$$\begin{aligned} &||y_n|| \geqslant e^n - 1 \text{ for all } n \geqslant 1 \\ \Rightarrow\ &||y_n|| \to +\infty \text{ as } n \to \infty. \end{aligned} \tag{5.5}$$

On the other hand, again from (5.3) and (5.4), we have

$$\limsup_{n\to\infty} \varphi(y_n) \leqslant c \text{ and } (1 + ||y_n||)||\varphi'(y_n)||_* \to 0 \text{ as } n \to \infty. \tag{5.6}$$

Since φ satisfies the C-condition and comparing (5.5) and (5.6) we reach a contradiction. □

Remark 5.1.13 If $\varphi \in C^1(\mathbb{R}^N)$ is bounded from below and coercive, then the PS-condition holds. Indeed, since φ is coercive and $\varphi(u_n) \to c$, it follows that $\{u_n\}_{n\geqslant 1} \subseteq \mathbb{R}$ is bounded and so it admits a convergent subsequence. On the other hand, if X is an infinite-dimensional Banach space, then it can happen that $\varphi \in C^1(X)$ is bounded below and coercive without satisfying the PS-condition. To see this, let $\xi : \mathbb{R}_+ \to \mathbb{R}$ be a smooth function such that $\xi(s) = 0$ for all $s \in [0, 2]$ and $\xi(s) = s$ for all $s \geqslant 3$. We set $\varphi(u) = \xi(||u||)$ for all $u \in X$. Then $\varphi \in C^1(X)$ and it is bounded below and coercive. However the PS_0-condition does not hold. Indeed, let $\{u_n\}_{n\geqslant 1} \subseteq X$ such that $||u_n|| = 1$ for all $n \geqslant 1$. Then $\varphi(u_n) = 0$ and $\varphi'(u_n) = 0$ for all $n \geqslant 1$. But $\{u_n\}_{n\geqslant 1}$ has no strongly convergent subsequence.

Using Proposition 5.1.12, we can compare the two compactness-type conditions for φ introduced in Definition 5.1.6.

Proposition 5.1.14 *Let X be a Banach space and $\varphi \in C^1(X)$ be a functional which is bounded from below. Then the PS-condition and the C-condition are equivalent.*

Proof Evidently, we only need to show that the C-condition implies the PS-condition. To this end, let $\{u_n\}_{n\geqslant 1} \subseteq X$ be a sequence such that

$$\{\varphi(u_n)\}_{n\geqslant 1} \subseteq \mathbb{R} \text{ is bounded and } \varphi'(u_n) \to 0 \text{ in } X^* \text{ as } n \to \infty.$$

Since φ satisfies the C-condition, we can use Proposition 5.1.12 and infer that $\{u_n\}_{n\geqslant 1} \subseteq X$ is bounded. Hence $(1 + ||u_n||)\varphi'(u_n) \to 0$ in X^* and so the C-condition implies that $\{u_n\}_{n\geqslant 1} \subseteq X$ admits a strongly convergent subsequence. Therefore φ satisfies the PS-condition. ☐

However, there is a situation where coercivity of the functional implies the PS-condition. This setting is encountered in the study of boundary value problems.

Proposition 5.1.15 *Let X be a reflexive Banach space and assume that $\varphi \in C^1(X)$ is coercive and*

$$\varphi' = A + K,$$

with $A : X \to X^*$ of type $(S)_+$ and $K : X \to X^*$ completely continuous. Then φ satisfies the PS-condition.

Proof Consider a sequence $\{u_n\}_{n\geqslant 1} \subseteq X$ such that

$$\{\varphi(u_n)\}_{n\geqslant 1} \subseteq \mathbb{R} \text{ is bounded and } \varphi'(u_n) \to 0 \text{ in } X^* \text{ as } n \to \infty. \tag{5.7}$$

Since φ is coercive, relation (5.7) implies that $\{u_n\}_{n\geqslant 1} \subseteq X$ is bounded. By virtue of the reflexivity of X, we may assume that $u_n \xrightarrow{w} u$ in X as $n \to \infty$. The complete continuity of K implies that $K(u_n) \to K(u)$ in X^* as $n \to \infty$. Then

$$\begin{aligned}
&\lim_{n\infty} \langle A(u_n), u_n - u\rangle = \lim_{n\to\infty} \left[\left\langle \varphi'(u_n), u_n - u\right\rangle - \langle K(u_n), u_n - u\rangle\right] = 0 \\
&\Rightarrow u_n \to u \text{ in } X \text{ as } n \to \infty \text{ (since } A \text{ is of type } (S)_+) \\
&\Rightarrow \varphi \text{ satisfies the } PS\text{-condition.}
\end{aligned}$$

The proof is now complete. ☐

Next, we derive some properties of the sublevel sets $\varphi^\lambda = \{u \in X : \varphi(u) \leqslant \lambda\}$ and see how they are related to the PS-condition.

Proposition 5.1.16 *Let X be a Banach space, $\varphi \in C^1(X)$, $c \in \mathbb{R}$ and assume that φ^λ is bounded for all $\lambda < c$ and unbounded for all $\lambda > c$. Then we can find a sequence $\{u_n\}_{n\geqslant 1} \subseteq X$ such that*

$$\varphi(u_n) \to c,\ ||u_n|| \to \infty \text{ and } \varphi'(u_n) \to 0 \text{ in } X^* \text{ as } n \to \infty\,.$$

In particular, the PS_c-condition fails.

Proof By hypothesis, for every $n \geqslant 1$ we can find $r_n \geqslant n$ such that

$$\varphi^{c-\frac{1}{n}} \subseteq B_{r_n}(0) = \{u \in X : ||u|| < r_n\},\ n \geqslant 1. \tag{5.8}$$

Let $D_n = X \backslash B_{r_n}(0)$. Then from (5.8) we have

$$c - \frac{1}{n} \leqslant \inf_{D_n} \varphi = c_n. \tag{5.9}$$

By hypothesis, the set $\varphi^{c+\frac{1}{n}}$ is unbounded. So, we can find $v_n \in X$ such that

$$v_n \in \varphi^{c+\frac{1}{n}} \text{ and } ||v_n|| \geqslant r_n + 1 + \frac{1}{\sqrt{n}} \qquad \text{for all } n \geqslant 1. \tag{5.10}$$

Evidently, $v_n \in D_n$ and so from (5.9) and (5.10) we have

$$c_n \leqslant \varphi(v_n) \leqslant c + \frac{1}{n} \leqslant c_n + \frac{2}{n} \qquad \text{for all } n \geqslant 1. \tag{5.11}$$

Applying Theorem 4.6.14 (the Ekeland variational principle) with $\epsilon = \frac{2}{n}$ and $\lambda = \frac{1}{\sqrt{n}}$, we produce $u_n \in D_n$ such that

- $c - \dfrac{1}{n} \leqslant c_n \leqslant \varphi(u_n) \leqslant \varphi(v_n) \leqslant c + \dfrac{1}{n} \leqslant c_n + \dfrac{2}{n}$ (see (5.1)); (5.12)
- $\varphi(u_n) \leqslant \varphi(u) + \dfrac{2}{\sqrt{2}}||u - u_n||$ for all $u \in D_n$; (5.13)
- $||u_n - v_n|| \leqslant \dfrac{1}{\sqrt{n}}$. (5.14)

From (5.10) and (5.13) we have

$$||u_n|| \geqslant r_n + 1 \text{ for all } n \geqslant 1. \tag{5.15}$$

This implies that $u_n \in \operatorname{int} D_n$ and so from (5.13) we have

$$||\varphi'(u_n)||_* \leqslant \frac{2}{\sqrt{n}} \qquad \text{for all } n \geqslant 1. \tag{5.16}$$

From (5.12), (5.15) and (5.16) we obtain

$$\varphi(u_n) \to c,\ ||u_n|| \to \infty \text{ and } \varphi'(u_n) \to 0 \text{ in } X^* \text{ as } n \to \infty,\ \text{respectively.}$$

The proof is now complete. □

An immediate consequence of the proposition is the following result.

Corollary 5.1.17 *Let X be a Banach space and assume that $\varphi \in C^1(X)$ satisfies the PS_c-condition and φ^λ is bounded for all $\lambda < c$. Then there exists an $\eta > 0$ such that $\varphi^{\lambda+\eta}$ is bounded.*

Remark 5.1.18 According to Proposition 5.1.16, in the above corollary we may replace the hypothesis that φ satisfies the PS_c-condition by the following weaker condition:

"Every sequence $\{u_n\}_{n\geqslant 1} \subseteq X$ such that

$$\varphi(u_n) \to c \text{ and } \varphi'(u_n) \to 0 \text{ in } X^* \text{ as } n \to \infty$$

admits a bounded subsequence"

.

Proposition 5.1.19 *Let X be a Banach space and assume that $\varphi \in C^1(X)$ is bounded from below and is not coercive. Define $\lambda_0 = \sup\left[\lambda \in \mathbb{R} : \varphi^\lambda \text{ is bounded}\right]$. Then φ does not satisfy the PS_{λ_0}-condition.*

Proof Let $\mathscr{L} = \{\lambda \in \mathbb{R} : \varphi^\lambda \text{ is bounded}\}$. Let $m = \inf\limits_X \varphi > -\infty$ (since by hypothesis φ is bounded below). Hence $\mathscr{L} \supseteq (-\infty, m]$ (recall that by definition the empty set is bounded) and so $\mathscr{L} \neq \emptyset$. Let $\lambda_0 = \sup \mathscr{L}$. Since φ is not coercive, we have $\lambda_0 < \infty$. Then φ^μ is unbounded for all $\mu > \lambda_0$. Invoking Proposition 5.1.16, we conclude that φ does not satisfy the PS_{λ_0}-condition. □

Remark 5.1.20 Note that $\lambda_0 = \sup \mathscr{L} = \inf\{\mu \in \mathbb{R} : \varphi^\mu \text{ is unbounded}\}$.

Next, we consider functionals $h : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ of the form

$$h = \varphi + \psi \tag{5.17}$$

with $\varphi \in C^1(X)$ and $\psi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ a convex and lower semicontinuous (as always, not identically $+\infty$). Such functionals are important in the study of problems with unilateral constraints (variational inequalities).

Definition 5.1.21 Let X be a Banach space and $h : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ be a functional of the form (5.17) with $\varphi \in C^1(X)$ and $\psi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ a convex and lower semicontinuous functional. We say that $u \in \operatorname{dom}\psi = \{v \in X : \psi(v) < \infty\}$ is a critical point of h, if

$$0 \in \varphi'(u) + \partial\psi(u).$$

Remark 5.1.22 So, according to this definition $u \in \operatorname{dom}\psi$ is a critical point of h if

$$-\varphi'(u) \in \partial\psi(u),$$

which is equivalent to saying that

$$0 \leqslant \left\langle \varphi'(u), y - u \right\rangle + \psi(y) - \psi(u) \text{ for all } y \in X.$$

If $\psi \equiv 0$, then we recover the classical definition of a critical point for φ, namely that $\varphi'(u) = 0$.

Definition 5.1.23 Let $h = \varphi + \psi$ as above. We say that h satisfies the generalized Palais–Smale condition (*GPS*-condition for short) if the following is true:

$$\text{“Every sequence } \{u_n\}_{n\geqslant 1} \subseteq X \text{ such that } h(u_n) \to c \in \mathbb{R} \text{ and}$$

$$-\epsilon_n ||y - u_n|| \leqslant \langle \varphi'(u_n), y - u_n \rangle + \psi(y) - \psi(u_n) \text{ for all } y \in X,$$

$$\text{with } \epsilon_n \to 0^+, \text{ admits a strongly convergent subsequence.”}$$

The next geometric lemma will help us to rephrase the above property in a form which is more convenient in applications.

Lemma 5.1.24 *Let X be a Banach space and assume that $g : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is convex, lower semicontinuous, $g(0) = 0$ and $-||u|| \leqslant g(u)$ for all $u \in X$. Then there exists a $\hat{v}^* \in X^*$ with $||\hat{v}^*||_* \leqslant 1$ such that*

$$\langle \hat{v}^*, u \rangle \leqslant g(u) \text{ for all } u \in X.$$

Proof Let $\hat{g}(u) = g(u) + ||u||$ for all $u \in X$. Then $\hat{g} : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is convex, lower semicontinuous and $\hat{g}(0) = 0$. By hypothesis, $u = 0$ is a global minimizer of g. Therefore

$$\begin{aligned} 0 \in \partial g(0) &= \partial g(0) + \partial || \cdot ||(0) \text{ (see Proposition 2.7.20)} \\ &= \partial g(0) + \bar{B}_1^* \text{ with } \bar{B}_1^* = \{u^* \in X^* : ||u^*||_* \leqslant 1\}. \end{aligned}$$

So, we can find $v^* \in \partial g(0)$ with $||v^*||_* \leqslant 1$. Since $g(0) = 0$, we have

$$\langle v^*, u \rangle \leqslant g(u) \text{ for all } u \in X.$$

The proof is now complete. □

Using this lemma, we deduce at once the following result.

Proposition 5.1.25 *If $h = \varphi + \psi$ is as above, then h satisfies the GPS-condition if and only if every sequence $\{u_n\}_{n\geqslant 1} \subseteq X$ such that $h(u_n) \to c \in \mathbb{R}$ and*

$$\langle u_n^*, y - u_n \rangle \leqslant \langle \varphi'(u_n), y - u_n \rangle + \psi(y) - \psi(u_n) \text{ for all } y \in X$$

with $u_n^* \to 0$ in X^* as $n \to \infty$ admits a strongly convergent subsequence.

Remark 5.1.26 Using this proposition, we can see easily that if $\psi \equiv 0$, then the GPS-condition reduces to the usual PS-condition (see Definition 5.1.6(a)).

Proposition 5.1.27 *If $h = \varphi + \psi$ is as above, it satisfies the GPS-condition, $\{u_n\}_{n\geqslant 1}$ is a GPS-sequence such that $h(u_n) \to c \in \mathbb{R}$ and u is an accumulation point for the sequence $\{u_n\}_{n\geqslant 1}$, then $u \in K_\varphi^c$ and so in particular $K_\varphi^c \subseteq X$ is compact.*

Proof By passing to a suitable subsequence if necessary, we may assume that $u_n \to u$ in X as $n \to \infty$. We have

$$-\epsilon_n ||y - u_n|| \leqslant \langle \varphi'(u_n), y - u_n \rangle + \psi(y) - \psi(u_n) \text{ for all } y \in X, \text{ all } n \geqslant 1. \tag{5.18}$$

Since $\psi(u) \leqslant \liminf_{n\to\infty} \psi(u_n)$, passing to the limit as $n \to \infty$ in (5.18), we obtain

$$0 \leqslant \langle \varphi'(u), y - u \rangle + \psi(y) - \psi(u) \text{ for all } y \in X$$
$$\Rightarrow u \text{ is a critical point of } h \text{ (see Definition 5.1.21).}$$

Therefore, we can conclude that K_φ^c is compact. ∎

5.2 Critical Points via Minimization—The Direct Method

We start with some results that do not involve any differential structure but which are basic in the direct method of the calculus of variations. We recall the following definition.

Definition 5.2.1 Let (V, τ) be a Hausdorff topological space and $\varphi : V \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$.

(a) We say that φ is τ-lower semicontinuous (τ-lsc for short) if for every $\lambda \in \mathbb{R}$, the sublevel set $\varphi^\lambda = \{v \in V : \varphi(v) \leqslant \lambda\}$ is τ-closed.

(b) We say that φ is sequentially τ-lower semicontinuous (seq τ-lsc for short) if for every sequence $\{v_n\}_{n\geqslant 1} \subseteq V$ such that $v_n \xrightarrow{\tau} v$ in V,

$$\varphi(v) \leqslant \liminf_{n\to\infty} \varphi(v_n).$$

An easy consequence of Definition 5.2.1(a) is the following result.

Proposition 5.2.2 *If (V, τ) is a Hausdorff topological space, then*

(a) $\varphi : V \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is τ-lsc if and only if epi $\varphi = \{(v, \lambda) \in V \times \mathbb{R} : \varphi(V) \leqslant \lambda\}$ *is closed in $V \times \mathbb{R}$;*

(b) $\varphi : V \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is τ-lsc if and only if for every net $\{v_\alpha\}_{\alpha\in J}$ such that $v_\alpha \xrightarrow{\tau} v$ in V, we have

$$\varphi(v) \leqslant \liminf_{\alpha\in J} \varphi(v_\alpha);$$

(c) given any family $\{\varphi_i\}_{i\in I}$ of τ-lsc functions, we have

$$\varphi = \sup_{i\in I} \varphi_i \textit{ is } \tau\textit{-lsc too};$$

(d) if $\{\varphi_k\}_{k=1}^N$ *are* τ*-lsc functions, then* $\varphi = \sum_{k=1}^{N} \varphi_k$ *is* τ*-lsc too.*

Remark 5.2.3 In fact, the notion of sequential τ-lower semicontinuity is topological. More precisely, let τ_{seq} be the topology on V whose closed sets are the sequentially τ-closed sets. Then we have

(a) τ_{seq} is the strongest topology on V for which the converging sequences are the τ-converging sequences.
(b) φ is seq τ-lsc if and only if φ is τ_{seq}-lsc.
(c) if (V, τ) is first countable, then $\tau_{seq} = \tau$.

Evidently, $\tau \subseteq \tau_{seq}$.

Proposition 5.2.4 *Let* (V, τ) *be a compact topological space and assume that* $\varphi : V \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ *is* τ*-lsc. Then there exists a* $v_0 \in V$ *such that* $\varphi(v_0) = \inf_V \varphi$*.*

Proof Let $\{v_n\}_{n \geqslant 1} \subseteq V$ be a minimizing sequence for φ (that is, $\varphi(v_n) \downarrow \inf_V \varphi$). The compactness of V implies that we can find a subnet $\{v_\alpha\}_{\alpha \in J}$ of $\{v_n\}_{n \geqslant 1}$ such that $v_\alpha \xrightarrow{\tau} v_0 \in V$ in V. The τ-lower semicontinuity of φ implies that

$$\begin{aligned} &\varphi(v_0) \leqslant \liminf_{\alpha \in J} \varphi(v_\alpha) = \inf_V \varphi \text{ (see Proposition 5.2.2(b))} \\ \Rightarrow\ &\varphi(v_0) = \inf_V \varphi. \end{aligned}$$

The proof is now complete. □

If on the space and the function we also introduce a linear structure, then we can say more.

Proposition 5.2.5 *If* X *is a Banach space and* $\varphi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ *is a convex function, then*

(a) φ *is strongly-lsc if and only if it is weakly-lsc.*
(b) If X^* *is separable, then* φ *is weakly-lsc if and only if it is seq weakly lsc.*
(c) If $X = Y^*$ *with* Y *being a separable Banach space, then* φ *is weakly***-lsc if and only if it is seq weakly***-lsc.*

Proof (a) For every $\lambda \in \mathbb{R}$, the sublevel set $\varphi^\lambda = \{u \in X : \varphi(u) \leqslant \lambda\}$ is convex. By the Mazur theorem, a convex set is closed if and only if it is weakly lower. Then according to Definition 5.2.1(a), φ is strongly-lsc if and only if it is weakly-lsc.

(b) Since X^* is separable, every bounded set in X equipped with the weak topology is metrizable (see, for example, Dunford and Schwartz [151, p. 426]). Moreover, from the Krein–Smulian theorem (see, for example, Megginson [295, p. 243]), for every $\lambda \in \mathbb{R}$ the set φ^λ is weakly closed if and only if $\varphi^\lambda \cap t\bar{B}_1$ is weakly closed for all $t > 0$ (here $\bar{B}_1 = \{u \in X : ||u|| \leqslant 1\}$). The latter is metrizable. Therefore, φ is weakly-lsc if and only if it is sequentially weakly lsc.

(c) Since Y is separable, every bounded set in Y^* equipped with the weak*-topology is metrizable (see, for example, Dunford and Schwartz [151, p. 426]). Then the result follows as above, using this time the Krein–Smulian theorem for convex sets in X^* (see, for example, Megginson [295, p. 242]). □

Theorem 5.2.6 *Let X be a reflexive Banach space, $C \subseteq X$ be nonempty, closed convex and assume that $\varphi : C \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is convex, lsc and coercive (see Definition 5.1.10). Then there exists a $u_0 \in C$ such that $\varphi(u_0) = \inf\limits_C \varphi$.*

Proof The coercivity of φ implies that we can find $R > 0$ such that $\varphi(0) \leqslant \varphi(u)$ for all $u \in C$ with $||u|| > R$. The closed ball $\bar{B}_R = \{u \in X : ||u|| \leqslant R\}$ is w-compact (since X is reflexive). Hence so is $C \cap \bar{B}_R$. Also, $\varphi : C \cap \bar{B}_R$ is convex, lsc, hence it is also weakly lsc. Thus by the Weierstrass theorem, we can find $u_0 \in C \cap \bar{B}_R$ such that

$$\varphi(u_0) = \inf \left\{\varphi(u) : u \in C \cap \bar{B}_R\right\} = \inf_C \varphi.$$

The proof is now complete. □

Proposition 5.2.7 *If the hypotheses of Theorem 5.2.6 hold and in addition φ is strictly convex, then the minimizer $u_0 \in C$ is unique.*

Proof Let $u_0, \hat{u}_0 \in C$ be two minimizers of φ. If $u_0 \neq \hat{u}_0$, then due to the strict convexity of φ, we have

$$\varphi\left(\frac{u_0 + \hat{u}_0}{2}\right) < \frac{1}{2}\varphi(u_0) + \frac{1}{2}\varphi(\hat{u}_0) = \inf_C \varphi.$$

But this is a contradiction, since $(u_0 + \hat{u}_0)/2 \in C$ (since C is convex). □

The next proposition provides useful characterizations of the minimizer $u_0 \in C$.

Proposition 5.2.8 *If X is a Banach space, $C \subseteq X$ is nonempty, convex and $\varphi : C \to \mathbb{R}$ is convex and Gâteaux differentiable, then the following statements are equivalent:*

(a) $u_0 \in C$ *and for all* $y \in C$ *we have* $\varphi(u_0) \leqslant \varphi(y)$*;*
(b) $u_0 \in C$ *and for all* $y \in C$ *we have* $\left\langle\varphi'(u_0), y - u_0\right\rangle \geqslant 0$.

Moreover, if for all $y, h \in C$ the mapping $\lambda \mapsto \varphi'(y + \lambda(h - y))$ is continuous on $[0, 1]$*, then the above statements* (a) *and* (b) *are equivalent to*

(c) $u_0 \in C$ *and for all* $y \in C$, $\left\langle\varphi'(y), y - u_0\right\rangle \geqslant 0$.

Proof (a) $\Longrightarrow$ (b) : Since $u_0 \in C$ is a minimizer of φ on C, we have

$$0 \leqslant \varphi(u_0 + \lambda(y - u_0)) - \varphi(u_0) \text{ for all } y \in C \text{ and all } \lambda \in (0, 1] \tag{5.19}$$

(by virtue of the convexity of C, $u_0 + \lambda(y - u_0) \in C$ for all $y \in C$ and all $\lambda \in (0, 1]$). Dividing (5.19) by $\lambda > 0$ and passing to the limit as $\lambda \to 0^+$, we obtain

$$0 \leqslant \left\langle \varphi'(u_0), y - u_0 \right\rangle \text{ for all } y \in C.$$

$(b) \Longrightarrow (a)$: Since φ is convex, we have

$$\begin{aligned} &\left\langle \varphi'(u_0), y - u_0 \right\rangle \leqslant \varphi(y) - \varphi(u_0) \text{ for all } y \in C \\ &\qquad\qquad\qquad\qquad \text{(see Definition 2.7.1 and Proposition 2.7.18)} \\ \Rightarrow\ &\varphi(u_0) \leqslant \varphi(y) \text{ for all } y \in C. \end{aligned}$$

Now assume that for all $y, h \in C$, the mapping $\lambda \longmapsto \varphi'(y + \lambda(h - y))$ is continuous.

$(b) \Longrightarrow (c)$: Exploiting the monotonicity of φ' (see Example 2.6.4(c)), we have

$$\begin{aligned} &\left\langle \varphi'(u_0), y - u_0 \right\rangle \leqslant \left\langle \varphi'(y), y - u_0 \right\rangle \text{ for all } y \in C \\ \Rightarrow\ &0 \leqslant \left\langle \varphi'(y), y - u_0 \right\rangle \text{ for all } y \in C. \end{aligned}$$

$(c) \Longrightarrow (b)$: For every $\lambda \in (0, 1]$ and every $y \in C$, we have

$$\left\langle \varphi'(u_0 + \lambda(y - u_0)), u_0 + \lambda(y - u_0) - u_0 \right\rangle \geqslant 0.$$

Dividing by $\lambda \in (0, 1]$, letting $\lambda \to 0^+$ and using the hypothesis that $\lambda \longmapsto \varphi'(u_0 + \lambda(y - u_0))$ is continuous, we obtain

$$\left\langle \varphi'(u_0), y - u_0 \right\rangle \geqslant 0 \text{ for all } y \in C.$$

The proof is now complete. □

Remark 5.2.9 Suppose that $C \subseteq X$ is a vector subspace and in statement (b) let $y = u_0 + \lambda h$ with $\lambda \in \mathbb{R}$ and $h \in C$. Then $\left\langle \varphi'(u_0), h \right\rangle = 0$ for all $h \in C$. Hence $\varphi'(u_0) \in C^{\perp} \subseteq X^*$. In particular, if $C = X$, then we recover the classical Fermat rule, namely that $\varphi'(u_0) = 0$.

As an illustration of these abstract results, we consider a calculus of variations problem. So, let $\Omega \subseteq \mathbb{R}^N$ be a bounded open set and $G : \mathbb{R}^N \to \mathbb{R}$ a convex function such that

$$|G(y)| \leqslant c_1(1 + ||y||^p) \text{ for all } y \in \mathbb{R}^N, \text{ with } c_1 > 0,\ 1 < p < \infty \tag{5.20}$$

$$G(y) \geqslant c_2||y||^p - c_3 \text{ for all } y \in \mathbb{R}^N, \text{ with } c_2 > 0,\ c_3 > 0. \tag{5.21}$$

Let $h \in L^{p'}(\Omega) \left(\frac{1}{p} + \frac{1}{p'} = 1\right)$ and consider the functional $\varphi : W^{1,p}(\Omega) \to \mathbb{R}$ defined by

$$\varphi(u) = \int_{\Omega} G(Du(z))dz - \int_{\Omega} h(z)u(z)dz \text{ for all } u \in W_0^{1,p}(\Omega).$$

Proposition 5.2.10 *If $G : \mathbb{R}^N \to \mathbb{R}$ is as above, then there exists a $u_0 \in W_0^{1,p}(\Omega)$ such that $\varphi(u_0) = \inf \left\{\varphi(u) : u \in W_0^{1,p}(\Omega)\right\}$.*

Proof Evidently, φ is convex. Also, we claim that φ is continuous on $W_0^{1,p}(\Omega)$. To see this, let $u_n \to u$ in $W_0^{1,p}(\Omega)$. By passing to a subsequence if necessary, we may assume that

$$Du_n(z) \to Du(z) \text{ a.e. in } \Omega \text{ and } ||Du_n(z)|| \leqslant \gamma(z) \text{ for a.a. } z \in \Omega \text{ and all } n \geqslant 1 \tag{5.22}$$

with $\gamma \in L^p(\Omega)$. The continuity of $G(\cdot)$ and (5.20) imply that G is continuous, hence $G(Du_n(z)) \to G(Du(z))$ for almost every in Ω. Then thanks to (5.20) and (5.22) we can apply the Lebesgue dominated convergence theorem and have that

$$\int_\Omega G(Du_n)dz \to \int_\Omega G(Du)dz$$
$$\Rightarrow \varphi(u_n) \to \varphi(u), \text{ that is, } \varphi \text{ is continuous.}$$

Also, using (5.21), we have

$$\varphi(u) \geqslant c_2||Du||_p^p - c_3|\Omega|_N \text{ for all } u \in W_0^{1,p}(\Omega), \tag{5.23}$$

where $|\cdot|_N$ denotes the Lebesgue measure on $\mathbb{R}^N$. From (5.23) and Poincaré's inequality (see Theorem 1.8.1), we infer that φ is also coercive. So, we can apply Theorem 5.2.6 and obtain $u_0 \in W_0^{1,p}(\Omega)$ such that

$$\varphi(u_0) = \inf \left\{\varphi(u) : u \in W_0^{1,p}(\Omega)\right\}.$$

The proof is now complete. □

Corollary 5.2.11 *If $G : \mathbb{R}^N \to \mathbb{R}$ is in addition strictly convex, then there exists a unique $u_0 \in W_0^{1,p}(\Omega)$ such that $\varphi(u_0) = \inf \left\{\varphi(u) : u \in W_0^{1,p}(\Omega)\right\}$.*

Proof The existence of u_0 follows from Proposition 5.2.10. Note that φ is strictly convex. So, according to Proposition 5.2.7, the minimizer u_0 is unique. □

By imposing differentiability conditions on $G(\cdot)$, we can show that every minimizer u_0 is the weak solution of a related boundary value problem.

So, the new hypothesis on $G(\cdot)$ is the following:

$H : G \in C^1(\mathbb{R}^N, \mathbb{R})$, satisfies growth conditions (5.20), (5.21) and

$$||\nabla G(y)|| \leqslant c_4(1 + ||y||^{p-1}) \text{ for some } c_4 > 0, \text{ all } y \in \mathbb{R}^N. \tag{5.24}$$

Proposition 5.2.12 *If hypotheses H hold and $u_0 \in W_0^{1,p}(\Omega)$ is a minimizer of φ, then $\int_\Omega (\nabla G(Du_0), Dy)_{\mathbb{R}^N}\, dz = \int_\Omega hydz$ for all $y \in W_0^{1,p}(\Omega)$.*

Proof We have $\varphi'(u_0) = 0$, hence

$$\langle \varphi'(u_0), y \rangle = 0 \text{ for all } y \in W_0^{1,p}(\Omega).$$

Let $\xi_1 : W_0^{1,p}(\Omega) \to L^p(\Omega, \mathbb{R}^N)$ and $\xi_2 : L^p(\Omega, \mathbb{R}^N) \to \mathbb{R}$ be defined by

$$\xi_1(u) = Du \text{ for all } u \in W_0^{1,p}(\Omega) \text{ and } \xi_2(v) = \int_\Omega G(v(z)dz) \text{ for all } v \in L^p(\Omega, \mathbb{R}^N).$$

Also let $L_h : L^p(\Omega) \to \mathbb{R}$ be defined by $L_h(u) = \int_\Omega hudz$ for all $u \in W_0^{1,p}(\Omega)$. Then

$$\varphi = \xi_2 \circ \xi_1 - L_h.$$

So, by the chain rule, we have

$$\begin{aligned} &\langle \varphi'(u_0), y \rangle = \int_\Omega (\nabla G(Du_0), Dy)_{\mathbb{R}^N} - \int_\Omega hydz = 0 \text{ for all } y \in W_0^{1,p}(\Omega) \\ \Rightarrow &\int_\Omega (\nabla G(Du_0), Dy)_{\mathbb{R}^N} dz = \int_\Omega hydz \text{ for all } y \in W_0^{1,p}(\Omega). \end{aligned}$$

The proof is now complete. □

Corollary 5.2.13 *If hypotheses H hold and $u_0 \in W_0^{1,p}(\Omega)$ is a minimizer of φ, then*

$$-\operatorname{div} \nabla G(Du_0(z)) = h(z) \text{ a.e. in } \Omega, \; u_0|_{\partial\Omega} = 0. \tag{5.25}$$

Proof From (5.24) and the representation theorem for $W^{-1,p'}(\Omega) = W_0^{1,p}(\Omega)^* \left(\frac{1}{p} + \frac{1}{p'} = 1\right)$, see Theorem 1.3.9, we have that $\operatorname{div}(\nabla G(Du_0)) \in W^{-1,p'}(\Omega)$. Also by integration by parts we have

$$\begin{aligned} &\int_\Omega (\nabla G(Du_0), Dy)_{\mathbb{R}^N} dz = \langle -\operatorname{div}(\nabla G(Du_0)), y \rangle \; \text{ for all } y \in W_0^{1,p}(\Omega) \\ \Rightarrow &\int_\Omega hydz = \langle h, y \rangle = \langle -\operatorname{div}(\nabla G(Du_0)), y \rangle \; \text{ for all } y \in W_0^{1,p}(\Omega) \\ &\qquad\qquad\qquad\qquad\qquad\qquad \text{(see Proposition 5.2.12)} \\ \Rightarrow &-\operatorname{div}(\nabla G(Du_0(z))) = h(z) \text{ almost every } z \text{ in } \Omega, \; u_0|_{\partial\Omega} = 0. \end{aligned}$$

The proof is now complete. □

If $G(\cdot)$ is also convex, then the minimizers of φ are all the solutions of the boundary value problem (5.25).

Proposition 5.2.14 *If hypotheses H hold and G is in addition convex, then every weak solution of (5.25) is a minimizer of φ.*

Proof Let $u \in W_0^{1,p}(\Omega)$ be a weak solution of (5.25).

From the convexity of $G(\cdot)$, we have

$$G(w) - G(y) \geqslant (\nabla G(y), w - y)_{\mathbb{R}^N} \text{ for all } w, y \in \mathbb{R}^N. \tag{5.26}$$

Let $v \in W_0^{1,p}(\Omega)$. Then from (5.26) we have

$$\int_\Omega (G(Dv) - G(Du))dz \geqslant \int_\Omega (\nabla G(Du), Dv - Du)_{\mathbb{R}^N} dz =$$
$$\int_\Omega h(v - u)dz \text{ (see Proposition 5.2.12)}$$
$$\Rightarrow \varphi(u) \leqslant \varphi(v) \text{ for all } v \in W_0^{1,p}(\Omega) \text{ (that is, } u \text{ is a minimizer of } \varphi).$$

The proof is now complete. □

Remark 5.2.15 If $G(y) = \frac{1}{p}||y||^p$ for all $y \in \mathbb{R}^N$, with $1 < p < \infty$, then $\nabla G(y) = ||y||^{p-2}y$ for all $y \in \mathbb{R}^N$ and the differential operator in (5.25) is the p-Laplacian. If $p = 2$, we have the standard Laplace differential operator.

5.3 Deformation Theorems

In the previous section, we produced critical points of $\varphi \in C^1(X)$ which are minimizers of φ (local or global). This was done using the direct method, which requires two types of properties for the functional φ. One is quantitative and requires that the sublevel sets $\{x \in X : \varphi(x) \leqslant \lambda\}$ are compact for some natural topology on X (coercivity property). The other one is qualitative and asks that $\varphi(\cdot)$ is lower semicontinuous for the same topology on X. We note that a coercive functional is bounded below. However, in many cases the functional φ need not be bounded below. So, we need to identify different kinds of critical points.

Definition 5.3.1 Let X be a Banach space and $\varphi \in C^1(X)$. We say that $u \in X$ is a "critical point of φ" if $\varphi'(u) = 0$. Then $\varphi(u) = c \in \mathbb{R}$ is a "critical value of φ". If $c \in \mathbb{R}$ is not a critical value of φ, then we say that it is a "regular value".

Given $\varphi \in C^1(X)$ and $D \subseteq \mathbb{R}$, we introduce the following sets:

$$K_\varphi = \{u \in X : \varphi'(u) = 0\} \text{ (the critical set of } \varphi),$$
$$K_\varphi^D = \{u \in K_\varphi : \varphi(u) \in D\}.$$

If $D = \{c\}$, then $K_\varphi^D = K_\varphi^c = \{u \in K_\varphi : \varphi(u) = c\}$ (the critical points of φ at the level $c \in \mathbb{R}$). The aim of this section is to generate information about the set K_φ. This will done using the so-called "deformation method".

Definition 5.3.2 A "deformation" is a continuous map $h : [0, 1] \times X \to X$ such that

$$h(0, u) = u \text{ for all } u \in X.$$

Let $\mathscr{D} \subseteq 2^X$. We say that $\mathscr{D}$ is "deformation invariant" if for every $A \in \mathscr{D}$ and every deformation $h(\cdot, \cdot)$, $h(1, A) \in \mathscr{D}$.

To produce information about K_φ we shall use deformations of the sublevel sets $\varphi^\lambda = \{u \in X : \varphi(u) \leqslant \lambda\}$ and try to spot critical values by consideration of minimax expressions of the form

$$c = \inf_{A \in \mathscr{D}} \sup_{u \in A} \varphi(u)$$

for various deformation invariant families $\mathscr{D} \subseteq 2^X$. The construction of suitable deformations is the most technical part of this method. Basically, we look for deformations $h(t, x)$ which exhibit the following properties:

(D1) For all $a, b \in \mathbb{R}$, $a < b$ with $\varphi^{-1}([a, b]) \cap K_\varphi = \emptyset$, there exists a $t_0 > 0$ such that

$$h(t, \varphi^b) \subseteq \varphi^a \text{ for all } t \geqslant t_0.$$

(D2) If $c \in \mathbb{R}$ and U is a neighborhood of K_φ^c, then there exist $t_0 > 0$ and $a, b \in \mathbb{R}$ with $c \in (a, b)$ such that $h(t_0, \varphi^b) \subseteq \varphi^a \cup U$.

Remark 5.3.3 Roughly speaking requirement (D1) says that effectively the deformation h decreases the values of φ on $X \setminus K_\varphi$. So, nothing topologically interesting can happen between the levels $a, b \in \mathbb{R}$ if the interval $[a, b]$ does not contain any critical values of φ. Condition (D2) says that if we start a little above a critical level $c \in \mathbb{R}$, then we will either bypass the "critical" neighborhood U and end up at a "harmless" level $a \in \mathbb{R}$ or we will land on U where topologically interesting things can happen.

The deformation method constructs deformations which satisfy (D1) and (D2). These deformations are produced using a kind of steepest descent method based on the pseudogradient vector field (see Definition 5.1.1), which we know always exists if $\varphi \in C^1(X)$ (see Theorem 5.1.4).

So, let us start by recalling the basic existence and uniqueness results for the Cauchy problem in Banach spaces. First we state a local existence result (see Cartan [104, p. 122]).

Proposition 5.3.4 *If X is a Banach space, $U \subseteq X$ is a nonempty open set and $V : U \to X$ is a locally Lipschitz vector field, then the Cauchy problem*

$$u'(t) = V(u(t)), \ u(0) = x, \tag{5.27}$$

has a unique C^1-local solution $u(x)(\cdot) = u(x, \cdot)$ defined on a maximal interval $(\eta_-(x), \eta_+(x))$ containing 0. The set

$$W = \{(x, t) : x \in U,\ t \in (\eta_-(x), \eta_+(x))\}$$

is open and the map $(x, t) \mapsto u(x, t)$ is locally Lipschitz from W into X. Moreover, if for some $x \in U$, $u(x, (\eta_-(x), \eta_+(x)))$ lies in a complete subset of U, then

$$\eta_+(x) < \infty \Rightarrow \int_0^{\eta_+(x)} ||V(u(t))||dt = +\infty. \tag{5.28}$$

If we impose a sublinear growth condition on $V(\cdot)$, then we have a global solution for the Cauchy problem (5.27).

Proposition 5.3.5 *If X is a Banach space, $U \subseteq X$ is a nonempty open set, $V : U \to X$ is a locally Lipschitz vector field satisfying*

$$||V(u)|| \leqslant a + c||u|| \textit{ for all } u \in X \textit{ and some } a, c > 0,$$

then the solution is global, that is,

$$\eta_-(x) = -\infty \textit{ and } \eta_+(x) = +\infty \textit{ for all } x \in U;$$

moreover, for every $t \in \mathbb{R}$ the mapping $x \to u(x, t)$ is a homeomorphism, while $(x, t) \mapsto u(x, t)$ is locally Lipschitz and maps bounded sets to bounded sets.

Proof Arguing by contradiction, suppose that we can find $x \in U$ such that $\eta_+(x) < \infty$. Then we have

$$\begin{aligned}
& ||u(x, t)|| \leqslant ||x|| + \int_0^t ||f(u(x, s))||ds \leqslant \\
& ||x|| + a\eta_+(x) + c\int_0^t ||u(x, s)||ds \text{ for all } t \in \big[0, \eta_+(x)\big) \\
\Rightarrow\ & ||u(x, t)|| \leqslant c_2 \text{ for some } c_1 = c_1(x) > 0,\ \text{all } t \in \big[0, \eta_+(x)\big) \\
& \qquad\qquad\qquad\qquad \text{(by Gronwall's inequality)} \\
\Rightarrow\ & ||f(u(t, x))|| \leqslant c_2 \text{ for some } c_2 = c_2(x) > 0,\ \text{all } t \in \big[0, \eta_+(x)\big) \\
\Rightarrow\ & \int_0^{\eta_+(x)} ||f(u(t, s))||ds \leqslant c_2\eta_+(x) < \infty,
\end{aligned}$$

which contradicts (5.28). This proves that $\eta_+(x) = +\infty$.

Reversing the time, in a similar fashion we show that $\eta_-(x) = -\infty$.

Because of the uniqueness of the solution, we have

$$\begin{aligned}&u^{-1}(x,t)=u(x,-t)\text{ for all }t\geqslant 0\text{ and all }x\in U\\ \Rightarrow\ &u(\cdot,t)\text{ is a homeomorphism for every }t\in\mathbb{R}.\end{aligned}$$

Finally, Proposition 5.3.4 and the Gronwall inequality imply that $(x,t)\mapsto u(x,t)$ is locally Lipschitz and bounded. $\square$

Remark 5.3.6 The above result implies that for every closed $C\subseteq U$ and for every compact $T\subseteq\mathbb{R}$ we have $u(C,T)\subseteq U$ is closed. Indeed, let $\{(x_n,t_n)\}_{n\geqslant 1}\subseteq C\times T$ and assume that

$$u(x_n,t_n)\to h\in U.$$

Since T is compact, we may assume that $t_n\to t$. From Proposition 5.3.5 we have

$$\begin{aligned}&x_n=u^{-1}(u(x_n,t_n),t_n)\to u^{-1}(h,t)\in C\text{ (since }C\text{ is closed)}\\ \Rightarrow\ &h=u(u(h,-t),t)\in u(C\times T).\end{aligned}$$

Now we can state and prove the "First Deformation Theorem" (or simply the "Deformation Theorem"), which as we will see in the next section, will lead to the minimax characterization of the critical values of φ.

Theorem 5.3.7 *Let X be a Banach space and assume that $\varphi\in C^1(X)$ satisfies the C_c-condition for some $c\in\mathbb{R}$. Then for every $\epsilon_0>0$, every neighborhood U of K_φ^c (if $K_\varphi^c=\emptyset$, then $U=\emptyset$) and every $\eta>0$, we find $\epsilon\in(0,\epsilon_0)$ and a deformation $h:[0,1]\times X\to X$ such that for all $(t,u)\in[0,1]\times X$ we have*

(a) $||h(t,u)-u||\leqslant\eta(1+||u||)t$;
(b) $\varphi(h(t,u))\leqslant\varphi(u)$;
(c) $h(t,u)\neq u\Rightarrow\varphi(h(t,u))<\varphi(u)$;
(d) $|\varphi(u)-c|\geqslant\epsilon_0\Rightarrow h(t,u)=u$;
(e) $h(1,\varphi^{c+\epsilon})\subseteq\varphi^{c-\epsilon}\cup U$ *and* $h(1,\varphi^{c+\epsilon}\setminus U)\subseteq\varphi^{c-\epsilon}$.

Proof Since φ satisfies the C_c-condition, we know that K_φ^c is compact, possibly empty (see Remark 5.1.7). So, we can find $r>0$ such that $(K_\varphi^c)_{3r}=\{u\in X: d(u,K_\varphi^c)<3r\}\subseteq U$.

Claim 1. There exist $\epsilon_1\in\left(0,\frac{\epsilon_0}{2}\right)$ and $\xi>0$ such that

$$c-2\epsilon_1\leqslant\varphi(u)\leqslant c+2\epsilon_1\text{ and }u\notin(K_\varphi^c)_r\Rightarrow(1+||u||)||\varphi'(u)||_*\geqslant\xi.$$

Suppose that Claim 1 is not true. Then we can find $\{u_n\}_{n\geqslant 1}\subseteq X$ such that

$$\varphi(u_n)\to 0,\ u_n\notin(K_\varphi^c)_r,\text{ for all }n\geqslant 1\text{ and }(1+||u_n||)\varphi'(u_n)\to 0\text{ in }X^*.\tag{5.29}$$

Since φ satisfies the C_c-condition, from (5.29) it follows that we may assume $u_n\to u$ in X. Then $\varphi(u)=c,\ \varphi'(u)=0$ and $u\notin(K_\varphi^c)_r$, a contradiction. This proves Claim 1.

Let $A = \{u \in X : |\varphi(u) - c| \geqslant 2\epsilon_1\} \cap \overline{(K_\varphi^c)_r}$ and $B = \{u \in X : |\varphi(u) - c| \leqslant \epsilon_1\} \cap (X \backslash (K_\varphi^c)_{2r})$. These sets are closed and $A \cap B = \emptyset$. So, we can find a locally Lipschitz function $\vartheta : X \to [0, 1]$ such that $\vartheta|_A = 0$ and $\vartheta|_B = 1$ (for example, we may take $\vartheta(u) = \frac{d(u,A)}{d(u,A)+d(u,B)}$). Also we choose $\mu \in (0, 1)$ such that $e^\mu \leqslant \eta + 1$. Let $V : X \backslash K_\varphi \to X$ be the pseudogradient vector field produced in Theorem 5.1.4. Let

$$\gamma(u) = \begin{cases} -\xi\mu\vartheta(u)\dfrac{V(u)}{||V(u)||^2} & \text{if } |\varphi(u) - c| \leqslant 2\epsilon_1 \text{ and } u \notin (K_\varphi^c)_r \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, $\gamma : X \to X$ is a locally Lipschitz vector field and it satisfies

$$||\gamma(u)|| \leqslant \mu(1 + ||u||), \tag{5.30}$$

$$\left\langle \varphi'(u), \gamma(u) \right\rangle \leqslant -\frac{1}{4}\xi\mu\vartheta(u) \text{ (see Definition 5.1.1).} \tag{5.31}$$

We consider the following abstract Cauchy problem

$$v'(t) = \gamma(v(t)) \text{ for all } t \in [0, 1],\ v(0) = u \in X. \tag{5.32}$$

Proposition 5.3.5 implies that problem (5.32) admits a unique global C^1-solution (see (3.30)) $v(u) : [0, 1] \to X$. We set

$$h(t, u) = v(u)(t) \text{ for all } (t, u) \in [0, 1] \times X.$$

Then $(t, u) \mapsto h(t, u)$ is continuous and from the definition of $\gamma(\cdot)$ we see that, if $|\varphi(u) - c| \geqslant 2\epsilon_0$, then $h(t, u) = u$ and so we have proved statement (d) of the theorem.

For every $(t, u) \in [0, 1] \times X$, we have

$$\begin{aligned} \frac{\partial}{\partial t}\varphi(h(t, u)) &= \left\langle \varphi'(h(t, u)), \frac{\partial}{\partial t}h(t, u) \right\rangle \text{ (by the chain rule)} \\ &= \left\langle \varphi'(h(t, u)),\ \gamma(h(t, u)) \right\rangle \leqslant 0 \text{ (see (5.31), (5.32))} \\ \Rightarrow \quad \varphi(h(t, u)) &\leqslant \varphi(u) \text{ for all } t \in [0, 1] \end{aligned} \tag{5.33}$$

and so we have proved statement (b) of the theorem.

Moreover, we see that

$$\frac{\partial}{\partial t}\varphi(h(t, u)) = \left\langle \varphi'(h(t, u)), \gamma(h(t, u)) \right\rangle \leqslant -\frac{1}{4}\xi\mu\vartheta(h(t, u)) \text{ (see (5.31)).} \tag{5.34}$$

Claim 2. If $h(t, u) \neq u$, then $\vartheta(h(t, u)) > 0$.

Suppose that $\hat{u} = h(t, u) \neq u$, but $\vartheta(\hat{u}) = 0$. Then $\gamma(\hat{u}) = 0$ and $v_1(t) = \hat{u}$ for all $s \in [0, 1]$ and $v_2(s) = h(s, u)$ for all $s \in [0, 1]$ are both solutions of the Cauchy problem

$$v'(s) = \gamma(v(s)) \text{ for all } s \in [0, 1], \ v(t) = \hat{u}.$$

The uniqueness of the solution of this Cauchy problem (see Proposition 5.3.5) implies that $v_1 = v_2$ and so $h(s, u) = \hat{u}$ for all $s \in [0, 1]$. In particular, $u = h(0, u) = \hat{u}$, a contradiction. This proves Claim 2.

From (5.34), (5.33) and Claim 2, we deduce that

$$h(t, u) \neq u \Rightarrow \varphi(h(t, u)) < \varphi(u).$$

This proves statement (c) of the theorem.

Integrating the Cauchy problem (5.32), we obtain

$$\begin{aligned} ||h(t, u) - u|| &\leqslant \int_0^t ||\gamma(h(s, u))||ds \\ &\leqslant \mu \int_0^t (1 + ||h(s, u)||)ds \text{ (see (5.30))} \\ &\leqslant \mu \int_0^t ||h(s, u) - u||ds + \mu(1 + ||u||)t \\ \Rightarrow \quad ||h(t, u) - u|| &\leqslant (1 + ||u||)(e^{\mu t} - 1) \text{ (by Gronwall's inequality)} \\ &\leqslant \eta(1 + ||u||)t \text{ (since } e^{\mu t} - 1 \leqslant \eta t \text{ for } t \in [0, 1]). \end{aligned}$$

This proves statement (a) of the theorem.

Next let $\epsilon \in (0, \epsilon_1)$ and $\rho > 0$ be such that

$$\overline{(K_\varphi^c)_{2r}} \subseteq B_\rho, \ 8\epsilon \leqslant \xi\mu \text{ and } 8\eta(1 + \rho)\epsilon \leqslant \xi\mu\rho. \tag{5.35}$$

To prove statement (e) of the theorem, we argue indirectly. So, suppose we can find $u \in \varphi^{c+\epsilon}$ such that $\varphi(h(1, u)) > c - \epsilon$ and $h(1, u) \notin U$. Since $h(t, u)$ is φ-decreasing (see (5.34)), we have

$$c - \epsilon < \varphi(h(1, u)) \leqslant \varphi(h(t, u)) \leqslant \varphi(u) \leqslant c + \epsilon \text{ for all } t \in [0, 1]. \tag{5.36}$$

Also, we have

$$h([0, 1], u) \cap (K_\varphi^c)_{2r} \neq \emptyset. \tag{5.37}$$

Indeed, if (5.37) is not true, then from (5.34) we have

$$\frac{1}{4}\xi\eta \leqslant \varphi(u) - \varphi(h(1, u)) < 2\epsilon,$$

which contradicts (5.35). Therefore (5.37) holds.

In view of relation (5.37), $h(1,u) \notin U$ and $(K_\varphi^c)_{3r} \subseteq U$. So, we can find $t_1, t_2 \in [0,1]$, $t_1 \neq t_2$ (and without any loss of generality we assume $t_1 < t_2$), such that

$$d(h(t_1,u), K_\varphi^c) = 2r,\ d(h(t_2,u), K_\varphi^c) = 3r \text{ and } 2r < d(h(t,u), K_\varphi^c) < 3r \quad \text{for all } t \in (t_1, t_2)$$

$$\Rightarrow \frac{1}{4}\xi\mu(t_2 - t_1) \leqslant \varphi(h(t_1,u)) - \varphi(h(t_2,u)) < 2\epsilon \text{ (see 5.36)}$$

$$\Rightarrow r \leqslant ||h(t_2,u) - h(t_1,u)|| \leqslant \eta(1 + ||h(t_1,u)||)(t_2 - t_1) \text{ (see part (a))}$$

$$< \eta(1+\rho)\frac{8\epsilon}{\xi\mu} \leqslant r \text{ (see (5.35))},$$

a contradiction. So, statement (e) of the theorem is proved. □

Another related deformation result is the following one. Its proof is along the lines of Theorem 5.3.7 and can be found in Gasinski and Papageorgiou [182, p. 627].

Theorem 5.3.8 *Let X be a Banach space, $\hat{\varphi} \in C^1(X)$, $\hat{c} \in \mathbb{R}$, $\hat{\varphi}$ satisfies the $C_{\hat{c}}$-condition, $A, C \subseteq X$ are two disjoint closed sets, $A \cap K_{\hat{\varphi}}^{\hat{c}} = \emptyset$ and*

$$\sup_A \hat{\varphi} \leqslant \hat{c} \leqslant \inf_C \hat{\varphi}.$$

Then there exist $\epsilon > 0$ and a $\hat{\varphi}$-decreasing, locally Lipschitz, parametric family $\{h(t,\cdot)\}_{t\in[0,1]}$ of homeomorphisms such that

$$h(1,A) \subseteq \hat{\varphi}^{\hat{c}-\epsilon}$$

and $h(t,u) = u$ for all $(t,u) \in [0,1] \times \big(C \cup \big(X \backslash \hat{\varphi}^{-1}([\hat{c} - 2\epsilon,\ \hat{c} + 2\epsilon])\big)\big)$.

Remark 5.3.9 If φ is even, then for all $t \in [0,1]$, $h(t,\cdot)$ is an odd homeomorphism.

We introduce the following topological concepts, which are also important in Morse theory (see Chap. 6).

Definition 5.3.10 Let Y be a Hausdorff topological space and $A \subseteq Y$ nonempty.

(a) We say that $h : [0,1] \times Y \to Y$ is a "deformation of Y into A" if h is a deformation and $h(1,Y) \subseteq A$. We say that Y is deformable into A.
(b) We say that A is a "deformation (resp. strong deformation) retract of Y" if there exists a deformation $h : [0,1] \times Y \to Y$ into A such that

$$h(1,\cdot)|_A = \mathrm{id}|_A \text{ (resp. } h(t,\cdot)|_A = \mathrm{id}|_A \text{ for all } t \in [0,1]).$$

Remark 5.3.11 Intuitively, A is a deformation retract of Y if Y can be continuously deformed into A in such a way that points in A end up where they started. Also, A is a strong deformation retract if the elements of A remain fixed during the deformation process. Consider $S^N = \partial B_1^{N+1} = \{u \in \mathbb{R}^{N+1} : |u| = 1\}$. It is well-known that S^N is a retract of $\mathbb{R}^{N+1}\backslash\{0\}$. In fact, it is a strong deformation retract. To see this, consider the deformation $h : [0,1] \times (\mathbb{R}^{N+1}\backslash\{0\}) \to \mathbb{R}^{N+1}\backslash\{0\}$ defined by

$$h(t,u) = (1-t)u + t\frac{u}{|u|} \text{ for all } (t,u) \in [0,1] \times (\mathbb{R}^{N+1}\backslash\{0\}).$$

It is easy to see that $A \subseteq Y$ is a deformation retract if and only if A is a retract of Y and Y is deformable into A.

The next result is known in the literature as the "Second Deformation Theorem". Note that in the statement of the theorem we allow $b = +\infty$, in which case we have $\varphi^b\backslash K_\varphi^b = X$.

Theorem 5.3.12 *Let X be a Banach space and suppose that $\varphi \in C^1(X)$, $a \in \mathbb{R}$, $b \in (a, +\infty]$, φ satisfies the C_c-condition for every $c \in [a, b)$, φ has no critical values in (a, b), and $\varphi^{-1}(a)$ contains at most a finite number of critical points of φ. Then there exists a deformation $h : [0,1] \times (\varphi^b\backslash K_\varphi^b) \to \varphi^b\backslash K_\varphi^b$ of $\varphi^b\backslash K_\varphi^b$ into φ^a such that*

(a) if $u \in \varphi^a$, then $h(t,u) = u$ for all $t \in [0,1]$ (that is, φ^a is strong deformation retract of $\varphi^b\backslash K_\varphi^b$);
(b) the deformation is φ-decreasing (that is, if $s, t \in [0,1]$ with $s \leqslant t$, then

$$\varphi(h(t,u)) \leqslant \varphi(h(s,u)) \text{ for all } u \in \varphi^b\backslash K_\varphi^b\Big).$$

Proof According to Theorem 5.1.4, there exists a pseudogradient vector field $V : X\backslash K_\varphi \to X$ (see Definition 5.1.1). For $x \in \varphi^{-1}([0,b])\backslash K_\varphi^b$ we consider the following abstract Cauchy problem

$$u'(t) = -\frac{V(u(t))}{(1+||u(t)||)||V(u(t))||^2} \text{ for } t \geqslant 0,\ u(0) = x. \tag{5.38}$$

In (5.38) the vector field is locally Lipschitz and so by Proposition 5.3.4 problem (5.38) admits a unique solution $u_x(\cdot)$ defined on a maximal interval $\big[0, \eta_+(x)\big)$. Moreover, from the properties of the pseudogradient vector field, we have

$$\frac{d}{dt}\varphi(u_x(t)) = \big\langle \varphi'(u_x(t)), u_x'(t)\big\rangle \leqslant -\frac{1}{4(1+||u_x(t)||)} \text{ for all } t \in \big[0, \eta_+(x)\big). \tag{5.39}$$

The construction of the desired deformation $h(t,x)$ will be based on a series of claims.

Claim 1. If $\varphi(u_x(t(x))) = a$ for some $t(x) < \eta_+(x)$, then $t(x)$ is unique and $x \mapsto t(x)$ is continuous.

The uniqueness of $t(x)$ is a direct consequence of (5.39). The time instant $t(x)$ is characterized by

$$\varphi(u_x(t)) < a < \varphi(u_x(s)) \text{ for all } s < t(x) < t < \eta_+(x). \tag{5.40}$$

We consider $x_n \to x$ in X. Then for $\epsilon > 0$ small, we have

$$\varphi(u_{x_n}(t(x_n) + \epsilon)) < a < \varphi(u_{x_n}(t(x_n) - \epsilon)) \text{ for all } n \geqslant 1 \text{ (see (5.40))}.$$

Since $x \mapsto u_x(t)$ is continuous (see Proposition 5.3.4, continuous dependence on the initial condition), we can find $n_0 = n_0(\epsilon) \geqslant 1$ such that

$$\varphi(u_x(t(x_n) + \epsilon)) < a < \varphi(u_x(t(x_n) - \epsilon)) \text{ for all } n \geqslant n_0.$$

From Bolzano's theorem and the uniqueness of $t(x)$, it follows that

$$\begin{aligned} &|t(x_n) - t(x)| \leqslant \epsilon \text{ for all } n \geqslant n_0 \\ \Rightarrow\ & t(x_n) \to t(x), \end{aligned}$$

which shows the continuity of the map $x \mapsto t(x)$. This proves Claim 1.

For $x \in \varphi^{-1}([a, b]) \backslash K_\varphi^b$, we set $t(x) = \eta_+(x)$ if $\varphi(u_x(t)) > a$ for all $t < \eta_+(x)$.

Claim 2. If $\{x_n\}_{n \geqslant 1} \subseteq \varphi^{-1}([a, b]) \backslash K_\varphi^b$, $v \in \varphi^{-1}(a)$ and $v = \lim_{n \to \infty} u_{s_n}(x_n)$ for $0 \leqslant s_n < t(x_n)$, then for every $\{t_n\}_{n \geqslant 1}$ such that $s_n \leqslant t_n < t(x_n)$ we have

$$v = \lim_{n \to \infty} u_{x_n}(t_n).$$

Recalling that $\varphi^{-1}(a)$ contains at most a finite number of critical points of φ, we can find $\epsilon > 0$ such that

$$K \cap \bar{B}_\epsilon(v) \cap \varphi^{-1}([a, b]) \subseteq \{v\}$$

with $\bar{B}_\epsilon(v) = \{v' \in X : ||v' - v|| \leqslant \epsilon\}$ and $b_1 = \sup \varphi(\bar{B}_\epsilon(v)) < b$. We show that we can find $n_0 \in \mathbb{N}$ such that

$$u_{x_n}(t_n) \in \bar{B}_\epsilon(v) \text{ for all } n \geqslant n_0. \tag{5.41}$$

Suppose that (5.41) does not hold. Then we can find a subsequence $\{(x_{n_k}, t_{n_k})\}_{k \geqslant 1}$ of $\{(x_n, t_n)\}_{n \geqslant 1}$ such that

$$||u_{x_{n_k}}(t_{n_k}) - v|| > \epsilon \text{ for all } k \geqslant 1.$$

By hypothesis we have

$$||u_{x_{n_k}}(s_{n_k}) - v|| < \frac{\epsilon}{2} \text{ for all } k \geqslant k_0.$$

Exploiting the continuity of $u_{x_{n_k}}(\cdot)$, we can find $\tau_{n_k}, \vartheta_{n_k} \in [s_{n_k}, t_{n_k}]$ with $\tau_{n_k} < \vartheta_{n_k}$ such that

$$||u_{x_{n_k}}(\tau_{n_k}) - v|| = \frac{\epsilon}{2},\ ||u_{x_{n_k}}(\vartheta_{n_k}) - v|| = \epsilon \text{ and}$$
$$u_{x_{n_k}}(t) \in \mathbb{R} = \left\{h \in X : \frac{\epsilon}{2} \leqslant ||h - v|| \leqslant \epsilon\right\} \text{ for all } t \in [\tau_{n_k}, \vartheta_{n_k}],\ \text{all } k \geqslant k_0. \quad (5.42)$$

By hypothesis φ satisfies the C_c-condition for all $c \in [a, b)$. So, we have

$$(1 + ||x||)||\varphi'(x)||_* \geqslant \delta > 0 \text{ for all } x \in R \cap \varphi^{-1}([a, b]). \quad (5.43)$$

Therefore

$$\begin{aligned}
\frac{\epsilon}{2}||u_{x_{n_k}}(\vartheta_{n_k}) - u_{x_{n_k}}(\tau_{n_k})|| &\leqslant \int_{\tau_{n_k}}^{\vartheta_{n_k}} ||u'_{x_{n_k}}(r)||dr \\
&\leqslant 2\int_{\tau_{n_k}}^{\vartheta_{n_k}} \frac{dr}{(1 + ||x_{n_k}||)||\varphi'(u_{x_{n_k}}(r))||_*} \\
&\qquad \text{(see (5.38) and Definition 5.1.1)} \\
&\leqslant 2\frac{\vartheta_{n_k} - \tau_{n_k}}{\delta} \text{ for } k \geqslant k_0 \text{ (see (5.42), (5.43))}. \quad (5.44)
\end{aligned}$$

Note that

$$\begin{aligned}
\varphi(u_{x_{n_k}}(\vartheta_{n_k})) - \varphi(u_{x_{n_k}}(\tau_{n_k})) &= \int_{\tau_{n_k}}^{\vartheta_{n_k}} \frac{d}{dt}\varphi(u_{x_{n_k}}(t))dt \\
&= \int_{\tau_{n_k}}^{\vartheta_{n_k}} \left\langle \varphi'(u_{x_{n_k}}(t)), u'_{x_{n_k}}(t)\right\rangle dt \\
&\qquad \text{(be the chain rule)} \\
&\leqslant -\frac{1}{4}(\vartheta_{n_k} - \tau_{n_k}) \text{ for all } k \geqslant 1 \text{ (see (5.39))} \\
\Rightarrow a \leqslant \varphi(u_{x_{n_k}}(\vartheta_{n_k})) &\leqslant \varphi(u_{x_{n_k}}(\tau_{n_k})) - \frac{1}{4}(\vartheta_{n_k} - \tau_{n_k}) \\
&\leqslant \varphi(u_{x_{n_k}}(s_{n_k})) - \frac{\epsilon}{\delta} \\
&\qquad \text{(see (5.44) and recall that } s_{n_k} \leqslant \tau_{n_k}).
\end{aligned}$$

Passing to the limit as $k \to \infty$ and using our hypothesis, we obtain

$$a \leqslant a - \frac{\epsilon}{\delta},$$

a contradiction. This proves Claim 2.

Claim 3. If $x \in \varphi^{-1}([a,b])\backslash K_\varphi^b$ is such that $t(x) = \eta_+(x)$, then there exists $v = \lim\limits_{t\to\eta_+(x)} u_x(t)$ and $v \in K_\varphi^a$.

We argue indirectly. So, suppose that the Claim is not true. Recall that K_φ^a is compact (see Remark 5.1.7). So, from Claim 2 (with $x_n = x$ for all $n \geqslant 1$), we see that we cannot find a sequence $\{s_n\}_{n\geqslant 1} \subseteq \big[0, \eta_+(x)\big)$ such that $d(u_{s_n}(x), K_\varphi^a) \to 0$ as $n \to \infty$. Therefore, we can find $\epsilon > 0$ and $\delta \in (0, \eta_+(x))$ such that

$$d(u_x(t), K_\varphi^a) > \epsilon \text{ for all } t \in \big[\delta, \eta_+(x)\big)\,. \tag{5.45}$$

The set $u_X([0,\delta])$ is compact in X and $u_x([0,\delta]) \cap K_\varphi^a = \emptyset$.
Choosing $\epsilon > 0$ even smaller if necessary (see (5.45)), we have

$$u_x(t) \in \varphi^{-1}([a, \varphi(x)]) \cap \{h \in X : d(h, K_\varphi^a) \geqslant \epsilon\} \text{ for all } t \in \big[0, \eta_+(x)\big)\,.$$

This set is complete and

$$\begin{aligned} & a < \varphi(u_x(t)) \leqslant \varphi(x) - t \text{ for all } t \in \big[0, \eta_+(x)\big) \\ \Rightarrow\ & \eta_+(x) \leqslant \varphi(x) - a < +\infty \\ \Rightarrow\ & +\infty = \int_0^{\eta_+(x)} 2\frac{dt}{(1+||u_x(t)||)||\varphi'(u_x(t))||_*} \\ & \text{(see (5.38) and Definition 5.1.1).} \end{aligned}$$

This implies that we can find a sequence $t_n \to \eta_+(x)^-$ such that

$$(1 + ||u_x(t_n)||)\varphi'(u_x(t_n)) \to 0 \text{ in } X^* \text{ as } n \to \infty\,. \tag{5.46}$$

Recall that

$$\varphi(u_x(t_n)) \leqslant \hat{b} \text{ for all } n \geqslant 1 \text{ and some } \hat{b} < b. \tag{5.47}$$

By hypothesis, φ satisfies the C_c-condition for all $c \in [a, b)$. Therefore, from (5.46), (5.47) and by passing to a subsequence if necessary, we may assume that

$$u_x(t_n) \to v \text{ in } X$$

and $v \in \varphi^{-1}([a, \hat{b}]) \cap K_\varphi$. Also, we have $\varphi(v) = a$, a contradiction. This proves Claim 3.

Claim 4. If $\{x_n\}_{n\geqslant 1} \subseteq \varphi^{-1}((a,b])\backslash K_\varphi^b$, $x \in \varphi^{-1}(a)$ and $x_n \to x$, then $x = \lim\limits_{n\to\infty} u_{x_n}(s_n)$ for every sequence $\{s_n\}_{n\geqslant 1}$ such that $0 \leqslant s_n \leqslant \eta_+(x)$.

From Claim 2, we see that we can assume that $s_n = \eta_+(x_n)$ for all $n \geqslant 1$. Then we can find $t_n < \eta_+(x_n)$ such that

$$||u_{x_0}(t_n) - u_{x_n}(\eta_+(x_n))|| \leqslant \frac{1}{n} \text{ for all } n \geqslant 1. \tag{5.48}$$

Since $x_n \to x$, from Claim 2 we have $u_{x_n}(t_n) \to x$ and so

$$u_{x_n}(\eta_+(x_n)) \to x \text{ in } X \text{ (see (5.48))}.$$

This proves Claim 4.

Claim 5. If $\{x_n\}_{n\geqslant 1} \subseteq \varphi^{-1}((a,b])\backslash K_\varphi^b$, $x_n \to x \in \varphi^{-1}((a,b])\backslash K_\varphi^b$ and $t(x) = \eta_+(x)$, then for every sequence $\{t_n\}_{n\geqslant 1}$ with $0 < t_n < t(x_n)$ and $\eta_+(x) \leqslant \liminf_{n\to\infty} t_n$ we have $u_x(t(x)) = \lim_{n\to\infty} u_{x_n}(t(x_n)) = \lim_{n\to\infty} u_{x_n}(t_n)$.

We set $v = u_x(t(x))$. Let $s_1 \in (0, \eta_+(x))$ be such that

$$\begin{aligned} & u_x(s_1) \in \bar{B}_{1/2}(v) \\ \Rightarrow\ & u_{x_n}(s_1) \in \bar{B}_1(v) \text{ for all } n \geqslant n_0. \end{aligned}$$

Because $\eta_+(x) \leqslant \liminf_{n\to\infty} t_n$, inductively we can produce a sequence $\{s_k\}_{k\geqslant 1}$ such that

$$u_{x_{n_k}}(s_k) \in \bar{B}_{1/k}(v) \text{ and } s_k < t_{n_k} \text{ for all } k \geqslant 1.$$

Claim 2 implies that

$$\begin{aligned} & u_{x_{n_k}}(s_k) \to v \\ \Rightarrow\ & u_{x_{n_k}}(t_k) \to v. \end{aligned}$$

Next, let $t_n < t(x_n)$ be such that

$$||u_{x_n}(t_n) - u_{x_n}(t(x_n))|| \to 0 \text{ and } \varphi(u_{x_n}(t_n)) \to a.$$

We cannot have $\liminf_{n\to\infty} t_n < \eta_+(x)$, because in that case we will have $t_{n_k} \to \tau < \eta_+(x)$ and so $\varphi(u_x(\tau)) = a$, which contradicts the hypothesis that $t(x) = \eta_+(x)$. Hence we have $\eta_+(x) \leqslant \liminf_{n\to\infty} t_n$ and this combined with the first part of the proof of the Claim implies

$$\begin{aligned} & u_{x_n}(t_n) \to v \\ \Rightarrow\ & u_{x_n}(t(x_n)) \to v. \end{aligned}$$

So, we have proved Claim 5.

Now, for each $x \in \varphi^a$, we set $t(x)$=0 and introduce the map $\gamma : \mathbb{R}_+ \times (\varphi^b \backslash K_\varphi^b) \to \varphi^b$ defined by

$$\gamma(t,x) = \begin{cases} x & \text{if } t(x) = 0 \\ u_x(t) & \text{if } 0 \leqslant t < t(x) \\ u_x(t(x)) & \text{if } 0 < t(x) \leqslant t. \end{cases} \tag{5.49}$$

Claim 6. $\gamma(\cdot, \cdot)$ is continuous.
Let $(t_n, x_n) \to (t, x)$ in $\mathbb{R}_+ \times X$ and assume that $a \leqslant \varphi(x)$.
First suppose that $t(x) = 0$ (that is, $\varphi(x) = a$). We have

$$\gamma(t_n, x_n) = u_{x_n}(s_n) \text{ with } s_n \leqslant t(x_n).$$

From Claim 4, we have

$$\gamma(t_n, x_n) \to x = \gamma(t, x) \text{ (see (5.49))}.$$

Next, suppose that $t(x) > 0$. If $t < t(x)$, we have

$$\begin{aligned} &\varphi(u_x(t)) > a \text{ (see (5.40))} \\ \Rightarrow\ &\varphi(u_{x_n}(t_n)) > a \text{ for all } n \geqslant n_0 \text{ (see Proposition 5.3.4)} \\ \Rightarrow\ &t_n < t(x_n) \text{ for all } n \geqslant n_0 \\ \Rightarrow\ &\gamma(t_n, x_n) = u_{x_n}(t_n) \text{ for all } n \geqslant n_0 \\ \Rightarrow\ &\gamma(t_n, x_n) \to u_x(t) = \gamma(t, x) \text{ (see (5.49))}. \end{aligned}$$

Finally, suppose that $0 < t(x) \leqslant t$. If $t(x) < \eta_+(x)$, then from Claim 1 we have

$$\begin{aligned} &t(x_n) \to t(x) \\ \Rightarrow\ &\gamma(t_n, x_n) = u_{x_n}(t(x_n)) \to u_x(t(x)) = \gamma(t, x) \text{ (see (5.49))}. \end{aligned}$$

If $t(x) = \eta_+(x)$, then we use Claim 5. So, we have proved Claim 6.

From Claim 3, for every $x \in \varphi^b \backslash K_\varphi^b$, the limit $\lim\limits_{t \to +\infty} \gamma(t, x) = \hat{\gamma}(x)$ exists. We introduce $h : [0, 1] \times (\varphi^b \backslash K_\varphi^b) \to \varphi^b$ defined by

$$h(t, x) = \begin{cases} \gamma\left(\frac{t}{1-t}, x\right) & \text{if } t \in [0, 1) \\ \hat{\gamma}(x) & \text{if } t = 1. \end{cases} \tag{5.50}$$

Clearly, h is φ-decreasing and we have

$$\begin{aligned} &h(0, \cdot) = \mathrm{id}_X,\ h(t, \cdot)|_{\varphi^a} = \mathrm{id}|_{\varphi^a} \text{ (see (5.49) and (5.50))}, \\ &h(1, \varphi^b \backslash K_\varphi^b) \subseteq \varphi^a. \end{aligned}$$

So, it remains to show that $h(\cdot, \cdot)$ is continuous.

Claim 7. If $x_n \to x$ and $t_n \to +\infty$, then $\gamma(t_n, x_n) \to \hat{\gamma}(x)$.
If $t(x) = 0$, then we reason as in the corresponding part of the proof of Claim 6.
If $0 < t(x) < \eta_+(x)$, then

$$\begin{aligned} &0 < t(x_n) < 2t(x) < +\infty \text{ for all } n \geqslant n_0 \\ \Rightarrow\ &\gamma(t_n, x_n) = u_{x_n}(t_n) \to u_x(t) = \hat{\gamma}(x) \text{ (see (5.49))}. \end{aligned}$$

If $t(x) = \eta_+(x)$, then from Claim 5, we have

$$\gamma(t_n, x_n) \to u_x(t(x)) = \hat{\gamma}(x).$$

This proves Claim 7.

From Claims 6 and 7 it follows that $h(\cdot, \cdot)$ is continuous, hence it is the desired deformation. $\square$

Corollary 5.3.13 *If X is a Banach space, $a \in \mathbb{R}$, $a < b \leqslant +\infty$, $\varphi \in C^1(X)$, φ satisfies the C_c-condition for every $c \in [a, b)$ and it has no critical values in $[a, b]$, then φ^a is a strong deformation retract of φ^b.*

Now we turn our attention to functions of the form

$$j = \varphi + \psi$$

with $\varphi \in C^1(X)$ and $\psi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ convex, lower semicontinuous and not identically $+\infty$. The cone of such functions ψ is denoted by $\Gamma_0(X)$. For such functionals we introduced critical points and a compactness condition (see Definition 5.1.23 and Proposition 5.1.25).

In this section, we will prove a deformation theorem for such functionals which, in the next section, will lead to minimax theorems for the critical values of the functionals. The presence of the term ψ in the definition of j makes such functionals suitable for the use of variational methods in problems with unilateral constraints.

Our setting is the following: X is a Banach space, $\varphi \in C^1(X)$ and $\psi \in \Gamma_0(X)$ (that is, $\psi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is convex, lower semicontinuous and not identically $+\infty$). We consider the functional $j = \varphi + \psi$. Let K_j be the set of critical points of j, that is,

$$K_j = \{u \in X : 0 \in \varphi'(u) + \partial\psi(u)\},$$

where $\partial\psi(u)$ is the convex subdifferential of ψ (see Definition 2.7.1).

First we prove two auxiliary propositions.

Proposition 5.3.14 *If $j = \varphi + \psi$ as above satisfies the GPS (see Definition 5.1.23) and U is a neighborhood of $K_j^c = \{u \in K_j : j(u) = c\}$, then for every $\epsilon_0 > 0$ there exist $\epsilon \in (0, \epsilon_0)$ and some $v_0 \in X$ such that*

$$\begin{gathered} u_0 \in (X \backslash U) \cap j^{-1}([c - \epsilon, c + \epsilon]) \\ \Downarrow \\ \langle \varphi'(u_0), v_0 - u_0 \rangle + \psi(v_0) - \psi(u_0) < -3\epsilon ||v_0 - u_0||. \end{gathered}$$

Proof Arguing by contradiction, suppose we could find $\{u_n\}_{n \geqslant 1} \subseteq X \backslash U$ such that

$$j(u_n) \to c \text{ and } \langle \varphi'(u_n), v - u_n \rangle + \psi(v) - \psi(u_n) \geqslant -\frac{1}{n} ||v - u_n|| \text{ for all } v \in X. \tag{5.51}$$

Since by hypothesis j satisfies the GPS (see Definition 5.1.23), from (5.51) and Proposition 5.1.27, we can find $u \in K_j^c$ and a subsequence $\{u_{n_k}\}_{k \geqslant 1}$ of $\{u_n\}_{n \geqslant 1}$ such that

$$u_{n_k} \to u \in K_j^c \text{ in } X.$$

But $\{u_{n_k}\}_{k \geqslant 1} \subseteq X \backslash U$, hence $u \in X \backslash U$, a contradiction. $\square$

Proposition 5.3.15 *If $j = \varphi + \psi$ is as above, it satisfies the GPS, U is a neighborhood of K_j^c and $\epsilon > 0$ is as in Proposition 5.3.14, then for every $u_0 \in j^{c+\epsilon} \backslash U$ there exist $v_0 \in X$ and an open neighborhood V_0 of u_0 such that*

$$\left\langle \varphi'(u), v_0 - u \right\rangle + \psi(v_0) - \psi(u) \leqslant ||v_0 - u|| \textit{ for all } u \in V_0, \tag{5.52}$$

$$\left\langle \varphi'(u), v_0 - u \right\rangle + \psi(v_0) - \psi(u) \leqslant -3\epsilon ||v_0 - u|| \quad \textit{for all } u \in V_0 \textit{ with } h(u) \geqslant c - \epsilon. \tag{5.53}$$

Moreover, if $u_0 \in K_j$, then $v_0 = u_0$. Otherwise, v_0, V_0 and $\delta_0 > 0$ can be chosen so that $v_0 \notin \overline{V}_0$ and

$$\left\langle \varphi'(u), v_0 - u \right\rangle + \psi(v_0) - \psi(u) \leqslant -\delta_0 ||v_0 - u|| \textit{ for all } u \in V_0. \tag{5.54}$$

Proof We first deal with the case $u_0 \in K_j$. Then according to Remark 5.1.22, we have

$$\begin{aligned} 0 \leqslant \; & \left\langle \varphi'(u_0), v - u_0 \right\rangle + \psi(v) - \psi(u_0) && \text{for all } v \in X \\ \Rightarrow \; & \psi(u_0) - \psi(v) \leqslant - \left\langle \varphi'(u_0), u_0 - v \right\rangle && \text{for all } v \in X \\ \Rightarrow \; & \left\langle \varphi'(v), u_0 - v \right\rangle + \psi(u_0) - \psi(v) && \leqslant \left\langle \varphi'(v) - \varphi'(u_0), u_0 - v \right\rangle \\ & && \leqslant ||\varphi'(v) - \varphi'(u_0)||_* ||u_0 - v|| \,. \end{aligned} \tag{5.55}$$

Since $\varphi \in C^1(X)$, choosing a suitably small neighborhood V_0 of u_0 we obtain

$$\begin{aligned} & ||\varphi'(v) - \varphi'(u_0)||_* \leqslant 1 \text{ for all } v \in V_0 \\ \Rightarrow \; & \left\langle \varphi'(v), u_0 - v \right\rangle + \psi(u_0) - \psi(v) \leqslant ||u_0 - v|| \text{ for all } v \in V_0 \text{ (see (5.55))}. \end{aligned}$$

So inequality (5.52) is satisfied with $v_0 = u_0$. Since $u_0 \in K_j \cap (j^{c+\epsilon} \backslash U)$, from Corollary 5.3.13 we have

$$j(u_0) < c - \epsilon.$$

If $j(u) < c - \epsilon$ for all u in a neighborhood of u_0, then we can choose V_0 inside this neighborhood and then (5.53) is empty. So, suppose that every neighborhood of u_0 has a point v such that $j(v) \geqslant c - \epsilon$. We have

$$\psi(v) - \psi(u_0) > \varphi(u_0) - \varphi(v).$$

Then the continuity of φ implies that we can find $m > 0$ and V_0 a small neighborhood of u_0 such that

$$\psi(v) - \psi(u_0) \geqslant m > 0 \text{ for all } v \in V_0 \text{ with } h(v) \geqslant c - \epsilon. \tag{5.56}$$

Then we have

$$\begin{aligned}\left\langle \varphi'(v), u_0 - v \right\rangle + \psi(u_0) - \psi(v) &\leqslant ||\varphi'(v)||_* ||u_0 - v|| - m \text{ (see (5.56))} \\ &\leqslant -3\epsilon ||u_0 - v||\end{aligned}$$

for all $v \in V_0 =$ small neighborhood of u_0 and $h(v) \geqslant c - \epsilon$.

This proves (5.53) for the case when $u_0 \in K_j$.

Now suppose that $u_0 \notin K_j$. We first assume that $j(u_0) < c - \epsilon$. Since u_0 is not a critical point of j, we can find $v_0 \in X$ such that

$$\left\langle \varphi'(u_0), v_0 - u_0 \right\rangle + \psi(v_0) - \psi(u_0) < 0 \text{ (see Remark 5.1.22).} \tag{5.57}$$

Let $y_0 = tv_0 + (1 - t)u_0$ with $t \in (0, 1)$. Exploiting the convexity of ψ, we have

$$\begin{aligned}&\left\langle \varphi'(u_0), y_0 - u_0 \right\rangle + \psi(y_0) - \psi(u_0) \\ &\leqslant t \left[\left\langle \varphi'(u_0), v_0 - u_0 \right\rangle + \psi(v_0) - \psi(u_0) \right] < 0.\end{aligned}$$

So, by letting $t \to 0^+$, we see that we can have v_0 arbitrarily close to u_0. As in the first part of the proof we have

$$\psi(v) - \psi(u_0) \geqslant m > 0 \text{ for all } v \text{ close to } u_0 \text{ and } j(v) \geqslant c - \epsilon.$$

Then using (5.57) and V_0, and choosing small $||v_0 - u_0||$,

$$\begin{aligned}&\psi(v_0) - \psi(v) \leqslant -\frac{m}{2} \text{ for all } v \in V_0 \text{ and } j(v) \geqslant c - \epsilon \\ \Rightarrow &\left\langle \varphi'(v), v_0 - v \right\rangle + \psi(v_0) - \psi(v) \leqslant -3\epsilon ||v_0 - v|| \\ &\qquad\qquad\qquad\qquad \text{for all } v \in V \text{ and } j(v) \geqslant c - \epsilon.\end{aligned}$$

So, (5.53) is satisfied. Since $v_0 \neq u_0$, we may assume that $v_0 \notin \overline{V}_0$. Since the right-hand side in (5.57) is negative, we can find a small $\delta_0 > 0$ such that

$$\left\langle \varphi'(u_0), v_0 - u_0 \right\rangle + \psi(v_0) - \psi(u_0) < -\delta_0 ||v_0 - u_0|| \,.$$

The continuity of φ' and the lower semicontinuity of ψ imply that by shrinking V_0 further if necessary, we can have

$$\left\langle \varphi'(u), v_0 - v \right\rangle + \psi(v_0) - \psi(v) < -\delta_0 ||v_0 - v|| \text{ for all } v \in V_0.$$

Finally, we treat the case $u_0 \notin K_j$ and $c - \epsilon \leqslant j(u_0)$. Let $v_0 \in X$ be as in Proposition 5.3.14. Then as before, the continuity of φ' and the lower semicontinuity imply that we can find a small neighborhood V_0 of u_0 such that $v_0 \notin \overline{U}_0$ such that

$$\left\langle \varphi'(u), v_0 - v \right\rangle + \psi(v_0) - \psi(v) < -3\epsilon||v_0 - v|| \text{ for all } v \in V_0.$$

So, (5.53) holds and from this follows (5.54). □

Now we are ready for the deformation theorem.

Theorem 5.3.16 *Let $j = \varphi + \psi$ be as above and satisfy the GPS-condition. Assume that $U \subseteq X$ is an open neighborhood of K_j^c and ϵ_0 is a positive number. Then there exists an $\epsilon \in (0, \epsilon_0)$ such that for every compact $K \subseteq X \backslash U$ with*

$$c \leqslant \sup_{u \in K} j(u) \leqslant c + \epsilon$$

we can find $C \subseteq X$ closed with $K \subseteq \operatorname{int} C$ and a deformation $h : [0, t_0] \times C \to X$ such that

(a) $||u - h(t, u)|| \leqslant t$ *for all* $u \in C$;
(b) $j(h(t, u)) - j(u) \leqslant 2t$ *for all* $u \in C$;
(c) $j(h(t, u)) - j(u) \leqslant -2\epsilon t$ *for all* $u \in C$ *with* $j(u) \geqslant c - \epsilon$ *and*

$$\sup_{u \in K} j(h(t, u)) - \sup_{u \in K} j(u) \leqslant -2\epsilon t.$$

Moreover, if $C_0 \subseteq X$ is a closed set and $C_0 \cap K_j = \emptyset$, we can construct C and the deformation $h(t, u)$ so that

$$j(h(t, u)) - j(u) \leqslant 0 \text{ for all } u \in C \cap C_0.$$

Proof Let $\epsilon > 0$ be as postulated by Proposition 5.3.14. For every $u_0 \in K$, let V_0 be the open neighborhood of u_0 produced in Proposition 5.3.15. By shrinking V_0 further if necessary, we may assume that $V_0 \cap C_0 = \emptyset$. Since $u_0 \in K$ is arbitrary, the collection $\{V_0\}_{u_0 \in K}$ forms an open cover of K. The compactness of K implies that there is a finite subcovering $\{V_i\}_{i \in F}$. Let $u_i \in V_i$ and $v_i \in X$ be the points postulated by Proposition 5.3.15 (in that proposition, they are denoted by u_0 and v_0). We may assume that if $i_0 \in F$ and $u_{i_0} \in K$, then $d(u_{i_0}, V_i) > 0$ for all $i \in F \backslash \{i_0\}$ (if this is not the case for the finite subcovering $\{V_i\}_{i \in F}$, and for some $u_{i_0} \in K$, we choose a closed neighborhood A of u_{i_0} such that $A \subseteq V_i$ and $u_i \notin A$ for all $i \in F \backslash \{i_0\}$ and produce a new finite subcovering by deleting A from V_i, $i \neq i_0$, so we have produced a refinement of $\{V_i\}_{i \in F}$).

We consider a partition of unity $\{\vartheta_i\}_{i \in F}$ subordinate to the covering $\{V_i\}_{i \in F}$.

Let $h : \mathbb{R}_+ \times X \to X$ be defined by

$$h(t, u) = u + t \sum_{\mathrm{i} \in \mathrm{F}} \vartheta_i(u) \frac{v_i - u}{||v_i - u||}. \tag{5.58}$$

This is a continuous map for all small $t \geqslant 0$ ($t \in [0, t_0]$). We have

$$||u - h(t, u)|| \leqslant t \text{ for all } u \in X.$$

So, statement (a) of the theorem holds. We write

$$h(t, u) = u + tw \text{ with } ||w|| \leqslant 1.$$

From the mean value theorem we have

$$\begin{aligned}
\varphi(u + tw) - \varphi(u) &= \langle \varphi'(u + \tau w), tw \rangle \text{ with } 0 < \tau < t \\
&= t \left\langle \varphi'(z), \sum_{\mathrm{i} \in \mathrm{F}} \vartheta_i(u) \frac{v_i - u}{||v_i - u||} \right\rangle \text{ with } z = u + \tau w \text{ (see (5.58))} \\
&= t \sum_{\mathrm{i} \in \mathrm{F}} \frac{\vartheta_i(u)}{||v_i - u||} \langle \varphi'(z), v_i - u \rangle \\
&= t \sum_{\mathrm{i} \in \mathrm{F}} \frac{\vartheta_i(u)}{||v_i - u||} \langle \varphi'(z), v_i - z \rangle + t \sum_{\mathrm{i} \in \mathrm{F}} \frac{\vartheta_i(u)}{||v_i - u||} \langle \varphi'(z), z - u \rangle \\
&\leqslant t \sum_{\mathrm{i} \in \mathrm{F}} \frac{\vartheta_i(u)}{||v_i - u||} \langle \varphi'(z), v_i - z \rangle + t \sum_{\mathrm{i} \in \mathrm{F}} \frac{\vartheta_i(u)}{||v_i - u||} k \, ||z - u|| \\
&\qquad \text{for some } k > 0 \\
&\leqslant t \sum_{\mathrm{i} \in \mathrm{F}} \frac{\vartheta_i(u)}{||v_i - u||} \langle \varphi'(z), v_i - z \rangle + t^2 \sum_{\mathrm{i} \in \mathrm{F}} \frac{\vartheta_i(u)}{||v_i - u||} k \\
&\leqslant t \sum_{\mathrm{i} \in \mathrm{F}} \frac{\vartheta_i(u)}{||v_i - u||} \langle \varphi'(z), v_i - z \rangle + kt^2 \text{ for } t \in [0, t_0] \text{ small.}
\end{aligned} \tag{5.59}$$

Note that

$$h(t, u) = \left(1 - t \sum_{\mathrm{i} \in \mathrm{F}} \frac{\vartheta_i(u)}{||v_i - u||} \right) u + t \sum_{\mathrm{i} \in \mathrm{F}} \frac{\vartheta_i(u)}{||v_i - u||} v_i \text{ (see (5.58))}.$$

For small t ($t \in [0, t_0]$), the coefficient of u is sublinear. Then by the convexity of ψ, we have

$$\psi(h(t, u)) \leqslant \left(1 - t \sum_{\mathrm{i} \in \mathrm{F}} \frac{\vartheta_i(u)}{||v_i - u||} \right) \psi(u) + t \sum_{\mathrm{i} \in \mathrm{F}} \frac{\vartheta_i(u)}{||v_i - u||} \psi(v_i). \tag{5.60}$$

From (5.59) and (5.60), we obtain

$$
\begin{aligned}
j(h(t,u)) \quad &\leqslant j(u) + t\sum_{\mathrm{i}\in\mathrm{F}} \frac{\vartheta_i(u)}{||v_i - u||}\left[\langle\varphi'(z), v_i - z\rangle + \psi(v_i) - \psi(u)\right] + kt^2 \\
&\leqslant j(u) + t + kt^2 \text{ (see Proposition 5.3.15)} \\
&\leqslant j(u) + 2t \text{ for } t \geqslant 0 \text{ small and all } u \in C.
\end{aligned}
$$

This proves statement (b) of the theorem.

Similarly, if $u \in C$ with $j(u) \geqslant c - \epsilon$, then using Proposition 5.3.15, we have

$$
j(h(t,u)) \leqslant j(u) - 2\epsilon t \sum_{\mathrm{i}\in\mathrm{F}} \vartheta_i(u) = j(u) - 2\epsilon t.
$$

From this it follows that

$$
\begin{aligned}
&\sup_{u\in K} j(h(t,u)) - \sup_{u\in K} j(u) \leqslant \\
&\sup_{u\in K}[j(h(t,u)) - j(u)] \leqslant -2\epsilon t
\end{aligned}
$$

and this completes the proof of the theorem. □

Remark 5.3.17 If both φ and ϕ are even and K is symmetric, then $h(t,\cdot)$ can be chosen to be odd.

5.4 Minimax Theorems

In this section, we use the deformation theorems of the previous section to produce minimax characterizations of the critical values of $\varphi \in C^1(X)$.

We start by introducing a notion which is central in this theory.

Definition 5.4.1 Let Y be a Hausdorff topological space, $E_0 \subseteq E$ and D be nonempty subsets of Y and $\gamma^* \in C(E_0, Y)$. We say that the pair $\{E_0, E\}$ is "linking with D in Y via γ^*" if the following conditions are satisfied:

(a) $E_0 \cap D = \emptyset$;
(b) For any $\gamma \in C(E, Y)$ with $\gamma|_{E_0} = \gamma^*$, we have $\gamma(E) \cap D \neq \emptyset$.

Remark 5.4.2 We say that $\{E_0, E, D\}$ are linking sets in Y via γ^*. If $\gamma^* = \mathrm{id}_{E_0}$ (which is usually the case), we simply say that the sets $\{E_0, E, D\}$ are linking sets.

Next we present some illustrative examples of linking sets, which we encounter often in applications.

Example 5.4.3 (a) Let X be a Banach space and $u_0, u_1 \in X$ with $u_0 \neq u_1$. We introduce the following sets.

$$E_0 = \{u_0, u_1\},\ E = \{(1-t)u_0 + tu_1 : t \in [0,1]\} \text{ and}$$
$$D = \partial U \text{ with } U \text{ an open neighborhood of } u_0 \text{ such that } u_1 \notin \overline{U}.$$

For example, we can take $D = \partial B_\rho(u_0) = \{u \in X : ||u - u_0|| = \rho\}$ with $0 < \rho < ||u_1 - u_0||$. Then the sets $\{E_0, E, D\}$ are linking sets. Indeed, let $\gamma \in C(E, X)$ such that $\gamma(u_0) = u_0$ and $\gamma(u_1) = u_1$. The set $\gamma(E) \subseteq X$ is connected. If $\gamma(E) \cap D = \emptyset$, then $\gamma(E) = U_1 \cup U_2$ with

$$U_1 = \gamma(C) \cap U \text{ and } U_2 = \gamma(C) \cap (X \backslash \overline{U}).$$

This contradicts the connectedness of $\gamma(C)$. Therefore $\gamma(E) \cap D \neq \emptyset$ and so $\{E_0, E, D\}$ are linking sets.

(*b*) Let X be a Banach space and assume that $X = Y \oplus V$ with $\dim Y < +\infty$. We introduce the following sets

$$E_0 = \partial B_\rho(0) \cap Y,\ E = \overline{B}_\rho(0) \text{ and } D = V.$$

Here $\overline{B}_\rho(0) = \{u \in X : ||u|| \leqslant \rho\}$ and $\partial B_\rho(0) = \{u \in X : ||u|| = \rho\}$. We claim that $\{E_0, E, D\}$ are linking sets. To this end, note that since Y is finite-dimensional there exists a projection operator $P_Y \in \mathscr{L}(X)$ onto Y. Let $\gamma \in C(E, X)$ such that $\gamma|_{E_0} = \mathrm{id}|_{E_0}$. We need to show that $\gamma(E) \cap D \neq \emptyset$. To achieve this it suffices to show that $0 \in P_Y(\gamma(E))$. So, we introduce the homotopy $h : [0,1] \times Y \to Y$ defined by

$$h(t, y) = tP_Y(\gamma(y)) + (1-t)y \text{ for all } (t, y) \in [0,1] \times Y.$$

Note that $h(1, \cdot) = P_Y \circ \gamma$ and for all $t \in [0,1]$ $h(t, \cdot)|_{E_0} = \mathrm{id}|_{E_0}$. Then the homotopy invariance and normalization properties of the Brouwer degree (see Theorem 3.1.25) imply that

$$\begin{aligned} & d(P_Y \circ \gamma, B_\rho(0) \cap Y, 0) = d(\mathrm{id}_Y, B_\rho(0) \cap Y, 0) = 1 \\ \Rightarrow\ & 0 \in P_Y(\gamma(E)) \ \text{ (by the solution property of Brouwer's degree,} \\ & \hspace{18em} \text{see Theorem 3.1.25)}. \end{aligned}$$

This proves that the sets $\{E_0, E, D\}$ are linking sets.

(*c*) Let X be a Banach space and assume that $X = Y \oplus V$ with $\dim Y < \infty$. Let $v_0 \in V$ with $||v_0|| = 1$ and $0 < \rho < r_1$ and $0 < r_2$ be given. We introduce the following sets

$$\begin{aligned} E_0 &= \{y + \lambda v_0 : y \in Y,\ (0 < \lambda < r_1,\ ||y|| = r_2) \text{ or } (\lambda \in \{0, r_1\}),\ ||y|| \leqslant r_2\}, \\ E &= \{y + \lambda v_0 : y \in Y,\ 0 \leqslant \lambda \leqslant r_1,\ ||y|| \leqslant r_2\}, \\ D &= \partial B_\rho(0) \cap V. \end{aligned}$$

Note that E is a cylinder with bottom basis $\partial B_{r_2}(0) \cap Y$ and height r_1 and E_0 is the boundary of this cylinder (the lateral surface and the bottom and top bases). We claim that $\{E_0, E, D\}$ are linking sets.

So, let $\gamma \in C(E, X)$ such that $\gamma|_{E_0} = \mathrm{id}|_{E_0}$. We need to show that $\gamma(E) \cap D \neq \emptyset$. For this it suffices to show that there exists a $u \in E$ such that

$$||\gamma(u)|| = \rho \text{ and } P_Y(\gamma(u)) = 0.$$

We consider the homotopy $h : [0, 1] \times \mathbb{R} \times Y \to \mathbb{R} \times Y$ defined by

$$h(t, (\lambda, y)) = (t||\gamma(u) - P_Y(\gamma(u))|| + (1-t)\lambda - \rho,\ tP_Y(\gamma(u)) + (1-t)y)$$

for all $t \in [0, 1]$, all $\lambda \in \mathbb{R}$ and all $u = y + \lambda v_0$ with $y \in Y$. Evidently, h is continuous and we have

$$h(0, (\lambda, y)) = (\lambda - \rho, y). \tag{5.61}$$

If $u = y + \lambda v_0 \in E_0$, then for all $t \in [0, 1]$ we have

$$h(t, (\lambda, y)) = (t||u - y|| + (1-t)\lambda - \rho, y) = (\lambda - \rho, y) \neq 0.$$

Identifying E with a subset of $\mathbb{R} \times Y$ by means of the decomposition $u = y + \lambda v_0$ and exploiting the homotopy invariance of the Brouwer degree (see Theorem 3.1.25), we have

$$d(h(1, \cdot),\ \mathrm{int}\, E,\ 0) = d(h(0, \cdot),\ \mathrm{int}\, E,\ 0) = 1 \text{ (see (5.61))}.$$

So, we can find $u \in E$ such that

$$\begin{aligned} &h(1, u) = 0 \\ \Rightarrow\ &||\gamma(u)|| = \rho \text{ and } P_Y(\gamma(u)) = 0. \end{aligned}$$

Using the notion of linking sets, we can prove a general minimax principle.

Theorem 5.4.4 *If X is a Banach space, $\{E_0, E, D\}$ are closed linking sets via γ^*, $\gamma^*(E_0) \subseteq X$ is closed, $\Gamma = \{\gamma \in C(E, X) : \gamma|_{E_0} = \gamma^*\}$, $\varphi \in C^1(X)$*

$$a = \sup_{\gamma^*(E_0)} \varphi \leqslant \inf_D \varphi = b, \tag{5.62}$$

$$c = \inf_{\gamma \in \Gamma} \sup_{u \in E} \varphi(\gamma(u)) \tag{5.63}$$

and φ satisfies the C_c-condition, then $c \geqslant b$ and c is a critical value of φ. Moreover, if $c = b$, then $D \cap K_\varphi^c \neq \emptyset$.

Proof From Definition 5.4.1 we know that for every $\gamma \in \Gamma$, $\gamma(E) \cap D$. Therefore $c \geqslant d$ (see (5.62) and (5.63)).

First suppose that $c > b$ and let $\epsilon_0 = c - b > 0$. Arguing by contradiction, suppose that $K_\varphi^c = \emptyset$. According to Theorem 5.3.7 with $\epsilon_0 = c - b > 0$ and $U = \emptyset$, we can find a deformation $h : [0, 1] \times X \to X$ satisfying the conditions of Theorem 5.3.7. The choice of $\epsilon_0 > 0$ implies that

$$h(t, \cdot)|_{E_0} = \mathrm{id}|_{E_0} \text{ for all } t \in [0, 1] \text{ (see Theorem 5.3.7}(d)).$$

From (5.63) we see that we can find $\gamma \in \Gamma$ such that

$$\varphi(\gamma(u)) \leqslant c + \epsilon \text{ for all } u \in E.$$

We define $\xi = h(1, \cdot) \circ \gamma \in C(E, X)$. If $u \in E_0$, then $\xi(u) = h(1, \gamma(u)) = h(1, \gamma^*(u)) = \gamma^*(u)$ and so $\xi \in \Gamma$. Moreover, from Theorem 5.3.7(e), with $\epsilon \in (0, \epsilon_0)$ we have

$$\varphi(\xi(u)) \leqslant c - \epsilon \text{ (recall that } U = \emptyset),$$

which contradicts the definition of c. So, $K_\varphi^c \neq \emptyset$ and c is a critical value of φ.

Next, we assume that $c = b \geqslant a$. We will show that $D \cap K_\varphi^c \neq \emptyset$. Again we argue indirectly. So, suppose that $D \cap K_\varphi^c = \emptyset$. We apply Theorem 5.3.8 with

$$\hat{\varphi} = -\varphi, \ \hat{c} = -c, \ A = D \text{ and } C = \gamma^*(E_0).$$

Then we can find $\epsilon > 0$ and a parametric family $\{h(t, \cdot)\}_{t \in [0,1]}$ of homeomorphisms of X into itself, satisfying the conditions of Theorem 5.3.8. As before, using (5.63) we can find $\gamma \in \Gamma$ such that

$$\varphi(\gamma(u)) < c + \epsilon \text{ for all } u \in E. \tag{5.64}$$

Let $\xi_1 = h(1, \cdot)^{-1} \circ \gamma \in C(E, X)$. For every $u \in E_0$, we have

$$\begin{aligned} &\xi_1(u) = h(1, \cdot)^{-1}(\gamma(u)) = h(1, \cdot)^{-1}(\gamma^*(u)) = \gamma^*(u) \\ &\qquad\qquad\qquad\qquad\qquad\qquad \text{(see Theorem 5.3.8)}, \\ \Rightarrow\ & \xi_1 \in \Gamma. \end{aligned}$$

Since by hypothesis $\{E_0, E, D\}$ are linking sets, we have

$$\xi_1(E) \cap D \neq \emptyset \text{ (see Definition 5.4.1).}$$

Therefore we can find $u_0 \in C$ such that $\xi_1(u_0) \in D$. Then we have

$$\varphi(\gamma(u_0)) = \varphi(h(1, \xi_1(u_0))) \geqslant c + \epsilon \text{ (see Theorem 5.3.8)},$$

which contradicts the choice of γ (see (5.64)). Therefore we obtain

$$D \cap K_\varphi^c \neq \emptyset ,$$

which concludes the proof. $\square$

Remark 5.4.5 We stress that in (5.62), equality is permitted. This is the so-called "limiting case" of the minimax principle.

Now, with suitable choices in the linking sets, we obtain some well-known minimax theorems of critical point theory. We start with the so-called "mountain pass theorem".

Theorem 5.4.6 *Let X be a Banach space and suppose $\varphi \in C^1(X),\ u_0, u_1 \in X$ with $||u_1 - u_0|| > \rho > 0$*

$$\max\{\varphi(u_0), \varphi(u_1)\} \leqslant \inf\{\varphi(u) : ||u - u_0|| = \rho\} = m_\rho,$$

$$c = \inf_{\gamma\in\Gamma} \max_{0\leqslant t\leqslant 1} \varphi(\gamma(t)) \text{ with } \Gamma = \{\gamma \in C([0, 1], X) : \gamma(0) = u_0, \gamma(1) = u_1\}$$

and φ satisfies the C_c-condition. Then $c \geqslant m_\rho$ and c is a critical value of φ. Moreover, if $c = m_\rho$, then

$$\partial B_\rho(u_0) \cap K_\varphi^c \neq \emptyset.$$

Proof Apply Theorem 5.4.4 to the linking sets

$$E_0 = \{u_0, u_1\},\ E_1 = \{(1 - t)u_0 + tu_1 : t \in [0, 1]\} \text{ and } D = \partial B_\rho(u_0)$$

(see Example 5.4.3(a)). $\square$

The next minimax theorem is known as the "saddle point theorem".

Theorem 5.4.7 *Let X be a Banach space, $X = Y \oplus V$ with $\dim Y < +\infty$ and $\varphi \in C^1(X)$. Assume that there exists a $\rho > 0$ such that*

$$\sup_{\partial B_\rho(0)\cap Y} \varphi \leqslant \inf_V \varphi,$$
$$c = \inf_{\gamma\in\Gamma} \max_{u\in\overline{B_\rho(0)}\cap Y} \varphi(\gamma(u)),$$

where $\Gamma = \left\{\gamma \in C(\overline{B_\rho(0)} \cap Y, X) : \gamma|_{\partial B_\rho(0)\cap Y} = id|_{\partial B_\rho(0)\cap Y}\right\}$ and φ satisfies the C_c-condition. Then $c \geqslant \inf_V \varphi$ and c is a critical value of φ. Moreover, if $c = \inf_V \varphi$, then

$$V \cap K_\varphi^c \neq \emptyset .$$

Proof Apply Theorem 5.4.4 to the linking sets

$$E_0 = \partial B_\rho(0) \cap Y,\ E = \overline{B_\rho(0)} \cap Y \text{ and } D = V$$

(see Example 5.4.3(b)). □

The third minimax theorem that we present is known as the "generalized mountain pass theorem".

Theorem 5.4.8 *Let X be a Banach space and suppose $X = Y \oplus V$ with $\dim Y < \infty$, $\varphi \in C^1(X)$, $v_0 \in V$ with $||v_0|| = 1$, $0 < \rho < r_1$, $0 < r_2$,*

$$\begin{aligned}
&E_0 = \{y + \lambda v_0 : y \in Y,\ (0 < \lambda < r_1,\ ||y|| = r_2) \text{ or } (\lambda \in \{0, r_1\},\ ||y|| \leqslant r_2)\}\,,\\
&E = \{y + \lambda v_0 : y \in Y,\ 0 \leqslant \lambda \leqslant r_1,\ ||y|| \leqslant r_2\},\\
&D = \partial B_\rho(0) \cap V,\\
&\sup_{E_0} \varphi \leqslant \inf_D \varphi,\\
&c = \inf_{\gamma\in\Gamma} \max_{u\in E} \varphi(\gamma(u)),
\end{aligned}$$

where $\Gamma = \left\{\gamma \in C(E, X) : \gamma|_{E_0} = id|_{E_0}\right\}$ and φ satisfies the C_c-condition. Then $c \geqslant \inf_D \varphi$ and c is a critical value of φ. Moreover, if $c = \inf_D \varphi$, then

$$D \cap K_\varphi^c \neq \emptyset.$$

Proof Apply Proposition 3.5.4 to the linking sets $\{E_0, E, D\}$ (see Example 5.4.3 (c)). □

Next, we present a general principle that includes as a special case the mountain pass theorem (see Theorem 5.4.6). We start with a definition.

Definition 5.4.9 Let X be a Banach space.

(a) For a curve $\gamma \in C([0, 1], X)$, its "geodesic length" $l(\gamma)$ is defined by

$$l(\gamma) = \int_0^1 \frac{||\gamma'(t)||}{1 + ||\gamma(t)||} dt.$$

Then given two points $u_0, u_2 \subseteq X$, their "geodesic distance" $\delta(u_0, u_1)$ is defined by

$$\delta(u_1, u_2) = \inf \left\{l(\gamma) : \gamma \in C^1([0, 1], X) : \gamma(0) = u_0, \gamma(1) = u_1\right\}.$$

(b) A closed set $C \subseteq X$ separates two distinct points $u_0, u_1 \in X$ if the two points belong to disjoint connected components of $X \setminus C$.

Remark 5.4.10 If we consider $\gamma_0(t) = (1-t)u_0 + tu_1$, with $t \in [0,1]$ (the linear path connecting u_0 and u_1), then we see that

$$\delta(u_0, u_1) \leqslant l(\gamma_0) \leqslant ||u_1 - u_0||\,.$$

On the other hand, given any bounded set $C \subseteq X$, there is a constant $\eta_C > 0$ such that $\delta(u_1, u_0) \geqslant \eta_C ||u_1 - u_0||$ for all $u_0, u_1 \in C$. So, on bounded sets the norm metric and the geodesic metric are equivalent. If $u_0 = 0$, then the infimum in the definition of the geodesic metric is achieved on the line segment connecting $u_0 = 0$ and u_1. Therefore we have

$$\delta(0, u_1) = \int_0^1 \frac{||u_1||}{1 + t||u_1||} dt = \ln(1 + ||u_1||).$$

It follows that norm bounded and δ-bounded sets of X coincide. Consequently, any δ-Cauchy sequence is norm Cauchy and hence it converges in both the norm and the δ-metric. Therefore (X, δ) is a complete metric space. In Definition 5.4.9(b), note that the set $X \backslash C$ is open and locally connected. Therefore the connected components of $X \backslash C$ are open sets. Being connected components of a Hausdorff topological spaces, they are also closed sets (a closed partition of $X \backslash C$). So, they are closed subsets of $X \backslash C$.

Theorem 5.4.11 *Let X be a Banach space, $\varphi : X \to \mathbb{R}$ a continuous Gâteaux differentiable functional, $\varphi'_G : X \to X^*$ continuous from X with the norm topology into X^* with the w^*-topology, $u_0, u_1 \in X$,*

$$\Gamma = \{\gamma \in C([0,1], X) : \gamma(0) = u_0, \gamma(1) = u_1\},$$
$$c = \inf_{\gamma \in \Gamma} \max_{t \in [0,1]} \varphi(\gamma(t))$$

and suppose there exists a closed set $C \subseteq X$ such that $C \cap \varphi_c$ separates u_0, u_1 (see Definition 5.4.9(b)) where $\varphi_c = \{u \in X : \varphi(u) \geqslant c\}$. Then there exists a sequence $\{u_n\}_{n \geqslant 1} \subseteq X$ such that

$$\delta(u_n, c) \to 0,\ \varphi(u_n) \to c \textit{ and } (1 + ||u_n||)\varphi'_G(u_n) \to 0 \textit{ in } X^*.$$

Proof Let $C_c = C \cap \varphi_c \subseteq X$. By hypothesis this is a closed set separating u_0 and u_1. Then from Remark 5.4.10, we can find open sets U_0, U_1 partitioning $X \backslash C_c$ such that $u_0 \in U_0$ and $u_1 \in U_1$.

Let $0 < \epsilon < \frac{1}{2} \min\{0, \delta(u_0, C_c), \delta(u_1, C_c)\}$ and choose $\gamma \in \Gamma$ such that

$$\max_{0 \leqslant t \leqslant 1} \varphi(\gamma(t)) < c + \frac{\epsilon^3}{4}\,. \tag{5.65}$$

We introduce the following two nonnegative numbers

$$t_0 = \sup\{t \in [0,1] : \gamma(t) \in U_0, \delta(\gamma(t), C_c) \geqslant \epsilon\},$$
$$t_1 = \inf\{t \in [t_0,1] : \gamma(t) \in U_1, \delta(\gamma(t), C_c) \geqslant \epsilon\}.$$

We have $0 \leqslant t_0 < t_1 < 1$ and $\delta(\gamma(t), C_c) \leqslant \epsilon$ for all $t \in [t_0, t_1]$.
We introduce the following set of paths

$$\Gamma_{t_0t_1} = \{\vartheta \in C([t_0,t_1], X) : \vartheta(t_0) = \gamma(t_0),\ \vartheta(t_1) = \gamma(t_1)\}.$$

We equip $\Gamma_{t_0t_1}$ with the metric

$$d(\vartheta_1, \vartheta_2) = \max_{t_0 \leqslant t \leqslant t_1} \delta(\vartheta_1(t), \vartheta_2(t)).$$

Evidently, $(\Gamma_{t_0t_1}, d)$ is a complete metric space. Let $\psi : X \to \mathbb{R}$ and $\xi : \Gamma_{t_0t_1} \to \mathbb{R}$ be defined by

$$\psi(u) = \max\{0, \epsilon^2 - \epsilon\delta(u, C_c)\} \tag{5.66}$$
$$\text{and}\quad \xi(\vartheta) = \max_{t_0 \leqslant t \leqslant t_1} \{\varphi(\vartheta(t)) + \psi(\vartheta(t))\}. \tag{5.67}$$

We have $\vartheta(t_0) = \gamma(t_0) = u_0 \in U_0$, $\vartheta(t_1) = \gamma(t_1) = u_1 \in U_1$ and U_0, U_1 are the connected components of $X \backslash C_c$. So, there exists a $\hat{t} \in (t_0, t_1)$ such that $\vartheta(\hat{t}) \in \partial U_0 \subseteq C_c$. It follows that

$$\xi(\vartheta) \geqslant \varphi(\vartheta(\hat{t})) + \psi(\vartheta(\hat{t})) \geqslant c + \epsilon^2 \text{ for all } \vartheta \in \Gamma_{t_0t_1} \text{ (see (5.65), (5.67))}. \tag{5.68}$$

Let $\hat{\gamma} = \gamma|_{[t_0,t_1]}$. Then $\hat{\gamma} \in \Gamma_{t_0t_1}$ and we have

$$\xi(\hat{\gamma}) \leqslant \max_{0 \leqslant t \leqslant 1} [\varphi(\gamma(t)) + \psi(\gamma(t))] \leqslant \left(c + \frac{\epsilon^2}{4}\right) + \epsilon^2 \text{ (see (5.65), (5.67))}. \tag{5.69}$$

Invoking Corollary 4.6.16, we can find $\hat{\vartheta} \in \Gamma_{t_0t_1}$ such that

$$\xi(\hat{\vartheta}) \leqslant \xi(\hat{\gamma}),\ d(\hat{\vartheta}, \hat{\gamma}) \leqslant \epsilon/2, \tag{5.70}$$
$$\xi(\vartheta) \geqslant \xi(\hat{\vartheta}) - \frac{\epsilon}{2} d(\vartheta, \hat{\vartheta}). \tag{5.71}$$

Let $M \subseteq [t_0, t_1]$ be the set defined by

$$M = \{t \in [t_0, t_1] : \varphi(\hat{\vartheta}(t)) + \psi(\hat{\vartheta}(t)) = \xi(\hat{\vartheta})\} \tag{5.72}$$

(that is, M is the set of points in $[t_0, t_1]$ where the maximum in (5.67) with $\vartheta = \hat{\vartheta}$ is realized). Evidently, $M \neq \emptyset$ and is closed (hence compact). We show that $t_0, t_1 \notin M$. To see this, note that $\delta(\gamma(t_0), C_c) = \delta(\gamma(t_1), C_c) = \epsilon$ and so $\psi(\gamma(t_0)) = \psi(\gamma(t_1)) = 0$. Hence

$$
\begin{aligned}
\varphi(\hat{\vartheta}(t_k)) + \psi(\hat{\vartheta}(t_k)) &\leqslant \varphi(\gamma(t_k)) + \psi(\gamma(t_k)) \\
&< c + \frac{\epsilon^2}{4} \text{ for } k = 0, 1 \text{ (see (5.65) and (5.69))} \\
&\leqslant \xi(\hat{\vartheta}) \text{ (see (5.68))}.
\end{aligned}
$$

So, indeed we have $t_0, t_1 \notin M$.

We claim that M (see (5.71)) contains a point $t \in [t_0, t_1]$ such that

$$
(1 + ||\hat{\vartheta}(t)||)||\varphi'_G(\hat{\vartheta}(t))||_* \leqslant \frac{3\epsilon}{2}. \tag{5.73}
$$

Arguing by contradiction, suppose that there is no $t \in M$ for which (5.73) holds. Then we have

$$
(1 + ||\hat{\vartheta}(t)||)||\varphi'_G(\hat{\vartheta}(t))||_* > \frac{3\epsilon}{2} \text{ for all } t \in M. \tag{5.74}
$$

So, for every $t \in M$, we can find $h(t) \in X$ such that

$$
||h(t)|| = (1 + ||\hat{\vartheta}(t)||)^{-1}, \tag{5.75}
$$

$$
\left\langle \varphi'_G(\hat{\vartheta}(t)), h(t) \right\rangle < -\frac{3\epsilon}{2}. \tag{5.76}
$$

Recall that $\varphi'_G : X \to X^*$ is by hypothesis norm to w^*-continuous. So, we can find a neighborhood $V(t)$ of t in $[t_0, t_1]$ such that

$$
\left\langle \varphi'_G(\hat{\vartheta}(s)), h(t) \right\rangle < -\frac{3\epsilon}{2} \text{ for all } s \in V(t) \text{ (see (5.76))}. \tag{5.77}
$$

The family $\{V(t)\}_{t \in M}$ is an open cover of M and M is compact. So, there is a finite subcover $\{V(t_k)\}_{k=1}^m$. Let $\{\beta_k\}_{k=1}^m$ be a corresponding partition of unity and define

$$
\hat{h}(t) = \sum_{\mathrm{k}=1}^{m} \beta_k(t) h(t_k).
$$

Then $\hat{h} : M \to X$ is a continuous map such that

$$
\left\langle \varphi'_G(\hat{\vartheta}(t)), \hat{h}(t) \right\rangle < -\frac{3\epsilon}{2} \text{ for all } t \in M \text{ (see (5.77))}, \tag{5.78}
$$

$$
||\hat{h}(t)|| \leqslant (1 + ||\hat{\vartheta}(t)||)^{-1} \text{ for all } t \in M \text{ (see (5.75))}. \tag{5.79}
$$

Since $t_0, t_1 \notin M \subseteq [t_0, t_1]$, we extend $\hat{h}$ to a continuous map on $[t_0, t_1]$ (for notational simplicity we continue to denote this extension by $\hat{h}$) such that

$$\hat{h}(t_0) = \hat{h}(t_1) = 0, \tag{5.80}$$

$$||\hat{h}(t)|| \leqslant (1 + ||\hat{\vartheta}(t)||)^{-1} \text{ for all } t \in [t_0, t_1] \text{ (see (5.79))}. \tag{5.81}$$

Because of (5.80), we see that $\hat{\vartheta} + \lambda\hat{h} \in \Gamma_{t_0 t_1}$ for all $\lambda > 0$. So, returning to (5.71) we have

$$\xi(\hat{\vartheta} + \lambda\hat{h}) \geqslant \xi(\hat{\vartheta}) - \frac{\epsilon}{2} d(\hat{\vartheta} + \lambda\hat{h}, \hat{\vartheta}) \text{ for all } \lambda > 0. \tag{5.82}$$

Let $t_\lambda \in [t_0, t_1]$ such that

$$\xi(\hat{\vartheta} + \lambda\hat{h}) = (\varphi + \psi)(\hat{\vartheta}(t_\lambda) + \lambda\hat{h}(t_\lambda)). \tag{5.83}$$

On the other hand from (5.67) we have

$$\xi(\hat{\vartheta}) \geqslant (\varphi + \psi)(\hat{\vartheta}(t_\lambda)). \tag{5.84}$$

Returning to (5.82) and using (5.83) and (5.84), we obtain

$$\begin{aligned}
&(\varphi + \psi)(\hat{\vartheta}(t_\lambda) + \lambda\hat{h}(t_\lambda)) \geqslant (\varphi + \psi)(\hat{\vartheta}(t_\lambda)) - \frac{\epsilon}{2} d(\hat{\vartheta} + \lambda\hat{h}, \hat{\vartheta}) \text{ for all } \lambda > 0 \\
\Rightarrow\ &\varphi(\hat{\vartheta}(t_\lambda) + \lambda\hat{h}(t_\lambda)) - \varphi(\hat{\vartheta}(t_\lambda)) \geqslant -\psi(\hat{\vartheta}(t_\lambda) + \lambda\hat{h}(t_\lambda)) \\
&\qquad\qquad + \psi(\hat{\vartheta}(t_\lambda)) - \frac{\epsilon}{2} d(\hat{\vartheta} + \lambda h, \hat{\vartheta}).
\end{aligned} \tag{5.85}$$

From (5.66) it is clear that

$$|\psi(u) - \psi(v)| \leqslant \epsilon ||u - v|| \text{ for all } u, v \in X. \tag{5.86}$$

Using (5.86) in (5.85) and recalling Remark 5.4.10, we obtain

$$\begin{aligned}
&\varphi(\hat{\vartheta}(t_\lambda) + \lambda\hat{h}(t_\lambda)) - \varphi(\hat{\vartheta}(t_\lambda)) \geqslant -\frac{3\epsilon}{2} d(\hat{\vartheta} + \lambda\hat{h}, \hat{h}) \text{ for all } \lambda > 0, \\
&\Rightarrow \langle \varphi'_G(\hat{\vartheta}(t_\lambda) + \tau_\lambda \lambda\hat{h}(t_\lambda)), \hat{h}(t_\lambda) \rangle \geqslant -\frac{3\epsilon}{2} \frac{1}{\lambda} d(\hat{\vartheta} + \lambda\hat{h}, \hat{h}) \text{ for some } \tau_\lambda \in (0, 1) \\
&\qquad\qquad \text{(use the mean value theorem).}
\end{aligned} \tag{5.87}$$

We may assume $t_\lambda \to t^* \in M$ as $\lambda \to 0^+$. So, if in (5.87) we pass to the limit as $\lambda \to 0^+$, then

$$\left\langle \varphi'_G(\hat{\vartheta}(t^*)), \hat{h}(t^*) \right\rangle > -\frac{3\epsilon}{2} \max_{t \in [t_0, t_1]} \frac{||\hat{h}(t)||}{1 + ||\hat{h}(t)||} = -\frac{3\epsilon}{2} \text{ (see (5.80))}.$$

But this contradicts (5.78). So, we conclude that (5.73) holds for some $t \in M$.

From (5.70) we have $d(\hat{\vartheta}, \hat{\gamma}) \leqslant \frac{\epsilon}{2}$. Hence for $t \in (t_0, t_1) \cap M$ as in (5.73), we have

$$\delta(\hat{\vartheta}(t), C_c) \leqslant \frac{\epsilon}{2} + \delta(\gamma(t), C_c) \leqslant \frac{3\epsilon}{2}.$$

Recall that $\xi(\hat{\vartheta}) \leqslant \xi(\hat{\gamma})$ (see (5.70)) and so

$$c + \epsilon^2 \leqslant \varphi(\hat{\vartheta}(t)) + \psi(\hat{\vartheta}(t)) \leqslant c + \frac{5}{4}\epsilon^2 \text{ (see (5.68), (5.69))}. \tag{5.88}$$

If we set $u = \hat{\vartheta}(t)$, then

$$\begin{aligned}
&(1 + ||u||)||\varphi'_G(u)||_* \leqslant \frac{3\epsilon}{2} \text{ (see (5.72))},\\
&\delta(u, C_c) \leqslant \frac{3\epsilon}{2} \text{ (since } t \in (t_0, t_1)),\\
&c \leqslant \varphi(u) \leqslant c + \frac{5}{4}\epsilon^2 \text{ (see (5.88))}.
\end{aligned}$$

Let $\epsilon = \frac{1}{n}$ and let $u_n = \hat{\vartheta}(t_n)$ as above for all $n \geqslant 1$. Then this is the desired sequence in X. □

As a corollary of this theorem, we obtain the following extension of the mountain pass theorem (see Theorem 5.4.6), if we exclude the limit case (see Remark 5.4.5).

Corollary 5.4.12 *If X is a Banach space, $\varphi : X \to \mathbb{R}$ is continuous, Gâteaux differentiable, $\varphi'_G : X \to X^*$ is continuous from X with the norm topology into X^* with the w^*-topology, $u_0, u_1 \in X,\ u_0 \neq u_1$*

$$\begin{aligned}
&c = \inf_{\gamma\in\Gamma} \max_{0\leqslant t\leqslant 1} \varphi(\gamma(t))\\
\textit{with}\quad &\Gamma = \{\gamma \in C([0,1], X) : \gamma(0) = u_0, \gamma(1) = u_1\},\\
&c > \max\{\varphi(u_0), \varphi(u_1)\}
\end{aligned} \tag{5.89}$$

and φ satisfies the C_c-condition, then c is a critical value of φ.

Proof Let $C = X$ and note that (5.89) implies that φ_c separates u_0, u_1. So, we can apply Theorem 5.4.11 and produce a C_c-sequence $\{u_n\}_{n\geqslant 1} \subseteq X$. Since φ satisfies the C_c-condition we may assume that $u_n \to u$ in X. Then

$$\varphi(u) = c \text{ and } \varphi'(u) = 0.$$

The proof is now complete. □

The next corollary deals with the limiting case.

Corollary 5.4.13 *If X is a Banach space, $\varphi : X \to \mathbb{R}$ is continuous, Gâteaux differentiable, $\varphi'_G : X \to X^*$ is continuous from X with the norm topology into X^* with the w^*-topology, $u_0, u_1 \in X,\ u_0 \neq u_1$*

$$
\begin{aligned}
& c = \inf_{\gamma\in\Gamma}\max_{0\leqslant t\leqslant 1}\varphi(\gamma(t)) \\
\textit{with}\quad & \Gamma = \{\gamma \in C([0,1], X) : \gamma(0) = u_0, \gamma(1) = u_1\}, \\
& c = \max\{\varphi(u_0), \varphi(u_1)\},
\end{aligned}
$$

φ satisfies the C_c-condition and there exists a nonempty closed $C \subseteq X$ separating u_0, u_1 and $\varphi(u) \geqslant c$ for all $u \in C$, then c is a critical value of φ.

Next we will prove a result producing multiple nontrivial critical points without imposing symmetry conditions on the functional φ. The most powerful and general results concerning multiple critical points can be obtained for functionals which are invariant under a group of symmetries. This case will be examined in Sect. 5.6.

For the next multiplicity theorem, we employ a splitting condition near the origin, known as "local linking at 0".

Definition 5.4.14 Let X be a Banach space such that $X = Y \oplus V$. We say that $\varphi \in C^1(X)$ has a "local linking at 0" if there exists a $\rho > 0$ such that

$$
\begin{aligned}
&\varphi(y) \leqslant 0 \text{ if } y \in Y, ||y|| \leqslant \rho, \\
&\varphi(v) \geqslant 0 \text{ if } v \in V, ||v|| \leqslant \rho.
\end{aligned}
$$

Remark 5.4.15 Note that the local linking condition implies that

$$
\left\langle \varphi'(0), h \right\rangle = 0 \text{ for all } h \in Y \text{ and } \left\langle \varphi'(0), h \right\rangle = 0 \text{ for all } h \in V.
$$

It follows that $\left\langle \varphi'(0), h \right\rangle = 0$ for all $h \in X$ and so $\varphi'(0) = 0$, that is, $u = 0$ is a critical point of the functional φ. If $X = H$ is a Hilbert space and $\varphi \in C^2(H)$ with $\varphi(0) = 0$ and $\varphi''(0) \in \mathscr{L}(H)$ is an isomorphism, then φ has a local linking at 0 (consider the positive negative spectrum of $\varphi''(0)$).

We start with an auxiliary result which we will need in the sequel. In what follows $V : N \to X$ ($N = \{u \in X : \varphi'(u) \neq 0\}$) denotes the locally Lipschitz pseudogradient vector field produced for $\varphi \in C^1(X)$ in Theorem 5.1.4.

Lemma 5.4.16 *If X is a Banach space, $\varphi \in C^1(X)$, φ satisfies the PS-condition, $u_0 \in X$ is a unique global minimizer of φ (that is, $\varphi(u_0) < \varphi(u)$ for all $u \neq u_0$), $v \neq u_0$ is such that $\varphi'(v) \neq 0$ and φ has no critical values in $(\varphi(u_0), \varphi(v)]$, then the negative pseudogradient flow defined by*

$$
u'(t) = -\frac{V(u(t))}{||V(u(t))||^2}, \ v(0) = v, \tag{5.90}
$$

is defined on a maximal finite interval $\left[0, \eta_+(v)\right)$, $t \mapsto \varphi(u(t))$ is decreasing on this interval and $\lim\limits_{t\to\eta_+(v)^-} u(t) = u_0$.

Proof By replacing φ with $\varphi_1(u) = \varphi(u + u_0) - \varphi(u_0)$ for all $u \in X$, without any loss of generality we may assume that $u_0 = 0$ and $\varphi(u_0) = 0$.

From Proposition 5.3.4 we know that problem (5.90) has a unique C^1-flow $u(\cdot)$ existing on a maximal interval $\big[0, \eta_+(v)\big)$. We have

$$\begin{aligned}
\frac{d}{dt}\varphi(u(t)) &= \langle \varphi'(u(t)), u'(t)\rangle \text{ (by the chain rule)} \\
&= \left\langle \varphi'(u(t)), -\frac{V(u(t))}{||V(u(t))||^2}\right\rangle \text{ (see (5.90))} \\
&\leqslant -\frac{||\varphi'(u(t))||_*^2}{||V(u(t))||^2} < -\frac{1}{4} \text{ (see Definition 5.1.1).}
\end{aligned} \tag{5.91}$$

Since $\varphi(u_0) = 0$ and u_0 is a global minimizer of φ we have $\eta_+(v) \leqslant 4\varphi(v)$. Also, from (5.91) we see that

$$0 < \varphi(u(t)) < \varphi(v) \text{ for all } t \in (0, \eta_+(v)).$$

We will show that $u(t) \to 0 = u_0$ as $t \to \eta_+(v)^-$.

First suppose that $\varphi'(\cdot)$ is bounded away from zero along the flow $u(\cdot)$. So, we can find $\xi_0 > 0$ such that

$$\begin{aligned}
&||\varphi'(u(t))||_* \geqslant \xi_0 \text{ for all } t \in (0, \eta_+(v)) \\
\Rightarrow\ & \xi_0 \leqslant ||V(u(t))|| \text{ for all } t \in (0, \eta_+(v)) \text{ (see Definition 5.1.1)} \\
\Rightarrow\ & \int_0^{\eta_+(v)} ||u'(t)||dt \leqslant \frac{1}{\xi_0}\eta_+(v) \leqslant \frac{4\varphi(v)}{\xi_0} < +\infty,
\end{aligned}$$

which contradicts Proposition 5.3.4.

Therefore the map $u \mapsto \varphi'(u)$ is not bounded away from zero and so we can find $t_n \to \eta_+(v)^-$ such that

$$\varphi'(u(t_n)) \to 0 \text{ in } X^* \text{ as } n \to \infty.$$

Since $\{\varphi(u(t_n))\}_{n\geqslant 1} \subseteq [0, \varphi(v)]$ and φ satisfies the PS-condition, by passing to a suitable subsequence if necessary, we may assume that

$$u(t_n) \to \hat{u} \text{ in } X \text{ with } \varphi'(\hat{u}) = 0.$$

But by hypothesis φ has no critical values in $(0, \varphi(v)]$ (recall $u_0 = 0$, $\varphi(u_0) = 0$). Hence $\hat{u} = 0$. Therefore we conclude that

$$\varphi(u(t)) \to \varphi(0) = 0 \text{ as } t \to \eta_+(v)^-.$$

Invoking Proposition 5.1.9, we have

$$u(t) \to u_0 = 0 \text{ as } t \to \eta_+(v)^-. \qquad \square$$

Now we are ready for the multiplicity result known as the "local linking theorem".

Theorem 5.4.17 *Let X be a Banach space such that $X = Y \oplus V$ with $\dim Y < \infty$, $\varphi \in C^1(X)$, $\varphi(0) = 0$, φ is bounded below, satisfies the PS-condition, $\inf_X \varphi < 0$ and has local linking at 0 (see Definition 5.4.14). Then φ has at least two nontrivial critical points.*

Proof From Proposition 5.1.8 (see also Proposition 5.1.14), we know that there exists a $u_0 \in X$ such that

$$\varphi(u_0) = \inf_X \varphi < 0 = \varphi(0). \tag{5.92}$$

Hence $u_0 \neq 0$ and it is a critical point of φ. From the local linking hypothesis we have that $u = 0$ is another critical point of φ. So, we already have two critical points u_0 and 0 and we are looking for a third.

Case 1. $0 < \dim Y,\ 0 < \dim V$.

Without any loss of generality, we may assume that $\rho = 1 < ||u_0||$. So, if $y \in Y$ with $||y|| = 1$, we have $\varphi'(y) \neq 0$. Invoking Lemma 5.4.16, we know that the flow $u(\cdot)$ of the Cauchy problem (5.90) exists for a maximal interval $\big[0, \eta_+(y)\big)$ and $\eta_+(y) \leqslant -4\varphi(u_0)$.

For every $\lambda \in \mathbb{R}$, let

$$\overset{\circ}{\varphi}{}^{\lambda} = \{u \in X : \varphi(u) < \lambda\}.$$

We claim that we may assume that there exists a $\delta > 0$ such that

$$\overset{\circ}{\varphi}{}^{\varphi(u_0)+\delta} \subseteq B_{\frac{||u_0||}{2}}(u_0) = \left\{u \in X : ||u - u_0|| < \frac{||u_0||}{2}\right\}. \tag{5.93}$$

If no such $\delta > 0$ exists, then from (5.92) we see that we can find a minimizing sequence $\{u_n\}_{n \geqslant 1} \subseteq X$ for φ such that

$$||u_n - u_0|| \geqslant \frac{||u_0||}{2} \text{ for all } n \geqslant 1.$$

From Corollary 4.6.16, we know that we can find another minimizing sequence $\{v_n\}_{n \geqslant 1} \subseteq X$ for φ such that

$$\varphi'(v_n) \to 0 \text{ in } X^* \text{ and } ||v_n - u_0|| \geqslant \frac{||u_0||}{4} \text{ for all } n \geqslant 1. \tag{5.94}$$

Since by hypothesis φ satisfies the PS-condition, we may assume that

$$v_n \to \hat{u} \text{ in } X.$$

Then $\varphi'(\hat{u}) = 0$ and $\varphi(\hat{u}) = \inf_X \varphi < 0 = \varphi(0)$. Hence $\hat{u} \neq 0$ and from (5.94) we see that $\hat{u} \neq u_0$. This is a second nontrivial critical point of φ and so we are done. Therefore, we may assume that we can find $\delta > 0$ such that (5.93) holds.

In (5.93) we choose $\delta > 0$ small such that we can find a unique $t(y) \in (0, \eta_+(y))$ for which we have

$$\varphi(u(t(y))) = \varphi(u_0) + \delta \leqslant 0. \tag{5.95}$$

The uniqueness is a consequence of (5.91).

Let $v_0 \in V$ with $||v_0|| = 1$. We introduce the set

$$E = \{u = y + \lambda v_0 : y \in Y, \lambda \geqslant 0, ||u|| \leqslant 1\} \text{ (a half ball)}.$$

Then ∂E is a hemisphere with a basis at the equator. We consider $\gamma_0 \in C(\partial E, X)$ defined as follows

$$\begin{aligned} &\gamma_0(y) = y \text{ for all } y \in Y \text{ with } ||y|| \leqslant 1 \\ &\qquad\qquad\qquad\qquad \text{(value of } \gamma_0 \text{ at the equator)}, \\ &\gamma_0(v_0) = u_0. \end{aligned}$$

If $u \in \partial E$ with $u \neq v_0$, then $||u|| = 1$ and it can be uniquely written as

$$u = \mu y + \lambda v_0$$

with $\lambda \in [0, 1]$, $y \in Y$, $||y|| = 1$, $\mu \in (0, 1]$ (μ, λ and y are unique). Then we set

$$\gamma_0(\mu y + \lambda v_0) = u(2\lambda t(y)) \text{ for all } \lambda \in \left[0, \frac{1}{2}\right].$$

So, we have

$$\gamma_0\left(\mu y + \frac{1}{2} v_0\right) = u(t(y)).$$

From (5.93) and (5.95) it follows that

$$\left\| \mu y + \frac{1}{2} v_0 - u_0 \right\| \leqslant \frac{||u_0||}{2}.$$

Finally we set

$$\gamma_0(\mu y + \lambda v_0) = (2\lambda - 1)u_0 + (2 - 2\lambda)u(t(y)) \text{ for all } \lambda \in \left[\frac{1}{2}, 1\right].$$

As λ moves from $\frac{1}{2}$ to 1, γ_0 spans the line segment from $u(t(y))$ to u_0 and so we have

$$||\gamma_0(\mu y + \lambda v_0) - u_0|| \leqslant \frac{||u_0||}{2} \text{ for all } \lambda \in \left[\frac{1}{2}, 1\right].$$

Clearly γ_0 is continuous and

$$\varphi|_{\gamma_0(\partial E)} \leqslant 0 \text{ (see (5.95))}.$$

Moreover, for some $r \leqslant 1$, we have

$$||\gamma_0(u)|| \geqslant r > 0 \text{ for all } ||u|| = 1.$$

We fix $\vartheta \in (0, r)$. As in Example 5.4.3(c), using a degree theoretic argument, we can see that $\{E_0 = \partial E, E, D = V \cap \partial B_\vartheta(0)\}$ are linking sets via γ_0.

Let $\Gamma = \{\gamma \in C(E, X) : \gamma|_{\partial E} = \gamma_0\}$ and set

$$c = \inf_{\gamma \in \Gamma} \max_{u \in E} \varphi(\gamma(u)) \geqslant 0 \text{ (see Theorem 5.4.4)}.$$

Invoking Theorem 5.4.4 we have that c is a critical value of φ. If $c > 0$, then it corresponds to a critical point of φ, distinct from u_0 and 0. If $c = 0$, then from Theorem 5.4.4 we know that

$$K_\varphi^c \cap D \neq \emptyset$$

and so again we produce a second nontrivial critical point for the functional φ. Therefore, in Case 1, it follows that there is a second nontrivial critical point of φ.

Case 2. $\dim Y = 0$.

From the beginning of the proof we already have a global minimizer $u_0 \in X,\ u_0 \neq 0$, of φ. By hypothesis, $u = 0$ is a local minimizer of φ and we may assume that it is a strict local minimizer or otherwise we have a whole sequence of distinct nontrivial critical points of φ. So, we can find $\rho_0 \in (0, ||u_0||)$ such that

$$\varphi(u) \geqslant m_0 > 0 \text{ for all } ||u|| = \rho_0. \tag{5.96}$$

Also, we have

$$\varphi(u_0) < 0 = \varphi(0). \tag{5.97}$$

Since φ satisfies the PS-condition, relations (5.96) and (5.97) permit the use of the mountain pass theorem (see Theorem 5.4.6) and so we have a second nontrivial critical point for φ and we are done. This takes care of Case 2.

Case 3. $\dim V = 0$ (in this case we may allow $\dim Y = +\infty$).

Again due to the local linking hypothesis, we may assume that $u = 0$ is a strict local minimizer of φ. So, we can find $\rho_0 > 0$ such that

$$\varphi(u) \leqslant -m_0 < 0 \text{ for all } ||u|| = \rho_0.$$

As before (see Case 2), applying the mountain pass theorem (see Theorem 5.4.6) to the functional $-\varphi$, we produce a second nontrivial critical point for φ. $\square$

We have a useful particular case of this theorem in the setting of Hilbert spaces and C^2-functionals. We start with a definition and a result, which is our first encounter with Morse theory, which will be the topic of investigation in Chap. 6.

Definition 5.4.18 Let H be a Hilbert space, $\varphi \in C^2(H)$ and u a critical point of φ

(a) We say that u is "nondegenerate" if $\varphi''(u) \in \mathscr{L}(H)$ is invertible.
(b) The "Morse index" of u is defined to be the supremum of the dimensions of all the vector subspaces of H on which $\varphi''(u)$ is negative definite.

The next result is known as the "Morse Lemma".

Proposition 5.4.19 *If H is a Hilbert space, $\varphi \in C^2(H)$ and $u = 0$ is a nondegenerate critical point of φ, then there exists a Lipschitz continuous homeomorphism h of a neighborhood W of 0 onto a neighborhood U of 0 such that*

$$h(0) = 0 \textit{ and } \varphi(h(u)) = \varphi(0) + \frac{1}{2}(\varphi''(0)u, u)_H \textit{ for all } u \in W.$$

Proof We consider the function

$$G(t, u) = (1-t)\left[\varphi(0) + \frac{1}{2}(\varphi''(0)u, u)_H\right] + t\varphi(u) \text{ for all } (t, u) \in [0, 1] \times H.$$

Using G, we introduce the vector field $g(t, u)$ defined by

$$g(t, u) = \begin{cases} -G_t'(t, u)\frac{G_u'(t,u)}{||G_u'(t,u)||^2} & \text{if } u \neq 0 \\ 0 & \text{if } u = 0. \end{cases} \tag{5.98}$$

We claim that for a suitable neighborhood B_r of the origin $g(t, \cdot)|B_r$ is Lipschitz. To this end let

$$\psi(u) = \varphi(u) - \varphi(0) - \frac{1}{2}(\varphi''(0)u, u)_H.$$

We see that $\psi(0) = 0$, $\psi'(0) = 0$ and $\psi''(0) = 0$. Therefore

$$\psi(u) = \int_0^1 (1-t)(\varphi''(tu)u, u)_H dt \text{ and } \psi'(u) = \int_0^1 \psi''(tu)u\, dt.$$

So, for every $\epsilon > 0$, there exists a $\delta(\epsilon) > 0$ such that

$$||u|| \leqslant \delta(\epsilon) \Rightarrow |\psi(u)| \leqslant \epsilon||u||^2 \text{ and} ||\psi'(u)|| \leqslant \epsilon||u||\,. \tag{5.99}$$

Since $u = 0$ is a nondegenerate critical point of φ, there is a $c \geqslant 1$ such that

$$\frac{||u||}{c} \leqslant ||\varphi''(0)(u)|| \leqslant c||u|| \text{ for all } u \in H. \tag{5.100}$$

Note that for $u \neq 0$, we have

$$g(t,u) = -\psi(u)\frac{\varphi''(0)u + t\psi'(u)}{||\varphi''(0)u + t\psi'(u)||^2}.$$

Let $\epsilon = \frac{1}{2c}$ and use (5.99) and (5.100) to infer that

$$||u|| \leqslant \delta(\epsilon) \Rightarrow |g(t,u)| \leqslant 2c(c+\epsilon)\epsilon||u||. \tag{5.101}$$

Clearly $g(t,0) = 0$. So, from (5.101) we infer that $g(t,\cdot)$ is continuous at $u = 0$ and then from (5.98) it follows that g is continuous. Let $r \in (0, \delta(\epsilon))$ be such that

$$||u|| \leqslant r \Rightarrow ||\psi'(u)|| \leqslant 1. \tag{5.102}$$

From (5.99), (5.100) and (5.102), we see that we can find $\hat{c} > 0$ such that

$$||g'_u(t,u)|| \leqslant \hat{c} \text{ for all } u \neq 0,\ ||u|| \leqslant r.$$

Then the mean value theorem implies

$$||g(t,u) - g(t,u')|| \leqslant c_0||u - u'|| \text{ for all } u, u' \in B_r \text{ for some } c_0 > 0.$$

We consider the following abstract Cauchy problem

$$\xi'(t) = g(t, \xi(t)),\ \xi(0) = v. \tag{5.103}$$

Proposition 5.3.4 implies that problem (5.103) admits a unique solution $\xi(t) = \xi(v)(t)$ for all $v \in W =$ neighborhood of the origin. We have

$$\begin{aligned}
&\frac{d}{dt}G(t,\xi(t)) = G'_t(t,\xi(t)) + (G'_u(t,\xi(t)), \xi'(t))_H = 0 \text{ (see (5.88) and (5.103))}\\
&\Rightarrow G(0,v) = G(1,\xi(1))\\
&\Rightarrow \varphi(0) + \frac{1}{2}(\varphi''(0)v, v)_H = \varphi(\xi(v)(1)).
\end{aligned}$$

Setting $h(v) = \xi(v)(1)$ we have the desired homeomorphism. ☐

Let H_+ be the subspace of H such that $\varphi''(0)|_{H_+} > 0$ and H_- be the subspace of H such that $\varphi''(0)|_{H_-} < 0$. Recalling that $u = 0$ is a nondegenerate critical point of φ, we have

$$H = H_- \oplus H_+.$$

Let P_+ be the orthogonal projection onto H_+. If on H we use the equivalent norm $|\cdot|$ defined by

$$|u| = |(\varphi''(0)u, u)_H| \text{ for all } u \in H \text{ (see (5.100))},$$

then we see that we can equivalently rewrite the conclusion of the Proposition 3.4.19 as

$$\varphi(h(u)) = \varphi(0) + \frac{1}{2}|P_+(u)|^2 + \frac{1}{2}|(I - P_+)u|^2.$$

Then it is clear that the splitting of local linking (see Definition 5.4.14) holds and so using Theorem 5.4.17, we have the following multiplicity result.

Theorem 5.4.20 *If H is a Hilbert space, $\varphi \in C^2(H)$ is bounded from below, satisfies the PS-condition, $u = 0$ is a nondegenerate critical point of φ with finite Morse index and* $\inf \varphi < 0 = \varphi(0)$*, then φ has at least two nontrivial critical points.*

We have two generalizations of Theorem 5.4.17. In the first we drop the requirement that $\dim Y < \infty$. In the second, we see what happens if φ is not bounded below. To achieve these generalizations, we need some preparation.

So, let X be a Banach space and assume that

$$X = Y \oplus V.$$

Let $\{Y_n\}_{n\geqslant 1}$ and $\{V_n\}_{n\geqslant 1}$ be increasing sequences of subspaces of Y and V respectively such that

$$Y = \overline{\bigcup_{n\geqslant 1} Y_n} \text{ and } V = \overline{\bigcup_{n\geqslant 1} V_n}.$$

Given a multi-index $\alpha = (\alpha_1, \alpha_2) \in \mathbb{N}^2$, and $\varphi \in C^1(X)$, we set

$$X_\alpha = Y_{\alpha_1} \oplus V_{\alpha_2} \text{ and } \varphi_\alpha = \varphi|_{X_\alpha}.$$

On $\mathbb{N}^2$ we consider the coordinate ordering, that is,

$$\alpha \leqslant \alpha' \text{ if and only if } \alpha_1 \leqslant \alpha_1' \text{ and } \alpha_2 \leqslant \alpha_2'.$$

Let $\alpha_n = (i_n, j_n)$, $n \geqslant 1$, be a sequence of multi-indices. If $i_n \to \infty$, $j_n \to \infty$, we say that the sequence is "admissible".

The next definition generalizes the PS-condition.

Definition 5.4.21 Let $\varphi \in C^1(X)$ and $c \in \mathbb{R}$. We say that φ satisfies the PS_c^*-condition if every sequence $\{u_{\alpha_n}\}_{n\geqslant 1} \subseteq X$ with $\{\alpha_n\}_{n\geqslant 1} \subseteq \mathbb{N}^2$ admissible such that

$$u_{\alpha_n} \in X_{\alpha_n},\ \varphi(u_{\alpha_n}) \to c \text{ and } \varphi'_{\alpha_n}(u_n) \to 0 \text{ in } X^*$$

admits a strongly convergent subsequence. If φ satisfies the PS_c^*-condition for every $c \in \mathbb{R}$, then we say that φ satisfies the PS^*-condition.

Remark 5.4.22 If $Y_n = X$ and $V_n = \{0\}$ for all $n \geqslant 1$, then Definition 5.4.21 reduces to the usual PS-condition (local and global, see Definition 5.1.6(a)).

Using this notation, we obtain the two generalizations of Theorem 5.4.17 method earlier. The results are due to Li and Willem [269], where the interested reader can find their proofs.

Theorem 5.4.23 *Let $\varphi \in C^1(X)$ be a functional that satisfies the following conditions:*

(i) φ has a local linking at 0;
(ii) φ satisfies the PS^*-condition;
(iii) φ is bounded (that is, maps bounded sets to bounded sets);
(iv) φ is bounded below and $c = \inf\limits_X \varphi < 0$.

Then φ has at least two nontrivial critical points.

Theorem 5.4.24 *Let $\varphi \in C^1(X)$ be a functional that satisfies the following conditions:*

(i) φ has a local linking at 0 and $Y \neq \{0\}$;
(ii) φ satisfies the PS^*-condition;
(iii) φ is bounded (that is, maps bounded sets to bounded sets);
(iv) for every $m \geqslant 1$, if $u \in Y_m \oplus V$ and $||u|| \to \infty$, then $\varphi(u) \to -\infty$.

Then φ has at least one nontrivial critical point.

Now we turn our attention to functionals of the form

$$j = \varphi + \psi,$$

where $\varphi \in C^1(X)$ and $\psi : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is convex, lower semicontinuous and not identically $+\infty$ (that is, $\psi \in \Gamma_0(X)$).

We start with a general minimax principle.

Theorem 5.4.25 *Let $j = \varphi + \psi$ be as above and assume that j satisfies the GPS-condition (see Definition 5.1.23). Suppose that $\{E_0 = \partial E, E, D\}$ are linking sets via $\gamma^* \in C(\partial E, X)$, with C compact and that the following conditions are fulfilled:*

$$\sup_{E_0} j < \inf_D j,$$
$$+\infty > c = \inf_{\gamma \in \Gamma} \max_{u \in C} j(\gamma(u)), \text{ where } \Gamma = \{\gamma \in C(E, X) : \gamma|_{E_0 = \partial E} = \gamma^*\}.$$

Then $c \geqslant \inf\limits_D j$ and c is a critical value of j (see Definition 5.1.21).

Proof From Definition 5.4.1 and the minimax expression for c, we have

$$c \geqslant \inf_D j.$$

On Γ we consider the supremum metric d_∞ defined by

$$d_\infty(\gamma, \xi) = \max_{u \in E} ||\gamma(u) - \xi(u)|| \text{ for all } \gamma, \xi \in \Gamma.$$

Then (Γ, d_∞) is a complete metric space. Let $\eta : \Gamma \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ be defined by

$$\eta(\gamma) = \sup_{u \in E} j(\gamma(u)).$$

Claim 1. η is lower semicontinuous.

We need to show that for every $\mu \in \mathbb{R}$, the set

$$\eta^\mu = \{\gamma \in \Gamma : \eta(\gamma) \leqslant \mu\}$$

is d_∞-closed. So, let $\{\gamma_n\}_{n \geqslant 1} \subseteq \eta^\mu$ and assume that $\gamma_n \xrightarrow{d_\infty} \gamma$. Then for all $v \in C$ we have

$$j(\gamma_n(v)) \leqslant \eta(\gamma_n) \leqslant \mu \text{ for all } n \geqslant 1.$$

Note that $\gamma_n(v) \to \gamma(v)$ and j is lower semicontinuous. So,

$$\begin{aligned} & j(\gamma(v)) \leqslant \liminf_{n \to \infty} j(\gamma_n(v)) \leqslant \mu \text{ for all } v \in C \\ \Rightarrow\ & \eta(\gamma) \leqslant \mu, \text{ hence } \gamma \in \eta^\mu. \end{aligned}$$

This proves Claim 1.

Let $\operatorname{dom} \eta = \{\gamma \in \Gamma : \eta(\gamma) < +\infty\}$ (the effective domain of η).

Claim 2. For every $\gamma \in \operatorname{dom} \eta$, $j \circ \gamma$ is continuous on E.

Note that ψ is bounded on $\gamma(E)$. Since ψ is convex and lower semicontinuous, it follows that it is bounded and lower semicontinuous on $\overline{\text{conv}}\,\gamma(E)$. It suffices to show that $\psi|_{\overline{\text{conv}}\,\gamma(E)}$ is continuous. Let $u_0 \in \overline{\text{conv}}\,\gamma(E)$ and let U be an open neighborhood of u_0 such that

$$\psi(u) \leqslant \mu < \infty \text{ for all } u \in U \cap \overline{\text{conv}}\,\gamma(C). \tag{5.104}$$

Without any loss of generality, we may assume that

$$u_0 = 0 \text{ and } \psi(u_0) = 0. \tag{5.105}$$

If $t \in (0, 1]$ and $u \in tU \cap \overline{\text{conv}}\,\gamma(E)$, then

$$\psi(u) \leqslant (1-t)\psi(0) + t\psi\left(\frac{u}{t}\right) = t\psi\left(\frac{u}{t}\right) \leqslant t\mu \text{ (see (5.104), (5.105))}$$
$$\Rightarrow \limsup_{\substack{u\to 0 \\ u\in\overline{\text{conv}}\,\gamma(E)}} \psi(u) \leqslant 0. \tag{5.106}$$

On the other hand, from the lower semicontinuity of ψ, we have

$$0 \leqslant \liminf_{\substack{u\to 0 \\ u\in\overline{\text{conv}}\gamma(E)}} \psi(u). \tag{5.107}$$

From (5.106) and (5.107) we conclude the continuity of $\psi|_{\overline{\text{conv}}\,\gamma(E)}$. Therefore $\psi \circ \gamma|_E$ is continuous and this proves Claim 2.

Arguing by contradiction, suppose c is not a critical value of j and let

$$\epsilon_0 = \inf_D j - \max_E j > 0.$$

Let $\epsilon \in (0, \epsilon_0)$ be as in Theorem 5.3.16 and choose $\epsilon' \in (0, \epsilon)$ such that

$$\max_{u\in E_0} j(u) < c - \epsilon'.$$

Using the Ekeland variational principle (see Theorem 4.6.14 and recall that by Claim 1, $\vartheta(\cdot)$ is lower semicontinuous), we can find $\gamma \in \Gamma$ such that

$$\vartheta(\gamma) \leqslant c + \epsilon' \text{ and } -\epsilon d_\infty(\gamma, \xi) \leqslant \vartheta(\xi) - \vartheta(\gamma) \text{ for all } \xi \in \Gamma. \tag{5.108}$$

Let $K = \gamma(E)$ and $K_0 = \{\gamma(u) : u \in E,\ j(\gamma(u)) \in [c-\epsilon', c+\epsilon']\}$.

The continuity of $j \circ \gamma$ on E (see Claim 2) implies that K_0 is compact and the choice of $\epsilon' > 0$ implies that

$$E_0 \cap K_0 = \emptyset.$$

Note that $K_0 \subseteq j^{c+\epsilon}\backslash j^{c-\epsilon}$. Let $\{h_t(\cdot) = h(t,\cdot)\}_{t\in[0,1]}$ be the deformation postulated by Theorem 5.3.16 and let $\xi = h_t \circ \gamma$. Then $\xi \in \Gamma$ and $d_\infty(\gamma, \xi) \leqslant t$. Note that

$$\vartheta(\xi) = \max_{u\in E} j(\xi(u)) > c - \epsilon'.$$

Therefore

$$\vartheta(\xi) = \max_{u\in E} j(\xi(u)) = \max_{u\in E} j((h_t \circ \gamma)(u)) = \max_{v\in K_0} j(h_t(v)).$$

Theorem 5.3.16(c) implies that

$$\vartheta(\xi) - \vartheta(\gamma) = \max_{v\in K_0} j(h_t(v)) - \max_{u\in E_0} j(u) \leqslant -2\epsilon t. \tag{5.109}$$

Then (5.108) and (5.109) lead to a contradiction (recall $d_\infty(\gamma, \xi) \leqslant t$). Therefore c is a critical value of φ. $\square$

As before, with suitable choices of the linking sets $\{E_0, E, D\}$, we have the classical minimax theorems for functionals of the form

$$j = \varphi + \psi \text{ with } \varphi \in C^1(X), \psi \in \Gamma_0(X). \tag{5.110}$$

Theorem 5.4.26 (Extended Mountain Pass Theorem) *Let $j = \varphi + \psi$ be as in (5.110). Assume that j satisfies the GPS-condition and there exist $u_1 \in X$ and $\rho \in (0, ||u_1||)$ such that*

$$\max\{j(u_1), j(0)\} < \inf\{j(u) : ||u|| = \rho\} = m_\rho$$

and $\quad c = \inf\limits_{\gamma\in\Gamma} \max\limits_{t\in[0,1]} j(\gamma(t)) \quad$ *with* $\quad \Gamma = \{\gamma \in C([0, 1], X) : \gamma(0) = 0, \gamma(1) = u_1\}$.
Then $c \geqslant m_\rho$ and c is a critical value of j.

Corollary 5.4.27 *If $j = \varphi + \psi$ is as in (5.110), φ satisfies the GPS-condition, 0 is a local minimizer of $j(\cdot)$ and there exists a $u \neq 0$ such that $j(u) \leqslant j(0)$, then $j(\cdot)$ has a critical point distinct from u and 0. In particular, if $j(\cdot)$ has two local minimizers, then $j(\cdot)$ has at least three distinct critical points.*

Theorem 5.4.28 (Extended Saddle Point Theorem) *Let $j = \varphi + \psi$ be as in (5.110). Assume that j satisfies the GPS-condition, $X = Y \oplus V$ with* $\dim Y < \infty$ *and there exists a $\rho > 0$ such that*

$$\max_{\partial B_r(0)\cap Y} j < \inf_V j$$

and $c = \inf\limits_{\gamma\in\Gamma} \max\limits_{u\in E} j(\gamma(u))$ *with* $E = \bar{B}_r(0) \cap Y$ *and*

$$\Gamma = \left\{\gamma \in C([0, 1], X) : \gamma|_{\partial B_r(0)\cap Y} = id|_{\partial B_r(0)\cap Y}\right\}.$$

Then $c \geqslant \inf\limits_V j$ and c is a critical value of j.

Theorem 5.4.29 (Extended Generalized Mountain Pass Theorem) *Let $j = \varphi + \psi$ be as in (5.110). Assume that j satisfies the GPS-condition and $X = Y \oplus V$ with* $\dim Y < +\infty$. *For $0 < \rho < r_1$, $0 < r_2$ and $v_0 \in V$ with $||v_0|| = 1$, we define*

$$E_0 = \{y + \lambda v_0 : y \in Y \text{ and } \lambda \in \{0, r_1\} \text{ or } ||y|| = r_2\},$$
$$E = \{y + \lambda v_0 : y \in Y, \lambda \in [0, r_1], ||y|| \leqslant r_2\},$$
$$D = \partial B_\rho(0) \cap V.$$

Assume that

$$\max_{E_0} j < \inf_D j$$

and set $c = \inf_{\gamma\in\Gamma} \max_{u\in E} j(\gamma(u))$ *with* $\Gamma = \{\gamma \in C(E, X) : \gamma|_{E_0} = id|_{E_0}\}$. *Then* $c \geqslant \inf_D j$ *and* c *is a critical value of* $j(\cdot)$.

5.5 Critical Points Under Constraints

In many situations, we have to find critical points of $\varphi \in C^1(X)$ in the presence of constraints, that is, critical points of $\varphi|_M$ where $M \subseteq X$ is a set of constraints. Typically M is a smooth manifold of X (= Banach space) of the form

$$M = \{u \in X : g_k(u) = 0, k = 1, \dots, m\} \text{ with } g_k \in C^1(X).$$

Definition 5.5.1 Let X be a Banach space and $M \subseteq X$. We say that M is a C^n-manifold of codimension m $(n, m \in \mathbb{N})$ if for every $u_0 \in M$ we can find an open neighborhood U of u_0 and a function $g \in C^n(U, \mathbb{R}^m)$ such that

(a) $g'(u)$ is surjective for every $u \in U$;
(b) $M \cap U = \{u \in U : g(u) = 0\}$.

Remark 5.5.2 Of special interest are manifolds modeled on a subspace of X of codimension 1. In this case $M = g^{-1}(0)$ with $g \in C^1(X)$ and $g'(u) \neq 0$ for all $u \in M$. For example, if X is a reflexive Banach space with locally uniformly convex dual (in particular, a Hilbert space), then the norm $||\cdot||$ is Fréchet differentiable on $X\backslash\{0\}$ and if $g(u) = ||u||^2 - r^2$, then $g \in C^1(X)$ and

$$M = g^{-1}(0) = \partial B_\rho(0) = \{u \in X : ||u|| = \rho\}$$

is a C^1-manifold of codimension 1. This situation arises in eigenvalue problems.

We need to explain what "critical point of $\varphi|_M$" means. To this end, we recall the following basic notion from Differential Geometry.

Definition 5.5.3 Let X be a Banach space, $M \subseteq X$ a C^n-manifold and $u_0 \in M$. The "tangent space to M at u_0", denoted by $T_{u_0}M$, is defined to be the set of all tangent vectors $\gamma'(0)$ to M at u_0, where $\gamma : (-\epsilon, \epsilon) \to M$ is a C^1-path such that $\gamma(0) = u_0$ and $\gamma'(t) \neq 0$ for all $t \in (-\epsilon, \epsilon)$.

Remark 5.5.4 In fact there are other equivalent ways to define the tangent space (the coordinate approach, the derivation approach and the ideal approach). The above definition permits us to think of the tangent vector as the velocity of a curve in the surface passing through the given point.

Now we can define what we mean by a critical point of $\varphi|_M$.

Definition 5.5.5 Let X be a Banach space, $M \subseteq X$ a C^1-manifold and $\varphi \in C^1(X)$. We say that $u_0 \in M$ is a "critical point of $\varphi|_M$" if and only if

$$\langle \varphi'(u_0), h \rangle = 0 \text{ for all } h \in T_{u_0}M.$$

Remark 5.5.6 So, according to this definition, $u_0 \in M$ is a critical of $\varphi|_M$ if and only if

$$\frac{d}{dt}\varphi(\gamma(t))|_{t=0} = 0$$

for every C^1-curve $\gamma : (-\epsilon, \epsilon) \to X$ passing through u_0 (see Definition 5.5.3).

One of the main results concerning critical points of constrained functionals is the so-called "Lagrange multiplier theorem", which says that the critical points of a constrained functional are critical points of a related unconstrained functional.

To prove the Lagrange multiplier rule, we will need the following simple result from linear functional analysis.

Lemma 5.5.7 *If X, Y are Banach spaces, $A \in \mathscr{L}(X, Y)$ is surjective,* $\ker A = N(A)$ *is complemented in X and $u^* \in X^*$, then the following statements are equivalent:*

(a) $\ker A \subseteq \ker u^*$.
(b) There exists a $y^* \in Y^*$ such that $u^* = y^* \circ A$.

Proof $(a) \Longrightarrow (b)$: Let $y \in Y$. Then the surjectivity of A implies that $A^{-1}(y) \neq \emptyset$ and it is closed. Moreover, if $u_1, u_2 \in A^{-1}(y)$, then $u_1 - u_2 \in \ker A \subseteq \ker u^*$ and so $\langle u^*, u_1\rangle_X = \langle u^*, u_2\rangle_X$. Let $y^*(y)$ be the common value of u^* on the closed set $A^{-1}(y)$. Note that $u^* = y^* \circ A$. So, it remains to show that y^* is linear and continuous (that is, $y^* \in Y^*$).

First we check linearity. If $y = A(u)$, then $y^*(y) = \langle u^*, u\rangle_X$ and so for $\lambda \in \mathbb{R}$, we have

$$\begin{aligned} &\lambda y = \lambda A(y) = A(\lambda y) \\ \Rightarrow\ &y^*(\lambda y) = \langle u^*, \lambda u\rangle_X = \lambda \langle u^*, u\rangle_X = \lambda y^*(y). \end{aligned}$$

Also, if $y_1 = A(u_1),\ y_2 = A(u_2)$, then

$$\begin{aligned} &y_1 + y_2 = A(u_1) + A(u_2) = A(u_1 + u_2) \\ \Rightarrow\ &y^*(y_1 + y_2) = \langle u^*, u_1 + u_2\rangle_X = \langle u^*, u_1\rangle_X + \langle u^*, u_2\rangle_X = y^*(y_1) + y^*(y_2). \end{aligned}$$

Therefore we conclude that $y \to y^*(y)$ is linear.

Let V be a topological complement of $\ker A$, that is,

$$V \text{ is a closed subspace of } X \text{ and } X = \ker A \oplus V.$$

Note that by the Banach theorem, $A|_V$ is an isomorphism. Also, we have

$$\begin{aligned} & y^*(y) = 0 \text{ if and only if } y = A(u) \text{ with } u \in \ker u^* \\ \Rightarrow\ & \ker y^* = A(\ker u^*). \end{aligned} \tag{5.111}$$

Clearly we have

$$A|_V(V \cap \ker u^*) \subseteq A(\ker u^*). \tag{5.112}$$

Let $y \in A(\ker u^*)$. We can find $u \in \ker u^*$ such that $y = A(u)$. Let $u = u_1 + u_2$ with $u_1 \in \ker A$ and $u_2 \in V$. Since by hypothesis $\ker A \subseteq \ker u^*$, we have $\langle u^*, u\rangle = \langle u^*, u_2\rangle$ and so

$$\begin{aligned} & u_2 \in V \cap \ker u^* \text{ and } y = A(u) = A(u_2) \\ \Rightarrow\ & A(\ker u^*) \subseteq A|_V(V \cap \ker u^*). \end{aligned} \tag{5.113}$$

From (5.112) and (5.113) we have

$$\begin{aligned} & A(\ker u^*) = A|_V(V \cap \ker u^*) \\ \Rightarrow\ & \ker y^* = A|_V(V \cap \ker u^*) \text{ (see (5.111))} \\ \Rightarrow\ & \ker y^* \text{ is closed (since } A|_V \text{ is an isomorphism and } V \cap \ker u^* \text{ is closed)} \\ \Rightarrow\ & y^* \in Y^*. \end{aligned}$$

$(b) \Longrightarrow (a)$: Obvious. □

Remark 5.5.8 The functional $y^* \in Y^*$ is unique. Indeed, suppose there is a $z^* \neq Y^*$, $z^* \neq y^*$, such that $u^* = y^* \circ A = z^* \circ A$. Let $y \in Y$ such that $\langle y^*, y\rangle_Y \neq \langle z^*, y\rangle_Y$. If $y = A(u)$ with $u \in X$, then

$$\langle u^*, u\rangle_X = \langle y^*, A(u)\rangle_Y \neq \langle z^*, A(u)\rangle_Y = \langle u^*, u\rangle_X$$

a contradiction. This result is an infinite-dimensional generalization of a result from linear algebra, which says that if $h, u_1, \cdots, u_n$ are linear functionals on a vector space X such that $\bigcap_{k=1}^{n} \ker u_k \subseteq \ker h$, then h is a linear combination of the $u_k's$. This result is a useful tool in the study of the weak topology on a Banach space.

Now we can state the "Lagrange Multipliers Theorem".

Theorem 5.5.9 *Let X, Y be Banach spaces, $U \subseteq X$ be a nonempty open set, $\varphi \in C^1(U)$, $g \in C^1(U, Y)$. $M = g^{-1}(0)$, $u_0 \in M$ is a critical point of $\varphi|_M$, $g'(u_0) \in \mathscr{L}(X, Y)$ is surjective and $\ker g'(u_0)$ is complemented in X. Then there exists a unique $y^* \in Y^*$ such that*

$$(\varphi - y^* \circ g)'(u_0) = 0.$$

Proof If $\ker g'(u_0) \subseteq \ker \varphi'(0)$, then from Lemma 5.5.7 (see also Remark 5.5.8), we know that there exists a unique $y^* \in Y^*$ such that

$$\varphi'(u_0) = y^* \circ g'(u_0) = (y^* \circ g)'(u_0).$$

So, we see that in order to prove the theorem, it suffices to show that

$$\ker g'(u_0) \subseteq \ker \varphi'(u_0). \tag{5.114}$$

We know that $\ker g'(u_0) \subseteq X$ is a closed subspace and so a Banach space too. By hypothesis $\ker g'(u_0)$ admits a complementary space V. This too is a Banach space. Let $\xi : \ker g'(u_0) \times V \to X$ be defined by

$$\xi(z, v) = z + v \text{ for all } z \in \ker \varphi'(u_0) \text{ and all } v \in V.$$

This map is a linear, continuous bijection, hence by the Banach theorem it is an isomorphism. Let $\hat{U} = \xi^{-1}(U) \subseteq \ker g'(u_0) \times V$ be open and $\hat{g} = g \circ \xi : \hat{U} \to Y$. We have

$$\begin{aligned}\hat{g}'(z, v)(h, w) &= g'(\xi(z, v)) \circ \xi'(z, v)(h, w)\\ &= g'(z + v)(h + w).\end{aligned} \tag{5.115}$$

Also, we know that

$$\hat{g}'(z, v)(h, w) = \hat{g}_1'(z, v)h + \hat{g}_2'(z, v)w \tag{5.116}$$

with $\hat{g}_k'(z, v),\ k = 1, 2$, being the partial derivatives with respect to the two variables. From (5.115) and (5.116), we have

$$\hat{g}_1(z, v) = g'(z + v)|_{\ker g'(u_0)} \text{ and } \hat{g}_2'(z, v) = g'(z + v)|_V.$$

So, if $u_0 = \bar{u}_0 + \hat{u}_0$ with $\bar{u}_0 \in \ker g'(u_0)$ and $\hat{u}_0 \in V$, we have

$$\hat{g}_1'(\bar{u}_0, \hat{u}_0) = g'(u_0)|_{\ker g'(u_0)} \text{ and } \hat{g}_2'(\bar{u}_0, \hat{u}_0) = g'(u_0)|_V$$
$$\Rightarrow \hat{g}_1'(\bar{u}_0, \hat{u}_0) = 0 \text{ and } \hat{g}_2'(\bar{u}_0, \hat{u}_0) \in \mathscr{L}(V, Y) \text{ is an isomorphism.}$$

So, we can apply the Implicit Function Theorem. According to this theorem, there exist an open set $\widehat{U}' \subseteq \widehat{U}$ such that $(\bar{u}_0, \hat{u}_0) \in \widehat{U}'$, an open neighborhood V of u_0 and a C^1-map $s : V \to V$ such that the following properties are equivalent:

$$\bullet (z, v) \in \widehat{U}' \text{ and } \hat{g}(z, v) = 0$$
$$\bullet z \in V \text{ and } v = s(z).$$

We consider the map $\eta : V \to \mathbb{R}$ defined by

$$\eta(z) = (\varphi \circ \xi)(z, h(z)).$$

Evidently, η has a local minimizer at $\bar{u}_0$. Hence

$$\eta'(\bar{u}_0) = 0. \tag{5.117}$$

Let $w \in \ker g'(u_0)$. Then

$$\begin{aligned} \left\langle \eta'(\bar{u}_0), w \right\rangle &= \varphi' \left(\xi(\bar{u}_0, s(\bar{u}_0)) \circ \xi'(\bar{u}_0, s(\bar{u}_0)) \right) \circ \left(id|_{\ker g'(u_0)}, s'(\bar{u}_0) \right) w \\ &= \left\langle \varphi'(u_0), \xi(w, s'(\bar{u}_0)w) \right\rangle \\ &= \varphi'(u_0)(w + s'(\bar{u}_0)w). \end{aligned} \tag{5.118}$$

But for $z \in V$, we have $\hat{g}(z, h(z)) = 0$ and so

$$\hat{g}_1'(\bar{u}_0, \hat{u}_0)h + \hat{g}_2(\bar{u}_0, \hat{u}_0) \circ s'(\bar{u}_0)w = 0.$$

Because $\hat{g}_1'(\bar{u}_0, \hat{u}_0) = 0$ and $\hat{g}_2'(\bar{u}_0, \hat{u}_0) \in \mathscr{L}(V, Y)$ is invertible, we have

$$\begin{aligned} & s'(\bar{u}_0) = 0 \\ \Rightarrow\ & \left\langle \eta'(\bar{u}_0), w \right\rangle = \left\langle \varphi'(u_0), w \right\rangle \text{ for all } w \in \ker g'(u_0) \text{ (see (5.118))} \\ \Rightarrow\ & \left\langle \varphi'(u_0), w \right\rangle = 0 \text{ for all } w \in \ker g'(u_0) \text{ (see (5.117))} \\ \Rightarrow\ & \text{inclusion (5.114) holds and this proves the theorem.} \end{aligned}$$

The proof is now complete. □

Remark 5.5.10 In a Hilbert space setting, $\ker g'(u_0)$ is always complemented. If $Y = \mathbb{R}$, then the surjectivity condition on $g'(u_0)$ is equivalent to the nondegeneracy condition $g'(u_0) \neq 0$. If $Y = \mathbb{R}^m$, then $M = \{u \in X : g_k(u) = 0 \text{ for all } k = 1, \cdots, m\}$ (with $g_k \in C^1(X)$) is a C^1-manifold of codimension m. According to Theorem 5.5.9 we can find $\{\lambda_k\}_{k=1}^m \subseteq \mathbb{R}$ such that

$$\varphi'(u_0) = \sum_{\text{k}=1}^{m} \lambda_k g_k'(u_0).$$

If in addition $X = \mathbb{R}^l$ with $l \in \mathbb{N}$, then

$$\nabla \varphi(u_0) = \sum_{\text{k}=1}^{m} \lambda_k \nabla g_k(u_0),$$

that is, $\nabla \varphi(u_0)$ is a linear combination of the $\nabla g_k(u_0) k = 1, \cdots, m$. The constants $\lambda_1, \cdots, \lambda_m$ are called "Lagrange multipliers". The surjectivity of $g'(u_0)$ is equivalent to the linear independence of the gradients $\{\nabla g_k(u_0)\}_{k=1}^m$.

Now let $M \subseteq X$ be a C^1-manifold, $\varphi \in C^1(X)$ and $u_0 \in M$ be a critical point of φ. Then u_0 is also a critical point of $\varphi|_M$ (see Definition 5.5.5). Indeed, if $\gamma : [0, 1] \to X$ is a C^1-path passing from u_0, then from the chain rule we have

$$\frac{d}{dt}\varphi(\gamma(t))|_{t=0} = \left\langle \varphi'(\gamma(t)), \gamma'(t) \right\rangle|_{t=0} = \left\langle \varphi'(u_0), \gamma'(0) \right\rangle = 0$$

(see Remark 5.5.6). The converse is not in general true. That is, a critical point of $\varphi|_M$ need not be a critical point of the unconstrained functional. This leads to the following definition.

Definition 5.5.11 Let X be a Banach space, $\varphi \in C^1(X)$ and $M \subseteq X$ a C^1-manifold. If a given critical point u_0 of $\varphi|_M$ is also a critical point of the unconstrained functional φ, then we say that M is a "natural constraint for the critical point $u_0 \in M$". If M is a natural constraint for every critical point u_0 of $\varphi|_M$, then we say that M is a "natural constraint for φ".

Remark 5.5.12 So, M is a natural constraint for φ if the "Lagrange multiplier" $y^* \in Y^*$ in Theorem 5.5.9 is zero. If K_ψ denotes the set of critical points of a functional ψ, then M is a natural constraint for ψ if and only if $K_{\psi|_M} = K_\psi \cap M$. A special case of interest is when φ is indefinite (that is, $\inf\limits_X \varphi = -\infty$ and $\sup\limits_X \varphi = +\infty$). In that case, it may be a good idea to look for critical points of φ on a natural constraint M, because it can happen that $\varphi|_M$ is bounded. Then we can try to find critical points of φ as extrema (minima or maxima) of $\varphi|_M$.

Definition 5.5.13 Let X be a Banach space and $\varphi \in C^1(X)$. The "Nehari manifold" of φ is the set $N_\varphi = \{u \in X : \left\langle \varphi'(u), u \right\rangle = 0, u \neq 0\}$.

Proposition 5.5.14 *Let X be a Banach space, $\varphi \in C^2(X)$ and N_φ is the Nehari manifold of φ which is assumed to be nonempty and satisfies the following conditions:*

(i) there exists a $\rho > 0$ such that $N_\varphi \cap B_\rho = \emptyset$ ($B_\rho(0) = \{u \in X : ||u|| < \rho\}$);
(ii) $\left\langle \varphi''(u)u, u \right\rangle \neq 0$ for all $u \in N_\varphi$.

Then M is a natural constraint for φ.

Proof Let $\psi(u) = \left\langle \varphi'(u), u \right\rangle$ for all $u \in X$. Then $\psi \in C^1(X)$ and $N_\varphi = \psi^{-1}(0)\backslash\{0\}$. For every $u \in N_\varphi$ we have

$$\left\langle \psi'(u), u \right\rangle = \left\langle \varphi''(u)u, u \right\rangle + \left\langle \varphi'(u), u \right\rangle = \left\langle \varphi''(u)u, u \right\rangle \neq 0 \tag{5.119}$$

(see Definition 5.5.13 and hypothesis(ii))

$$\Rightarrow \psi'(u) \neq 0 \text{ for all } u \in N_\varphi.$$

This fact and hypothesis (i) imply that N_φ is a C^1-manifold of codimension 1 (see Definition 5.5.1 and Remark 5.5.2). If u_0 is a critical point of $\varphi|_M$, then from Theorem 5.5.9 we can find $\lambda \in \mathbb{R}$ such that

$$\begin{aligned}
&\varphi'(u_0) = \lambda\psi'(u_0) \\
\Rightarrow\ &\langle\varphi'(u_0), u_0\rangle = \lambda\langle\psi'(u_0), u_0\rangle \\
\Rightarrow\ &0 = \psi(u_0) = \lambda\langle\psi'(u_0), u_0\rangle \text{ and } \langle\psi'(u_0), u_0\rangle \neq 0 \text{ (see (5.119))} \\
\Rightarrow\ &\lambda = 0 \\
\Rightarrow\ &\varphi'(u_0) = 0 \text{ and so } u_0 \text{ is a critical point of the unconstrained functional } \varphi.
\end{aligned}$$

This proves that N_φ is a natural constraint for φ. ☐

Related to the problem of finding critical points under constraints is the so-called "fibering method". This method is based on the representation of a solution u of an equation in a Banach space X, in the form

$$u = tv \tag{5.120}$$

with t a parameter, $t \neq 0$ and $t \in I \subseteq \mathbb{R}$ an open set and $v \in X\backslash\{0\}$ satisfying

$$\xi(t, v) = c \neq 0 \tag{5.121}$$

with ξ being a function satisfying a sufficiently general condition. One important special fibering function is $\xi(t, v) = ||v||$, and then $c = 1$, that is, the constraint set is $\partial B_1(0) = \{v \in X : ||v|| = 1\}$. Then we look for a solution of our problem in the form (5.120) with $t \in \mathbb{R}$ and $v \in \partial B_1(0)$. So, in the fibering method, we embed the space X of the original problem into the larger space $\mathbb{R} \times X$ and investigate the new problem of solvability of the equation under the constraint (5.121).

So, let X be a Banach space and $\varphi \in C^1(X\backslash\{0\})$. Let

$$\hat{\varphi}(t, v) = \varphi(tv) \text{ for all } t \in \mathbb{R},\ v \in X. \tag{5.122}$$

We consider $\hat{\varphi}$ restricted on $I \times \partial B_1(0)$ with $I \subseteq \mathbb{R}$ open.

Proposition 5.5.15 *If X has a Fréchet differentiable norm on $X\backslash\{0\}$ and $(t, v) \in (I\backslash\{0\}) \times \partial B_1(0)$ is a critical point of $\hat{\varphi}|_{I\times\partial B_1(0)}$ then $u = tv$ is a critical point of φ.*

Proof From the Lagrange multiplier rule (see Theorem 5.5.9), we can find $\lambda, \mu \in \mathbb{R}$ with $(\lambda, \mu) \neq (0, 0)$ such that

$$\hat{\varphi}'_t(t, v) = 0 \text{ and } \lambda\hat{\varphi}'_v(t, v) = \mu\frac{J(v)}{||v||} \tag{5.123}$$

with $J : X \to X^*$ being the duality map (see Proposition 2.7.32). From (5.123) we have

$$\begin{aligned}
\lambda\langle\hat{\varphi}'_v(t, v), v\rangle = \lambda t\langle\varphi'(tv), v\rangle = \lambda t\hat{\varphi}'_t(t, v) = \mu = 0& \\
\text{(recall that } \langle J(v), v\rangle = ||v||^2 \text{ and see (5.123))}.&
\end{aligned}$$

Hence $\lambda \neq 0$ and from (5.123) we have

$$\begin{aligned} 0 &= \hat{\varphi}'_v(t, v) = t\varphi'(tv) \\ \Rightarrow \varphi'(u) &= 0 \text{ with } u = tv \text{ (since } t \neq 0). \end{aligned}$$

The proof is now complete. $\square$

Remark 5.5.16 We can go beyond the norm fibering functional and consider differentiable functions $\xi(t, v)$ such that $\langle \xi'_v(t, v), v \rangle \neq t\xi_t(t, v)$ for all $(t, v) \in (I\backslash\{0\}) \times (X\backslash\{0\})$ such that $\xi(t, v) = c \neq 0$.

5.6 Critical Points Under Symmetries

In this section we look for multiple critical points of a functional φ. Such multiplicity results can be obtained if φ exhibits symmetries, that is, if there is a topological group acting continuously on X and the functional φ is invariant under this group action.

Definition 5.6.1 Let X be a Banach space and G a topological group. A "representation" of G over X is a family $\{S(g)\}_{g \in G} \subseteq \mathscr{L}(X)$ such that

(a) $S(e) = \mathrm{id}$ (here e denotes the identity element of G);
(b) $S(g_1 g_2) = S(g_1)S(g_2)$ for all $g_1, g_2 \in G$;
(c) the map $(g, u) \mapsto S(g)u$ is continuous from $G \times X$ into X.

We say that the representation is "isometric" if for each $g \in G$, $S(g)$ is an isometry, that is, $||S(g)u|| = ||u||$ for all $g \in G$ and all $u \in X$.

A set $C \subseteq X$ is said to be "invariant" (or "G-invariant") if

$$S(g)C \subseteq C \text{ for all } g \in G.$$

A functional $\varphi : X \to \mathbb{R}$ is said to be "invariant" (or "G-invariant") if

$$\varphi \circ S(g) = \varphi \text{ for all } g \in G.$$

A map $h : X \to X$ is said to be "equivariant" (or "G-equivariant") if

$$S(g) \circ h = h \circ S(g) \text{ for all } g \in G.$$

The set of invariant (or fixed) points of X is the set

$$\Sigma = X^G = \{u \in X : S(g)u = u \text{ for all } g \in G\}.$$

Remark 5.6.2 Often we identify $S(g)$ with g, write gu instead of $S(g)u$ and speak of the "(linear) action" of G on X. For notational simplicity, we follow this practice here.

First we establish some useful consequences of the action of G on X, on the derivative of φ and on the corresponding pseudogradient vector field.

In what follows, X is a Banach space and G is a topological group acting on X. Additional hypotheses will be introduced as needed.

Proposition 5.6.3 *If $\varphi \in C^1(X)$ is invariant, then*

(a) $\langle \varphi'(gu), h\rangle = \langle \varphi'(u), g^{-1}h\rangle$ *for all $g \in G$ and all $u, h \in X$;*
(b) *if the action of G is isometric, then $||\varphi'(gu)||_* = ||\varphi'(u)||_*$ for all $g \in G$ and all $u \in X$.*

Proof (a) From Definition 5.5.1 we see that

$$S(g)^{-1} = S(g^{-1}) \text{ for all } g \in G.$$

We have

$$\begin{aligned}
\langle \varphi'(gu), h\rangle &= \lim_{t\to 0} \frac{1}{t}\left[\varphi(g(u)+th) - \varphi(g(u))\right] \\
&= \lim_{t\to 0} \frac{1}{t}\left[\varphi(g(u+tg^{-1}h)) - \varphi(g(u))\right] \\
&= \lim_{t\to 0} \frac{1}{t}\left[\varphi(u+tg^{-1}h) - \varphi(u)\right] \text{ (since } \varphi \text{ is invariant)} \\
&= \langle \varphi'(u), g^{-1}h\rangle \text{ for all } g \in G, \text{ all } u, h \in X.
\end{aligned}$$

(b) Just take the supremum over $h \in \bar{B}_1(0) = \{v \in X : ||v|| \leqslant 1\}$ of both sides in (a) and recall that $||h|| = ||g^{-1}h||$ since the action of G on X is isometric. $\square$

Proposition 5.6.4 *If $\varphi \in C^1(X)$ is invariant, the topological group G is compact and the linear action of G on X is isometric, then we can find an equivariant pseudogradient vector field $V : X\backslash K_\varphi \to X$ (recall that $K_\varphi = \{u \in X : \varphi'(u) = 0\}$, the critical set of φ).*

Proof Recall that a compact topological group admits a unique (up to scalar multiplication) G-invariant, finite regular Borel measure μ, known as the "Haar measure". For simplicity we assume that $\mu(G) = 1$ (see, for example, Dunford and Schwartz [151, p. 460]). From Theorem 5.1.4 we know that there exists a pseudogradient vector field $V : X\backslash K_\varphi \to X$. Let

$$\hat{V}(u) = \int_G gV(g^{-1}u)d\mu \text{ for all } u \in X\backslash K_\varphi.$$

From Proposition 5.6.3 we see that $X\backslash K_\varphi$ is invariant and so $\hat{V}$ is well-defined. Also, for every fixed $g' \in G$, we have

$$\begin{aligned}
\hat{V}(g'u) &= \int_G g\, V(g^{-1}g'u)d\mu \\
&= \int_G g'(g')^{-1}g\, V(((g')^{-1}g)^{-1}u)d\mu \\
&= g'\int_G (g')^{-1}g\, V(((g')^{-1}g)^{-1}u)d\mu \\
&= g'\,\hat{V}(u) \\
\Rightarrow \quad \hat{V} &\quad \text{is equivariant (see Definition 5.6.1).}
\end{aligned}$$

Next, we check that $\hat{V}$ is a pseudogradient vector field for φ (see Definition 5.1.1). We have

$$\begin{aligned}
||\hat{V}(u)|| &\leqslant \int_G ||V(g^{-1}u)||d\mu \text{ (recall that the linear action of } G \text{ is isometric)} \\
&\leqslant 2\int_G ||\varphi'(g^{-1}u)||_* d\mu \text{ (see Definition 5.1.1)} \\
&= 2\int_G ||\varphi'(u)||_* d\mu \text{ (see Proposition 5.6.3(b))} \\
&= 2||\varphi'(u)||_* .
\end{aligned}$$

In addition, we have

$$\begin{aligned}
\left\langle \varphi'(u), \hat{V}(u)\right\rangle &= \int_G \left\langle \varphi'(u), gV(g^{-1}u)\right\rangle d\mu \\
&= \int_G \left\langle \varphi'(g^{-1}u), V(g^{-1}u)\right\rangle d\mu \text{ (see Proposition 5.6.3(a))} \\
&\geqslant \int_G ||\varphi'(g^{-1}(u))||_*^2 d\mu \text{ (see Definition 5.1.1)} \\
&= \int_G ||\varphi'(u)||_*^2 d\mu \text{ (see Proposition 5.6.3(b))} \\
&= ||\varphi'(u)||_*^2.
\end{aligned}$$

Finally, we show that $\hat{V}(\cdot)$ is locally Lipschitz. For $u \in X\backslash K_\varphi$ let $O(u) = \{gu : g \in G\}$ (the orbit of u). The compactness of G and the continuity of the map $g \mapsto gu$ (see Definition 5.6.1) imply that $O(u) \subseteq X$ is compact. Therefore we can find $\delta > 0$ such that $V|_{O(u)_\delta}$ is Lipschitz continuous with Lipschitz constant $k > 0$ (here $O(u)_\delta = \{v \in X\backslash K_\varphi : d(v, O(u)) \leqslant \delta\}$). Evidently, the set $O(u)_\delta$ is invariant. So, for all $v, h \in \bar{B}_\delta(u) \cap (X\backslash K_\varphi)$ we have

$$
\begin{aligned}
||\hat{V}(v) - \hat{V}(h)|| &\leqslant \int_G ||g(V(g^{-1}v) - V(g^{-1}h))||d\mu \\
&= \int_G ||V(g^{-1}v) - V(g^{-1}h)||d\mu \text{ (the action is isometric)} \\
&\leqslant k||g^{-1}v - g^{-1}h|| \text{ (since } V|_{\overline{B}_\delta(u)\cap(X\setminus K_\varphi)} \text{ is } k\text{-Lipschitz)} \\
&= k||v-h|| \text{ (again the isometry of the action).}
\end{aligned}
$$

So, $\hat{V}(\cdot)$ is locally Lipschitz and we conclude that $\hat{V}(\cdot)$ is a pseudogradient vector field for φ. $\square$

This proposition leads to an equivalent version of Theorem 5.3.7 (the Deformation Theorem).

Theorem 5.6.5 *If G is a compact topological group with an isometric linear action on a Banach space X, $\varphi \in C^1(X)$ and φ satisfies the C_c-condition for some $c \in \mathbb{R}$, then for every $\epsilon_0 > 0$, every open invariant neighborhood U of K_φ^c (if $K_\varphi^c = \emptyset$, then $U = \emptyset$) and every $\eta > 0$, we can find a deformation $h : [0, 1] \times X \to X$ which satisfies properties (a)–(e) in Theorem 5.3.7 and in addition (f) $h(t, \cdot)$ is equivariant for every $t \in [0, 1]$.*

Now let $\mathscr{C}$ be the family of all closed and invariant subsets of X, that is,

$$\mathscr{C} = \{C \subseteq X : C \text{ is closed and } gC = C \text{ for all } g \in G\}. \tag{5.124}$$

Definition 5.6.6 An index (or G-index) on X is a map $i : \mathscr{C} \to \mathbb{N} \cup \{0, \infty\}$ which satisfies the following properties:

(a) $i(A) = 0$ if and only if $A = \emptyset$;
(b) if $C_1, C_2 \in \mathscr{C}$ and $h : C_1 \to C_2$ is continuous and equivariant, then $i(C_1) \leqslant i(C_2)$ (monotonicity property);
(c) if $C_1, C_2 \in \mathscr{C}$, then $i(C_1 \cup C_2) \leqslant i(C_1) + i(C_2)$ (subadditivity property);
(d) if $C \in \mathscr{C}$ is compact, then there exists a closed neighborhood D of C such that

$$D \in \mathscr{C} \text{ and } i(D) = i(C) \text{ (continuity property).}$$

Now, we consider the Banach space X together with a topological group G which acts on X isometrically and an index $i : \mathscr{C} \to \mathbb{N} \cup \{+\infty\}$. We can classify the compact, invariant sets of X as follows:

$$\mathscr{C}_k = \{C \in \mathscr{C} \text{ is compact and } i(A) \geqslant k\} \text{ for all } k \in \mathbb{N}.$$

Given $\varphi \in C^1(X)$, we define

$$c_k = \inf_{C \in \mathscr{C}_k} \max_{u \in C} \varphi(u). \tag{5.125}$$

Since the family $\{\mathscr{C}_k\}_{k\geqslant 1}$ is decreasing, we have

$$-\infty \leqslant c_1 \leqslant c_2 \leqslant \cdots$$

We have the following basic multiplicity result, known also as the "Ljusternik–Schnirelmann Multiplicity Theorem".

Theorem 5.6.7 *If G is a compact topological group with a linear isometric action on a Banach space X, i is an index on X, $\varphi \in C^1(X)$ is invariant, satisfies the C-condition and $c_k > -\infty$ (see (5.125)) for some $k \in \mathbb{N}$, then c_k is a critical value of φ; more precisely, if $c_m = c_k = c > -\infty$ for some $m \geqslant k$, then $i(K_\varphi^c) \geqslant m - k + 1$.*

Proof Note that since φ is invariant, K_φ^c is an invariant set (see Proposition 5.6.3), and since φ satisfies the C-condition, K_φ^c is also compact. Therefore $K_\varphi^c \in \mathscr{C}$ (see (5.124)).

We show that if $c_m = c_k = c > -\infty$ for some $m \geqslant k$, then $i(K_\varphi^c) \geqslant m - k + 1$.

From Definition 5.6.6(d), we can find a closed, invariant neighborhood D of K_φ^c such that $i(D) = i(C)$ (continuity property of the index). Let $U = \operatorname{int} D$. Then U is an open invariant neighborhood of K_φ^c and so we can apply Theorem 5.6.5 and produce $\epsilon > 0$ and a deformation $h : [0, 1] \times X \to X$ which satisfies properties $(a) \to (f)$ (see also Theorem 5.3.7).

From (5.125) we see that we can find $c \in \mathscr{C}_m$ such that

$$\sup_C \varphi \leqslant c_m + \epsilon \text{ (recall } c = c_m = c_k).$$

Let $E = C \setminus U \subseteq X$. This is a compact set. Using Definition 5.6.6, we have

$$\begin{aligned} m &\leqslant i(C) \\ &\leqslant i(E) + i(D) \text{ (since } C \subseteq E \cup D) \\ &= i(E) + i(K_\varphi^c) \text{ (recall the choice of } D). \end{aligned} \tag{5.126}$$

Note that $E \subseteq \varphi^{c+\epsilon} \setminus U$. So, by Theorem 5.6.5 we have that $A = h(1, E) \subseteq \varphi^{c-\epsilon}$. Since $h(1, \cdot)$ is equivariant (see Theorem 5.6.5(f)) and E is compact and invariant, we infer that A is compact invariant too and we have $\max\limits_A \varphi \leqslant c - \epsilon$. From (5.125) it follows that

$$i(A) \leqslant k - 1. \tag{5.127}$$

From the monotonicity property of the index (see Definition 5.6.6(b)), we have

$$i(E) \leqslant i(A) \leqslant k - 1 \text{ (see (5.127))}. \tag{5.128}$$

Using (5.128) in (5.126), we obtain

$$\begin{aligned} & m \leqslant k - 1 + i(K_\varphi^c) \\ \Rightarrow\ & m - k + 1 \leqslant i(K_\varphi^c). \end{aligned}$$

So, $i(K_\varphi^c) \neq 0$ and the Definition 5.6.6(a) implies $K_\varphi^c \neq \emptyset$. $\square$

There are two particular indices which are useful in applications. These are the "Krasnoselskii genus" and the "Ljusternik–Schnirelmann category".

We start by recalling the following fundamental topological notion.

Definition 5.6.8 Let Y be a Hausdorff topological space and $C \subseteq Y$. We say that C is "contractible in Y" if the inclusion map $i : C \to Y$ is homotopic to a constant $y_0 \in Y$, that is, there exists a continuous map $h : [0, 1] \times C \to Y$ such that

$$h(0, \cdot) = \mathrm{id}|_C \text{ and } h(1, y) = y_0 \text{ for all } y \in C.$$

Remark 5.6.9 Clearly Y is contractible if and only if some point $y_0 \in Y$ is a deformation retract (see Definition 5.3.10). Every convex subset or more generally every star-shaped set of a topological vector space is contractible (recall that a subset C of a topological vector space is star-shaped if there exists a $y_0 \in C$ such that for all $y \in C$ the "interval" $[y_0, y] = \{(1 - t)y_0 + ty : t \in [0, 1]\}$ is in C). Contractibility is preserved by homeomorphisms and every contractible space is simply connected (that is, it is path connected and the fundamental group is trivial). The unit sphere of a Euclidean space $S^{m-1} = \{y \in \mathbb{R}^m : |y| = 1\}$ is not contractible in itself. In contrast, the unit sphere of an infinite-dimensional Banach space is contractible in itself.

Now we are ready to introduce the Krasnoselskii genus and the Ljusternik–Schnirelmann category.

Definition 5.6.10 (a) Let X be a Banach space and let

$$\mathscr{S} = \{C \subseteq X \backslash \{0\} : C \text{ is closed and symmetric}\}.$$

The "Krasnoselskii genus" $\gamma : \mathscr{S} \to \mathbb{N} \cup \{0, \infty\}$ is defined by

$$\begin{aligned} & \gamma(\emptyset) = 0, \\ & \gamma(D) = \inf\left\{k \in \mathbb{N} : \text{ there exists an odd } \varphi \in C(D, \mathbb{R}^k \backslash \{0\})\right\}, \\ & \gamma(D) = +\infty \text{ if no such map } \varphi \text{ can be found.} \end{aligned}$$

(b) Let X be a Hausdorff topological space. The "Ljusternik–Schnirelmann category" $\mathrm{cat}_X : 2^X \to \mathbb{N} \cup \{0, \infty\}$ is defined by

$$\begin{aligned} & \mathrm{cat}_X(\emptyset) = 0, \\ & \mathrm{cat}_X(D) = \min\left\{k \in \mathbb{N} : D \subseteq \bigcup_{i=1}^k D_i,\ D_i \text{ is closed, contractible for } i = 1, \cdots, m\right\}, \\ & \mathrm{cat}_X(D) = +\infty \text{ if no such finite cover can be found.} \end{aligned}$$

Remark 5.6.11 In the definition of the Krasnoselskii genus, we can always assume that $\varphi \in C(X, \mathbb{R}^k)$. Indeed by Proposition 2.1.9 (the Dugundji extension theorem), given $\varphi \in C(D, \mathbb{R}^l \setminus \{0\})$ we can always find $\hat{\varphi}_* \in C(X, \mathbb{R}^k \setminus \{0\})$ such that $\hat{\varphi}_*|_D = \varphi$. If $\hat{\varphi}$ is the odd part of $\hat{\varphi}_*$, that is, $\hat{\varphi}(u) = \frac{1}{2}[\hat{\varphi}_*(u) - \hat{\varphi}_*(-u)]$, then $\hat{\varphi} : X \to \mathbb{R}^k$ is an odd map such that $0 \notin \varphi(D)$. Evidently, $\mathrm{cat}_X(D) = \mathrm{cat}_X(\bar{D})$. Also, since closedness and contractibility are preserved by homeomorphisms, we get the same values for homeomorphic X or homeomorphic closed D. However, the space X in cat_X is essential, since if $X \subseteq X_1$, sets may be contractible in X_1 but not in X (that is, if X is embedded continuously in X_1, then $\mathrm{cat}_{X_1}(D) \leqslant \mathrm{cat}_X(D)$ for any $D \subseteq X$).

Next we verify that the Krasnoselskii genus $\gamma(\cdot)$ and the Ljusternik–Schnirelmann category cat_X are both G-indices for suitable choices of the group G.

First we deal with the Krasnoselskii genus. In this case the group is

$$G = \mathbb{Z}_2 = \{\mathrm{id}, -\mathrm{id}\}.$$

Our first observation shows that the Krasnoselskii genus generalizes the dimension of the linear space.

Proposition 5.6.12 *If $U \subseteq \mathbb{R}^N$ is a bounded, symmetric neighborhood of the origin, then $\gamma(\partial U) = N$.*

Proof If in Definition 5.6.10(a) we choose $\varphi = \mathrm{id}$, we see that $\gamma(\partial U) \leqslant N$. Arguing by contradiction, suppose that $\gamma(\partial U) = m < N$. Then we can find an odd map $\varphi \in C(\mathbb{R}^N, \mathbb{R}^m)$ such that $0 \notin \varphi(\partial U)$. But this contradicts the Borsuk–Ulam Theorem (see Theorem 3.1.45). So, $\gamma(\partial U) = N$. $\square$

Corollary 5.6.13 *If $S^{N-1} = \{u \in \mathbb{R}^N : |u| = 1\}$, then $\gamma(S^{N-1}) = N$.*

Moreover, if X is an infinite-dimensional separable Banach space, then $\gamma(\partial B_1) = +\infty$, where $\partial B_1 = \{u \in X : ||u|| = 1\}$.

The next proposition shows that the Krasnoselskii genus is indeed an index (for $G = \mathbb{Z}_2 = \{\mathrm{id}, -\mathrm{id}\}$).

Proposition 5.6.14 *If X is a Banach space and $\mathscr{S} = \{C \subseteq X : C$ is closed and symmetric $\}$, then*

(a) $C_1, C_2 \in \mathscr{S}, C_1 \subseteq C_2 \Rightarrow \gamma(C_1) \leqslant \gamma(C_2)$;
(b) *for* $C_1, C_2 \in \mathscr{S}$, *we have* $\gamma(C_1 \cup C_2) \leqslant \gamma(C_1) + \gamma(C_2)$;
(c) *for* $C \in \mathscr{S}$ *and odd* $h \in C(C, X)$ *we have*

$$\gamma(C) \leqslant \gamma(h(C));$$

(d) *for* $C \in \mathscr{S}$ *compact with* $0 \notin C$, *we have* $\gamma(C) < \infty$ *and there exists a symmetric neighborhood* U *of* C *such that* $\gamma(\bar{U}) = \gamma(C)$.

Proof (*a*) This follows immediately from Definition 5.6.10(a).

(*b*) Let $\gamma(C_1) = m$ and $\gamma(C_2) = n$. Then we can find $\varphi_1 \in C(X, \mathbb{R}^m)$ and $\varphi_2 \in C(X, \mathbb{R}^n)$ such that $0 \notin \varphi_1(C_1)$ and $0 \notin \varphi_2(C_2)$ (see Remark 5.6.11). Let $\psi : X \to \mathbb{R}^m \times \mathbb{R}^n$ be defined by

$$\psi(u) = (\varphi_1(u), \varphi_2(u)) \text{ for all } u \in X.$$

Evidently, ψ is continuous, odd and $\psi(u) \neq 0$ for all $u \in C_1 \cup C_2$. Therefore

$$\gamma(C_1 \cup C_2) \leqslant m + n = \gamma(C_1) + \gamma(C_2).$$

(*c*) Suppose $\gamma(h(C)) = m$. Then we can find an odd $\varphi \in C(X, \mathbb{R}^m)$ such that $0 \notin \varphi(h(C))$. Let $\psi = \varphi \circ h$. Then $\psi \in C(C, \mathbb{R}^m \setminus \{0\})$ and it is odd. Therefore $\gamma(C) \leqslant m = \gamma(h(C))$.

(*d*) Let $u \in C$ and let $\epsilon > 0$ be such that $B_\epsilon(u) \cap B_\epsilon(-u) = \emptyset$ (recall $u \neq 0$). Let $C_u = B_\epsilon(u) \cup B_\epsilon(-u)$. We claim that $\gamma(C_u) = 1$. Indeed, since $B_\epsilon(u) \cap B_\epsilon(-u) = \emptyset$, we can find an odd function $\varphi \in C(X, \mathbb{R})$ such that $\varphi(u) = \lambda \neq 0$ for all $u \in B_\epsilon(u)$, hence $\varphi|_{B_\epsilon(u) \cup B_\epsilon(-u)} \neq 0$ and so from Remark 5.6.11 we infer that $\gamma(C_u) = 1$. Note that $\{C_u\}_{u \in C}$ is an open cover of C. The compactness of C implies that we can find a finite set $\{u_k\}_{k=1}^m$ such that $C \subseteq \bigcup\limits_{\mathrm{k}=1}^{m} C_{u_k}$. From parts (*a*) and (*b*), we have

$$\gamma(C) \leqslant \sum_{\mathrm{k}=1}^{m} \gamma(C_{u_k}) = m.$$

Suppose that $\gamma(C) = m$. Then we find odd $\varphi \in C(X, \mathbb{R}^m)$ such that $0 \notin \varphi(C)$. Exploiting the continuity of φ, we can find $\epsilon > 0$ such that $\varphi(u) \neq 0$ for all $u \in \bar{U}_\epsilon = \{u \in X : d(u, C) \leqslant \epsilon\}$. Then $\gamma(\bar{U}_\epsilon) \leqslant m$. But $C \subseteq \bar{U}_\epsilon$ and so from part (*a*) we have $m = \gamma(C) \leqslant \gamma(U_\epsilon)$. We conclude that $\gamma(\bar{U}_\epsilon) = m$. ☐

Now we turn our attention to the Ljusternik–Schnirelmann category. The next proposition shows that cat_X is an index for $G = \{\mathrm{id}\}$ and $\mathscr{C} = \{C \subseteq X : C \text{ is closed}\}$.

Proposition 5.6.15 *If X is a Banach space and $\mathscr{C} = \{C \subseteq X : C \text{ is closed}\}$, then*

(a) $C_1, C_2 \in \mathscr{C}$, $C_1 \subseteq C_2 \Rightarrow \mathrm{cat}_X(C_1) \leqslant \mathrm{cat}_X(C_2)$;
(b) *for* $C_1, C_2 \in \mathscr{C}$, *we have* $\mathrm{cat}_X(C_1 \cup C_2) \leqslant \mathrm{cat}_X(C_1) + \mathrm{cat}_X(C_2)$;
(c) *for* $C \in \mathscr{C}$ *and* $h \in C(C, X)$ *homotopic to the identity, we have*

$$\mathrm{cat}_X(C) \leqslant \mathrm{cat}_X(\overline{h(C)});$$

(d) *for* $C \in \mathscr{C}$ *compact, we have* $\mathrm{cat}_X(C) < \infty$ *and there is a neighborhood* V *of* C *such that* $\mathrm{cat}_X(\bar{V}) = \mathrm{cat}_X(C)$.

Proof (a) If $C_1 \subseteq C_2$, then any cover of C_2 is automatically a cover of C_1 and so directly from Definition 5.6.10(b), we have $\mathrm{cat}_X(C_1) \leqslant \mathrm{cat}_X(C_2)$.

(b) If $C_1 \subseteq \bigcup_{k=1}^{m} C_k^1$ and $C_2 \subseteq \bigcup_{j=1}^{n} C_j^2$ with C_k^1, C_j^2 $(k = 1, \ldots, m,\ j = 1, \ldots, n)$ closed and contractible. Then

$$C_1 \cup C_2 \subseteq \left(\bigcup_{k=1}^{m} C_k^1\right) \cup \left(\bigcup_{j=1}^{n} C_j^2\right)$$
$$\Rightarrow \mathrm{cat}_X(C_1 \cup C_2) \leqslant m + n$$
$$\Rightarrow \mathrm{cat}_X(C_1 \cup C_2) \leqslant \mathrm{cat}_X(C_1) + \mathrm{cat}_X(C_2).$$

(c) Let $m = \mathrm{cat}_X(h(C))$. Then $\eta(C) \subseteq \bigcup_{k=1}^{m} D_k$ with each D_k closed and contractible in X and let $C_k = h^{-1}(D_k)$, $k = 1, \ldots, m$. Each C_k is closed in C and because C is closed in X, C_k is also closed in X. Moreover, since h is homotopic to the identity and D_k is contractible, we have that C_k is contractible. We have

$$C \subseteq \bigcup_{k=1}^{m} C_k$$
$$\Rightarrow \mathrm{cat}_X(C) \leqslant m = \mathrm{cat}_X(h(C)).$$

(d) Every $u \in C$ has a convex (hence contractible, see Remark 5.6.9) open neighborhood V_u. Then $\mathrm{cat}_X(\bar{V}_u) = 1$. The family $\{V_u\}_{u \in C}$ is an open cover of C. The compactness of C implies that we can find a finite subcover $\{V_{u_k}\}_{k=1}^n$, $u_k \in C$ for all $k = 1, \ldots, n$. Then

$$C \subseteq \bigcup_{k=1}^{n} V_{u_k}$$
$$\Rightarrow \mathrm{cat}_X(C) \leqslant \sum_{k=1}^{n} \mathrm{cat}_X(\bar{V}_{u_k}) = n \text{ (see parts } (a),\ (b)).$$

So, we have proved that $\mathrm{cat}_X(C)$ is finite. To prove the second part of (d), suppose $\mathrm{cat}_X(C) = m$. Then $C \subseteq \bigcup_{k=1}^{m} C_k$ with each C_k being closed and contractible. We can always assume that each C_k is in fact compact (just consider $C \cap C_k$ if necessary). Then we can find a neighborhood V_k of C_k such that each $\bar{V}_k$ is contractible. Let $V = \bigcup_{k=1}^{m} V_k$. Then $\bar{V} \subseteq \bigcup_{k=1}^{m} \bar{V}_k$ and so $\mathrm{cat}_X(\bar{V}) \leqslant m = \mathrm{cat}_X(C)$. But from part (a) we have $\mathrm{cat}_X(C) \leqslant \mathrm{cat}_X(V)$. Therefore we conclude that $\mathrm{cat}_X(\bar{V}) = \mathrm{cat}_X(C)$. $\square$

Remark 5.6.16 In fact, to be more in line with Definition 5.6.10, where X is a Hausdorff topological space (no linear structure is assumed, since after all $\mathrm{cat}_X(C)$ is a purely topological notion), we can have properties (a)–(d) of the above proposition for all elements in 2^X, provided we assume that X is path connected and has the neighborhood extension property (for short, X is an NES), if for any normal space Y, a closed set $B \subseteq Y$ and any $\psi \in C(B, X)$, there exist a neighborhood V of B and a map $\hat{\psi} \in C(V, X)$ such that $\hat{\psi}|_B = \psi$. These properties have as a consequence that any point $u \in X$ has a neighborhood which is contractible in X. A careful reading of part (d) of the above proposition reveals that this property was crucial in the proof.

Finally, we mention a result of Rabinowitz [345, Theorem 3.7] relating the Krasnoselskii genus and the Ljusternik–Schnirelmann category.

Proposition 5.6.17 *If $C \subseteq \mathbb{R}^N \setminus \{0\}$ is compact symmetric and $\widetilde{C} = C/_{\mathbb{Z}_2}$ (the antipodal points are identified), then $\gamma(A) = \operatorname{cat}_{\mathbb{R}^N \setminus \{0\}/_{\mathbb{Z}_2}}(\widetilde{C})$.*

Now we consider some consequences of the abstract multiplicity result Theorem 5.6.7.

Proposition 5.6.18 *If X is a Banach space, $\varphi \in C^1(X)$ is even and bounded below, it satisfies the C-condition and there exists a compact symmetric set C such that $\gamma(C) = m$ and $\sup\limits_C \varphi < \varphi(0)$, then φ has at least m-distinct pairs $\{\pm u_k\}_{k=1}^m$ of critical points with*

$$\varphi(-u_k) = \varphi(u_k) < \varphi(0).$$

Proof Let $\{c_k\}_{k \geqslant 1}$ be as in (5.125). Since φ is bounded below, $\gamma(C) = m$ and using (5.125), we have

$$-\infty < \inf_X \varphi \leqslant c_1 \leqslant c_2 \leqslant \ldots \leqslant c_m \leqslant \max_C \varphi < \varphi(0).$$

From Theorem 5.6.7, we have that for each $k \in \{1, \ldots, m\}$, $K_\varphi^{c_k} \neq \emptyset$, it is symmetric (recall that φ is even) and $0 \notin K_\varphi^{c_k}$. If the values c_k are distinct, then the sets $K_\varphi^{c_k}$ $(k \in \{1, \ldots, m\})$ are pairwise disjoint and we have the desired conclusion. If for some $k \in \{2, \ldots, m\}$ we have $c_{k-1} = c_k$, then Theorem 5.6.7 implies that $\gamma(K_\varphi^{c_k}) \geqslant 2$ and so $K_\varphi^{c_k}$ is infinite (otherwise we must have $\gamma(K_\varphi^{c_k}) = 1$, see the proof of Proposition 5.6.14(d)). So, again we have reached the desired conclusion. □

An analogous result is also true for functionals of the form

$$j = \varphi + \psi \text{ with } \varphi \in C^1(X) \text{ and } \psi \in \Gamma_0(X).$$

Theorem 5.6.19 *If X is a Banach space, $j = \varphi + \psi$ is as above with φ, ψ even, $j(0) = 0$, j satisfies the GPS-condition and*

$$-\infty < c_k < 0 \textit{ for all } k \in \{1, \ldots, m\} \ (c_k \textit{ as in (5.125)}),$$

then $j(\cdot)$ has at least m-distinct pairs $\{\pm u_k\}$ of nontrivial critical points.

We can also have a $\mathbb{Z}_2$-version (symmetric version) of the mountain pass theorem. We present the result for the more general setting of functionals $j = \varphi + \psi$ with $\varphi \in C^1(X),\ \psi \in \Gamma_0(X)$. If $\psi \equiv 0$, then we recover the standard symmetric mountain pass theorem.

Theorem 5.6.20 *If X is a Banach space, $j = \varphi + \psi$ with $\varphi \in C^1(X)$, $\psi \in \Gamma_0(X)$, $j(0) = 0$, φ and ψ are even, j satisfies the GPS-condition and*

(i) there exists a subspace V of X of finite codimension and $\beta, \rho > 0$ such that $j|_{\partial B_\rho(0)\cap V} \geqslant \beta$;
(ii) there exists a finite-dimensional subspace Y of X with $\dim Y > \operatorname{codim} V$ such that $j|_Y$ is anticoercive (that is, $j(u) \to -\infty$ as $||u|| \to \infty,\ u \in Y$),

then j has at least $\dim Y - \operatorname{codim} V$ distinct pairs of nontrivial critical points.

Proof Let $\vartheta > 0$ be such that $j^{-\vartheta} = \{u \in X : j(u) \leqslant -\vartheta\}$ contains no critical points of $j(\cdot)$ (if no such $\vartheta > 0$ exists, then we already have an infinite sequence of antipodal pairs of critical points of j and so we are done). From hypothesis (ii) we can find $\rho_0 > \rho$ such that

$$j|_{\partial B_{\rho_0}(0)\cap Y} \leqslant -\vartheta.$$

Let $m = \operatorname{codim} V,\ k = \dim Y$ and $B = \bar{B}_{\rho_0}(0) \cap Y = \{u \in Y : ||u|| \leqslant \rho_0\}$.
For $i \in \{1, \dots, m\}$ we introduce the following items

$\mathscr{D} = \big\{h \in C(B, X) : h$ is odd, $h|_{\partial B}$ is homotopic to $\mathrm{id}|_{\partial B}$ in $j^{-\vartheta}$ by an odd homotopy$\}$,
$\Gamma_i = \{h(B\backslash U) : h \in \mathscr{D},\ U$ is open in B and symmetric, $U \cap \partial B = \emptyset$ and for each $C \subseteq X,\ 0 \notin C$, closed symmetric $\gamma(C) \leqslant k - i\ \}$,
$\Delta_i = \{C \subseteq X : 0 \notin C,\ C$ is closed, symmetric (that is, $C \in \mathscr{S}$, see Definition 5.6.10) and for each open set W with $C \subseteq W$, there exists a $D_0 \in \Gamma_i$ with $D_0 \subseteq W\ \}$.

Note that $B \in \Delta_i$ with $D_0 = B,\ U = \emptyset$ and $h = \mathrm{id}|_B$.
We introduce the following minimax values

$$C_i = \inf_{C\in\Delta_i} \sup_{u\in C} \varphi(u) \text{ for all } i \in \{1, \dots, m\}.$$

With standard topological arguments (see Szulkin [397], Lemmas 4.5 and 4.6) we can show the following simple facts.

(a) $\beta \leqslant c_i$ for all $i \in \{m + 1, \dots, k\}$;
(b) $\Delta_{i+1} \subseteq \Delta_i$ for all $i \in \{1, \dots, k - 1\}$;
(c) if $C \in \Delta_i$, D is a closed and symmetric set such that $C \subseteq \operatorname{int} D$ and $\psi : D \to X$ an odd map such that $\psi|_{D\cap j^{-\vartheta}}$ is homotopic to $\mathrm{id}|_{D\cap j^{-\vartheta}}$ by an odd homotopy, then $\psi(C) \in \Delta_i$;

(d) if $D \in \mathscr{S}$ (see Definition 5.6.10), $\gamma(D) \leqslant l$ and $j|_D > -\vartheta$, then there exists a number $\delta > 0$ such that for every $C \in \Delta_i$, we have

$$C \setminus \text{int}\,(\bar{D})_\delta \in \Delta_i.$$

From properties (a) and (b) above, we have

$$\beta \leqslant c_{m+1} \leqslant \ldots \leqslant c_k.$$

Suppose that

$$c_{i+1} = \ldots = c_{i+1+l} = c \text{ for some } l \in \{0, \ldots, k-m+1\}.$$

Note that $c \geqslant \beta > 0$ and so $0 \notin K_j^c$. Also, since j is even, we see that K_j^C is symmetric and also compact (see Proposition 5.1.27). It follows that $K_j^c \in \mathscr{S}$ (see Definition 5.6.10).

We show that $l + 1 \leqslant \gamma(K_j^c)$. Arguing by contradiction, suppose that $\gamma(K_j^c) \leqslant l$. From Proposition 5.6.14 (d), we know that there exists a $\delta > 0$ such that

$$\gamma(K_j^c) = \gamma(\overline{(K_j^c)_\delta}).$$

Let $\epsilon_0 > 0$ and $U = \text{int}(\overline{(K_j^c)_\delta})$. Then according to Theorem 5.3.16 (see also Remark 5.3.17), we can find $\epsilon \in (0, \epsilon_0)$ and a deformation $h : [0,1] \times X \to X$ satisfying the properties of that theorem. So, we have

$$h(s,\cdot) \circ h(t,\cdot) = h(t+s,\cdot) \text{ for all } t, s \in [0,1],\, t+s \leqslant 1 \text{ (semigroup property);}$$
$$h(t,\cdot) \text{ is an odd homeomorphism for all } t \in [0,1].$$

Theorem 5.3.16 implies that

$$h(1,u) = u \text{ for all } u \in j^{c+\epsilon_0} \setminus j^{c-\epsilon_0},$$
$$h(1, j^{c+\epsilon} \setminus U) \subseteq j^{c-\epsilon}.$$

We can find $C \in \Delta_{i+1}$ such that

$$\sup_C j \leqslant c + \epsilon.$$

Since K_j^c is compact, $\gamma(K_j^c) \leqslant l$ and $j|_{K_j^c} > -\vartheta$, from (d) above, we have that by choosing $\delta > 0$ small, we can say that

$$C \setminus U \in \Delta_i \text{ (recall the definition of } U).$$

In addition we have

$$C \backslash U \subseteq j^{c+\epsilon} \backslash U$$
$$\Rightarrow h(1, C \backslash U) \subseteq j^{c-\epsilon}.$$

So, if $\epsilon_0 > 0$ is sufficiently small, then

$$h(1,u) = u \text{ for all } u \in j^{-\vartheta}$$
$$\Rightarrow h(1, C \backslash U) \in \Delta_i \text{ (see } (c) \text{ above).}$$

Then we have

$$\sup_{h(1,C\backslash U)} j \leqslant c - \epsilon,$$

which contradicts the fact that $\sup\limits_{h(1,C\backslash U)} j \geqslant c$. This proves that

$$\gamma(K_j^c) \geqslant l + 1$$
$$\Rightarrow \gamma(K_j^c) \geqslant 1$$
$$\Rightarrow K_j^c \text{ has at least an antipodal pair } \{\pm u_i\}.$$

This produces the claimed number of critical point pairs if the $c_i' s$ are distinct. If they are not, then $l > 0$ for some i and so $\gamma(K_j^{c_i}) > 1$ and so j has an infinity of critical antipodal pairs. □

Corollary 5.6.21 *Let X be a Banach space and suppose $j = \varphi + \psi$ with $\varphi \in C^1(X)$, $\psi \in \Gamma_0(X)$, $j(0) = 0$, φ and ψ are even, j satisfies the GPS-condition and*

(i) there exists a subspace V of X of finite codimension and $\beta, \rho > 0$ such that $j|_{\partial B_\rho(0) \cap V} \geqslant \beta$;
(ii) for any $k \in \mathbb{N}$, there is a k-dimensional subspace Y of X such that $j|_Y$ is anti-coercive.

Then j has infinitely many distinct antipodal critical pairs with an unbounded sequence of corresponding critical values.

When a variational problem exhibits symmetry which is expressed as invariance under the action of a group G, then it is natural and important to look for critical points which are invariant under the action of G.

Our setting is the following. Let X be a Banach space and G be a topological group acting on X (that is, G admits a representation on X, see Definition 5.6.1). The action of G on X naturally leads to an action of G on X^*, by setting

$$\langle gu^*, u \rangle = \langle u^*, g^{-1}u \rangle \text{ for all } g \in G, \text{ all } u \in X, \text{ and all } u^* \in X^*.$$

We consider the linear subspaces of G-invariant (fixed) points of X and X^*:

$$\Sigma = X^G = \{u \in X : gu = u \text{ for all } g \in G\},$$
$$\Sigma_* = (X^*)^G = \{u^* \in X^* : gu^* = u^* \text{ for all } g \in G\}.$$

Also, let $C^1_G(X)$ be the subset of all G-invariant elements of $C^1(X)$. We consider the following principle, known in the literature as the "principle of symmetric criticality".

$$(PSC): \text{ For all } \varphi \in G^1_G(X), \text{ we have "} (\varphi|_\Sigma)'(u) = 0 \Rightarrow \varphi'(u) = 0\text{"}. \quad (5.129)$$

Remark 5.6.22 Using the language of Sect. 5.5, this principle says that for every G-invariant C^1-functional φ, Σ is a natural constraint for φ (see Definition 5.5.11).

Theorem 5.6.23 *The* (PSC) *is valid if and only if* $\Sigma_* \cap \Sigma^\perp = \{0\}$, *where* $\Sigma^\perp = \{u^* \in X^* : \langle u^*, u\rangle = 0$ *for all* $u \in \Sigma\}$.

Proof $\underline{\Rightarrow}$: Suppose that the (PSC) holds and arguing by contradiction suppose that $\Sigma_* \cap \Sigma^\perp \neq \{0\}$. Let $u^* \in \Sigma_* \cap \Sigma^\perp$, $u^* \neq 0$. Let $\varphi_*(u) = \langle u^*, u\rangle$ for all $u \in X$. Then $\varphi_* \in C^1(X)$ and

$$\begin{aligned}\varphi_*(gu) = \langle u^*, gu\rangle = \left\langle g^{-1}u^*, u\right\rangle &= \langle u^*, u\rangle \text{ (since } u^* \in \Sigma^*)\\ &= \varphi_*(u) \text{ for all } g \in G\end{aligned}$$
$$\Rightarrow \varphi_* \in C^1(X).$$

We have $\varphi_*'(\cdot) = u^*(\cdot) \neq 0$. Hence $K_{\varphi_*} = \emptyset$. But $u^* \in \Sigma^\perp$, so $\left\langle \varphi_*'(u), u\right\rangle = \langle u^*, u\rangle = 0$ for all $u \in \Sigma$, hence $(\varphi_*|_\Sigma)'(u) = 0$ for all $u \in \Sigma$, which contradicts the (PSC) (see (5.129)).

$\underline{\Leftarrow}$: We assume that $\Sigma_* \cap \Sigma^\perp = \{0\}$.

Let $u_0 \in K_{\varphi|_\Sigma}$. Since Σ is a linear subspace of X, we have

$$\begin{aligned}&\varphi|_\Sigma(u_0 + h) = \varphi(u_0 + h) \text{ for all } h \in \Sigma\\ \Rightarrow &\left\langle (\varphi|_\Sigma)'(u_0), h\right\rangle_\Sigma = \left\langle \varphi'(u_0), h\right\rangle \text{ for all } h \in \Sigma.\end{aligned} \quad (5.130)$$

Here by $\langle\cdot,\cdot\rangle_\Sigma$ we denote the duality brackets for the pair (Σ^*, Σ) and by $\langle\cdot,\cdot\rangle$ the duality brackets for the pair (X^*, X). From (5.130) and since $u_0 \in K_{\varphi|_\Sigma}$ we have

$$\begin{aligned}&\left\langle \varphi'(u_0), h\right\rangle = 0 \text{ for all } h \in \Sigma\\ \Rightarrow &\varphi'(u_0) \in \Sigma^\perp.\end{aligned} \quad (5.131)$$

From Proposition 5.6.3, we have

$$\begin{aligned}
&\langle \varphi'(gu), h\rangle = \langle \varphi'(u), g^{-1}h\rangle = \langle g\varphi'(u), h\rangle \text{ for all } g \in G, \text{ all } u, h \in X \\
&\Rightarrow \varphi' \text{ is } G\text{-equivariant (see Definition 5.6.1)} \\
&\Rightarrow g\varphi'(u_0) = \varphi'(u_0) \text{ for all } g \in G \text{ (since } u_0 \in \Sigma) \\
&\Rightarrow \varphi'(u_0) \in \Sigma_* \\
&\Rightarrow \varphi'(u_0) \in \Sigma_* \cap \Sigma^\perp \text{ (see (5.131), hence } \varphi'(u_0) = 0, \text{ that is } u_0 \in K_\varphi).
\end{aligned}$$

The proof is now complete. □

Next, we want to produce conditions which guarantee that $\Sigma_* \cap \Sigma^\perp = \{0\}$. We consider two distinct cases. In the first, we assume that the action of G on X is isometric. In the second, we assume that the group G is compact.

First we deal with the case of isometric action.

Proposition 5.6.24 *If the action of G on X is isometric, then the induced action of G on X^* is isometric too.*

Proof For every $g \in G$ and every $u^* \in X^*$, we have

$$\begin{aligned}
||gu^*||_* = \sup_{||h||\leqslant 1} |\langle gu^*, h\rangle| &= \sup_{||h||\leqslant 1} |\langle u^*, g^{-1}h\rangle| \\
&\leqslant \sup_{||h||\leqslant 1} ||u^*||_* ||g^{-1}h|| \\
&= \sup_{||h||\leqslant 1} ||u^*||_* ||h|| \text{ (since the action on } X \text{ is isometric)} \\
&= ||u^*||_*. \qquad (5.132)
\end{aligned}$$

Similarly, we show that

$$||u^*||_* = ||g^{-1}(gu^*)||_* \leqslant ||gu^*||_* \text{ for every } g \in G \text{ and every } u^* \in X^*. \qquad (5.133)$$

From (5.132) and (5.133), we have

$$\begin{aligned}
&||u^*||_* = ||gu^*||_* \text{ for every } g \in G \text{ and every } u^* \in X^* \\
\Rightarrow\ &\text{the induced action of } G \text{ on } X^* \text{ is isometric.}
\end{aligned}$$

The proof is now complete. □

Proposition 5.6.25 *If X is a reflexive and strictly convex Banach space, the action of G on X is isometric and $J : X \to X^*$ is the duality map for X, then $J^{-1}(\Sigma_*) \subseteq \Sigma$.*

Proof From Definition 2.7.21 and Theorem 2.8.10, we know that $J(\cdot)$ is maximal monotone. Moreover, since $||J(u)||_* = ||u||$ for all $u \in X$, we have that J is coercive. So, Corollary 2.8.7 implies that J is surjective. Since X is strictly convex, the map $J^{-1} : X^* \to X$ is single-valued and is the duality map of X^* (see Proposition 2.7.27). Let $u^* \in \Sigma_*$. For all $g \in G$, we have

$$\begin{aligned}||gJ^{-1}(u^*)|| &= ||J^{-1}(u^*)|| \text{ since the action of } G \text{ on } X \text{ is isometric)}\\ &= ||u^*||_* \text{ (since } J^{-1} \text{ is the duality of } X^*) \qquad (5.134)\\ \text{and } \langle u^*, gJ^{-1}(u^*)\rangle &= \langle g^{-1}u^*, J^{-1}(u^*)\rangle = \langle u^*, J^{-1}(u^*)\rangle \text{ (since } u^* \in \Sigma)\\ &= ||u^*||_*^2\end{aligned}$$

$$\begin{aligned}&\Rightarrow gJ^{-1}(u^*) = J^{-1}(u^*) \text{ for all } g \in G \text{ (see (5.134))}\\ &\Rightarrow J^{-1}(u^*) \in \Sigma \text{ for all } u^* \in \Sigma_*.\end{aligned}$$

The proof is now complete. □

Theorem 5.6.26 *If X is a reflexive Banach space with strictly convex dual and the action of G on X is isometric, then the (PSC)-holds.*

Proof According to Theorem 5.6.23, we need to show that $\Sigma_* \cap \Sigma^\perp = \{0\}$. Let $u^* \in \Sigma_* \cap \Sigma^\perp$. Then by Proposition 5.6.25 we have $J^{-1}(u^*) \in \Sigma$. Hence

$$\begin{aligned}&||u^*||_*^2 = \langle u^*, J^{-1}(u^*)\rangle = 0 \text{ (since } u^* \in \Sigma^\perp)\\ &\Rightarrow \Sigma_* \cap \Sigma^\perp = \{0\} \text{ and so the } (PSC) \text{ holds.}\end{aligned}$$

The proof is now complete. □

Now we consider the case where the group G is compact. To this end, let μ be the Haar probability measure on G. For each $u \in X$, there exists a unique $A(u) \in X$ such that

$$\langle u^*, A(u)\rangle = \int_G \langle u^*, gu\rangle \, d\mu \text{ for all } u^* \in X^*.$$

The map A is a continuous linear projection of X onto Σ and if $C \subseteq X$ is closed, convex and G-invariant, then $A(C) \subseteq C$. Note that $A(u)$ is the barycenter of μ (see Rudin [366]).

Theorem 5.6.27 *If X is a Banach space and G is a compact topological group acting on X, then the (PSC) holds.*

Proof Again Theorem 5.6.23 says that it suffices to show that $\Sigma_* \cap \Sigma^\perp = \{0\}$. So, let $u^* \in \Sigma_* \cap \Sigma^\perp$ and suppose that $u^* \neq 0$. We consider the hyperplane

$$H = \{u \in X : \langle u^*, u\rangle = 1\}. \qquad (5.135)$$

This is a nonempty, closed, convex and G-invariant subset of X. So, for all $u \in H$ we have $A(u) \in H \cap \Sigma$. Then $\langle u^*, A(u)\rangle = 0$ for all $u \in H$ (since $u^* \in \Sigma^\perp$). But this contradicts the fact that $A(u) \in H$ (see (5.135)). □

We can extend this principle to functionals of the form

$$j = \varphi + \psi \text{ with } \varphi \in C^1(X), \psi \in \Gamma_0(X).$$

The "nonsmooth principle of symmetric criticality" has the following form:

$(NPSC)$: If $j = \varphi + \psi$ with $\varphi \in C^1(X)$, $\psi \in \Gamma_0(X)$ is G-invariant and $0 \in (\varphi|_\Sigma)'(u) + \partial(\psi|_\Sigma)(u)$, then $0 \in \varphi'(u) + \partial\psi(u)$ (that is, $u \in K_j$).

The following theorem is due to Kobayashi and Otani [241].

Theorem 5.6.28 *Assume that X is a Banach space, G is a topological group acting on X and one of the following conditions holds:*

(i) X is reflexive, both X and X^* are strictly convex and the action of G on X is isometric; or
(ii) G is a compact topological group.

Then the (NPSC) holds.

We conclude our discussion of critical point theory under a linear group action, by stating the so-called "Fountain Theorem". The result uses the following notion.

Definition 5.6.29 Let Y be a finite-dimensional Banach space and G a compact topological group acting on G. For every $k \in \mathbb{N}$, $k \geqslant 2$, let $Y^k = Y \times \cdots \times Y$ (k times) and suppose that G acts on Y^k diagonally, that is,

$$g(y_1, \ldots, y_k) = (gy_1, \ldots, gy_k).$$

We say that this action is "admissible" if for all $k \in \mathbb{N}$, $k \geqslant 2$, and each U an open, bounded, invariant neighborhood of the origin in Y^k, we have that every continuous, equivariant map $h : \partial U \to Y^{k-1}$ has a zero.

Remark 5.6.30 From the Borsuk–Ulam theorem (see Theorem 3.1.45), we know that the action of $G = \mathbb{Z}_2 = \{\mathrm{id}, -\mathrm{id}\}$ on $Y = \mathbb{R}$ is admissible.

The setting for the "fountain theorem" is the following. Let X be a Banach space and G a compact topological group acting on X. We assume that the action of G on X is isometric. Assume that

$$X = \overline{\underset{\mathrm{k}\geqslant 1}{\oplus} X_k}$$

where each space X_k is invariant and there exists a finite-dimensional space Y such that G acts on Y and the action is admissible, and for each $k \in \mathbb{N}$ there is an equivariant isomorphism $i_k : Y \to X_k$.

We set

$$Z_m = \overset{m}{\underset{\mathrm{k}=1}{\oplus}} X_k \text{ and } V_m = \overline{\underset{\mathrm{k}\geqslant \mathrm{m}}{\oplus} X_k}.$$

Using Theorem 5.6.5 we can have the following multiplicity theorem due to Bartsch [35] (see also Willem [415, p. 58]). The result is known as the "fountain theorem".

Theorem 5.6.31 *If the above setting holds, $\varphi \in C^1(X)$ is invariant and satisfies the $(PS)_c$-condition for every $c > 0$ and for every $m \in \mathbb{N}$, we can find $\rho_m, r_m > 0$ such that*

(i) $\max[\varphi(u) : u \in Z_m, ||u|| \leqslant \rho_m] \leqslant 0$;
(ii) $\inf\{\varphi(u) : u \in V_m, ||u|| = r_m\} \to +\infty$ *as* $m \to \infty$,

then φ has an unbounded sequence of critical values.

5.7 The Structure of the Critical Set

In this section we investigate the fine structure of the functional near a critical point. Different types of critical points can be distinguished by the topological structure of their neighborhoods in the sublevel sets.

Let X be a Banach space, $\varphi \in C^1(X)$ and $c \in \mathbb{R}$. We recall the following notation, which will be used extensively in the sequel:

$K_\varphi = \{u \in X : \varphi'(u) = 0\}$ (the critical set of φ),
$K_\varphi^c = \{u \in K_\varphi : \varphi(u) = c\}$ (the critical points of φ at the level $c \in \mathbb{R}$),
$\varphi^c = \{u \in X : \varphi(u) \leqslant c\}$ (the sublevel set at c),
$\varphi_c = \{u \in X : \varphi(u) \geqslant c\}$ (the superlevel set at c),
$\overset{\circ}{\varphi}{}^c = \{u \in X : \varphi(u) < c\}$ (the strict sublevel set at c).

Remark 5.7.1 Note that φ^c and φ_c are (possibly empty) closed subsets of X, while $\overset{\circ}{\varphi}{}^c$ is a (possibly empty) open set in X. We have

$$\overline{\overset{\circ}{\varphi}{}^c} \subseteq \overset{\circ}{\varphi}{}^c \cup \varphi^{-1}(c) \text{ and } \partial\, \overset{\circ}{\varphi}{}^c \subseteq \varphi^{-1}(c)$$

and the inclusions can be strict. Indeed, let $\varphi(u) = u(1-u)^2$ for all $u \in \mathbb{R}$. Then

$$\varphi^0 = (-\infty, 0]\,,\ \overline{\overset{\circ}{\varphi}{}^0} = (-\infty, 0]\,,\ \partial\, \overset{\circ}{\varphi}{}^0 = \{0\},$$
$$\varphi^{-1}(0) = \{0, 1\},\ \overset{\circ}{\varphi}{}^0 \cup \varphi^{-1}(0) = (-\infty, 0] \cup \{1\}.$$

However, if c is not a critical value of φ (that is, c is a regular value of φ), then by the inverse function theorem, we have

$$\partial\, \overset{\circ}{\varphi}{}^c = \varphi^{-1}(c) \text{ and } \overline{\overset{\circ}{\varphi}{}^c} = \overset{\circ}{\varphi}{}^c \cup \varphi^{-1}(c).$$

We introduce the following two distinct kinds of critical points of φ.

Definition 5.7.2 Let $u_0 \in K_\varphi^c$ (not necessarily isolated in K_φ).

(a) We say that u_0 is a "local minimizer of φ" if there exists an open neighborhood of u_0 such that $\varphi(u_0) \leqslant \varphi(u)$ for all $u \in U$.

(b) We say that u_0 is of "mountain pass type" if for any open neighborhood U of u_0, we have

$$U \cap \overset{\circ}{\varphi^c} \neq \emptyset \text{ and } U \cap \overset{\circ}{\varphi^c} \text{ is not path connected}$$

where $c = \varphi(u_0)$.

We will need the following topological lemma.

Lemma 5.7.3 *If (E, d) is a metric space and $K, V \subseteq E$ are nonempty sets with K compact, V open, $K \subseteq \bar{V}$, $\{U(u)\}_{u\in K}$ is an open cover of K with $u \in U(u)$ and $U(u) \cap V$ is path connected for every $u \in K$, then there exists a finite, disjoint open cover $\{U_i\}_{i=1}^m$ of K such that for each $i \in \{1, \ldots, m\}$ the set $U_i \cap V$ is contained in a path-component of* $\left(\bigcup_{\mathrm{u}\in\mathrm{K}} U(u)\right) \cap V$.

Proof The compactness of K implies that we can find a finite subcover $\{U(u_i)\}_{i=1}^N$ of the open cover $\{U(u)\}_{u\in K}$ ($u_i \in K$ for all $i \in \{1, \ldots, N\}$). We set

$$\delta = \min_{u\in K} \max_{i\in\{1,\ldots,N\}} d(u, E \backslash U(u_i)).$$

We claim that $\delta > 0$. If $\delta = 0$, then for every $k \in \mathbb{N}$, we can find $u_k \in K$

$$d(u_k, E \backslash U(u_i)) \leqslant \frac{1}{k} \text{ for all } i \in \{1, \ldots, N\}. \tag{5.136}$$

The compactness of K implies that by passing to a suitable subsequence if necessary we may assume that

$$u_k \to u \text{ in } E,\ u \in K. \tag{5.137}$$

Then from (5.136) we have

$$\begin{aligned}
&d(u, E \backslash U(u_i)) = 0 \text{ for all } i \in \{1, \ldots, N\} \\
\Rightarrow\ &u \in E \backslash U(u_i) \text{ for all } i \in \{1, \ldots, N\} \\
\Rightarrow\ &u \in E \backslash K, \text{ a contradiction (see (5.137)).}
\end{aligned}$$

So, we have $\delta > 0$. We have

$$B_\delta(u) \subseteq U(\bar{u}) \text{ for all } u \in K, \text{ some } \bar{u} \in K \tag{5.138}$$

(so $\delta > 0$ is a kind of Lebesgue number for the cover $\{U(u)\}_{u\in K}$).

We define the following equivalence relation on K

$$u \sim u'$$
if and only if

there are finitely many points $\{u_i\}_{i=0}^{k+1} \subseteq K$ such that $u_0 = u,\ u_{k+1} = u'$.

The compactness of K implies that we have a finite number of equivalence classes for N. Let us denote them by $K_1, \ldots, K_M$. We define

$$W_i = \{u \in E : d(u, K_i) < \frac{\delta}{4}\} \text{ for all } i \in \{1, \ldots, M\}. \tag{5.139}$$

We have

$$W_i \cap W_j = \emptyset \text{ for } i \neq j \text{ and } K \subseteq \bigcup_{\mathrm{i}=1}^{M} W_i.$$

It remains to show that each set $W_i \cap V$ is contained in a path component of $\left(\bigcup_{\mathrm{u}\in\mathrm{C}} U(u)\right) \cap V = D$.

To this end, we introduce a second equivalence relation $\overset{*}{\sim}$, this time on D. So, we define

$$u \sim u'$$
if and only if
u and u' belong to the same path-component of D.

Fix $i \in \{1, \ldots, M\}$ and let $u, u' \in W_i \cap V$. We need to show $u \overset{*}{\sim} u'$. From (5.139), we see that we can find a finite chain $u_i \in K_i$, $i = 0, \ldots, m+1$, such that

$$d(u, u_0) < \frac{\delta}{4}, d(u', u_{m+1}) < \frac{\delta}{4} \text{ and } d(u_i, u_{i+1}) < \delta \text{ for all } i = 0, \ldots, m+1. \tag{5.140}$$

Let $\epsilon = \delta - \max_{0 \leqslant i \leqslant m+1} d(u_i, u_{i+1}) > 0$ (see (5.140)). Since $K \subseteq \bar{U}$, for every $i \in \{0, \ldots, m+1\}$, we can find a $v_i \in V$ such that

$$\begin{aligned} & d(v_i, u_i) < \frac{\epsilon}{2} < \delta \\ \Rightarrow\ & d(v_i, v_{i+1}) < \delta \text{ for all } i \in \{0, \ldots, m\} \\ \Rightarrow\ & u, v_0 \in B_\delta(u_0) \cap V \subseteq U(\bar{u}_0) \cap V \text{ (see (5.138)) for some } \bar{u}_0 \in K \\ \Rightarrow\ & u \overset{*}{\sim} v_0. \end{aligned} \tag{5.141}$$

Similarly we show that $u' \overset{*}{\sim} v_{m+1}$. Finally, for $i \in \{0, \ldots, m\}$ we have

$$\begin{aligned} & d(u_i, v_{i+1}) \leqslant d(u_i, u_{i+1}) + d(u_{i+1}, v_{i+1}) < \delta - \epsilon + \frac{\epsilon}{2} < \delta \text{ (see (5.141))} \\ \Rightarrow\ & v_i, v_{i+1} \in B_\delta(u_i) \cap V \subseteq U(\bar{u}_i) \cap V \text{ for some } \bar{u}_i \in K \text{ (see (5.138))} \\ \Rightarrow\ & v_i \overset{*}{\sim} v_{i+1} \text{ for all } i = 0, \ldots, m. \end{aligned}$$

Therefore

$$u \overset{*}{\sim} v_0 \overset{*}{\sim} v_1 \overset{*}{\sim} \dots \overset{*}{\sim} v_{m+1} \overset{*}{\sim} u'.$$

The proof is now complete. □

We also have the following straightforward variant of the Deformation Theorem (see Theorem 5.3.7).

Proposition 5.7.4 *If $\varphi \in C^1(X)$ satisfies the C-condition, $\epsilon_0 > 0$, $c \in \mathbb{R}$, and U and V are open neighborhoods of K_φ^c such that*

$$\bar{U} \subseteq V \text{ and } d(\partial V, U) > 0,$$

then there exist $\epsilon \in (0, \epsilon_0]$ and a φ-decreasing locally Lipschitz homotopy of homeomorphisms $h : [0, 1] \times X \to X$ such that

(a) $h(1, \varphi^{c+\epsilon} \setminus U) \subseteq \varphi^{c-\epsilon}$;
(b) $h([0, 1] \times \bar{U}) \subseteq V$;
(c) $h(t, u) = u$ *for all* $t \in [0, 1]$ *and all* $u \in \varphi^{c-\epsilon_0} \cup \varphi_{c+\epsilon_0}$.

Now we have the first structural result for the critical set K_φ^c.

Theorem 5.7.5 *If $\varphi \in C^1(X)$, satisfies the C-condition, $u_0, u_1 \in X$, $u_0 \neq u_1$,*

$$\begin{aligned}
&\Gamma = \{\gamma \in C([0, 1], X) : \gamma(0) = u_0, \gamma(1) = u_1\},\\
&c = \inf_{\gamma \in \Gamma} \max_{t \in [0,1]} \varphi(\gamma(t)), \text{ and}\\
&c > \max\{\varphi(u_0), \varphi(u_1)\} = \xi
\end{aligned}$$

then $K_\varphi^c \neq \emptyset$ and one of the following holds:

(a) K_φ^c *contains a local minimizer of* φ*; or*
(b) K_φ^c *contains a critical point of mountain pass type.*

Proof Theorem 5.4.6 (the mountain pass theorem) guarantees that $K_\varphi^c \neq \emptyset$. To prove the theorem, we argue by contradiction. So, suppose K_φ^c contains no local minimizers nor critical points of mountain pass type. Hence for any $u \in K_\varphi^c$ we can find an open neighborhood $U(u)$ of u such that $U(u) \cap \overset{\circ}{\varphi}{}^c$ is path connected. Note that K_φ^c is compact and since it does not have local minimizers, we have

$$K_\varphi^c = \overline{\overset{\circ}{\varphi}{}^c}.$$

The family $\{U(u)\}_{u \in K_\varphi^c}$ is an open cover of the compact set K_φ^c. So, we can find a finite subcover $\{U_k\}_{k=1}^m$ of pairwise disjoint sets such that $U_k \cap \overset{\circ}{\varphi}{}^c$ is contained in a path-component of $\left(\bigcup_{u\in K_\varphi^c} U(u)\right) \cap \overset{\circ}{\varphi^c}$ (see Lemma 5.7.3). Let $V = \bigcup_{k=1}^{m} U_k$. Then $d(\partial V, K_\varphi^c) > 0$. Also, let $\widehat{U} = \bigcup_{u\in K_\varphi^c} U(u)$ and let

$$\epsilon_0 = \frac{1}{2}(c-\xi), \qquad \delta = \frac{1}{8}\min\{d(\partial\widehat{U} \cup \{u_0, u_1\}, K_\varphi^c), d(\partial V, K_\varphi^c)\} \text{ and}$$
$$U = \{u \in X : d(u, K_\varphi^c) < \delta\}.$$

Using these items, we apply Proposition 5.7.4 and obtain $\epsilon \in (0, \epsilon_0]$ and a φ-decreasing, locally Lipschitz homotopy of homeomorphisms $\{h_t(\cdot) = h(t, \cdot)\}_{t\in[0,1]}$ satisfying the properties of that proposition. From the definition of c, we can find $\gamma \in \Gamma$ such that

$$\varphi(\gamma(t)) \leqslant c + \epsilon \text{ for all } t \in [0, 1].$$

We define

$$\begin{aligned} &\Sigma = \{t \in [0, 1] : \gamma(t) \notin U\}, \\ &D = (\widehat{U} \cap K_\varphi^c) \cup h(1, \gamma(\Sigma)). \end{aligned} \tag{5.142}$$

Evidently, $u_0, u_1 \in D$ and let D_0 be the path component of D containing u_0. We will show that $u_1 \in D_0$. If $\Sigma = [0, 1]$, then this is clearly true. So, suppose $\Sigma \neq [0, 1]$. The set Σ is closed. We define

$$t^* = \sup\{t \in \Sigma : h(1, \gamma(t)) \in D_0)\}.$$

Suppose $t^* < 1$. Then 0 belongs to the relative interior of Σ in $[0, 1]$. So $t^* \in (0, 1)$. Let $[t_1^*, t_2^*]$ be the component of Σ containing t^*.

If $t_1^* < t^*$, then $t^* = t_2^*$. So, let $t^* = t_1^*$. Then $\gamma(t_1^*) \in \partial U$ and so $h(1, \gamma(t_1^*)) \in \operatorname{int} D$. So, we can find $\hat{\epsilon} > 0$ such that

$$\begin{aligned} &B_{\hat{\epsilon}}(h(1, \gamma(t_1^*))) \subseteq D \\ \Rightarrow\ &B_\epsilon(h, (1, \gamma(t_1^*))) \cap D_0 \neq \emptyset \\ \Rightarrow\ &h(1, \gamma(t_1^*)) = h(1, \gamma(t^*)) \in D_0 \\ \Rightarrow\ &h(1, \gamma(t_2^*)) \in D_0 \text{ and so } t^* \geqslant t_2^* > t_1^*, \text{ a contradiction.} \end{aligned}$$

Therefore $t^* = t_2^*$ and we have

$$\begin{aligned} &h(1, \gamma(t^*)) \in D_0 \\ \Rightarrow\ &\gamma(t^*) \in \partial U. \end{aligned}$$

Let $k_0 \in \{1, \ldots, m\}$ such that

$$d(\gamma(t^*), U_{k_0} \cap K_\varphi^c) = \delta.$$

Let $\hat{t} = \sup\{t \in [0, 1] : \gamma(t) \in \overline{U \cap U_{k_0}}\}$. Since $t^* = t_2^*$, we have $\hat{t} \in (t^*, 1)$. Also, we have $\gamma(\hat{t}) \in \partial(U \cup U_{k_0})$ and so $\gamma(\hat{t}) \in \Sigma$ (see (5.142)). Then

$$z = h(1, \gamma(\hat{t})) \in U_{k_0} \cap \overset{\circ}{\varphi^c},$$
$$w = h(1, \gamma(t^*)) \in U_{k_0} \cap \overset{\circ}{\varphi^c}.$$

Recall that $U_{k_0} \cap \overset{\circ}{\varphi^c}$ is contained in a path-component of $\left(\bigcup\limits_{u \in K_\varphi^c} U(u)\right) \cap \overset{\circ}{\varphi^c}$ and we have $z \overset{*}{\sim} w$ (see the proof of Lemma 5.7.3). Also, $w \in D_0$. Hence

$$u_0 \overset{*}{\sim} w \overset{*}{\sim} z$$
$$\Rightarrow t^* \geqslant \hat{t} > t^*, \text{ a contradiction.}$$

This proves that $t^* = t$ and so $u_1 \in D_0$. But this contradicts the definition of c since $D_0 \subseteq D \subseteq \overset{\circ}{\varphi^c}$. This proves the theorem. □

The next result concerns the nature of local minimizers of φ.

Theorem 5.7.6 *If $\varphi \in C^1(X)$, it satisfies the C-condition and u_0 is a local minimizer of φ, then one of the following two statements holds:*

(a) There exists a small $\rho > 0$ such that

$$\inf\{\varphi(u) : ||u - u_0|| = \rho\} > \varphi(u_0).$$

(b) For every small $\rho > 0$, φ has a local minimizer u_ρ such that

$$||u_\rho - u_0|| = \rho, \ \varphi(u_\rho) = \varphi(u_0).$$

Proof Suppose that (a) does not hold. So, given any small $\rho > 0$, we have

$$\inf\{\varphi(u) : ||u - u_0|| = \rho\} = \varphi(u_0). \tag{5.143}$$

Since u_0 is a local minimizer of φ, we can find an open neighborhood U of u_0 such that $\varphi(u_0) \leqslant \varphi(u)$ for all $u \in U$ (see Definition 5.7.2). So, by taking $\rho > 0$ even smaller if necessary, we can find $\delta \in (0, \rho)$ such that

$$R = \{u \in X : \rho - \delta \leqslant ||u - u_0|| \leqslant \rho + \delta\} \subseteq U.$$

Consider φ restricted to R and let $u_n \in R$ such that

$$||u_n - u_0|| = \rho \text{ and } \varphi(u_n) \leqslant \varphi(u_0) + \frac{1}{n}, \quad n \geqslant 1 \text{ (since (5.143))}.$$

Using the generalized Ekeland variational principle (see Proposition 4.8.7) we can find $\{v_n\}_{n\geqslant 1} \subseteq R$ such that

$$\varphi(v_n) \leqslant \varphi(u_n), ||v_n - u_n|| \leqslant \frac{1}{n} \text{ and } \varphi(v_n) \leqslant \varphi(u) + \frac{1}{n}\frac{||u - v_n||}{1 + ||v_n||} \text{ for all } u \in R.$$

So, $u_n \in \operatorname{int} R$ for large $n \geqslant 1$, hence $(1 + ||v_n||)||\varphi'(v_n)||_* \leqslant \frac{1}{n}$ for large $n \geqslant 1$ (just take $u = v_n + th$ with $h \in X$, $||h|| = 1$ and $t > 0$ small). Since φ satisfies the C-condition, we may assume that

$$v_n \to v_\rho \text{ in } X.$$

Then $||v_\rho - u_0|| = \rho$, $\varphi(v_\rho) = \varphi(u_0)$, $\varphi'(v_\rho) = 0$, which is statement (b). $\square$

Combining Theorems 5.7.5 and 5.7.6, we obtain the following property.

Theorem 5.7.7 *If $\varphi \in C^1(X)$, it satisfies the C-condition, $u_0, u_1 \in X$, $u_0 \neq u_1$,*

$$\begin{aligned} &\Gamma = \{\gamma \in C([0,1], X) : \gamma(0) = u_0, \gamma(1) = u_1\},\\ &c = \inf_{\gamma\in\Gamma} \max_{0\leqslant t 1} \varphi(\gamma(t)),\\ &c > \max\{\varphi(u_0), \varphi(u_1)\} = \xi,\\ \textit{and}\quad &K_\varphi^c \textit{ is discrete}, \end{aligned}$$

then K_φ^c contains a critical point of mountain pass type.

Another consequence of Theorem 5.7.6 is the following result.

Theorem 5.7.8 *If $\varphi \in C^1(X)$, it satisfies the C-condition and it has two distinct local minimizers u_0 and u_1, then φ has at least one more critical point distinct from u_0, u_1.*

In fact, we can elaborate this last result further.

Theorem 5.7.9 *If $\varphi \in C^1(X)$, it satisfies the C-condition and it has two distinct local minimizers u_0 and u_1, then one of the following two statements holds:*

(a) φ has a critical point u which is a saddle point (that is, there exists an open neighborhood U of u such that contains point $y, u \in U$ such that

$$\varphi(y) < \varphi(u) < \varphi(v)).$$

(b) u_0 and u_1 can be path-connected in any neighborhood of the set of local minimizers u of φ with $\varphi(u) = \varphi(v_0)$ (hence $\varphi(u_1) = \varphi(u_0)$).

Proof Without any loss of generality, we may assume that $u_0 = 0$ and $\varphi(u_0) = \varphi(0) = 0$.

Let $\Gamma = \{\gamma + C([0, 1], X) : \gamma(0) = 0, \gamma(1) = u_1\}$ and define

$$c = \inf_{\gamma\in\Gamma} \max_{0\leqslant t\leqslant 1} \varphi(\gamma(t)).$$

Suppose K_φ^c contains only local minimizers of φ. So, given any $u \in K_\varphi^c$, we can find an open neighborhood $U(u)$ of u such that

$$\varphi(u) = c \leqslant \varphi(v) \text{ for all } v \in U(u). \tag{5.144}$$

Let $\widehat{U} = \bigcup_{u\in K_\varphi^c} U(u)$. For any neighborhood $\widetilde{U}$ of K_φ^c, let $\epsilon > 0$ and $\{h_t(\cdot) = h(t, \cdot)\}_{t\in[0,1]}$ be as postulated by the Deformation Theorem (see Theorem 5.3.7) with $\epsilon_0 = 1,\ U = \widehat{U} \cap \widetilde{U}$. Choose $\gamma \in \Gamma$ such that

$$\varphi(\gamma(t)) \leqslant c + \epsilon \text{ for all } t \in [0, 1]. \tag{5.145}$$

Set $\gamma_0 = h_1 \circ \gamma$. Then $\gamma_0 \in \Gamma$ and we have

$$\gamma_0([0, 1]) \subseteq h(1, \varphi^{c+\epsilon}) \subseteq \varphi^{c-\epsilon} \cup U \subseteq \varphi^{c-\epsilon} \cup \widehat{U}$$
$$\text{(see (5.145) and Theorem 5.3.7)}.$$

But the sets $\varphi^{c-\epsilon}$ and $\widehat{U}$ are disjoint (see (5.144)). So, we must have

$$\gamma_0([0, 1]) \subseteq \varphi^{c-\epsilon} \text{ or } \gamma_0([0, 1]) \subseteq U. \tag{5.146}$$

The first inclusion in (5.146) contradicts the definition of c. So, we must have

$$\begin{aligned} &\gamma_0([0, 1]) \subseteq U \subseteq \widetilde{U} \\ \Rightarrow\ &u_0 = 0 \text{ and } u_1 \text{ can be path connected in } \widetilde{U}. \end{aligned}$$

The proof is now complete. □

5.8 Remarks

5.1: The deformation approach described in Sect. 5.1 is based on the study of the asymptotic properties of the gradient flow associated to a C^2 or more generally to a C^{2-0} functional φ. So, we need to find conditions on φ which guarantee that a sequence of "almost critical points" of φ leads to a critical point. Similarly if instead of the deformation approach, we employ one based on the Ekeland variational principle (see Theorem 4.6.14). Such a condition is, by its nature, a compactness-type condition. This leads us to the "Palais–Smale condition" (see Definition 5.1.6(a)). The original definition of Palais and Smale [328] (see also Palais [325] and Smale [383]), was not that one. In Palais and Smale [328] the authors consider functionals $\varphi \in C^2(H, \mathbb{R})$ with H a separable Hilbert space (or more generally a C^2-Riemannian manifold M without boundary modeled on a separable Hilbert space H) which satisfy the following condition:

Definition 5.8.1 Let H be a separable Hilbert space and $\varphi \in C^2(H, \mathbb{R})$. We say that φ satisfies "condition (C)" if the closure of any nonempty set $D \subseteq H$ such that $\varphi|_D$ is bounded, but $\nabla\varphi|_D$ is not bounded away from zero, contains a critical point of φ.

Palais and Smale [328] have the following result which justifies condition (C).

Proposition 5.8.2 *If H is a separable Hilbert space, $\varphi \in C^2(H, \mathbb{R})$ and it is bounded below and satisfies condition (C), then for any $x \in H$, the gradient flow exists for all $t \geqslant 0$ and has a critical point as limit point when $t \to +\infty$.*

How is "condition (C)" related to the "PS-condition" as formulated in Definition 5.1.6(a)? The PS-condition is stronger than condition (C). Indeed, the zero functional satisfies condition (C) but not the PS-condition. We mention that in the books of Berger [44] and Chow and Hale [123], the PS-condition is called condition (C), while in Schwartz [376] condition (C) is called the "PS-condition". Of course the PS-condition makes sense in the more general context of Banach spaces and for functionals which are only C^1. The difficulty with this generalization is when we implement the deformation approach. Since $\varphi'(x) \in X^*$, we need to replace the derivative of φ at X with a vector field with values in X, in order to have a flow in X. This leads to the notion of a pseudogradient vector field (see Definition 5.1.1). This fruitful concept was introduced by Palais [325], who also proved Theorem 5.1.4. This theorem shows that the deformation approach works in Banach spaces for functionals which are only C^1. No further regularity on the functional is necessary. The "Cerami condition" ("C-condition" for short, see Definition 5.1.6(b)) was introduced by Cerami [115] as a more convenient compactness condition for functionals which are defined on an unbounded Riemannian manifold modeled on a separable Hilbert space. Aubin and Ekeland [21] (see Proposition 3, p. 270) proved the following result:

Proposition 5.8.3 *If X is a reflexive Banach space and $\varphi : X \to \mathbb{R}$ is a Gâteaux differentiable, convex, lower semicontinuous and coercive functional, then φ satisfies condition (C).*

Propositions 5.1.8 and 5.8.3 suggest that there is a link between the property of coercivity of φ and the PS-condition. This issue was investigated by Čaklović, Li and Willem [98] and Costa and Silva [130]. The extension of the PS-condition to functionals $\varphi + \psi$ with $\varphi \in C^1(X, \mathbb{R})$ and ψ convex and lower semicontinuous (see Definition 5.1.21 and 5.1.23) is due to Szulkin [397, 398].

Struwe [392] motivated by applications to the Plateau problem, developed a critical point theory for C^1-functionals φ restricted to a closed, convex set D of a Banach space X. In doing this, Struwe [392] extended the notion of a PS-condition. So, let X be a Banach space, $D \subseteq X$ a nonempty closed, convex set and $\varphi : D \to \mathbb{R}$ a functional which admits a C^1-extension on all of X. We set

$$\sigma(u) = \sup\{\langle u - y, \varphi'(u)\rangle : y \in D, ||y - u|| < 1\}, \; u \in D, \tag{5.147}$$

and say that $u \in D$ is a critical point of φ on D, if $\sigma(u) = 0$. Then the corresponding PS-condition on closed convex sets has the following form:

Definition 5.8.4 Every sequence $\{u_n\}_{n\geqslant 1} \subseteq D$ such that $\{\varphi(u_n)\}_{n\geqslant 1}$ is bounded and

$$\sigma(u_n) \to 0 \text{ as } n \to \infty$$

admits a strongly convergent subsequence.

Remark 5.8.5 Id $D = X$, then $\sigma(u) = ||\varphi'(u)||_*$ for all $u \in X$ and so the above definition coincides with Definition 5.1.6(a).

The notions of this section and many of the results can be extended to nonsmooth functionals φ which are locally Lipschitz (using the Clarke subdifferential) and to nonsmooth functionals of the form $\varphi + \psi$ with φ locally Lipschitz and ψ convex, lower semicontinuous. For details we refer to Gasinski and Papageorgiou [180–182], Kourogenis and Papageorgiou [247, 248], Kyritsi and Papageorgiou [256], Papageorgiou and Kyritsi [329], and Rădulescu [348, 350].

5.2: The notion of lower semicontinuity (see Definition 5.2.1) was introduced by Borel [55], but a systematic use of it in the study of variational problems was made by Tonelli [404], who developed the so-called "Direct Method of the Calculus of Variations" (see Theorem 5.2.6). Lower semicontinuity is studied in the books of Attouch et al. [20], Butazzo [97], Cesari [116], Dal Maso [134], Denkowski et al. [143], Ekeland and Temam [161], Ioffe and Tichomirov [221]. Moreover, more results on the minimization of the integral functionals and detailed expositions of the theory of the Calculus of Variations can be found in the books of Buttazzo [97], Dacorogna [133], Ekeland and Temam [161], Giaquinta [186] and Morrey [306].

5.3: In the literature we find two approaches to critical point theory. The first is based on the deformation properties of the negative gradient or pseudogradient flow. In this volume we follow this approach. The second approach is based on the Ekeland variational principle and can be found in the works of Cuesta [132], Ekeland [158, 159] and de Figueiredo [168]. Earlier forms of the deformation theorem (see Theorem 5.3.7) were obtained by Browder [85], Palais [325, 326], and Schwartz [375, 376]. However, the deformation result close in form to Theorem 5.3.7 was proved by Clark [126], who developed a Ljusternik–Schnirelmann theory for even functions defined on a Banach space based on the Krasnoselskii genus (see Definition 5.6.10(a)). Clark's theory was well adapted to the study of the existence and multiplicity of semilinear elliptic boundary value problems. The "second deformation theorem" (see Theorem 5.3.12) is due to Rothe [361], Marino and Prodi [288] and Chang [117, 118]. Usually the deformation theorems are formulated in terms of the

PS-condition. The first to employ the C-condition were Bartolo et al. [34]. The deformation theorem for Szulkin functionals of the form $\varphi + \psi$ with φ a C^1-functional and ψ a convex and lower semicontinuous functional (see Theorem 5.3.16) was proved by Szulkin [397]. Also, Szulkin [398] used the Ekeland variational principle to extend the Palais version of the infinite-dimensional Ljusternik–Schnirelmann theory to C^1-functionals bounded from below and defined on a C^1-Finsler manifold.

The deformation approach to critical point theory can be found in the books of Ambrosetti and Malchiodi [16], Chang [118], Costa [129], Gasinski and Papageorgiou [181, 182], Ghoussoub [184], Jabri [223], Kavian [230], Motreanu et al. [310], and Willem [415]. In Willem [415] we can find the following quantitative deformation theorem (no compactness condition is used).

Theorem 5.8.6 *If X is a Banach space, $\varphi \in C^1(X, \mathbb{R})$, $D \subseteq X$, $c \in \mathbb{R}$ and $\epsilon, \delta > 0$ are such that for all $u \in \varphi^{-1}([c - 2\epsilon, c + 2\epsilon]) \cap D_{2\delta}$ we have $||\varphi'(u)||_* \geqslant \frac{8\epsilon}{\delta}$ where $D_{2\delta} = \{v \in X : d(v, D) \leqslant 2\delta\}$, then there exists an $h \in C([0, 1] \times X, X)$ such that*

(a) $h(t, u) = u$ *if* $t = 0$ *or if* $u \notin \varphi^{-1}([c - 2\epsilon, c + 2\epsilon]) \cap D_{2\delta}$;
(b) $h(1, \varphi^{c+\epsilon} \cap D) \subseteq \varphi^{c-\epsilon}$;
(c) *for every* $t \in [0, 1]$, $h(t, \cdot)$ *is a homeomorphism on* X;
(d) $||h(t, u) - u|| \leqslant \delta$ *for all* $(t, u) \in [0, 1] \times X$;
(e) $t \to \varphi(h(t, u))$ *is nonincreasing for all* $u \in X$;
(f) $\varphi(h(t, u)) < c$ *for all* $t \in [0, 1]$ *and all* $u \in \varphi^c \cap D_\delta$.

5.4: As is evident from the results of this section, the notion of linking sets (see Definition 5.4.1) is very important in critical point theory. It was introduced by Benci and Rabinowitz [43]. Extensions of this notion can be found in the books of Schechter [379], Zou and Schechter [431] and in the paper of Schechter [378]. Slightly more restrictive versions of Theorem 5.4.4 can be found in Ekeland [159], Mawhin and Willem [293] and Struwe [393]. The mountain pass theorem (see Theorem 5.4.6) is due to Ambrosetti and Rabinowitz [17] and marks the first major breakthrough in the minimax approach to critical point theory. The mountain pass theorem was the first result to give a simple (minimax) procedure to find a critical point which is not a local minimum. After the mountain pass theorem followed the saddle point theorem (see Theorem 5.4.7) and the generalized mountain pass theorem (see Theorem 5.4.8), both due to Rabinowitz [346]. The lecture notes of Rabinowitz [347] give a nice overview of some basic aspects of critical point theory with applications to semilinear elliptic equations. Theorem 5.4.11 is due to Ghoussoub and Preiss [185], who in their work were able to localize the results of the Ekeland variational principle (see Theorem 4.6.14). The notion of local linking (see Definition 5.4.14) was first used by Liu and Li [279] under the stronger conditions that

$$\dim Y < +\infty \text{ and } \varphi(v) \geqslant r > 0 \text{ for all } v \in V \text{ with } ||v|| = r.$$

Theorem 5.4.17 is due to Brezis and Nirenberg [69], who were also the first to consider the more general notion of local linking in Definition 5.4.14. The Morse lemma (see Proposition 5.4.19) was first proved by Morse [307] for functions defined on $\mathbb{R}^N$. Later Morse [308] extended his theory to compact, smooth, finite-dimensional manifolds. The theory of Morse was extended to Hilbert spaces by Rothe [363] and to infinite-dimensional manifolds modeled over Hilbert spaces by Palais [325], Palais and Smale [328] and Smale [383]. To deal with the fact that the underlying ambient space is no longer locally compact, a compactness-type condition was introduced on the functional φ itself (the PS-condition, see Definition 5.1.6). The minimax results for functionals of the form $j = \psi + \varphi$ with $\varphi \in C^1(X)$ and $\psi \in \Gamma_0(X)$ (see Theorems 5.4.25, 5.4.26, 5.4.28 and 5.4.29) are essentially due to Szulkin [397].

5.5: The method of Lagrange multipliers in constrained optimization is discussed in the books of Alexeev et al. [8], Gasinski and Papageorgiou [182], Papageorgiou and Kyritsi [329], Tichomirov [403] and Zeidler [425, 427]. In Theorem 5.5.9 the assumption that $R(g'(u_0)) = Y$ (surjectivity of $g'(u_0)$) is crucial. A helpful result in this direction is the following proposition, which can be found in Yosida [418], p. 208.

Proposition 5.8.7 *If X, Y are Banach spaces and $A : D(A) \subseteq X \to Y$ is a closed and densely defined linear operator, then A is surjective (that is, $R(A) = Y$) if and only if A^* has a continuous inverse (that is, there exists a $c > 0$ such that*

$$||A^* y^*||_{X^*} \geqslant c||y^*||_{Y^*} \textit{ for all } y^* \in D(A^*)).$$

Remark 5.8.8 If $X = H$ is a Hilbert space and $A : D(A) \subseteq H \to H$ is a closed, densely defined linear operator such that there exists a $c > 0$ for which we have

$$c||u||^2 \leqslant (A(u), u)_H \text{ for all } u \in D(A)$$

(so, A is strongly monotone), then $R(A) = H$ (that is, A is surjective).

Natural constraints are discussed in the book of Ambrosetti and Malchiodi [16]. The Nehari manifold (see Definition 5.5.13) was first introduced by Nehari [316]. The fibering method is used in the book of Kuzin and Pohozaev [255].

5.6: Symmetry plays an important role in proving theorems which produce multiple critical points for a C^1-functional. The best known examples are when the acting group is $G = \mathbb{Z}_2$ or $G = S^1$. In the first case, we have the Krasnoselskii genus (see Krasnoselskii [250] and Coffman [127]), while in the second case we have the cohomological index due to Fadell and Rabinowitz [165]. The definition of genus given in Definition 5.6.10(a) is due to Coffman [127]. The study of critical points of not necessarily quadratic functionals started with Ljusternik [280], who considered C^2-functionals defined on a finite-dimensional manifold and for that purpose introduced the notion of what we call here Ljusternik–Schnirelmann category (see Definition 5.6.10(b)). Soon thereafter his ideas were pursued further by Ljusternik and Schnirelmann [281, 282]. They exploited the fact that a compact set has a neighborhood of the same category, in order to compute categories by means of elementary concepts of combinatorial topology. A notion closely related to the Krasnoselskii genus was introduced by Yang [416] under the name B-index. Let $i(\cdot)$ denote the B-index. We have

$$i(C) \leqslant \gamma(C) \text{ and } i(S^n) = n \ (S^n = \partial B_1(0) \subseteq R^{n+1}).$$

The Ljusternik–Schnirelmann theory is discussed in the books of Ambrosetti and Malchiodi [16], Deimling [142], Gasinski and Papageorgiou [181], Struwe [393], and Zeidler [425]. Extensions to Banach manifolds with Finsler structure can be found in Ghoussoub [184], Palais [326], and Szulkin [398]. Theorem 5.6.20 with $\psi \equiv 0$ is due to Rabinowitz [347] and for $\psi \in \Gamma_0(X)$ is due to Szulkin [398]. The principle of symmetric criticality (see Theorem 5.6.23) is due to Palais [327]. This principle is not valid in general, as shown by counter-examples produced by Palais [327] (see also Kobayashi and Otani [241]). Extensions to nonsmooth functionals were proved by Kobayashi and Otani [241]. Theorem 5.6.31 (the "fountain theorem") is due to Bartsch [35]. An extension of this result can be found in Zou [430]. A more detailed discussion of the issue of symmetry versus multiplicity of critical points can be found in the book of Bartsch [36].

5.7: Theorem 5.7.6 can be found in de Figueiredo and Solimini [169]. For an alternative proof we infer to Aizicovici et al. [5] (proof of Proposition 29). Theorems 5.7.7 and 5.7.9 are essentially due to Hofer [210] (see also Hofer [211]). Additional related results can be found in Chang [117, 118], Hofer [212] and Pucci and Serrin [343, 344].

Chapter 6
Morse Theory and Critical Groups

To think freely is great, but to think rightly is greater.
Thomas Thorild, engraved in golden letters at the entrance to the Grand Auditorium, Uppsala University

Let H be a Hilbert space with inner product $(\cdot,\cdot)_H$ and let $\varphi \in C^2(H)$. By $\varphi'(\cdot)$ we denote the Fréchet derivative of φ and by $\nabla\varphi(\cdot)$ its gradient, that is, $\nabla\varphi(u) \in H$ for every $u \in H$ and

$$(\nabla\varphi(u), h)_H = \left\langle \varphi'(u), h \right\rangle \text{ for all } h \in H, \tag{6.1}$$

where by $\langle\cdot,\cdot\rangle$ we denote the duality brackets for the pair (H^*, H). Recall that u_0 is a critical point of φ if $\varphi'(u_0) = 0$, which by (6.1) is equivalent to saying that $\nabla\varphi(u_0) = 0$. We say that $c \in \mathbb{R}$ is a critical level if $\varphi^{-1}(c)$ contains critical points. Otherwise $c \in \mathbb{R}$ is said to be a regular level.

Let $a < b$ be two regular values and let $M = \varphi^{-1}([a, b])$. The object of Morse theory is the relation between the local topological structure of the level sets of $\varphi|_M$ near a critical point and the topological structure of the manifold M. More precisely, suppose that $u_0 \in H$ is an isolated critical point of φ. Then the local behavior of φ near u_0 and the topological type of u_0 are described by a sequence of abelian groups $\{C_k(\varphi, u_0)\}_{k\in\mathbb{N}_0}$, known as the "critical groups" of φ at u_0 and defined using homology theory. If the critical point u_0 is nondegenerate (that is, $\varphi''(u_0) \in \mathscr{L}(H, H)$ is invertible), then the critical groups can be computed by linearization using the Morse lemma (see Proposition 5.4.19). In fact, if u_0 is nondegenerate, then

$$\operatorname{rank} C_k(\varphi, u_0) = \delta_{k,m} \text{ for all } k \in \mathbb{N}_0,$$

with m being the Morse index of u_0 (see Definition 5.4.18(b)). In the degenerate case, no such simple relation exists. Nevertheless, we can still have some results in the degenerate case provided $\varphi''(u)$ is a Fredholm operator. The critical groups are

N. S. Papageorgiou et al., *Nonlinear Analysis–Theory and Methods*,
Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-030-03430-6_6

invariant under small perturbations of the function φ. The global aspect of Morse theory is expressed by the so-called "Morse inequalities", which relate the critical groups of $\varphi|_M$ to the homology groups $H_k(\varphi^b, \varphi^a)$, $k \in \mathbb{N}_0$, which are isomorphic to the homology groups $H_k(M, \varphi^{-1}(a))$, $k \in \mathbb{N}_0$, by the excision property of homology theory.

Since Morse theory and critical groups use homology theory, in Sect. 6.1 we conduct a quick review of those tools from "Algebraic Topology" which we will use in the sequel. We present all the relevant notions and derive some fundamental consequences of these definitions. Special attention is given to singular homology theory, because this is the homology theory which we will use to define the critical groups, which are defined in Sect. 6.2, and the case of nondegenerate and of degenerate critical points are examined. We also derive the Morse relations which express the global aspects of Morse theory. In Sect. 6.3, we establish the invariance properties of critical groups. So, we show their C^1-invariance and their homotopical invariance. As in degree theory, these properties are very prolific tools in the computation of critical groups of a given functional. In Sects. 6.4 and 6.5, we consider the case of minimizers, maximizers and of saddle points (critical points of mountain pass type). In Sect. 6.6, we introduce homological counterparts of the notions of linking sets (see Definition 5.4.1) and of local linking (see Definition 5.4.14) and compute the critical groups for these more general settings. In Sects. 6.7 and 6.8 we use critical groups to prove the existence of multiple critical points. After all, the importance of critical groups lies in the fact that they provide very efficient tools to generate additional critical points and also to distinguish between critical points.

6.1 Elements of Algebraic Topology

In this section we review some basic definitions and facts of algebraic topology which will be used in the sequel.

Definition 6.1.1 (a) A "pair of spaces" (X, A) is a Hausdorff topological space X together with a subspace $A \subseteq X$. We write $(X, A) \subseteq (Y, B)$ if $X \subseteq Y$ and $A \subseteq B$.

(b) A "map of pairs" (X, A), (Y, B) is a continuous map $\varphi : X \to Y$ such that $\varphi(A) \subseteq B$. We denote the collection of all such maps by

$$C((X, A), (Y, B)).$$

Also, by $\mathrm{id}_{(X,A)} : (X, A) \to (X, A)$, we denote the identity map seen as a map of pairs.

(c) A map φ is a "homeomorphism of pairs" (X, A), (Y, B), if $\varphi : X \to Y$ is a homeomorphism and φ^{-1} is a map of pairs (Y, B), (X, A) (that is, $\varphi|_A : A \to B$ is a homeomorphism and $\varphi(A) = B$).

Remark 6.1.2 A space X can be regarded as the pair of spaces $(X, \emptyset)$. If A is a singleton (that is, $A = \{u_0\}$), then the pair $(X, A) = (X, \{u_0\})$ is denoted by (X, u_0)

and it is usually called a "pointed space". The composition of two maps of pairs is still a map of pairs.

In Definition 3.1.13 we introduced the notion of homotopy between two continuous maps, which played a central role in degree theory. Sometimes it is necessary to consider homotopies between maps of pairs. Then Definition 3.1.13 is extended easily to the following one:

Definition 6.1.3 Given pairs (X, A) and (Y, B), two maps of pairs $\varphi, \psi : (X, A) \to (Y, B)$ are said to be "homotopic" if there exists a map of pairs

$$h : ([0, 1] \times X, [0, 1] \times A) \to (Y, B)$$

such that

$$h(0, u) = \varphi(u) \text{ and } h(1, u) = \psi(u) \text{ for all } u \in X.$$

We write $\varphi \simeq \psi$ to indicate that φ, ψ are homotopic in the above sense. If $\varphi, \psi : (X, A) \to (Y, B)$ are maps of pairs and $\varphi|_A = \psi|_A$, we say that φ and ψ are "homotopic relative to A" if there exists a homotopy

$$h : ([0, 1] \times X, [0, 1] \times A) \to (Y, B)$$

such that

$$h(t, \cdot)|_A = \varphi|_A = \psi|_A \text{ for all } t \in [0, 1]$$

(that is, the homotopy h is fixed on A). In this case we write

$$\varphi \simeq_A \psi.$$

Remark 6.1.4 In the above definition it is said that in continuously deforming φ to ψ, it is required that at each time instant $t \in [0, 1]$, the set A is mapped into B.

Using Definitions 3.1.13 and 6.1.3 we are led to the fundamental notion of "homotopy equivalence" of topological spaces.

Definition 6.1.5 (a) Two Hausdorff topological spaces X, Y are said to be "homotopy equivalent" (or of the "same homotopy type") if there exist maps $\varphi : X \to Y$ and $\psi : Y \to X$ such that $\psi \circ \varphi$ is homotopic to id_X and $\varphi \circ \psi$ is homotopic to id_Y. In this case the map φ is a "homotopy equivalence" and ψ is the "homotopy inverse" of φ. If X, Y are homotopy equivalent, then we write

$$X \sim Y.$$

(b) Two pairs of spaces (X, A) and (Y, B) are "homotopy equivalent" if there exist maps of pairs $\varphi : (X, A) \to (Y, B)$ and $\psi : (Y, B) \to (X, A)$ such that $\psi \circ \varphi \simeq \mathrm{id}_X$ and $\varphi \circ \psi \simeq \mathrm{id}_Y$ (see Definition 6.1.3; so the homotopies are homotopies of pairs). If the pairs (X, A), (Y, B) are homotopy equivalent, then we write

$$(X, A) \sim (Y, B).$$

Remark 6.1.6 As the names suggest, both notions are equivalence relations. In general homotopy equivalence of X, Y (respectively of $(X, A), (Y, B)$), roughly speaking, means that X (respectively (X, A)) can be deformed continuously to Y (respectively (Y, B)). It is easy to see that two homeomorphic spaces are homotopy equivalent but the converse is not true in general. So, the classification of Hausdorff topological spaces up to homeomorphism is more refined than the classification up to homotopy equivalence. Deformation retracts and strong deformation retracts (see Definition 5.3.10(b)) are homotopy equivalences which are easy to visualize. Indeed, if $A \subseteq X$ is a retract of X, $r : X \to A$ is a retraction map (that is, $r(\cdot)$ is continuous and $r|_A = \mathrm{id}|_A$), and $i : A \to X$ is the inclusion map, then $r \circ i = \mathrm{id}_A$ and $i \circ r \simeq \mathrm{id}_X$ (for deformation retracts) and $i \circ r \simeq_A \mathrm{id}_X$ (for strong deformation retracts). So, if A is a deformation retract or strong deformation retract of X, then $A \sim X$.

Using homotopies we can introduce the following fundamental topological notion.

Definition 6.1.7 A Hausdorff topological space X is said to be "contractible" if the identity map $\mathrm{id}_X : X \to X$ is homotopic to a constant map $\varphi : X \to *$ (that is, there exists continuous map $h : [0, 1] \times X \to X$ such that $h(0, u) = u$ for all $u \in X$ and $h(1, u) = *$ for all $u \in X$).

Remark 6.1.8 Clearly, X is contractible if and only if it is homotopy equivalent to a singleton if and only if every point of X is a deformation retract. A contractible space is simply connected and any two maps into a contractible space are homotopic. Evidently, every convex set of a Banach space or more generally any star-shaped set is contractible (recall that a subset X of a Banach space is star-shaped if there exists a $u_0 \in X$ such that for all $u \in X$, $[u_0, u] = (1-t)u_0 + tu$, $t \in [0, 1]$, lies in X).

Example 6.1.9 (a) Let $X = S^1 = \{u \in \mathbb{R}^2 : |u| = 1\}$ (the unit sphere in $\mathbb{R}^2$) and $Y = S^1 \cup [(1, 0), (2, 0)]$ (recall $[(1, 0), (2, 0)] = (1-t)(1, 0) + t(2, 0)$ for all $t \in [0, 1]$), that is, $[(1, 0), (2, 0)]$ is a closed line segment on the horizontal axis, joining the points $(1, 0)$ and $(2, 0)$. We claim that X and Y are not homeomorphic. Indeed, if from X we remove any point, the remaining set is still connected. On the other hand, if from Y we remove the point $(1, 0)$, the remaining set is disconnected. However, the sets X and Y are homotopy equivalent. To see this, let $\varphi : X \to Y$ and $\psi : Y \to X$ be the following maps

$$\varphi(u) = u \text{ for all } u \in X \text{ and } \psi(v) = \begin{cases} v & \text{if } v \in S^1 \\ (1, 0) & \text{if } v \in [(1, 0), (2, 0)] \end{cases} \text{ for all } v \in Y.$$

We have $\psi \circ \varphi = \mathrm{id}_X$ and $\varphi \circ \psi \simeq \mathrm{id}_Y$ since $\varphi \circ \psi = \psi$.

(b) Any convex set in a Euclidean space is homotopy equivalent to a point (just recall that a convex set is contractible, see Remark 6.1.8).

(c) $S^{N-1} = \{u \in \mathbb{R}^N : |u| = 1\}$ $(N \geqslant 2)$ is homotopy equivalent to $\mathbb{R}^N \setminus \{0\}$ (just recall that S^{N-1} is a strong deformation retract of $\mathbb{R}^N \setminus \{0\}$).

(d) Let $B^2 = \{u \in \mathbb{R}^2 : |u| < 1\}$ and consider the solid torus $S^1 \times B^2$. This space is homotopy equivalent to S^1. More generally, if V is a vector bundle over a topological space X, the zero section is a strong deformation retract of V, hence homotopy equivalent to it.

Since for our purposes the role of algebraic topology is auxiliary, to avoid a lengthy presentation, we will follow the axiomatic approach of homology theory (naive homology theory).

Definition 6.1.10 Let $\{G_k\}_{k\in I}$ be a family of abelian groups and $\{j_k\}_{k\in I}$ a corresponding family of homomorphisms

$$\ldots \to G_{k+1} \xrightarrow{j_{k+1}} G_k \xrightarrow{j_k} G_{k-1} \to \ldots \tag{6.2}$$

We say that the sequence (chain) (6.2) is exact if and only if

$$\text{im } j_{k+1} = \ker j_k \text{ for all } k \in I.$$

Remark 6.1.11 If G_1, G_2 are two abelian groups and we consider the chain

$$0 \to G_1 \xrightarrow{j} G_2 \to 0, \tag{6.3}$$

then (6.3) is exact if and only if j is an isomorphism.

More generally, suppose that G_1, G_2, G_3 are three abelian groups and consider the following exact chain

$$0 \to G_1 \xrightarrow{j_1} G_2 \xrightarrow{j_2} G_3 \to 0. \tag{6.4}$$

From the exactness of (6.4), we see that j_1 is injective and $j_1(G_1)$ is isomorphic to G_1 and equal to $\ker j_2$. Moreover, j_2 is surjective and we have that $\ker j_2 \oplus \text{im } j_2$ is isomorphic to $G_1 \oplus G_3$.

Next, we introduce a "homology theory" by listing a number of axioms which must hold. They are usually called the "Eilenberg–Steenrod axioms".

Definition 6.1.12 A "homology theory" on a family of pairs of spaces (X, A) consists of:

(a) A sequence $\{H_k(X, A)\}_{k\in\mathbb{N}_0}$ of abelian groups known as "homology groups" for the pair (X, A) (note that for the pair $(X, \emptyset)$, we write $H_k(X)$, $k \in \mathbb{N}_0$).

(b) To every map of pairs $\varphi : (X, A) \to (Y, B)$ is associated a homomorphism

$$\varphi_* : H_k(X, A) \to H_k(Y, B) \text{ for all } k \in \mathbb{N}_0.$$

(c) To every $k \in \mathbb{N}_0$ and every pair (X, A) is associated a homomorphism

$$\partial : H_k(X, A) \to H_{k-1}(A) \text{ for all } k \in \mathbb{N}.$$

These items satisfy the following axioms:

Axiom 1: If $\varphi = \mathrm{id}_X$, then $\varphi_* = \mathrm{id}|_{H_k(X,A)}$.
Axiom 2: If $\varphi : (X, A) \to (Y, B)$ and $\psi : (Y, B) \to (Z, C)$ are maps of pairs, then $(\psi \circ \varphi)_* = \psi_* \circ \varphi_*$.
Axiom 3: If $\varphi : (X, A) \to (Y, B)$ is a map of pairs, then $\partial \circ \varphi_* = (\varphi|_A)_* \circ \partial$.
Axiom 4: If $i : A \to X$ and $j : (X, \emptyset) \to (X, A)$ are inclusion maps, then the following sequence is exact

$$\ldots \xrightarrow{\partial} H_k(A) \xrightarrow{i_*} H_k(X) \xrightarrow{j_*} H_k(X, A) \xrightarrow{\partial} H_{k-1}(A) \to \ldots$$

Axiom 5: If $\varphi, \psi : (X, A) \to (Y, B)$ are homotopic maps of pairs, then $\varphi_* = \psi_*$.
Axiom 6 (Excision): If $U \subseteq X$ is an open set with $\bar{U} \subseteq \mathrm{int}\, A$ and $i : (X\backslash U, A\backslash U) \to (X, A)$ is the inclusion map, then $i_* : H_k(X\backslash U, A\backslash U) \to H_k(X, A)$ is an isomorphism.
Axiom 7: If $X = \{*\}$, then $H_k(\{*\}) = 0$ for all $k \in \mathbb{N}$.

Remark 6.1.13 If an abelian group G is isomorphic to $H_0(X)$ for every singleton X, then we say that G is the group of coefficients of the homology theory. Note that $H_k(X, A) = 0$ for all $k \in -\mathbb{N}$. The excision axiom (see Axiom 6) can be equivalently reformulated as follows:

Axiom 6': If $A, B \subseteq X$ and $X = \mathrm{int}\, A \cup \mathrm{int}\, B$, then the inclusion map $i : (A, A \cap B) \to (X, A)$ induces an isomorphism $i_* : H_k(A, A \cap B) \to H_k(X, A)$.

Next we derive some useful consequences of the above axioms.

Proposition 6.1.14 *If the pairs (X, A) and (Y, B) are homotopy equivalent, then $H_k(X, A) = H_k(Y, B)$ for all $k \in \mathbb{N}_0$ (hereafter, the symbol = denotes that the groups are isomorphic).*

Proof Let $\varphi : (X, A) \to (Y, B)$ be a homotopy equivalence and ψ its homotopy inverse. According to Definition 6.1.5(b) we have $\psi \circ \varphi \simeq \mathrm{id}_X$. Then Axioms 1 and 2 imply that $\psi_* \circ \varphi_* = \mathrm{id}_{H_k(X,A)}$. Similarly we show that $\varphi_* \circ \psi_* = \mathrm{id}_{H_k(Y,B)}$. It follows that $\varphi_* : H_k(X, A) \to H_k(Y, B)$ is an isomorphism and $\varphi_*^{-1} = \psi_*$. $\square$

Proposition 6.1.15 *If $A \subseteq X$ is a deformation retract of X, then $H_k(X, A) = 0$ for all $k \in \mathbb{N}_0$.*

Proof From Remark 6.1.6 we know that X and A are homotopy equivalent. Hence $H_k(X) = H_k(A)$ for all $k \in \mathbb{N}_0$. Using Axiom 4 we have the exact chain

$$\ldots \to H_k(A) \xrightarrow{i_*} H_k(X) \xrightarrow{j_*} H_k(X, A) \xrightarrow{\partial} H_{k-1}(A) \xrightarrow{i_*} H_{k-1}(X) \to \ldots \tag{6.5}$$

The exactness of (6.5) and the equality $H_k(A) = H_k(X)$ for all $k \in \mathbb{N}_0$ (that is, i_* is an isomorphism) imply that $H_k(X, A) = 0$ for all $k \in \mathbb{N}_0$. $\square$

Corollary 6.1.16 $H_k(X, X) = 0$ *for all* $k \in \mathbb{N}_0$.

Proposition 6.1.17 *If A is a retract of X, then $H_k(X) = H_k(X, A) \oplus H_k(A)$ for all $k \in \mathbb{N}_0$.*

Proof Let $r : X \to A$ be the retraction and $i : A \to X$ the inclusion map. From Definition 3.1.30 we know that $r \circ i = \mathrm{id}_A$. Then from Axioms 1 and 2, we have

$$r_* \circ i_* = \mathrm{id}_{H_k(A)} \text{ for all } k \in \mathbb{N}_0,$$

and i_* is an injection onto a direct summand of $H_k(X)$. The other summand is the kernel of r_*. Let $j : (X, \emptyset) \to (X, A)$ be the inclusion map and consider the sequence

$$\ldots \to H_{k+1}(X, A) \xrightarrow{\partial} H_k(A) \xrightarrow{i_*} H_k(X) \xrightarrow{j_*} H_k(X, A) \xrightarrow{\partial} \ldots \tag{6.6}$$

This is an exact sequence (see Axiom 4) and since i_* is injective, we have $\ker i_* = 0$. So, from (6.6) it follows that ∂ is the trivial map. From the exactness of (6.6) it follows that j_* is surjective. Since $\ker j_* = \operatorname{im} i_*$ [see (6.6)], j_* is an isomorphism of $\ker r_*$ onto $H(X, A)$, Therefore we conclude that $H_k(X) = H_k(X, A) \oplus H_k(A)$ for all $k \in \mathbb{N}_0$. $\square$

A map of pairs $\varphi : (X, A) \to (Y, B)$ defines the maps

$$\varphi_1 : X \to Y \text{ and } \varphi_2 : A \to B.$$

Evidently, $\varphi_2 = \varphi|_A$ (see Definition 6.1.1(b)).

We introduce the homomorphisms φ_*, $(\varphi_1)_*$ and $(\varphi_2)_*$ induced by these maps and consider the following diagram:

$$\begin{array}{ccccccccc} \ldots \to & H_{k+1}(X, A) & \xrightarrow{\partial} & H_k(A) & \xrightarrow{i_*} & H_k(X) & \xrightarrow{j_*} & H_k(X, A) & \to \ldots \\ & \downarrow \varphi_* & & \downarrow (\varphi_1)_* & & \downarrow (\varphi_2)_* & & \downarrow \varphi_* & \\ \ldots \to & H_{k+1}(Y, B) & \xrightarrow{\partial} & H_k(B) & \xrightarrow{i'_*} & H_k(Y) & \xrightarrow{j'_*} & H_k(Y, B) & \to \ldots \end{array} \tag{6.7}$$

where i, j, i', j' are the appropriate inclusions.

Proposition 6.1.18 *Diagram (6.7) is commutative.*

Proof We must verify the equalities

$$\varphi_* \circ j_* = j'_* \circ (\varphi_1)_*, \ (\varphi_1)_* \circ i_* = i'_* \circ (\varphi_2)_*, \ (\varphi_2)_* \circ \partial = \partial \circ \varphi_*.$$

The first two equalities follows from Axiom 2 since

$$\varphi \circ j = j' \circ \varphi_1 \text{ and } \varphi_1 \circ i = i' \circ \varphi_2.$$

The third equality is actually Axiom 3. $\square$

To continue we will need an auxiliary result known as the "five lemma".

Lemma 6.1.19 *If we have a commutative diagram of abelian groups and homomorphisms*

$$\begin{array}{ccccccccc} G_1 & \xrightarrow{h_1} & G_2 & \xrightarrow{h_2} & G_3 & \xrightarrow{h_3} & G_4 & \xrightarrow{h_4} & G_5 \\ \downarrow \alpha & & \downarrow \beta & & \downarrow \gamma & & \downarrow \delta & & \downarrow e \\ \hat{G}_1 & \xrightarrow{\hat{h}_1} & \hat{G}_2 & \xrightarrow{\hat{h}_2} & \hat{G}_3 & \xrightarrow{\hat{h}_3} & \hat{G}_4 & \xrightarrow{\hat{h}_4} & \hat{G}_5 \end{array} \tag{6.8}$$

in which each row is exact and α, β, δ, e are isomorphisms, then γ is an isomorphism.

Proof First we show that γ is injective. So, suppose that $\gamma(u_3) = 0$ with $u_3 \in G_3$. Then from the commutativity of (6.8), we have

$$\begin{aligned} &(\delta \circ h_3)(u_3) = (\hat{h}_3 \circ \gamma)(u_3) = 0 \\ \Rightarrow\ & h_3(u_3) = 0 \text{ (since by hypothesis } \delta \text{ is an isomorphism).} \end{aligned} \tag{6.9}$$

The exactness of the top row in (6.8) and (6.9) imply that we can find $u_2 \in G_2$ such that $h_2(u_2) = u_3$. Then exploiting once again the commutativity of (6.8), we have

$$(\hat{h}_2 \circ \beta)(u_2) = 0$$

and there exists a $\hat{u}_1 \in \hat{H}_1$ such that

$$\hat{h}_1(\hat{u}_1) = \beta(u_2).$$

Let $u_1 \in G_1$ such that $\alpha(u_1) = \hat{u}_1$. We have

$$\begin{aligned} &(\beta \circ h_1)(u_1) = \beta(u_2) \\ \Rightarrow\ & h_1(u_1) = u_2 \text{ (since by hypothesis } \beta(\cdot) \text{ is an isomorphism)} \\ \Rightarrow\ & (h_2 \circ h_1)(u_1) = u_3. \end{aligned}$$

But $(h_2 \circ h_1)(u_1) = 0$ [by the exactness of the row in (6.8)]. Therefore

$$\begin{aligned} &u_3 = 0 \\ \Rightarrow\ &\gamma \text{ is injective.} \end{aligned}$$

Next we show that γ is surjective. So, let $\hat{u}_3 \in \hat{G}_3$. Then there is a $u_4 \in G_4$ such that $\delta(u_4) = \hat{h}_3(\hat{u}_3)$. From the commutativity of (6.8) we have

$$(e \circ h_4)(u_4) = (\hat{h}_4 \circ \delta)(u_4) = (\hat{h}_4 \circ \hat{h}_3)(\hat{u}_3) \tag{6.10}$$

and the exactness of the lower row in (6.8) implies that

$$\begin{aligned} &(\hat{h}_4 \circ \hat{h}_3)(\hat{u}_3) = 0 \\ \Rightarrow\ &h_4(u_4) = 0 \text{ (see (6.9) and recall that } e \text{ is an isomorphism).} \end{aligned}$$

The exactness of the top row in (6.8) implies that there exists a $u_3 \in G_3$ such that

$$\begin{aligned} &h_3(u_3) = u_4 \\ \Rightarrow\ &\hat{h}_3(\hat{u}_3 - \gamma(u_3)) = 0. \end{aligned}$$

So, there exists a $\hat{u}_2 \in \hat{G}_2$ such that

$$\hat{h}_2(\hat{u}_2) = \hat{u}_3 - \gamma(u_3).$$

Let $u_2 \in G_2$ be such that $\beta(u_2) = \hat{u}_2$. Then

$$\begin{aligned} &u_3 + h_2(u_2) \in G_3 \text{ and } \gamma(u_3 + h_2(u_2)) = \gamma(u_3) + \hat{h}_2(\hat{u}_2) = \hat{u}_3 \\ \Rightarrow\ &\gamma \text{ is surjective, hence an isomorphism.} \end{aligned}$$

The proof is now complete. $\square$

Using this lemma, we can prove the following result.

Proposition 6.1.20 *If* $(X, A) = \bigcup_{i=1}^{n} (X_i, A_i)$ *with* $\{X_i\}_{i=1}^n$ *nonempty, closed and pairwise disjoint subsets of* X*, then* $H_k(X, A) = \bigoplus_{i=1}^{n} H_k(X_i, A_i)$ *for all* $k \in \mathbb{N}_0$*.*

Proof We do the proof for $n = 2$, the general case following by induction.

Consider the inclusion maps $i_1 : X_1 \to X$ and $i_2 : X_2 \to X$. We show that these maps yield an isomorphism $(i_1)_* \oplus (i_2)_* : H_k(X_1) \oplus H_k(X_2) \to H_k(X)$ for all $k \in \mathbb{N}_0$. To this end it suffices to show that $(i_1)_*$, $(i_2)_*$ are injective and im $(i_1)_* \oplus$ im $i_2)_* = H_k(X)$. Let $j_1 : (X, \emptyset) \to (X, X_1)$ be the inclusion map. Then using Axioms 2 and 6, we have that $(j_1 \circ i_2)_* = (j_1)_* \circ (i_2)_*$ is an isomorphism. Hence $(i_2)_*$ is injective and

$$H_k(X) = \ker\,(j_1)_* \oplus \operatorname{im}\,(i_2)_*\,. \tag{6.11}$$

Similarly we prove that $(i_1)_*$ is injective. Moreover, from Axiom 4 we have

$$\ker (j_1)_* = \operatorname{im} (i_1)_*$$
$$\Rightarrow H_k(X) = \operatorname{im} (i_1)_* \oplus \operatorname{im} (i_2)_* \text{ as claimed (see (6.11)).}$$

In a similar fashion, we show that the inclusion maps $i_1^A : A_1 \to A$ and $i_2^A : A_2 \to A$ produce an isomorphism $i_1^A \oplus i_2^A : H_k(A_1) \oplus H_k(A_2) \to H_k(A)$ for all $k \in \mathbb{N}_0$. Then for every $k \in \mathbb{N}_0$, Axiom 4 gives us a commutative diagram

$$\begin{array}{ccccccccc}
\overset{2}{\underset{i=1}{\oplus}} H_k(A_i) & \longrightarrow & \overset{2}{\underset{i=1}{\oplus}} H_k(X_i) & \longrightarrow & \overset{2}{\underset{i=1}{\oplus}} H_k(X_i, A_i) & \longrightarrow & \overset{2}{\underset{i=1}{\oplus}} H_{k-1}(A_i) & \longrightarrow & \overset{2}{\underset{i=1}{\oplus}} H_{k-1}(X_i) \\
\downarrow \alpha & & \downarrow \beta & & \downarrow \gamma & & \downarrow \delta & & \downarrow e \\
H_k(A) & \longrightarrow & H_k(X) & \longrightarrow & H_k(X, A) & \longrightarrow & H_{k-1}(A) & \longrightarrow & H_{k-1}(X)
\end{array}$$

with the rows being exact and from the previous considerations, we have that α, β, δ, e are all isomorphisms. Invoking Lemma 6.1.19, we infer that γ is an isomorphism too and so we conclude that

$$H_k(X, A) = \overset{2}{\underset{i=1}{\oplus}} H_k(X_i, A_i) \text{ for all } k \in \mathbb{N}_0.$$

The proof is now complete. ☐

Corollary 6.1.21 *If $X = \bigcup_{i \in I} X_i$ is the decomposition of the space into its path components X_i, then $H_k(X) = \underset{i \in I}{\oplus} H_k(X_i)$ for all $k \in \mathbb{N}_0$.*

Remark 6.1.22 For any Hausdorff topological space X, $H_0(X)$ is a free abelian group with a basis consisting of an arbitrary point in each path component. Hence $H_0(X)$ is a direct sum of G's, one for each path component of X. If X is path-connected, then $H_0(X) = G$.

The next proposition generalizes the long exact sequence in Axiom 4.

Proposition 6.1.23 *If $C \subseteq A \subseteq X$ and $i : (A, C) \to (X, C)$, $j : (X, C) \to (X, A)$, $\hat{j} : (A, \emptyset) \to (A, C)$ are the inclusion maps, then the sequence*

$$\ldots \xrightarrow{\hat{j}_* \circ \partial} H_k(A, C) \xrightarrow{i_*} H_k(X, C) \xrightarrow{j_*} H_k(X, A) \xrightarrow{\hat{j}_* \circ \partial} H_{k-1}(A, C) \longrightarrow \ldots, \; k \in \mathbb{N}_0$$

is exact.

Proof We first show that

$$\operatorname{im} (\hat{j}_* \circ \partial) \subseteq \ker i_*, \; \operatorname{im} i_* \subseteq \ker j_*, \; \operatorname{im} j_* \subseteq \ker (\hat{j}_* \circ \partial). \tag{6.12}$$

From Axiom 2 it follows that the next diagram is commutative

$$\begin{array}{ccccc} H_{k+1}(X,A) & \xrightarrow{\partial} & H_k(A) & \xrightarrow{\hat{j}_*} & H_k(X,A) \\ & & \downarrow{\scriptstyle (i_1)_*} & & \downarrow{\scriptstyle i_*} \\ & & H_k(X) & \xrightarrow{\tilde{j}_*} & H_k(X,C) \end{array}$$

with $i_1 : A \to X$ and $\tilde{j} : (X, \emptyset) \to (X, C)$ being the inclusion maps. From Axiom 4 we have that $(i_1)_* \circ \partial = 0$. Hence $i_* \circ (\hat{j}_* \circ \partial) = 0$ and this proves the first inclusion in (6.12). The other two inclusions in (6.12) are verified in a similar way.

Next we show the opposite inclusions from those in (6.12), namely we show that

$$\ker i_* \subseteq \operatorname{im} \hat{j}_* \circ \partial),\ \ker j_* \subseteq \operatorname{im} i_*,\ \ker (\hat{j} \circ \partial) \subseteq \operatorname{im} j_*. \tag{6.13}$$

Let $u \in \ker j_*$. Using Axioms 2 and 3, we introduce the following commutative diagram of homology groups and homomorphisms

$$\begin{tikzcd} H_k(A) \arrow[r, "(i_1)_*"] \arrow[d, "\hat{j}_*"] & H_k(X) \arrow[d, "\tilde{j}_*"] \arrow[dr, "(j_1)_*"] & \\ H_k(A,C) \arrow[r, "i_*"] \arrow[dr, "\hat{\partial}"'] & H_k(X,C) \arrow[r, "j_*"] \arrow[d, "\tilde{\partial}"] & H_k(X,A) \arrow[d, "\hat{\partial}"] \\ & H_{k-1}(C) \arrow[r, "(i_2)_*"] & H_{k-1}(A) \end{tikzcd} \tag{6.14}$$

where $j_1 : (X, \emptyset) \to (X, A)$, $i_2 : C \to A$ are the inclusion maps and $\hat{\partial}, \tilde{\partial}$ are the boundary maps guaranteed by Definition 6.1.12. The argument is simple and follows the diagram (6.14). It involves four steps:

Step 1: We have $j_*(u) = 0$, hence $(i_2)_* \circ \tilde{\partial}(u) = \hat{\partial} \circ j_*(u) = 0$ [see (6.14)]. From the exactness of (6.14), we have $\ker (i_2)_* = \operatorname{im} \hat{\partial}$ and so we can find $y \in H_k(A, C)$ such that $\hat{\partial}(y) = \tilde{\partial}(u)$.

Step 2: We have $\tilde{\partial}(i_*(y) - u) = 0$ (see (6.14) and Step 1). We know that $\ker \tilde{\partial} = \operatorname{im} \tilde{j}_*$. So, we can find $x \in H_k(X)$ such that $\tilde{j}_*(x) = i_*(y) - u$.

Step 3: Since $j_* \circ i_* = 0$ (see the second inclusion in (6.12)) and $j_*(u) = 0$ (recall that $u \in \ker j_*$), we have

$$(j_1)_*(x) = j_*(i_*(y) - u) \text{ (see (14) and Step 2).}$$

But from the exactness of (6.14), we have $\ker (j_1)_* = \operatorname{im} (i_1)_*$. So, we can find $v \in H_k(A)$ such that $x = (i_1)_*(v)$.

Step 4: From the previous three steps we have

$$\begin{aligned} & i_*(y) - u = \tilde{j}_*((i_1)_*(v)) = i_*(\hat{j}_*(v)) \\ \Rightarrow & u \in \operatorname{im} i_*. \end{aligned}$$

This proves the second inclusion in (6.13). The other two inclusions in (6.13) are proved similarly.

From (6.12) and (6.13), we conclude that the sequence of the proposition is exact. □

Corollary 6.1.24 *Suppose that* $C \subseteq A \subseteq X$.

(a) If C *is a deformation retract of* A*, then* $H_k(X, A) = H_k(X, C)$ *for all* $k \in \mathbb{N}_0$.
(b) If A *is a deformation retract of* X*, then* $H_k(X, C) = H_k(A, C)$ *for all* $k \in \mathbb{N}_0$.

Next, we focus on homology groups of the form

$$H_k(X, *) \text{ with } * \in X,\ \ k \in \mathbb{N}_0.$$

We start by establishing the precise relation between the homology groups $H_k(X)$ and $H_k(X, *)$.

Proposition 6.1.25 $H_k(X, *) = \ker r_*$ *where* $r : X \to \{*\}$ *is the map* $r(u) = *$ *for all* $u \in X$ *and we have* $H_k(X) = H_k(X, *) \oplus H_k(*)$ *for all* $k \in \mathbb{N}_0$.

Proof We know that $\{*\} \subseteq X$ is a retract of X. So, from Proposition 6.1.17 we have

$$H_k(X) = H_k(X, *) \oplus H_k(*) \text{ for all } k \in \mathbb{N}_0.$$

The proof is now complete. □

Remark 6.1.26 From Axiom 7 and Proposition 6.1.25, we see that

$$H_0(X) = H_0(X, *) \oplus G \text{ and } H_k(X) = H_k(X, *) \text{ for all } k \in \mathbb{N}.$$

It is often convenient to have a slightly modified version of homology, for which a point has trivial homology groups in all dimensions, including zero. This is done in the next definition.

Definition 6.1.27 The "reduced homology groups" of X are defined by

$$\tilde{H}_k(X) = H_k(X, *) \text{ for all } k \in \mathbb{N}_0,\ \text{ with } * \in X.$$

Remark 6.1.28 Evidently, $H_0(X) = \tilde{H}_0(X) \oplus G$ and $H_k(X) = \tilde{H}_k(X)$ for all $k \in \mathbb{N}$.

The next result, known as the "reduced exact homology sequence", is a particular case of Proposition 6.1.23.

Proposition 6.1.29 *If* (X, A) *is a pair of space and* $* \in A$*, then the long sequence of homology groups*

$$\ldots \to H_k(A, *) \to H_k(X, *) \to H_k(X, A) \to H_{k-1}(A, *) \to \ldots$$

is exact.

Reduced homology groups are simple when the space is contractible (see Definition 6.1.7).

Proposition 6.1.30 *If X is a contractible Hausdorff topological space, then $H_k(X, *) = 0$ for all $k \in \mathbb{N}_0$ and all $* \in X$.*

Proof From Remark 6.1.8 we know that since X is contractible, every singleton $\{*\}$ with $* \in X$ is a deformation retract of X . Invoking Proposition 6.1.15, we have.

$$H_k(X, *) = 0 \text{ for all } k \in \mathbb{N}_0.$$

The proof is now complete. □

Proposition 6.1.31 *If $A \subseteq X$ is a subspace which is contractible in itself, then $H_k(X, A) = H_k(X, *)$ for all $* \in A$ and all $k \in \mathbb{N}_0$.*

Proof By Propositions 6.1.29 and 6.1.30, we have the following exact chain

$$0 = H_k(A, *) \to H_k(X, *) \to H_k(X, A) \to H_{k-1}(A, *) = 0.$$

The exactness of this chain implies that

$$H_k(X, *) = H_k(X, A) \text{ for all } k \in \mathbb{N}_0.$$

The proof is now complete. □

The next theorem is a basic tool for computing homology groups. It gives a recipe for computing the homology groups of a space which is the union of two open sets in terms of the homology groups of the two open sets and those of their intersection. This global result is known as the "Mayer–Vietoris theorem".

We will need the following general result about exact sequences, known in the literature as the "Whitehead–Barratt Lemma".

Lemma 6.1.32 *If the commutative diagram of abelian groups and homomorphisms*

$$\begin{array}{ccccccccc}
\ldots \to & A_k & \xrightarrow{\varphi_k} & B_k & \xrightarrow{\psi_k} & C_k & \xrightarrow{w_k} & A_{k-1} & \xrightarrow{\varphi_{k-1}} \ldots \\
& \big\downarrow \alpha_k & & \big\downarrow \beta_k & & \big\downarrow \gamma_k & & \big\downarrow \alpha_{k-1} & \\
\ldots \to & \hat{A}_k & \xrightarrow{\hat{\varphi}_k} & \hat{B}_k & \xrightarrow{\hat{\psi}_k} & \hat{C}_k & \xrightarrow{\hat{w}_k} & \hat{A}_{k-1} & \xrightarrow{\hat{\varphi}_{k-1}} \ldots
\end{array}$$

has exact rows and γ_k is an isomorphism for all $k \in \mathbb{N}_0$, then the sequence

$$\ldots \to A_k \xrightarrow{(\alpha_{k_1} - \varphi_k)} \hat{A}_k \oplus B_k \xrightarrow{\hat{\varphi}_k + \beta_k} \hat{B}_k \xrightarrow{w_k \circ \gamma_k^{-1} \circ \hat{\psi}_k} A_{k-1} \to \ldots$$

is exact.

The “Mayer–Vietoris theorem” reads as follows.

Theorem 6.1.33 *If X is a Hausdorff topological space, $A, B \subseteq X$ are two nonempty sets whose interiors cover X and $* \in A \cap B$, then there is an exact sequence*

$$\ldots \to H_k(A \cap B, *) \to H_k(A, *) \oplus H_k(B, *) \to H_k(A \cup B, *) \to H_k(A \cap B, *) \to \ldots$$

Proof From Proposition 6.1.23 we have the following commutative diagram

$$\begin{array}{ccccccccc}
\ldots \to H_k(A \cap B, *) & \xrightarrow{(i_1)_*} & H_k(A, *) & \xrightarrow{(j_1)_*} & H_k(A, A \cap B) & \xrightarrow{\partial} & H_{k-1}(A \cap B, *) \to \ldots \\
\downarrow \alpha_k & & \downarrow \beta_k & & \downarrow \gamma_k & & \downarrow \alpha_{k-1} \\
\ldots \to H_k(B, *) & \xrightarrow{(i_2)_*} & H_k(A \cup B, *) & \xrightarrow{(j_2)_*} & H_k(A \cup B, B) & \xrightarrow{\hat{\partial}} & H_{k-1}(B, *) \to \ldots .
\end{array}$$

Here $i_1, j_1, j_2, \alpha_k, \beta_k, \gamma_k, \alpha_{k-1}$ are the suitable inclusion maps. From Axiom 6’ (see Remark 6.1.13) we know that γ_k is an isomorphism. Then the theorem is a consequence of Lemma 6.1.32. $\square$

We will use the previous results to compute the homology groups of the ball $\bar{B}^n$ and of the sphere S^n in any homology theory.

So, let

$$\bar{B}^n = \{u \in \mathbb{R}^n : |u| \leqslant 1\}, \; B^n = \{u \in \mathbb{R}^N : |u| < 1\} \text{ and}$$
$$S^n = \{u \in \mathbb{R}^{n+1} : |u| = 1\}.$$

Example 6.1.34 *(a)* Since $\bar{B}^n$ is contractible, from Proposition 6.1.30 we have

$$H_k(\bar{B}^n, *) = 0 \text{ for all } k \in \mathbb{N}_0 \text{ and all } * \in \bar{B}^n.$$

(b) In contrast S^n is not contractible. To see this we argue by contradiction. So, suppose that S^n is contractible. According to Definition 6.1.7 we can find a function $h \in C([0, 1] \times S^n, S^n)$ such that

$$h(0, u) = u \text{ for all } u \in S^n \text{ and } h(1, u) = u_0 \text{ with } u_0 \in S^n.$$

Using the Tietze extension theorem, we can find $\hat{h} \in C([0, 1] \times \bar{B}^{n+1}, \mathbb{R}^{n+1})$ such that $\hat{h}|_{[0,1]\times S^n} = h$. We set

$$\hat{\varphi}(\cdot) = \hat{h}(0, \cdot) \text{ and } \hat{\psi}(\cdot) = \hat{h}(1, \cdot).$$

From the homotopy invariance of the Brouwer degree (see Proposition 3.1.14), we have

$$\begin{aligned} & d(\hat{\varphi}, B^{n+1}, 0) = d(\hat{\psi}, B^{n+1}, 0) \\ \Rightarrow \; & d(\mathrm{id}_{\mathbb{R}^{n+1}}, B^{n+1}, 0) = d(u_0, B^{n+1}, 0). \end{aligned} \tag{6.15}$$

Here by u_0 we mean the constant function $\psi(u) = u_0$ for all $u \in S^n$. But from Theorem 3.1.25(a), we have

$$\begin{aligned} & d(\mathrm{id}_{\mathbb{R}^{n+1}}, B^{n+1}, 0) \neq 0 \\ \Rightarrow\ & d(u_0, B^{n+1}, 0) \neq 0 \text{ (see (6.15))}. \end{aligned}$$

On the other hand since $u_0 \in S^n$, we have $d(u_0, B^{n+1}, 0) = 0$, a contradiction. This proves that S^n is not contractible (see also Proposition 3.1.32).

So, to compute the reduced homology groups of S^n, we proceed as follows.

First note that the homology groups $H_k(S^n, *)$ depend only on the homotopy type of S^n (see Proposition 6.1.14). So, without any loss of generality we can take the Euclidian norm on $\mathbb{R}^{n+1}$.

If $n = 0$, then from Corollary 6.1.16 and Proposition 6.1.17, we have

$$H_k(S^0, *) = H_k(*) \oplus H_k(*, *) = H_k(*) \text{ for all } k \in \mathbb{N}_0.$$

Now let $n \geqslant 1$ and let $u_N \in S^n$ and $u_S \in S^n$ be the north and south poles respectively. Set

$$S_1^n = S^n \backslash \{u_N\} \text{ and } S_2^n = S^n \backslash \{u_S\}.$$

Then $S^n = S_1^n \cup S_2^n$ and so by Theorem 6.1.33, we have the exact sequence

$$\bigoplus_{i=1}^{2} H_k(S_i^n, *) \to H_k(S^n, *) \to H_{k-1}(S_1^n \cap S_2^n, *) \to \bigoplus_{i=1}^{2} H_{k-1}(S_i^n, *). \tag{6.16}$$

Note that the spaces S_1^n, S_2^n are contractible. Therefore

$$H_k(S_1^n, *) = H_k(S_2^n, *) = 0 \text{ for all } k \in \mathbb{N}_0. \tag{6.17}$$

Also, the pair $(S_1^n \cap S_2^n, *)$ is clearly homotopically equivalent to $(S^{n-1}, *)$ and so, by Proposition 6.1.14

$$H_k(S_1^n \cap S_2^n, *) = H_k(S^{n-1}, *) \text{ for all } k \in \mathbb{N}_0\,. \tag{6.18}$$

From (6.16), (6.17), (6.18) we obtain

$$H_k(S^n, *) = H_{k-1}(S^{n-1}, *) \text{ for all } k \in \mathbb{N}_0.$$

Then by induction we have

$$H_k(S^n, *) = \begin{cases} H_0(*) & \text{if } k = n \\ 0 & \text{if } k \neq n. \end{cases}$$

So, in any homology theory, $H_n(S^n, *)$ is the only reduced homology group of S^n which is nontrivial and it coincides with $H_0(*)$ (the group of coefficients of the homology theory).

(c) From the previous two examples and the reduced exact homology sequence (see Proposition 6.1.29), we have

$$H_k(\bar{B}^n, S^{n-1}) = H_{k-1}(S^{n-1}, *) = \begin{cases} H_0(*) & \text{if } k = n \\ 0 & \text{if } k \neq 0. \end{cases}$$

Remark 6.1.35 If X is an infinite-dimensional Banach space and $\partial B_1 = \{u \in X : ||u|| = 1\}$, then ∂B_1 is contractible (compare with Example 6.1.34(b)).

Proposition 6.1.36 *If* $X_1 \subseteq \cdots \subseteq X_{k+1}$ *are Hausdorff topological spaces, then* $\operatorname{rank} H_n(X_{k+1}, X_1) \leqslant \sum_{i=1}^{k} \operatorname{rank} H_n(X_{i+1}, X_i)$ *for all* $n \in \mathbb{N}_0$.

Proof Consider the triple (X_{k+1}, X_k, X_1) and the long exact sequence corresponding to it according to Proposition 6.1.23. We have

$$\ldots \to H_n(X_k, X_1) \xrightarrow{i_*} H_n(X_{k+1}, X_1) \xrightarrow{j_*} H_n(X_{k+1}, X_k) \to \ldots \tag{6.19}$$

Then from the rank theorem we have

$$\begin{aligned} \operatorname{rank} H_n(X_{k+1}, X_1) &\leqslant \operatorname{rank} \ker j_* + \operatorname{rank} \operatorname{im} j_* \\ &= \operatorname{rank} \operatorname{im} i_* + \operatorname{rank} \operatorname{im} j_* \text{ (by the exactness of (6.19))} \\ &\leqslant \operatorname{rank} H_n(X_k, X_1) + \operatorname{rank} H_n(X_{k+1}, X_k). \end{aligned} \tag{6.20}$$

Since for $k = 1$, the inequality claimed by the proposition is in fact an equality the result follows from (6.20) and induction on $k \in \mathbb{N}$. □

Proposition 6.1.37 *If* $X_1 \subseteq X_2 \subseteq X_3 \subseteq X_4$ *are Hausdorff topological spaces then for all* $n \in \mathbb{N}$,

$$\operatorname{rank} H_n(X_3, X_2) - \operatorname{rank} H_n(X_4, X_1) \leqslant \operatorname{rank} H_{n-1}(X_2, X_1) + \operatorname{rank} H_{n+1}(X_4, X_3).$$

Proof We consider the triple (X_3, X_2, X_1) and the long exact sequence corresponding to it according to Proposition 6.1.23. We have

$$\ldots \to H_n(X_3, X_1) \xrightarrow{i_*} H_n(X_3, X_2) \xrightarrow{\partial} H_{n-1}(X_2, X_1) \to \ldots \tag{6.21}$$

From the rank theorem we have

$$
\begin{aligned}
&\text{rank } H_n(X_3, X_2) = \text{rank ker } \partial + \text{rank im } \partial = \\
&\text{rank im } i_* + \text{rank im } \partial \text{ (from the exactness of (6.21))} \\
\Rightarrow\ &\text{rank } H_n(X_3, X_2) - \text{rank im } i_* \leqslant \text{rank } H_{n-1}(X_2, X_1) \\
\Rightarrow\ &\text{rank } H_n(X_3, X_2) - \text{rank } H_n(X_3, X_1) \leqslant \text{rank } H_{n-1}(X_2, X_1). \qquad (6.22)
\end{aligned}
$$

Similarly, if we consider the triple (X_4, X_3, X_1), then from Proposition 6.1.23 we have the long exact sequence

$$\ldots \to H_{n+1}(X_4, X_3) \xrightarrow{\partial} H_n(X_3, X_1) \xrightarrow{i_*} H_n(X_4, X_1) \to \ldots . \qquad (6.23)$$

As above, via the rank theorem and the exactness of (6.23), we obtain

$$
\begin{aligned}
\text{rank } H_n(X_3, X_1) &= \text{rank ker } i_* + \text{rank im } i_* \\
&= \text{rank im } \partial + \text{rank im } i_* \\
&\leqslant \text{rank } H_{n+1}(X_4, X_3) + \text{rank } H_n(X_4, X_1). \qquad (6.24)
\end{aligned}
$$

Using (6.22) in (6.24), we obtain

$$
\begin{aligned}
&\text{rank } H_n(X_3, X_2) - \text{rank } H_{n-1}(X_2, X_1) \leqslant \\
&\qquad\qquad \text{rank } H_{n+1}(X_4, X_3) + \text{rank } H_n(X_4, X_1) \\
\Rightarrow &\text{rank } H_n(X_3, X_2) - \text{rank } H_n(X_4, X_1) \leqslant \\
&\qquad\qquad \text{rank } H_{n-1}(X_2, X_1) + \text{rank } H_{n+1}(X_4, X_3).
\end{aligned}
$$

The proof is now complete. □

Next we introduce a concrete homology theory which we will use in the sequel and which is known as "singular homology theory".

Singular homology theory extended simplicial homology theory to general topological spaces. Simplicial homology theory is defined for a special kind of spaces, namely compact polyhedra and the complexes resulting from them.

Let Δ_n be the standard n-simplex defined by

$$\Delta_n = \{(\lambda_n)_{k=0}^n \in \mathbb{R}^{n+1} : \sum_{\text{k}=0}^{n} \lambda_k = 1, \lambda_k \geqslant 0\}.$$

For $k \in \{0, \ldots, n\}$ we set $e_k = (0, \ldots, 0, 1, 0, \ldots, 0)$ (with 1 located at the $k+1$-entry).

Definition 6.1.38 Let X be a Hausdorff topological space.

(a) A "singular n-simplex" is a continuous map $\sigma : \Delta_n \to X$.
(b) The free abelian group with the singular n-simplexes as generators and coefficients in $\mathbb{Z}$ is called the "$n \stackrel{th}{=}$ singular chain group" and is denoted by $C_n(X)$. For $n < 0$, $C_n(X) = 0$ and if $c \in C_n(X)$, then c is called a "singular n-chain".

Remark 6.1.39 The world "singular" is used here to reflect the fact that the map $\sigma(\cdot)$ need not be a homeomorphism and can have "singularities". Moreover, its image $\sigma(\Delta_n)$ in general does not look at all like a simplex. A singular 0-simplex is a map from the singleton Δ_0 into X. Hence it can be identified with a point of X. A singular 1-simplex is a continuous map $\sigma : \Delta_1 \simeq [0, 1] \to X$, hence it is a path in X.

Definition 6.1.40 (a) Let X, Y be Hausdorff topological space and $\varphi : X \to Y$ a continuous map. If $\sigma : \Delta_n \to X$ is a singular n-simplex in X, then the composition $\varphi \circ \sigma : \Delta_n \to Y$ is a singular n-simplex in Y, denoted by $\varphi\sigma$. Suppose that $c = \sum_{k=1}^{n} a_k\sigma_k$, where $a_k \in \mathbb{Z}$ is an n-chain X (that is, $c \in C_n(X)$). Then

$$\varphi_*(c) = \sum_{k=1}^{n} a_k\varphi\sigma_k \in C_n(Y)$$

and the homomorphism $\varphi_* : C_n(X) \to C_n(Y)$ is the "homomorphism induced by φ".

(b) For each $k \in \{0, \dots, n\}$, let $d_k : \Delta_{n-1} \to \Delta_n$ be the affine function defined by

$$d_k(x_0, \dots, x_{n-1}) = (x_0, \dots, x_{k-1}, 0, x_{k+1}, \dots, x_{n-1})$$

is called the "k-face function in dimension n". For every singular n-simplex $\sigma : \Delta_n \to X$, the "boundary of σ" is defined to be the singular $(n-1)$-chain $\partial\sigma$ defined by

$$\partial\sigma = \sum_{k=0}^{n} (-1)^k \sigma \circ d_k.$$

This extends uniquely to a homomorphism $\partial : S_n(X) \to S_{n-1}(X)$ known as the "boundary operator".

Remark 6.1.41 Sometimes we write ∂_n instead of ∂ in order to indicate the chain group on which the boundary operator is acting.

The next proposition gives the most important feature of the boundary operator. Its proof is straightforward but it involves a tedious calculation and so it is omitted.

Proposition 6.1.42 $\partial_n \circ \partial_{n+1} = 0$ *for all* $n \in \mathbb{N}_0$.

Definition 6.1.43 (a) A singular n-chain c is said to be an "n-cycle" if $\partial c = 0$.

(b) A singular n-chain c is said to be an "n-boundary" if there is an $(n+1)$-chain b such that $\partial b = c$.

(c) By $Z_n(X)$ we denote the set of all n-cycles and by $B_n(X)$ the set of all n-boundaries. Both are abelian subgroups of $C_n(X)$.

Example 6.1.44 (a) Recall that a singular 1-simplex is a path $\sigma : [0, 1] \to X$ and $\partial\sigma$ corresponds to the formal difference $\sigma(1) - \sigma(0)$. Hence a 1-cycle is a formal $\mathbb{Z}$-linear combination of paths with the property that the set of initial points counted with multiplicities is the same as the set of terminal points with multiplicities.

(b) In the case of a singular 2-simplex $\sigma : \Delta_2 \to X$, the boundary is the sum of three paths with signs. Consider $\sigma : i_{\Delta^2} : \Delta^2 \to \mathbb{R}^3$ the inclusion map. Then

$$\partial i_{\Delta_2} = a(e_1, e_2) - a(e_0, e_2) + a(e_0, e_1).$$

So, ∂i_{Δ_2} is the sum of the singular 1-simplexes in the boundary of Δ_2 with appropriate signs.

Now we can define the singular homology groups.

Definition 6.1.45 Let X be a Hausdorff topological space. The collection $\{C_n(X), \partial_n\}_{n\geqslant 0}$ is called a "singular chain complex for X". We set

$$\begin{aligned} &Z_n(X) = \ker \partial_n \text{ for all } n \in \mathbb{N},\ Z_0(X) = C_0(X),\\ &B_n(X) = \operatorname{im} \partial_{n+1} \text{ for all } n \in \mathbb{N}_0 \text{ (see Definition 6.1.43).} \end{aligned}$$

Both are abelian subgroups of $C_n(X)$ and by Proposition 6.1.42 we have

$$B_n(X) \subseteq Z_n(X) \text{ for all } n \in \mathbb{N}_0.$$

So, we can define the quotient groups

$$H_n(X) = Z_n(X)/_{B_n(X)} = \begin{cases} \ker \partial_n/_{\operatorname{im} \partial_{n+1}} & \text{if } n \in \mathbb{N} \\ C_0(X)/_{\operatorname{im} \partial_1} & \text{if } n = 0. \end{cases}$$

This is the "n-th singular homology group of X". The singular homology of X is the collection

$$H_*(X) = \{H_n(X)\}_{n\in\mathbb{N}_0}.$$

Remark 6.1.46 The elements of $H_n(X)$ are called singular homology classes, the coset $u + B_n(X)$ being the class for the singular n-cycle u. Two n-cycles u and u' are said to be homologous if they belong to the same singular homology class. Clearly, u and u' are homologous if and only if $u - u' = \partial_{n+1}c$ for some singular $(n+1)$-chain c. If $H_n(X)$ is finitely generated, then rank $H_n(X)$=the n-th Betti number of X. Since $Z_n(X)$, $B_n(X)$ are subgroups of the abelian group $S_n(X)$, they are normal subgroups.

We can also define relative singular homology groups.

Definition 6.1.47 Let (X, A) be a pair of spaces. We set

$$C_n(X, A) = C_n(X)/C_n(A) \text{ for all } n \in \mathbb{N}_0.$$

This is the "relative n-singular chain group of X mod A", which is a free abelian group with generators those singular n-simplexes $\sigma : \Delta_n \to X$ whose images are not completely contained in A. The elements of $C_n(X, A)$ are called "relative singular n-chains of X mod A". Because $\partial_n : C_n(X) \to C_{n-1}(X) (n \in \mathbb{N})$ is a homomorphism and $\partial_n(C_n(A)) \subseteq C_{n-1}(A)$, there exists a unique homomorphism

$$\partial_n : C_n(X, A) \to C_{n-1}(X, A)$$

(for notational economy we use the same symbol). This is the "boundary operator" for the relative singular homology groups. As before (see Proposition 6.1.42), we have

$$\partial_{n-1} \circ \partial_n = 0 \text{ for all } n \in \mathbb{N}.$$

We set

$$Z_n(X, A) = \ker \partial_n \text{ for all } n \in \mathbb{N}_0$$

(the subgroup of relative singular n-cycles of X mod A),

$$B_n(X, A) = \operatorname{im} \partial_{n+1} \text{ for all } n \in \mathbb{N}_0$$

(the subgroup of relative singular n-boundaries of X mod A).

We have $B_n(X, A) \subseteq Z_n(X, A)$ and so we can define

$$H_n(X, A) = Z_n(X, A)/B_n(X, A) \text{ for all } n \in \mathbb{N}_0.$$

This is the "n-th relative singular homology group of X mod A". This is a free abelian group and if it is finitely generated, then rank $H_n(X, A)$ is the "n-th Betti number" of the pair (X, A).

Remark 6.1.48 We have

$$Z_n(X, A) = \begin{cases} \{c \in C_n(X) : \partial_n c \in C_{n-1}(A)\} & \text{if } n \in \mathbb{N} \\ C_0(X) & \text{if } n = 0 \end{cases}$$

and $B_n(X, A) = B_n(X) + C_n(A)$ (that is, the subgroup generated by $B_n(X)$ and $C_n(A)$). If $A = \emptyset$, then $H_n(X, \emptyset) = H_n(X)$.

Proposition 6.1.49 *The relative singular homology introduces a homology theory in the sense of Definition 6.1.12 on the collection of all pairs of spaces.*

Remark 6.1.50 We have defined singular homology theory using $\mathbb{Z}$ as the group of coefficients, because this is the most standard singular homology. However, in some cases, in order to avoid torsion phenomena, we replace $\mathbb{Z}$ by a field $\mathbb{F}$. In this case $H_n(X, A)$, $n \in \mathbb{N}_0$, is a vector space. Recall that in the presence of torsion, we may have rank $H_n(X, A) = 0$ although $H_n(X, A) \neq 0$. Finally, we mention that rank $H_0(X)$ coincides with the number of path components of X. More generally, if

$A \subseteq X$, then rank $H_0(X, A)$ coincides with the number of path components $C \subseteq X$ which do not intersect A. So, if each $u \in X$ can be connected to an element of A by a path in X, then $H_0(X, A) = 0$.

6.2 Critical Groups, Morse Relations

In Morse theory the local behavior of a smooth function φ near an isolated critical point is described by a sequence of abelian groups, known as "critical groups".

So, let X be a Banach space, $\varphi \in C^1(X)$ and $c \in \mathbb{R}$. From Sects. 5.2 and 5.3 we recall the following notation:

$$\varphi^c = \{u \in X : \varphi(u) \leqslant c\} \text{ (the sublevel set of } \varphi \text{ at } c \in \mathbb{R}),$$
$$K_\varphi = \{u \in X : \varphi'(u) = 0\} \text{ (the critical set of } \varphi),$$
$$K_\varphi^c = \{u \in K_\varphi : \varphi(u) = c\} \text{ (the critical set of } \varphi \text{ at the level } c \in \mathbb{R}).$$

Definition 6.2.1 Suppose that $u \in K_\varphi$ is isolated. The "critical groups" of φ at u are defined by

$$C_k(\varphi, u) = H_k(\varphi^c \cap U, \varphi^c \cap U \setminus \{u\}) \text{ for all } k \in \mathbb{N}_0,$$

where H_* denotes the relative singular homology group with $\mathbb{Z}$ being the group of coefficients and U is a neighborhood of u such that $K_\varphi \cap U = \{u\}$.

Remark 6.2.2 The excision property (see Definition 6.1.12, Axiom 6) implies that the above definition is independent of the choice of the neighborhood U for which we have $K_\varphi \cap U = \{u\}$ (recall u is isolated). If we choose the elements of a field $\mathbb{F}$ as coefficients for the homology groups, then the critical groups are $\mathbb{F}$-vector spaces. From the above definition it is clear that the critical groups depend only on the behavior of φ near u. Evidently, they can be defined, even if φ is defined only in a neighborhood of u. This will become even more evident in the next section. Finally, recall that $C_k(\varphi, u) = 0$ for all $k \in -\mathbb{N}$.

Proposition 6.2.3 *If $u \in X$ is a local minimizer of $\varphi \in C^1(X)$ which is an isolated critical point, then $C_k(\varphi, u) = \delta_{k,0}\mathbb{Z}$ (recall $\delta_{k,m} = \begin{cases} 1 \text{ if } k = m \\ 0 \text{ if } k \neq m \end{cases}$ for all $k, m \in \mathbb{N}_0$, the "Kronecker symbol").*

Proof Since u is a local minimizer and an isolated critical point of φ, we can find a neighborhood U of u such that

$$K_\varphi \cap U = \{u\} \text{ and } c = \varphi(u) < \varphi(v) \text{ for all } v \in U \setminus \{u\}. \tag{6.25}$$

Therefore according to Definition 6.2.1, we have

$$\begin{aligned}
& C_k(\varphi, u) = H_k(\{u\}, \emptyset) \text{ for all } k \in \mathbb{N}_0 \text{ (see (6.25))} \\
\Rightarrow\ & C_k(\varphi, 0) = \delta_{k,0}\mathbb{Z} \text{ for all } k \in \mathbb{N}_0 \\
& \text{(see Definition 6.1.12, Axiom 7 and Remark 6.1.13).}
\end{aligned}$$

The proof is now complete. □

The situation is more involved with local maximizers.

Proposition 6.2.4 *If $u \in X$ is a local maximizer of $\varphi \in C^1(X)$ which is an isolated critical point, then when* $\dim X = m < \infty$, *we have $C_k(\varphi, u) = \delta_{k,m}\mathbb{Z}$ for all $k \in \mathbb{N}_0$, when* $\dim X = \infty$, *we have $C_k(\varphi, u) = 0$ for all $k \in \mathbb{N}_0$.*

Proof Since u is a local maximizer and an isolated critical point of φ, we can find $r > 0$ such that

$$K_\varphi \cap \bar{B}_r(u) = \{u\} \text{ and } \varphi(v) < \varphi(u) = c \text{ for all } v \in \bar{B}_r(u)\backslash\{u\}. \tag{6.26}$$

Here $\bar{B}_r(u) = \{y \in X : ||y - u|| \leqslant r\}$.

First assume that $\dim X = m < \infty$. Consider the deformation

$$h(t, y) = u + (1 - t)(y - u) + tr\frac{y - u}{||y - u||} \text{ for all } (t, u) \in [0, 1] \times \bar{B}_r(u)\backslash\{u\}.$$

Evidently, we have

$$h(0, y) = y \text{ for all } y \in \bar{B}_r(u)\backslash\{u\} \text{ and } h(1, \cdot)|_{\partial B_r(u)} = \mathrm{id}|_{\partial B_r(u)}$$

with $\partial B_r(u) = \{y \in X : ||y - u|| = r\}$. Therefore $\partial B_r(u)$ is a deformation retract of $\bar{B}_r(u)\backslash\{u\}$ (see Definition 5.3.10(b)). Then Definition 6.2.1 and Corollary 6.1.24(a), together with (6.26), imply that

$$C_k(\varphi, u) = H_k(\bar{B}_r(u), \bar{B}_r(u)\backslash\{u\}) = H_k(\bar{B}_r(u), \partial B_r(u)) \text{ for all } k \in \mathbb{N}_0. \tag{6.27}$$

But from Example 6.1.34(c) we know that

$$\begin{aligned}
& H_k(\bar{B}_r(u), \partial B_r(u)) = \delta_{k,m}\mathbb{Z} \text{ for all } k \in \mathbb{N}_0 \\
\Rightarrow\ & C_k(\varphi, u) = \delta_{k,m}\mathbb{Z} \text{ for all } k \in \mathbb{N}_0.
\end{aligned}$$

Next, assume that $\dim X = \infty$. In this case we know that both $\bar{B}_r(u)$ and $\bar{B}_r(u)\backslash\{u\}$ are contractible (see Remark 6.1.35) and so from Proposition 6.1.30 and 6.1.31, we conclude that

$$\begin{aligned}
& H_k(\bar{B}_r(u), \bar{B}_r(u)\backslash\{u\}) = 0 \text{ for all } k \in \mathbb{N}_0 \\
\Rightarrow\ & C_k(\varphi, u) = 0 \text{ for all } k \in \mathbb{N}_0 \text{ (see (6.27)).}
\end{aligned}$$

The proof is now complete. □

When $X = \mathbb{R}$, the critical groups at any isolated critical point can be completely described.

Proposition 6.2.5 *If $X = \mathbb{R}$, $\varphi \in C^1(\mathbb{R})$ and $u \in K_\varphi$ is isolated, then one of the following three situations can occur:*

(a) if u is a local minimizer of φ, then $C_k(\varphi, u) = \delta_{k,0}\mathbb{Z}$ for all $k \in \mathbb{N}_0$;
(b) if u is a local maximizer of φ, then $C_k(\varphi, u) = \delta_{k,1}\mathbb{Z}$ for all $k \in \mathbb{N}_0$;
(c) for all other cases $C_k(\varphi, u) = 0$ for all $k \in \mathbb{N}_0$.

Proof Evidently, (a) and (b) follow from Propositions 6.2.3 and 6.2.4, respectively.
Since u is isolated, we can find $\epsilon > 0$ such that

$$K_\varphi \cap [u - \epsilon, u + \epsilon] = \{u\}.$$

By hypothesis u is not a local extremum of φ. So, φ is either increasing or decreasing on $[u - \epsilon, u + \epsilon]$. To fix things, we assume that φ is increasing (the reasoning in the same if φ is decreasing). Then

$$\begin{aligned} &\varphi^{\varphi(u)} \cap [u - \delta, u + \delta] = [u - \delta, u] \\ \Rightarrow\ &C_k(\varphi, u) = H_k([u - \delta, u], [u - \delta, u)) \text{ for all } k \in \mathbb{N}_0 \text{ (see Definition 6.2.1)} \\ \Rightarrow\ &C_k(\varphi, u) = 0 \text{ for all } k \in \mathbb{N}_0 \text{ (see Propositions 6.1.30 and 6.1.31).} \end{aligned}$$

The proof is now complete. □

Now we pass to a Hilbert space setting, where Morse theory is more effective. So, let $X = H =$ a Hilbert space and $\varphi \in C^2(H)$. Recall that $u \in K_\varphi$ is "nondegenerate" if the self-adjoint operator $\varphi''(u) \in \mathscr{L}(H, H)$ is invertible. The dimension of the negative space of $\varphi''(u)$ is called the "Morse index of u" and is denoted by $m = m(u) \in \mathbb{Z} \cup \{+\infty\}$ (see Definition 5.4.18). Using the Morse lemma (see Proposition 5.4.19), we can compute the critical groups of $\varphi \in C^2(H)$ at an isolated critical point u which is nondegenerate.

Proposition 6.2.6 *If H is a Hilbert space, $\varphi \in C^2(H)$ and $u \in K_\varphi$ is isolated and nondegenerate, with Morse index m, then $C_k(\varphi, u) = \delta_{k,m}\mathbb{Z}$ for all $k \in \mathbb{N}_0$.*

Proof By replacing φ with $\psi(u) = \varphi(u + y) - \varphi(u)$ for all $y \in H$ if necessary, we see that without any loss of generality we may assume that

$$u = 0 \text{ and } c = \varphi(u) = 0.$$

Invoking Proposition 5.4.19 (the Morse lemma), we can find a Lipschitz continuous homeomorphism of a neighborhood W of 0 onto a neighborhood U of 0 such that

$$h(0) = 0 \text{ and } \varphi(h(u)) = \frac{1}{2}(\varphi''(0)u, u)_H \text{ for all } u \in W \tag{6.28}$$

(here, by $(\cdot,\cdot)_H$ we denote the inner product of H). If $\psi(u) = \varphi(h(u))$ and $B \subseteq W$ is a closed ball centered at 0, we have

$$C_k(\varphi, 0) = H_k(\varphi^\circ \cap h(B), \varphi^\circ \cap h(B)\backslash\{0\} = H_k(\psi^\circ \cap B, \psi^\circ \cap B\backslash\{0\})) \quad \text{for all } k \in \mathbb{N}_0. \tag{6.29}$$

Since u is nondegenerate, $\varphi''(0)$ is invertible and so we have

$$H = H_- \oplus H_+$$

with $\psi''(0)$ positive (respectively negative) definite on H_+ (respectively H_-). So, any $v \in U$ can be decomposed in a unique way as $v = v_- + v_+$ with $v_- \in H_-$, $v_+ \in H_+$. We consider the deformation $\xi : [0,1] \times B \to B$ defined by

$$\xi(t, v) = v_- + (1-t)v_+ \text{ for all } (t, v) \in [0,1] \times B.$$

Then from (6.28) and exploiting the orthogonality of the component spaces, we have

$$\psi(\xi(t, v)) = \psi(v_-) + (1-t)^2\psi(v_+).$$

This shows that $V_- \cap B\backslash\{0\}$ is a deformation retract of $\psi^\circ \cap B\backslash\{0\}$ and $V_- \cap B$ is a deformation retract of $\psi^\circ \cap B$. Then we have

$$\begin{aligned} H_k(\psi^\circ \cap B, \psi^\circ \cap B\backslash\{0\}) &= H_k(V_- \cap B, V_- \cap B\backslash\{0\}) \text{ (see Corollary 6.1.24)} \\ &= H_k(\bar{B}^m, S^{m-1}) \\ &= \delta_{k,m}\mathbb{Z} \text{ for all } k \in \mathbb{N}_0 \text{ (see Example 6.1.34(c))} \\ \Rightarrow \qquad & C_k(\varphi, 0) = \delta_{k,m}\mathbb{Z} \text{ for all } k \in \mathbb{N}_0. \end{aligned}$$

The proof is now complete. □

What about degenerate critical points? For this case we have the so-called "Shifting Theorem", which says that for a degenerate critical point, the critical groups depend on the Morse index and on the "degenerate part" of the functional. Thus the computation of the critical groups is reduced to a finite-dimensional problem. The Shifting Theorem will be proved with the help of an extension of the Morse lemma (see Proposition 5.4.19), which we prove first.

We start with a definition.

Definition 6.2.7 Let X and Y be two Banach spaces and let $L \in \mathscr{L}(X, Y)$. We say that L is a "Fredholm operator" if $\ker L$ is finite-dimensional and $R(L) = L(X)$ is finite codimensional (that is, $\dim(Y/\ker L) < \infty$). The number

$$i(L) = \dim \ker L - \dim(Y/\ker L)$$

is called the "index" of L. The set of all Fredholm operators $L : X \to Y$ is denoted by Fred (X, Y).

Remark 6.2.8 If $L \in \text{Fred}\,(X, Y)$, then $R(L) \subseteq Y$ is closed. Moreover, we have

$$X = \ker L \oplus V$$

and $L|_V$ is an isomorphism of V onto $L(X)$. The set Fred (X, Y) is open in $\mathscr{L}(X, Y)$ and the map $L \to i(L)$ is continuous (hence, it is constant on each connected component of Fred (X, Y)). Every $L \in \text{Fred}\,(X, Y)$ is invertible modulo finite rank operators, that is, there exists an $S \in \mathscr{L}(Y, X)$ such that both

$$L \circ S - \text{id}_Y \text{ and } S \circ L - \text{id}_X$$

are finite rank operators. Finally, if $K \in \mathscr{L}_c(X, X)$, then $\lambda \text{id}_X - K$ is a Fredholm operator for every $\lambda \neq 0$.

Next we state and prove an extension of the Morse lemma which we will need in the proof of the shifting theorem.

Proposition 6.2.9 *If H is a Hilbert space, U is an open neighborhood of the origin $\varphi \in C^2(U)$, $0 \in K_\varphi$ with* $\dim\ker \varphi''(0) > 0$, *$L = \varphi''(0)$ is a Fredholm operator hence*

$$H = \ker L \oplus R(L)$$

and so every $u \in H$ admits a unique decomposition

$$u = w + v \textit{ with } w \in \ker L,\ v \in R(L),$$

then there exists an open neighborhood V of the origin, an open neighborhood W of the origin in $\ker L$, *a homeomorphism h from V into U and a function $\hat{\varphi} \in C^2(W)$ such that*

$$h(0) = 0, \hat{\varphi}'(0) = 0, \hat{\varphi}''(0) = 0$$

and $\varphi(h(u)) = \frac{1}{2}(Lv, v)_H + \hat{\varphi}(w)$ for all $u \in V$.

Proof Let $P \in \mathscr{L}(H, H)$ be the orthogonal projection onto $R(L)$. The implicit function theorem implies that we can find $\rho_1 > 0$ and a C^1-function

$$\sigma : B_{\rho_1} \cap \ker L \to R(L) \ \ (B_{\rho_1} = \{u \in H : ||u|| < \rho_1\})$$

such that $\sigma(0) = 0$, $\sigma'(0) = 0$ and

$$P(\nabla\varphi(w + \sigma(w))) = 0. \tag{6.30}$$

We let $W = B_{\rho_1} \cap \ker L$ (an open neighborhood of the origin in $\ker L$) and consider $\hat{\varphi} : W \to \mathbb{R}$ defined by

$$\hat{\varphi}(w) = \varphi(w + \sigma(w)) \text{ for all } w \in W.$$

Evidently, $\hat{\varphi} \in C^1(W)$ and using (6.30) we have

$$\nabla\hat{\varphi}(w) = (\mathrm{id}_H - P)\nabla\varphi(w + \sigma(w))$$
$$\text{and } \varphi''(w) = (\mathrm{id}_H - P)\varphi''(w + \sigma(w))(w + \sigma'(w)).$$

So, we have

$$\nabla\hat{\varphi}(0) = (\mathrm{id}_H - P)\nabla\varphi(0) = 0,$$
$$\varphi''(0) = (\mathrm{id}_H - P)\varphi''(0) = 0.$$

On $[0, 1] \times U$, we define the function

$$\xi(t, v, w) = (1 - t)[\hat{\varphi}(w) + \frac{1}{2}(Lv, v)_H] + t\varphi(v + w + \sigma(w))$$

and the vector field

$$g(t, v, w) = \begin{cases} 0 & \text{if } v = 0 \\ -\xi_t'(t, v, w)||\xi_v'(t, v, w)||^{-2}\xi_v(t, v, w) & \text{if } v \neq 0. \end{cases}$$

We consider the following abstract Cauchy problem

$$\gamma'(t) = g(t, \gamma(t), w),\ t \in [0, 1],\ \gamma(0) = v. \tag{6.31}$$

We will establish the existence of a local flow for (6.31). To this end let

$$\psi(v, w) = \varphi(v + w + \sigma(w)) - \hat{\varphi}(w) - \frac{1}{2}(Lv, v)_H.$$

Then using (6.30) we see that

$$\psi(0, w) = 0,\ \psi_v'(0, w) = 0,\ \psi_v''(0, 0) = 0.$$

It follows that

$$\psi(v, w) = \int_0^1 (1 - s)(\psi_v''(sv, w)v, v)_H ds, \tag{6.32}$$

$$\psi_v(v, w) = \int_0^1 \psi_v''(sv, w)v ds. \tag{6.33}$$

From (6.32) and (6.33) we infer that for every $\epsilon > 0$, there exists a $\delta(\epsilon) \in (0, \rho_1)$ such that

$$|\psi(v, w)| \leqslant \epsilon ||v||^2 \text{ and } ||\psi_v'(v, w)|| \leqslant \epsilon ||v|| \text{ when } ||v + w|| \leqslant \delta(\epsilon). \tag{6.34}$$

Recall that $L|_{R(L)}$ is invertible. So, we can find $c > 0$ such that

$$\frac{1}{c}||v|| \leqslant ||L(v)|| \leqslant c||v|| \text{ for all } v \in R(L). \tag{6.35}$$

For $v \neq 0$, we have

$$g(t, v, w) = -\psi(v, w)||L(v) + t\psi_v'(v, w)||^{-2}(L(v) + t\psi_v'(v, w)).$$

Let $\epsilon = \frac{1}{2c}$. Using (6.34) and (6.35), we see that

$$|g(t, v, w)| \leqslant 2c(c + \epsilon)\epsilon||v|| \text{ for } ||v + w|| \leqslant \delta(\epsilon). \tag{6.36}$$

By definition $g(t, 0, w) = 0$ and so we see that g is continuous. Let $\rho \in (0, \delta(\epsilon))$ be such that

$$||\psi_v''(v, w)||_{\mathscr{L}} \leqslant 1 \text{ for } ||v + w|| \leqslant \rho \text{ with } v \neq 0. \tag{6.37}$$

Using (6.34), (6.35) and (6.37), we see that we can find $c_1 > 0$ such that

$$||g_v'(t, v, w)|| \leqslant c_1 \text{ for all } ||v + w|| \leqslant \rho \text{ with } v \neq 0.$$

Now from (6.36) and the mean value theorem, we see that we can find $c_2 > 0$ such that

$$|g(t, v_1, w) - g(t, v_2, w)| \leqslant c_2||v_1 - v_2|| \text{ for } ||v_i + w|| \leqslant \rho, \ i = 1, 2.$$

So, the flow $\gamma(\cdot)$ of (6.31) exists locally. Since $\gamma(t, 0, w) = 0$, the flow γ is well-defined on $[0, 1] \times V$ with V a neighborhood of the origin in H. We set

$$h(u) = h(v, w) = w + \sigma(w) + \gamma(1, v, w) \text{ for all } u \in V.$$

The invertibility of h follows from the invertibility of the flow $\gamma(1, \cdot, w)$. Then h is the desired local homeomorphism. $\square$

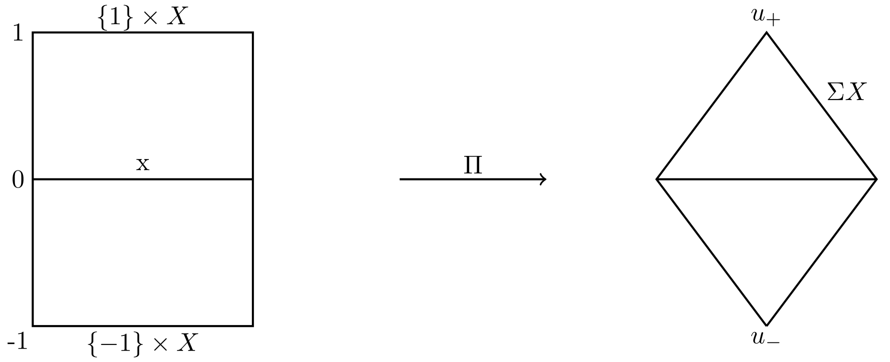
To prove the shifting theorem, we will need one more auxiliary result. First a definition.

Definition 6.2.10 For a Hausdorff topological space X, the quotient space

$$\Sigma X = [-1, 1] \times X/_{\{-1\}\times X, \{1\}\times X}$$

is called the "suspension of X" or "double cone over X".

Remark 6.2.11 So, the suspension ΣX of X is obtained from $[-1,1] \times X$ by identifying each of the subsets $\{-1\} \times X$ and $\{+1\} \times X$ with two different points. The following figure explains this notion and justifies its name



Proposition 6.2.12 *If* $A \subseteq \mathbb{R}^n$, $0 \in A$ *and* B^m *is the* m*-ball, then* $H_k(B^m \times A, (B^m \times A)\backslash\{0\}) = H_{n-m}(A, A\backslash\{0\})$.

Proof Let $m \geqslant 2$ and recall that B^m is homeomorphic to $[-1,1]^m$. Then we have

$$(B^m \times A, (B^m \times A)\backslash\{0\}) = (B^{m-1} \times [-1,1] \times A, (B^{m-1} \times [-1,1] \times A)\backslash\{0\}).$$

Then the result follows by induction from the case $m = 1$.

From the excision property, we have

$$H_k([-1,1] \times A, ([-1,1] \times A)\backslash\{0\}) = H_k(\Sigma A, \Sigma A\backslash\{0\}) \text{ for all } k \in \mathbb{N}_0. \quad (6.38)$$

We introduce the sets

$$\hat{A}_+ = \Sigma A\backslash\{u_-\} \text{ and } \hat{A}_- = \Sigma A\backslash\{u_+\},$$

where u_+ and u_- are the two points which are identified with $\{1\} \times A$ and $\{-1\} \times A$ respectively. Also, set

$$V_+ = \hat{A}_+\backslash(\{0\} \times [-1,0]) \text{ and } V_- = \hat{A}_-\backslash(\{0\} \times [0,1]).$$

We have

$$\hat{A}_+ \cup \hat{A}_- = \Sigma A, \ \hat{A}_+ \cap \hat{A}_- = [-1,1] \times A, \quad (6.39)$$

$$V_+ \cup V_- = \Sigma A\backslash\{0\}, \ V_+ \cap V_- = [-1,1] \times (A\backslash\{0\}). \quad (6.40)$$

Evidently,

$$\hat{A}_+ \text{ and } V_+ \text{ are contractible to } u_+,$$
$$\hat{A}_- \text{ and } V_- \text{ are contractible to } u_-.$$

So, we have

$$H_k(\hat{A}_\pm, V_\pm) = H_k(u_\pm, u_\pm) = 0 \text{ for all } k \in \mathbb{N}.$$

Using Theorem 6.1.33 (the Mayer–Vietoris theorem), we obtain

$$\begin{aligned}
& H_k(\hat{A}_+ \cup \hat{A}_-, V_+ \cup V_-) = H_{k-1}(\hat{A}_+ \cap \hat{A}_-, V_+ \cap V_-) \text{ for all } k \in \mathbb{N} \\
\Rightarrow\ & H_k(\Sigma A, \Sigma A \backslash \{0\}) = H_{k-1}([-1,1] \times A, [-1,1] \times (A \backslash \{0\})) \text{ (see (6.39), (6.40))} \\
\Rightarrow\ & H_k([-1,1] \times A, ([-1,1] \times A) \backslash \{0\}) = H_{k-1}(A, A \backslash \{0\}) \text{ for all } k \in \mathbb{N} \text{ (see (6.38))}.
\end{aligned}$$

This proves the proposition for $m = 1$ and then by induction for every $m \in \mathbb{N}$. ☐

Now we are ready to state and prove the shifting theorem, which takes care of the degenerate case.

Theorem 6.2.13 *If H is a Hilbert space, $U \subseteq H$ is open, $\varphi \in C^2(U)$ and $u \in K_\varphi$ is isolated with finite Morse index m and* dimker $\varphi''(u)$ *is finite too, then $C_k(\varphi, u) = C_{k-m}(\hat{\varphi}, 0)$ for all $k \in \mathbb{N}_0$, with $\hat{\varphi}$ as in Proposition 6.2.9.*

Proof Without any loss of generality, we may assume that $u = 0$. Let $C \subseteq U$ be a closed neighborhood of the origin. Using Proposition 6.2.9, we set

$$c = \varphi(0) = \hat{\varphi}(0) \text{ and } \psi(u) = \varphi(v + w) = \frac{1}{2}(Lv, v)_H + \hat{\varphi}(w) \text{ for all } u \in V$$

(we have kept the notation introduced in Proposition 6.2.9). We have

$$\begin{aligned}
C_k(\varphi, 0) &= H_k(\varphi^c \cap h(C), \varphi^c \cap h(C) \backslash \{0\}) \\
&= H_k(\psi^c \cap C, \psi^c \cap C \backslash \{0\}) = C_k(\psi, 0) \text{ for all } k \in \mathbb{N}_0.
\end{aligned}$$

By hypothesis, $0 \in \ker L$ is the only critical point of $\hat{\varphi} \in C^2(W)$. Since dimker L is finite, the Palais–Smale condition is satisfied over any closed ball $B_r \subseteq W$. From the deformation theorem (see Theorem 5.3.7), we can find $\epsilon > 0$ and $E \subseteq W$ closed positively invariant for the negative gradient flow such that $\hat{\varphi}^c \cap E$ is a strong deformation retract of $\hat{\varphi}^{c+\epsilon} \cap E$ and $\hat{\varphi}$ is nondecreasing along this deformation h. We set

$$\hat{h}(t, v, w) = v_- + (1-t)v_+ + h(t, w) \text{ for } t \in [0,1], u \in C = R(L) \times (\hat{\varphi}^{c+\epsilon} \cap E)$$

(see the proof of Proposition 6.2.6). We can easily check that $H_- \cap \hat{\varphi}^c \cap E$ is a strong deformation retract of $\psi^c \cap C$ and $(H_- \times (\hat{\varphi}^c \cap E)) \backslash \{0\}$ is a strong deformation retract of $(\psi^c \cap C) \backslash \{0\}$. Therefore we have

$$C_k(\psi, 0) = H_k(\psi^c \cap C, (\psi^c \cap C)\setminus\{0\})$$
$$= H_k(H_- \times (\hat{\varphi}^c \cap E), (H_- \times (\hat{\varphi}^c \cap E))\setminus\{0\}) \text{ for all } k \in \mathbb{N}_0. \quad (6.41)$$

If $m = \dim H_- = 0$, then

$$C_k(\psi, 0) = H_k(\hat{\varphi}^c \cap E, (\hat{\varphi}^c \cap E)\setminus\{0\}) = C_k(\hat{\varphi}, 0) \text{ for all } k \in \mathbb{N}_0$$

which is the result of the theorem.

If $m = \dim H_- \geqslant 1$, then using Proposition 6.2.12 we have

$$\begin{aligned} C_k(\psi, 0) &= H_k(\mathbb{R}^m \times (\hat{\varphi}^c \cap E), (\mathbb{R}^m \times (\hat{\varphi}^c \cap E))\setminus\{0\}) \\ &= H_k(B^m \times (\hat{\varphi}^c \cap E), (B^m \times (\hat{\varphi}^c \cap E))\setminus\{0\}) \\ &= H_{k-m}(\hat{\varphi}^c \cap E, (\hat{\varphi}^c \cap E)\setminus\{0\}) = C_{k-m}(\hat{\varphi}, 0). \end{aligned}$$

The proof is now complete. □

Let $m^*(u) = m(u) + \dim \ker L$ (the extended Morse index of u). Then from Theorem 6.2.13 we infer the following result.

Corollary 6.2.14 *If everything is as in Theorem 6.2.13 and $C_k(\varphi, u) \neq 0$, then $m(u) \leqslant k \leqslant m^*(u)$.*

We return to the more general setting of a Banach space X. We show that nontrivial singular homology groups imply the presence of a critical level between two levels $a < b$. More precisely, we have the following property.

Proposition 6.2.15 *If X is a Banach space, $\varphi \in C^1(X)$ satisfies the C-condition and there exist $k_0 \in \mathbb{N}_0$ and levels $a, b \in \mathbb{R}$ such that $a < b$ and*

$$H_{k_0}(\varphi^b, \varphi^a) \neq 0,$$

then $K_\varphi \cap \varphi^{-1}([a, b]) \neq \emptyset$.

Proof We argue indirectly. So, suppose that $K_\varphi \cap \varphi^{-1}([a, b]) = \emptyset$. Then Corollary 5.3.13 implies that φ^a is a strong deformation retract of φ^b. Proposition 6.1.15 implies that

$$H_k(\varphi^b, \varphi^a) = 0 \text{ for all } k \in \mathbb{N}_0,$$

a contradiction to our hypothesis that $H_{k_0}(\varphi^b, \varphi^a) \neq 0$. □

We can be more precise and relate the change in the topology of sublevel sets across a critical level to the critical groups of the critical points for that level.

Proposition 6.2.16 *If X is a Banach space, $\varphi \in C^1(X)$, $a, b \in \mathbb{R}$ with $a < b$, φ satisfies the $C_{c'}$-condition at every level $c' \in [a, b)$, $K_\varphi \cap [a, b] = \{c\}$ with $c \notin \{a, b\}$ and $K_\varphi^c = \{u_i\}_{i=1}^n$ is finite, then $H_k(\varphi^b, \varphi^a) = \bigoplus_{i=1}^{n} C_k(\varphi, u_i)$ for all $k \in \mathbb{N}_0$; in particular*

$$\text{rank } H_k(\varphi^b, \varphi^a) = \sum_{i=1}^{n} \text{rank } C_k(\varphi, u_i) \textit{ for all } k \in \mathbb{N}_0.$$

Proof Using Corollary 5.3.13 and Corollary 6.1.24, we see that

$$H_k(\varphi^b, \varphi^a) = H_k(\varphi^c, \varphi^c \backslash K_\varphi^c) \text{ for all } k \in \mathbb{N}_0. \tag{6.42}$$

Let $\{U_i\}_{i=1}^n$ be pairwise disjoint open neighborhoods of the critical points $\{u_i\}_{i=1}^n$ such that

$$V = \bigcup_{i=1}^{n} U_i \subseteq \varphi^{-1}([a, b]).$$

We have that $U_i \cap K_\varphi = \{u_i\}$ and so from Definition 6.2.1 it follows that

$$H_k(\varphi^c \cap U_i, \varphi^c \cap U_i \backslash \{0\}) = C_k(\varphi, u_i) \text{ for all } k \in \mathbb{N}_0 \text{ and all } i \in \{1, \dots, n\}.$$

From the excision property of singular homology (see Definition 6.1.12 and Proposition 6.1.49) and using Proposition 6.1.20, we obtain

$$H_k(\varphi^c, \varphi^c \backslash K_\varphi^c) = H_k(\varphi^c \cap V, (\varphi^c \backslash K_\varphi^c) \cap V) = \bigoplus_{i=1}^{n} C_k(\varphi, u_i) \text{ for all } k \in \mathbb{N}_0. \tag{6.43}$$

From (6.42) and (6.43) we conclude that

$$H_k(\varphi^b, \varphi^a) = \bigoplus_{i=1}^{n} C_k(\varphi, u_i) \text{ for all } k \in \mathbb{N}_0.$$

In particular, from the above isomorphism, we infer that

$$\text{rank } H_k(\varphi^b, \varphi^a) = \sum_{i=1}^{n} \text{rank } C_k(\varphi, u_i) \text{ for all } k \in \mathbb{N}_0.$$

The proof is now complete. ☐

What can be said about the change in the topology when we cross multiple critical levels? In this direction, we have the following result.

Proposition 6.2.17 *If X is a Banach space, $\varphi \in C^1(X)$, $-\infty < a < b < \infty$ are regular values of φ, $\varphi^{-1}([a, b]) \cap K_\varphi$ is finite and φ satisfies the C_c-condition for every $c \in [a, b]$, then* $\text{rank } H_k(\varphi^b, \varphi^a) \leqslant \sum_{u \in K_\varphi^{[a,b]}} \text{rank } C_k(\varphi, u)$ *where* $K_\varphi^{[a,b]} = \varphi^{-1}([a, b]) \cap K_\varphi$.

Proof Let $\{c_i\}_{i=1}^n$ be the critical values of φ in (a, b) in increasing order (that is, $c_1 < \dots < c_n$). Let $\{a_i\}_{i=1}^{n+1} \subseteq [a, b]$ be such that

$$a = a_1 < c_1 < a_2 < c_2 < \ldots < c_{n-1} < a_n < c_n < a_{n+1} = b.$$

Using Proposition 6.1.36 with $X_i = \varphi^{a_i}$ we have

$$\begin{aligned}
\operatorname{rank} H_k(\varphi^b, \varphi^a) &\leqslant \sum_{i=1}^{n} \operatorname{rank} H_k(\varphi^{a_{i+1}}, \varphi^{a_i}) \\
&= \sum_{i=1}^{n} \sum_{u \in K_\varphi^{c_i}} C_k(\varphi, u) \text{ (see Proposition 6.2.16)} \\
&= \sum_{u \in K_\varphi^{[a,b]}} C_k(\varphi, u) \text{ for all } k \in \mathbb{N}_0.
\end{aligned}$$

The proof is now complete. □

We can make the above result more precise, with the so-called "Morse relation". The next definition introduces some algebraic quantities which are important in this direction.

Definition 6.2.18 Let X be a Banach space, $\varphi \in C^1(X)$, $a, b \in \mathbb{R} \backslash \varphi(K_\varphi)$, $a < b$, and suppose that $\varphi^{-1}((a, b))$ contains a finite number of critical points $\{u_i\}_{i=1}^n$.

(a) The "Morse-type numbers" of φ for (a, b) are defined by

$$M_k(a, b) = \sum_{i=1}^{n} \operatorname{rank} C_k(\varphi, u_i) \text{ for all } k \in \mathbb{N}_0.$$

Suppose that $M_k(a, b)$ is finite for every $k \in \mathbb{N}_0$ and vanishes for all large $k \in \mathbb{N}_0$. We define

$$M(a, b)(t) = \sum_{k \geqslant 0} M_k(a, b) t^k \text{ for all } t \in \mathbb{R}.$$

Then $M(a, b)(\cdot)$ is called the "Morse polynomial" of φ for (a, b).

(b) The "Betti-type numbers" of φ for (a, b) are defined by

$$\beta_k(a, b) = \operatorname{rank} H_k(\varphi^b, \varphi^a) \text{ for all } k \in \mathbb{N}_0.$$

Suppose that $\beta_k(a, b)$ is finite for all $k \in \mathbb{N}_0$ and vanishes for all large $k \in \mathbb{N}_0$.

We define

$$P(a, b)(t) = \sum_{k \geqslant 0} \beta_k(a, b) t^k \text{ for all } t \in \mathbb{R}.$$

Then $P(a, b)(\cdot)$ is called the "Poincaré polynomial" of φ for (a, b).

To prove the "Morse relation", we will need the following simple lemma.

Lemma 6.2.19 *If $D_0 \subseteq D_1 \subseteq \ldots \subseteq D_n (n \geqslant 2)$ are Hausdorff topological spaces and* rank $H_k(D_i, D_{i-1})$ *is finite for all $k \in \mathbb{N}_0$ and all $i \in \{1, \ldots, n\}$, and vanishes for all large $k \in \mathbb{N}_0$, then*

$$\sum_{k \geqslant 0} \Big(\sum_{i=1}^{n} \text{rank } H_k(D_i, D_{i-1})\Big) t^k = \sum_{k \geqslant 0} \text{rank } H_k(D_n, D_0) t^k + (1+t) Q(t),$$

where $Q(t)$ is a polynomial with nonnegative integer coefficients.

Proof We prove the statement for $n = 2$, the general case following by induction. For the triple (A_2, A_1, A_0) we consider the corresponding long exact sequence. We have

$$\begin{aligned} \ldots \to H_{k+1}(A_2, A_1) \xrightarrow{\partial_k} H_k(A_1, A_0) \xrightarrow{i_*} H_k(A_2, A_0) \xrightarrow{j_*} \\ H_k(A_2, A_1) \xrightarrow{\partial_{k-1}} H_{k-1}(A_1, A_0) \to \ldots . \end{aligned} \tag{6.44}$$

From Proposition 6.1.36 we have

$$\text{rank } H_k(A_2, A_0) \leqslant \text{rank } H_k(A_1, A_0) + \text{rank } H_k(A_2, A_1). \tag{6.45}$$

Let $r_k = \text{rank im } \partial_k$. From (6.45) and the exactness of (6.44), we have

$$\begin{aligned} &\text{rank } H_k(A_2, A_0) + r_k + r_{k-1} \\ &= (r_k + \text{rank im } i_*) + (t_{k-1} + \text{rank im } j_*) \\ &= (\text{rank ker } \ i_* + \text{rank im } i_*) + (r_{k-1} + \text{rank ker } \partial_{k-1}) \\ &= \text{rank } H_k(A_1, A_0) + \text{rank } H_k(A_2, A_1) \text{ (by the rank theorem).} \end{aligned}$$

Evidently, $Q(t) = \sum_{k \geqslant 0} r_k t^k$, $t \in \mathbb{R}$, is the desired polynomial. ☐

The next theorem establishes the so-called "Morse relation".

Theorem 6.2.20 *If X is a Banach space, $\varphi \in C^1(X)$, $a, b \in \mathbb{R} \backslash \varphi(\{K_\varphi)$, $a < b$, $\varphi^{-1}((a, b))$ contains a finite number of critical points $\{u_i\}_{i=1}^n$ and φ satisfies the C_c-condition for every $c \in [a, b)$, then*

(a) for all $k \in \mathbb{N}_0$, we have $M_k(a, b) \geqslant \beta_k(a, b)$;
(b) if the Morse-type numbers $M_k(a, b)$ are finite for all $k \in \mathbb{N}_0$ and vanish for all large $k \in \mathbb{N}_0$, then so do the Betti numbers $\beta_k(a, b)$ and we have

$$\sum_{k \geqslant 0} M_k(a, b) t^k = \sum_{k \geqslant 0} \beta_k(a, b) t^k + (1+t) Q(t) \text{ for all } t \in \mathbb{R},$$

where $Q(t)$ is a polynomial in $t \in \mathbb{R}$ with nonnegative integer coefficients.

Proof (a) Let $c_k = \varphi(u_k)$ for all $k \in \{1, \dots, n\}$ and pick $\{\vartheta_k\}_{k=0}^n \subseteq [a, b] \setminus \varphi(K_\varphi)$ such that

$$a = \vartheta_0 < c_1 < \vartheta_1 < \cdots < \vartheta_{i-1} < c_i < \vartheta_i < \cdots < c_n < \vartheta_n = b.$$

Then from Definition 6.2.18 and Propositions 6.2.16 and 6.2.17, we have

$$\beta_k(a, b) \leqslant \sum_{\mathrm{i}=1}^{n} \beta_k(\vartheta_{i-1}, \vartheta_i) = \sum_{\mathrm{i}=1}^{n} M_k(\vartheta_{i-1}, \vartheta_i) = M_k(a, b) \text{ for all } k \in \mathbb{N}_0. \quad (6.46)$$

(b) If $M_k(a, b)$ is finite for all $k \in \mathbb{N}_0$ and vanishes for large $k \in \mathbb{N}_0$, then from (6.46) it is clear that so do the Betti numbers $\beta_k(a, b)$, $\beta_k(\vartheta_{i-1}, \vartheta_i)$. Then using Lemma 6.2.19 we have

$$\sum_{\mathrm{k}\geqslant 0} \left(\sum_{\mathrm{i}=1}^{n} \beta_k(\vartheta_{i-1}, \vartheta_i) \right) t^k = \sum_{\mathrm{k}\geqslant 0} \beta_k(a, b) t^k + (1 + t) Q(t), \quad (6.47)$$

where $Q(t)$ is a polynomial in $t \in \mathbb{R}$ with nonnegative integer coefficients. From (6.46) and (6.47) we conclude that

$$\sum_{\mathrm{k}\geqslant 0} M_k(a, b) t^k = \sum_{\mathrm{k}\geqslant 0} \beta_k(a, b) t^k + (1 + t) Q(t) \text{ for all } t \in \mathbb{R}. \quad (6.48)$$

The proof is now complete. □

Remark 6.2.21 If in (6.48) we choose $t = -1$, then

$$\sum_{\mathrm{k}\geqslant 0} (-1)^k M_k(a, b) = \sum_{\mathrm{k}\geqslant 0} (-1)^k \beta_k(a, b)$$

and this equality is known as the “Poincaré–Hopf formula”.

When the functional $\varphi \in C^1(X)$ has critical values which are bounded from below and satisfy the C-condition, then the global behavior of φ can be described by the critical groups of φ at infinity.

Definition 6.2.22 Let $\varphi \in C^1(X)$ and assume that φ satisfies the C-condition and $\inf \varphi(K_\varphi) > -\infty$. The “critical groups of φ at infinity” are defined by

$$C_k(\varphi, \infty) = H_k(X, \varphi^c) \text{ for all } k \in \mathbb{N}_0,$$

with $c < \inf \varphi(K_\varphi)$.

Remark 6.2.23 Corollary 5.3.13 reveals that the above definition is independent of the choice of the level $c < \inf \varphi(K_\varphi)$. Indeed, if $d < c < \inf \varphi(K_\varphi)$, then from

Corollary 5.3.13 we know that φ^d is a strong deformation retract of φ^c. So, Corollary 6.1.24 (a) implies that $H_k(X, \varphi^c) = H_k(X, \varphi^d)$ for all $k \in \mathbb{N}_0$.

Proposition 6.2.24 *If X is a Banach space, $\varphi \in C^1(X)$ and φ satisfies the C-condition then*

(a) for $\varphi(\cdot)$ bounded from below, we have

$$C_k(\varphi, \infty) = \delta_{k,0}\mathbb{Z} \text{ for all } k \in \mathbb{N}_0;$$

(b) for $\varphi(\cdot)$ unbounded from below and $\inf \varphi(K_\varphi) > -\infty$, we have $C_k(\varphi, \infty) = \tilde{H}_{k-1}(\varphi^c)$ for all $k \in \mathbb{N}_0$ and all $c < \inf \varphi(K_\varphi)$.

Proof *(a)* Let $c < \inf \varphi(X)$. Then $\varphi^c = \emptyset$ and so by Definition 6.2.22 we have

$$C_k(\varphi, \infty) = H_k(X, \varphi^c) = H_k(X) = \delta_{k,0}\mathbb{Z} \text{ for all } k \in \mathbb{N}_0$$

(see Definition 6.1.12, Axiom 7 and Remark 6.1.13).

(b) From Definition 6.2.22 and since X is contractible, we see that the reduced homology groups of X are trivial for all $k \in \mathbb{N}_0$. We consider the following long exact sequence

$$\ldots \to \tilde{H}_k(X) \to H_k(X, \varphi^c) \to \tilde{H}_{k-1}(\varphi^c) \to \tilde{H}_{k-1}(X) \to \ldots \tag{6.49}$$

From the exactness of (6.49) and since $\tilde{H}_k(X) = \tilde{H}_{k-1}(X)$ for all $k \in \mathbb{N}_0$, we infer that

$$\begin{aligned} &H_k(X, \varphi^c) = \tilde{H}_{k-1}(\varphi^c) \\ \Rightarrow\ &C_k(\varphi, \infty) = \tilde{H}_{k-1}(\varphi^c) \text{ for all } k \in \mathbb{N}_0. \end{aligned}$$

The proof is now complete. □

Remark 6.2.25 In particular, in the setting of part (b), we have $C_0(\varphi, \infty) = 0$.

Proposition 6.2.26 *If X is a Banach space, $\varphi \in C^1(X)$, φ satisfies the C-condition and K_φ is finite, then* $\operatorname{rank} C_k(\varphi, \infty) \leqslant \sum_{u \in K_\varphi} \operatorname{rank} C_k(\varphi, u)$ *for all $k \in \mathbb{N}_0$.*

Proof This proposition is an immediate consequence of Proposition 6.2.17 with $b > \sup \varphi(K_\varphi)$ and $a < \inf \varphi(K_\varphi)$. □

Corollary 6.2.27 *If X is a Banach space, $\varphi \in C^1(X)$, φ satisfies the C-condition, K_φ is finite and $C_{k_0}(\varphi, \infty) \neq 0$ for some $k_0 \in \mathbb{N}_0$, then there exists a $u \in K_\varphi$ such that $C_{k_0}(\varphi, u) \neq 0$.*

Proposition 6.2.28 *If X is a Banach space, $\varphi \in C^1(X)$ and satisfies the C-condition,* $\inf \varphi(K_\varphi) > -\infty$*, then*

(a) *for* $-\infty < a < \inf \varphi(K_\varphi) \leqslant \sup \varphi(K_\varphi) < b$, *we have*

$$C_k(\varphi, \infty) = H_k(\varphi^b, \varphi^a) \text{ for all } k \in \mathbb{N}_0;$$

(b) $C_k(\varphi, \infty) = 0$ *for all* $k \in \mathbb{N}_0$ *when* $K_\varphi = \emptyset$;
(c) $C_k(\varphi, \infty) = C_k(\varphi, u)$ *for all* $k \in \mathbb{N}_0$ *when* $K_\varphi = \{u\}$.

Proof (a) This is a consequence of Proposition 6.2.16.
(b) Follows from part (a) with $a = b$ (see Corollary 6.1.16).
(c) Follows from part (a) and Proposition 6.2.16.

□

Also, from Theorem 6.2.20(b) and Proposition 6.2.28(a), we infer a global version of the Morse relation.

Theorem 6.2.29 *If* X *is a Banach space,* $\varphi \in C^1(X)$, φ *satisfies the* C*-condition,* K_φ *is finite,* $C_k(\varphi, u)$ *has a finite rank for all* $k \in \mathbb{N}_0$ *and all* $u \in K_\varphi$, *and vanishes for large* $k \in \mathbb{N}_0$, *then*

$$\sum_{u \in K_\cdot} \left(\sum_{k \geqslant 0} \operatorname{rank} C_k(\varphi, u) t^k \right) = \sum_{k \geqslant 0} C_k(\varphi, \infty) + (1 + t) Q(t) \ \ for\ all\ t \in \mathbb{R}$$

with $Q(t)$ *a polynomial with nonnegative integer coefficients.*

Next, we discuss the critical groups at infinity in more detail. First we consider functionals which exhibit some kind of local linking at infinity.

Proposition 6.2.30 *If* X *is a Banach space with* $X = Y \oplus V$ *with* $\dim Y < \infty$, $\varphi \in C^1(X)$, φ *satisfies the* C*-condition,* $\inf \varphi(K_\varphi) > -\infty$, $\varphi|_V$ *is bounded from below and* $\varphi|_Y$ *is anticoercive (that is, if* $y \in Y$, $||y|| \to \infty$, *then* $\varphi(y) \to -\infty$*), then* $C_d(\varphi, \infty) \neq 0$ *with* $d = \dim Y$.

Proof Let $c < \min\{\inf \varphi|_V, \inf \varphi(K_\varphi)\}$. Since by hypothesis $\varphi|_Y$ is anticoercive, we can find large $r > 0$ such that

$$\partial B_r^Y = \{y \in Y : ||y|| = r\} \subseteq \varphi^c.$$

So, we have

$$\partial B_r^Y \subseteq \varphi^c \subseteq X \setminus V \subseteq X.$$

Consider the deformation $h : [0, 1] \times (X \setminus V) \to X \setminus V$ defined by

$$h(t, u - v) = (1 - t)(u - v) + t\rho \frac{u - v}{||u - v||} \text{ for all } t \in [0, 1],\ \text{all } u \in X,\ v \in V.$$

It follows that ∂B_ρ^Y is a strong deformation retract of $X \setminus V$. Hence

$$H_k(X \setminus V, \partial B_\rho^Y) = 0 \text{ for all } k \in \mathbb{N}_0 \tag{6.50}$$

(see Proposition 6.1.15). We consider the following commutative diagram

$$\begin{tikzcd} H_k(\varphi^c, \partial B_\rho^Y) \arrow[r, "i_*"] \arrow[dr, "\vartheta_*"'] & H_k(X, \partial B_\rho^Y) \arrow[r, "j_*"] & H_k(X, \varphi^c) = C_k(\varphi, \infty) \\ & H_k(X \setminus V, \partial B_\rho^Y) \arrow[u, "\eta_*"] & \end{tikzcd} \tag{6.51}$$

with i_*, j_*, ϑ_*, η_* being the group homomorphisms induced by the corresponding inclusion maps. In (6.51) the top row is exact (see Proposition 6.1.23). We have $i_* = \eta_* \circ \vartheta_*$ and from (6.50) we see that $i_* = 0$. The exactness of the top row implies that j_* is injective for all $k \in \mathbb{N}_0$. From the reduced exact homology sequence (see Proposition 6.1.29) we have

$$H_d(X, \partial B_\rho^Y) = H_{d-1}(\partial B_\rho^Y, *) \text{ with } d = \dim Y. \tag{6.52}$$

Since Y is finite-dimensional, we have

$$\begin{aligned} & H_{d-1}(\partial B_\rho^Y, *) = \mathbb{Z} \text{ (see Example 6.1.34(c))} \\ \Rightarrow\ & H_d(X, \partial B_\rho^Y) = \mathbb{Z} \text{ (see (6.52))} \\ \Rightarrow\ & C_d(\varphi, \infty) \neq 0 \text{ (since } j_* \text{ is injective).} \end{aligned}$$

The proof is now complete. □

The next two results provide some further information about critical groups at infinity in the context of Hilbert spaces. The proofs of these results can be found in Bartsch and Li [38].

So, let H be a Hilbert space and $\varphi \in C^1(H)$. We introduce the following condition on φ.

(A_∞) $\varphi(u) = \frac{1}{2}(A(u), u)_H + \psi(u)$ for all $u \in H$, with $A \in \mathscr{L}(H, H)$ self adjoint, 0 is isolated in the spectrum of A, $\psi \in C^1(H)$, $\lim\limits_{||u|| \to \infty} \frac{\psi(u)}{||u||^2} = 0$ (subquadratic), ψ and ψ' are bounded (that is, map bounded sets to bounded sets) and φ is bounded from below and satisfies the C-condition.

Remark 6.2.31 If A_∞ holds, then we set

$$Y = \ker A \text{ and } V = Y^\perp.$$

The space V admits an orthogonal direct sum decomposition

$$V = V_- \oplus V_+$$

with V_+, V_- being A-invariant, $A|_{V_-} < 0$ and $A|_{V_+} > 0$. So, we can find $c_0 > 0$ such that

$$\pm\frac{1}{2}(A(u), u)_H \geqslant c_0||u||^2 \text{ for all } u \in V\pm.$$

Let $m = \dim V_-$ [the Morse index of φ at infinity, compare with Proposition 3.4.18(b)] and $\nu = \dim Y$ (known as the nullity of φ at infinity).

The results of Bartsch–Li [38] mentioned earlier read as follows.

Theorem 6.2.32 *If H is a Hilbert space and $\varphi \in C^1(H)$ satisfies condition (A_∞) above, then $C_k(\varphi, \infty) = 0$ for all $k \notin \{m, m+1, \ldots, m+\nu\}$.*

Remark 6.2.33 In this theorem we do not require that m, ν are finite. If $m \in \mathbb{N}$ and $\nu = 0$, then $C_k(\varphi, \infty) = \delta_{k,m}\mathbb{Z}$ for all $k \in \mathbb{N}_0$.

Imposing the so-called angle conditions on φ, we derive more information concerning the critical groups at infinity.

Theorem 6.2.34 *If H is a Hilbert space, $\varphi \in C^1(H)$, φ satisfies (A_∞) and $m, \nu \in \mathbb{N}$, then*

(a) $C_k(\varphi, \infty) = \delta_{k,m}\mathbb{Z}$ *for all $k \in \mathbb{N}_0$, when φ satisfies the following "angle condition":*

$$(A_\infty^+)\text{“there exist } M > 0 \text{ and } \vartheta \in (0, 1) \text{ such that}$$

$$\langle\varphi'(u), y\rangle \geqslant 0 \text{ for all } u = y + v \in H,\ \ y \in Y, v \in V$$

$$\text{with } ||u|| \geqslant M \text{ and } ||v|| \leqslant \vartheta||u||''.$$

(b) $C_k(\varphi, \infty) = \delta_{k,m+\nu}\mathbb{Z}$ *for all $k \in \mathbb{N}_0$, when φ satisfies the following angle condition*

$$(A_\infty^-)\text{“there exist } M > 0 \text{ and } \vartheta \in (0, 1) \text{ such that}$$

$$-\langle\varphi'(u), y\rangle \geqslant 0 \text{ for all } u = y + v \in H, y \in Y, v \in V$$

$$\text{with } ||u|| \geqslant M \text{ and } ||v|| \leqslant a||u||.''$$

We will derive similar information for the critical groups at an isolated critical point. To this end, we prove three auxiliary results.

Lemma 6.2.35 *If X is a reflexive Banach space, $\varphi \in C^1(X)$, φ satisfies the C-condition and $u_0 \in K_\varphi$ is isolated with $c = \varphi(u_0)$ isolated in $\varphi(K_\varphi)$, then there exist $\psi \in C^1(X)$, $U \subseteq X$ open with $u_0 \in U$ and $\delta > 0$ such that*

(a) ψ *satisfies the C-condition;*
(b) $\varphi \leqslant \psi$ *and* $\varphi|_U = \psi|_U$;
(c) $K_\varphi = K_\psi$;
(d) $K_\psi \cap \psi^{-1}([c-\delta, c+\delta]) = \{u_0\}$;

(e) if $X = H =$ a Hilbert space and $\varphi \in C^\rho(H)$ with $\rho \geqslant 2$, then we have $\psi \in C^\rho(H)$ too.

Proof Thanks to the Troyanski renorming theorem (see Theorem 2.7.36), we may assume that X and X^* are locally uniformly convex with Fréchet differentiable norms (except at the origin). Then the map $h : X \to \mathbb{R}_+$ defined by

$$h(u) = \frac{1}{2}||u||^2 \text{ for all } u \in X$$

is of class C^1 and we have

$$h'(u) = J(u) \text{ for all } u \in X,$$

with $J : X \to X^*$ being the duality map (see Definition 2.7.21 and Proposition 2.7.33). Then given $0 < \rho_1 < \rho_2$ such that $\bar{B}_{\rho_2}(u_0) \cap K_\varphi = \{u_0\}$ and φ, φ' restricted to $\bar{B}_{\rho_2}(u_0)$ are bounded, we can find $\eta \in C^1(X)$ such that

$$\eta(u) = \begin{cases} 0 \text{ if } ||u|| \leqslant \rho_1 \\ 1 \text{ if } ||u|| \geqslant \rho_2 \end{cases}, 0 \leqslant \eta \leqslant 1, M = \sup_{u \in X} ||\eta'(u)||_* < \infty. \tag{6.53}$$

The existence of such a function is easily seen if we recall that the smoothness of X implies the existence of a C^1-bump function (recall that a bump function on X is a function on X with nonempty bounded support).

Let $U = B_{\rho_1}(u_0)$. Because φ satisfies the C-condition, we can find $\gamma > 0$ such that

$$\gamma \leqslant ||\varphi'(u)||_* \text{ for all } u \in X \text{ with } \rho_1 \leqslant ||u|| \leqslant \rho_2. \tag{6.54}$$

Recall that $c = \varphi(u_0)$ is isolated in the critical values $\varphi(K_\varphi)$. Then we can find $c_0 \in \left(c - \frac{\gamma}{2M}, c\right)$ and $\delta > 0$ such that

$$[c_0 - \delta, c_0 + \delta] \subseteq \mathbb{R} \backslash \varphi(K_\varphi)$$

(that is, $[c_0 - \delta, c_0 + \delta]$ is a regular interval). We set

$$\psi(u) = \varphi(u) + (c - c_0)\eta(u) \text{ for all } u \in X. \tag{6.55}$$

Evidently, $\psi \in C^1(X)$. We claim that (ψ, U, δ) as above is the desired triple postulated by the lemma

(a) Suppose that $\{u_n\}_{n \geqslant 1} \subseteq X$ is a sequence such that

$$\{\psi(u_n)\}_{n \geqslant 1} \subseteq \mathbb{R} \text{ is bounded and } (1 + ||u_n||)\varphi'(u_n) \to 0 \text{ in } X^* \text{ as } n \to \infty. \tag{6.56}$$

We have

$$\begin{aligned}||\psi'(u)||_* &\geqslant ||\varphi'(u)||_* - (c-c_0)||\eta'(u)||_* \text{ (see (6.55))}\\ &\geqslant \gamma - \frac{\gamma}{2M}M = \frac{\gamma}{2} \text{ if } \rho_1 \leqslant ||u|| \leqslant \rho_2 \text{ see (6.54).}\end{aligned} \tag{6.57}$$

From (6.56) and (6.57), we see that we can find $n_0 \in \mathbb{N}$ such that

$$\begin{aligned}&||u_n| \notin [\rho_1, \rho_2] \text{ for all } n \geqslant n_0\\ \Rightarrow\ &\varphi'(u_n) = \psi'(u_n) \text{ for all } n \geqslant n_0 \text{ (see (6.53) and (6.55)).}\end{aligned} \tag{6.58}$$

From (6.55) we see that $\varphi \leqslant \psi$. So, because of (6.56) we infer that

$$\{\varphi(u_n)\}_{n\geqslant 1} \subseteq \mathbb{R} \text{ is bounded.} \tag{6.59}$$

From (6.56), (6.58), (6.59) and since φ satisfies the C-condition, we conclude that $\{u_n\}_{n\geqslant 1} \subseteq X$ admits a strongly convergent subsequence. Therefore ψ satisfies the C-condition.

(b) This follows at once from (6.53) and (6.55).

(c) Recall that

$$\begin{aligned}&\varphi'(u) = \psi'(u) \text{ when } ||u|| \notin [\rho_1, \rho_2]\\ \text{and }\ &\varphi'(u) \neq 0, \psi'(u) \neq 0 \text{ when } \rho_1 \leqslant ||u|| \leqslant \rho_2\end{aligned}$$

[see (6.53), (6.55) and (6.57)]. Therefore we see that

$$K_\varphi = K_\psi.$$

(d) Let $u \in K_\varphi \backslash \{u_0\}$. Then from part (c) we have $u \in K_\varphi$. Recalling the choice of ρ_2, we see that $||u|| > \rho_2$. Because of (6.53) we have

$$\psi(u) = \varphi(u) + (c - c_0). \tag{6.60}$$

From the choices of c_0 and δ it follows that

$$\begin{aligned}&\varphi(u) \notin [c_0 - \delta, c_0 + \delta]\\ \Rightarrow\ &\psi(u) \notin [c - \delta, c + \delta] \text{ (see (6.60)).}\end{aligned}$$

(e) Since $h(u) = \frac{1}{2}||u||^2$ for all $u \in X = H$ is C^∞, we have $\eta \in C^\infty(H)$. So, if $\varphi \in C^\rho(H)$, $\rho \geqslant 2$, then $\psi \in C^\rho(H)$ [see (6.55)]. □

Remark 6.2.36 For every $\vartheta > 0$, the set $(c - \vartheta, c)\backslash\varphi(K_\varphi)$ is open. So, in Lemma 6.2.35 we can replace the hypothesis that $c = \varphi(u_0)$ is isolated in $\varphi(K_\varphi)$ by a weaker one which says that there is a sequence $\{c_n\}_{n\geqslant 1} \subseteq \mathbb{R}\backslash\varphi(K_\varphi)$ with $c_n < c$ and $c_n \to c$.

Lemma 6.2.37 *If $\varphi \in C^2(\mathbb{R}^N)$, $U \subseteq \mathbb{R}^N$ is open and bounded and $K \subseteq U$ is compact such that $K_\varphi \cap \overline{(U \setminus K)} = \emptyset$, then for every $\epsilon > 0$, we can find $\psi \in C^2(\mathbb{R}^N)$ such that*

(a) $|\varphi(u) - \psi(u)| + ||\varphi'(u) - \psi'(u)||_* \leqslant \epsilon$ *for all $u \in \mathbb{R}^N$;*
(b) $\varphi|_{\mathbb{R}^N \setminus U} = \psi|_{\mathbb{R}^N \setminus U}$*;*
(c) $K_\psi \cap \bar{U}$ *is finite and all its elements are nondegenerate critical points.*

Proof By hypothesis, we have

$$0 < \gamma = \inf\{||\varphi'(u)||_* : u \in \overline{U \setminus K}\}.$$

Choose $\eta \in C^\infty(\mathbb{R}^N)$ such that

$$\eta(u) = \begin{cases} 1 \text{ if } u \in K \\ 0 \text{ if } u \notin K. \end{cases} \tag{6.61}$$

Also, let $\rho > 0$ and $\vartheta > 0$ such that

$$u \subseteq B_\rho(0), \ \vartheta\rho\, ||\eta||_\infty \leqslant \frac{\epsilon}{2} \text{ and } \vartheta\, ||\eta||_\infty + \vartheta_\rho ||\eta'||_\infty \leqslant \frac{1}{2} \min\{\epsilon, 1\}.$$

By Sard's theorem (see Theorem 3.1.16), we can find $e \in \mathbb{R}^N$ such that

$$|e| \leqslant \vartheta \text{ and } -e \text{ is not a critical value of } \varphi'$$

(that is, $\varphi''(u)$ is nondegenerate whenever $\varphi'(u) = -e$). We consider $\psi \in C^2(\mathbb{R}^N)$ defined by

$$\psi(u) = \varphi(u) + \eta(u)(u, e)_{\mathbb{R}^N} \text{ for all } u \in \mathbb{R}^N.$$

We have

$$\psi'(u) = \varphi'(u) + \eta'(u)(u, e)_{\mathbb{R}^N} + \eta(u)e.$$

From the choice of η and e, we see that ψ defined above satisfies statements (a) and (b) of the lemma. Also, we have

$$\frac{\gamma}{2} \leqslant \inf\{||\psi'(u)||_* : u \in \overline{U \setminus K}\}. \tag{6.62}$$

Let $u \in K_\psi \cap \bar{U}$. Then from (6.62) it follows that $u \in \operatorname{int} K$. Hence from (6.61) it follows that

$$0 = \psi'(u) = \varphi'(u) + e.$$

Since e is not a critical value of φ', we infer that $\varphi''(u) = \psi''(u)$ is invertible. Therefore the elements of $K_\psi \cap \bar{U}$ are nondegenerate, hence isolated (by the inverse function theorem) and located in the compact set K. Therefore $K_\psi \cap \bar{U}$ is also finite. This proves part (c). $\square$

Lemma 6.2.38 *If $\varphi \in C^2(\mathbb{R}^N)$, $u_0 \in K_\varphi$ is isolated and $c = \varphi(u_0)$, then we can find $\psi \in C^2(\mathbb{R}^N)$ such that*

(a) $\varphi = \psi$ is a neighborhood of u_0;
(b) K_ψ is finite;
(c) $K_\psi^c = \{u_0\}$;
(d) ψ is coercive (hence it satisfies the C-condition).

Proof By modifying φ outside a ball centered at u_0, if necessary, we may assume that φ is coercive and there exists a $\rho > 0$ such that $K_\varphi \cap (\mathbb{R}^N \backslash B_\rho(u_0)) = \emptyset$. Let $r \in (0, \rho)$ be such that $K_\varphi \cap \overline{B_r(u_0)} = \{u_0\}$. Let $\hat{\psi} \in C^2(\mathbb{R}^N)$ be the function obtained in Lemma 6.2.37, with $U = B_{2\rho}(u_0)\backslash\overline{B_{\frac{r}{2}}(u_0)}$, $K = \bar{B}_\rho(u_0)\backslash B_r(u_0)$ and any $\epsilon > 0$. Evidently, $\hat{\psi}$ satisfies parts (a),(b),(c) of the lemma. Finally apply Lemma 6.2.35 to $\hat{\psi}$ and denote by ψ_0 the function we obtain in this way. Then ψ_0 satisfies (a)–(d) in the lemma. □

We will use these lemmata to derive some useful consequences concerning the critical groups of isolated critical points for C^2-functions.

Proposition 6.2.39 *If $\varphi \in C^2(\mathbb{R}^N)$ and $u_0 \in K_\varphi$ is isolated, then* rank $C_k(\varphi, u_0) < \infty$ *for all $k \in \mathbb{N}_0$ and $C_k(\varphi, u_0) = 0$ for all $k \notin \{0, 1 \ldots, N\}$.*

Proof From Definition 6.2.1 and Remark 6.2.2, we know that the critical groups $C_k(\varphi, u_0)$, $k \in \mathbb{N}_0$, depend only the local structure of φ. So, using Lemma 6.2.38 we see that without any loss of generality, we may assume that φ is coercive (hence it satisfies the C-condition), K_φ is finite and $K_\varphi^{c_0} = \{u_0\}$, where $c_0 = \varphi(u_0)$. Let $a, b \in \mathbb{R}$ such that $a < c_0 = \varphi(u_0) < b$ and $K_\varphi \cap \varphi^{-1}([a, b]) = \{u_0\}$. Invoking Proposition 6.2.16, we have

$$C_k(\varphi, u_0) = H_k(\varphi^b, \varphi^a) \text{ for all } k \in \mathbb{N}_0. \tag{6.63}$$

Let $r > 0$ be such that

$$\overline{B_r(u_0)} \subseteq \{u \in \mathbb{R}^N : a < \varphi(u) < b\}.$$

Let $U = B_r(u_0)$ and $K = \overline{B_{r/2}(u_0)}$. Then

$$a < c = \inf_U \varphi \text{ and } m = \sup_U \varphi < b.$$

Pick $\epsilon > 0$ such that

$$\epsilon < \min\{c - a, b - d\}$$

and let $\psi \in C^2(\mathbb{R}^N)$ be as postulated by Lemma 6.2.37 for the aforementioned choices of U, K and $\epsilon > 0$. Then from Lemma 6.2.37, we have

$$\psi^b = \varphi^b \text{ and } \psi^a = \varphi^a. \tag{6.64}$$

Then from (6.63), (6.64) and Theorem 6.2.20 (the Morse relation), we have

$$\operatorname{rank} C_k(\varphi, u_0) = \operatorname{rank} H_k(\psi^b, \psi^a) \leqslant \sum_{i=1}^{n} \operatorname{rank} C_k(\psi, u_i) \text{ for all } k \in \mathbb{N}_0,$$

with $\{u_i\}_{i=1}^n = K_\psi \cap U$. Each u_i is a nondegenerate critical point for ψ and so from Proposition 6.2.6, we obtain

$$\begin{aligned} &\operatorname{rank} C_k(\psi, u_i) \in \{0, 1\} \text{ for all } k \in \mathbb{N}_0, \ \text{all } i \in \{1, \ldots, N\}, \\ &\operatorname{rank} C_k(\psi, u_i) = 0 \text{ for all } k \in \mathbb{N}_0, \ \text{all } i \notin \{1, \ldots, N\}. \end{aligned}$$

This proves the proposition. □

As a consequence of Proposition 6.2.39 and of Theorem 6.2.13 (the shifting theorem), we have:

Corollary 6.2.40 *If H is a Hilbert space, $\varphi \in C^2(H)$ and $u_0 \in K_\varphi$ is isolated with finite Morse index m and $\nu = \operatorname{dimker} \varphi''(u_0) < +\infty$, then $\operatorname{rank} C_k(\varphi, u_0)$ is finite for all $k \in \mathbb{N}_0$ and $C_k(\varphi, u_0) = 0$ for all $k \notin \{m, \ldots, m + \nu\}$.*

The next proposition is useful in obtaining nontrivial critical points with a nontrivial critical group.

Proposition 6.2.41 *If X is a Banach space, $\varphi \in C^1(X)$, φ satisfies the C-condition, K_φ is finite with $0 \in K_\varphi$ and for some $k \in \mathbb{N}_0$ we have*

$$C_k(\varphi, 0) = 0 \textit{ and } C_k(\varphi, \infty) \neq 0,$$

then there exists a $u \in K_\varphi \backslash \{0\}$ such that $C_k(\varphi, u) \neq 0$.

Proof From Corollary 6.2.27 we know that there exists a $u \in K_\varphi$ such that

$$C_k(\varphi, u) \neq 0. \tag{6.65}$$

On the other hand, by hypothesis,

$$C_k(\varphi, 0) = 0. \tag{6.66}$$

Comparing (6.65) and (6.66), we conclude that $u \neq 0$. □

Proposition 6.2.42 *If X is a Banach space, $\varphi \in C^1(X)$, φ satisfies the C-condition, K_φ is finite with $0 \in K_\varphi$ and for some $k \in \mathbb{N}_0$ we have*

$$C_k(\varphi, 0) \neq 0 \textit{ and } C_k(\varphi, \infty) = 0,$$

then there exists a $u \in K_\varphi$ such that

$$\varphi(u) < 0 \textit{ and } C_{k-1}(\varphi, u) \neq 0$$
$$\textit{or } \; \varphi(u) > 0 \textit{ and } C_{k+1}(\varphi, u) \neq 0.$$

Proof Without any loss of generality we may assume that $\varphi(0) = 0$. Choose $\epsilon > 0$ small such that $\varphi(K_\varphi) \cap [-\epsilon, \epsilon] = \{0\}$. Let $c < \min\left\{-\epsilon, \inf \varphi(K_\varphi)\right\}$. From Proposition 6.2.17 we have

$$\operatorname{rank} C_k(\varphi, 0) \leqslant \operatorname{rank} H_k(\varphi^\epsilon, \varphi^{-\epsilon}), \tag{6.67}$$

while from Definition 6.2.22 and the choice of c we have

$$H_k(X, \varphi^c) = C_k(\varphi, \infty). \tag{6.68}$$

We consider the sets $\varphi^c \subseteq \varphi^{-\epsilon} \subseteq \varphi^\epsilon \subseteq X$ and use Proposition 6.1.37. Then

$$\begin{aligned} 0 < \operatorname{rank} C_k(\varphi, 0) - \operatorname{rank} C_k(\varphi, \infty) \leqslant \operatorname{rank} H_{k-1}(\varphi^{-\epsilon}, \varphi^c) + \\ \operatorname{rank} H_{k+1}(X, \varphi^\epsilon) \text{ see (6.67), (6.68)} \\ \Rightarrow H_{k-1}(\varphi^{-\epsilon}, \varphi^c) \neq 0 \text{ or } H_{k+1}(X, \varphi^\epsilon) \neq 0. \end{aligned}$$

In the first case, by Proposition 6.2.15 we can find $u \in K_\varphi$ such that

$$\varphi(u) < 0 \text{ and } C_{k-1}(\varphi, u) \neq 0.$$

Similarly, in the second case, we can find $\tilde{u} \in K_\varphi$ such that

$$\varphi(\tilde{u}) > 0 \text{ and } C_{k+1}(\varphi, \tilde{u}) \neq 0.$$

The proof is now complete. □

We conclude this section by relating critical groups with the Leray–Schauder degree. We start with the definition of the Leray–Schauder index.

Definition 6.2.43 Let X be a Banach space, $\varphi = i - f : X \to X$ with f compact and $u_0 \in X$ be an isolated solution of the equation $\varphi(u) = 0$. Let $r > 0$ be such that u_0 is the only solution of the equation in $\bar{B}_r(u_0)$. The "Leray–Schauder index of φ at u_0" is defined by

$$i_{LS}(\varphi, u_0) = d_{LS}(\varphi, B_r(u_0), 0).$$

Suppose that $X = H =$ a Hilbert space, $\varphi \in C^1(H)$ and $\nabla\varphi = i - f$ with $f : H \to H$ compact (here by $\nabla\varphi(\cdot)$ we denote the gradient of φ). Note that both $C_k(\varphi, u_0)$ and $i_{LS}(\nabla\varphi, u_0)$ are topological invariants describing the local behavior at an isolated critical point $u_0 \in K_\varphi$. So, it is reasonable to expect that the two quantities are related. The precise relation is given in the next proposition, the proof of which can be found in Chang [119] (Theorem 3.2, p. 100).

Proposition 6.2.44 *If H is a Hilbert space, $\varphi \in C^2(H)$, φ satisfies the C-condition, $\nabla\varphi = i - f$ with $f : H \to H$ a compact map and $u_0 \in K_\varphi$ is isolated, then* $i_{LS}(\nabla\varphi, u_0) = \sum\limits_{k \in \mathbb{N}_0} (-1)^k \operatorname{rank} C_k(\varphi, u_0)$.

Remark 6.2.45 This proposition reveals that for potential compact vector fields, the critical groups provide more information than the Leray–Schauder index.

6.3 Continuity and Homotopy Invariance of Critical Groups

In this section we show that critical groups are continuous with respect to the C^1-topology and are invariant under homotopies which preserve the isolation of the critical point.

We start with a definition.

Definition 6.3.1 Let X be a Banach space, $\varphi \in C^1(X)$, $C \subseteq X$ a nonempty closed subset. We say that φ satisfies the "PS-condition over C" if every sequence $\{u_n\}_{n\geqslant 1} \subseteq C$ such that $\{\varphi(u_n)\}_{n\geqslant 1}$ is bounded and $\varphi'(u_n) \to 0$ in X^* as $n \to \infty$ admits a strongly convergent subsequence.

The next lemma is a property of the negative pseudogradient flow. So, let $V : X\backslash K_\varphi \to X$ be a pseudogradient vector field corresponding to $\varphi \in C^1(X)$ (see Theorem 5.1.4). We consider the abstract Cauchy problem

$$\sigma'(t) = -V(\sigma(t)),\ \sigma(0) = x. \tag{6.69}$$

Let $\big[0, \eta_+(x)\big)$ be the maximal interval of existence for (6.69).

Lemma 6.3.2 *If X is a Banach space, $\varphi \in C^1(X)$, $u \in K_\varphi$ is isolated and φ satisfies the PS-condition over a closed neighborhood C of u, then there exists $\epsilon > 0$ and a neighborhood D of u such that if $x \in D$, then either $\sigma(t, x) \in C$ for all $t \in (0, \eta_+(x))$ or $\sigma(t, x) \in C$ until $\varphi(\sigma(t, x))$ becomes less than $\varphi(u) - \epsilon$.*

Proof Let $r > 0$ be such $\bar{B}_r(u) \subseteq C$, $\varphi|_{\bar{B}_r(u)}$ is bounded and if $A = \{v \in X : r/2 \leqslant ||v - u|| \leqslant r\}$, then $A \cap K_\varphi = \emptyset$. The PS-condition and the definition of the pseudogradient vector field (see Definition 5.1.1) imply that

$$\vartheta = \inf\{||V(v)|| : v \in A\} > 0. \tag{6.70}$$

We set $D = \bar{B}_{r/2}(u) \cap \varphi^{c+\vartheta\frac{r}{4}}$, where $c = \varphi(u)$. Let $x \in D$ be such that $\sigma(t, x)$ does not stay in C for all $t \in (0, \eta_+(x))$. So, there exist $0 \leqslant t_1 < t_2 < \eta_+(x)$ such that

$$\begin{aligned} &\sigma(t, x) \in A \text{ for } t \in [t_1, t_2], \\ &||\sigma(t_1, x) - u|| = \frac{r}{2},\ ||\sigma(t_2, x) - u|| = r. \end{aligned} \tag{6.71}$$

Then we have

$$\begin{aligned}
\varphi(\sigma(t_2)) &= \varphi(\sigma(t_1)) + \int_{t_1}^{t_2} \frac{d}{d\tau}\varphi(\sigma(\tau))d\tau \\
&\leqslant \varphi(u) + \int_{t_1}^{t_2} \frac{d}{d\tau}\varphi(\sigma(\tau))d\tau \text{ (since the flow } \sigma(\cdot) \text{ is } \varphi\text{-decreasing)} \\
&= c + \int_{t_1}^{t_2} \left\langle \varphi'(\sigma(\tau)), \sigma'(\tau)\right\rangle d\tau \text{ (by the chain rule)} \\
&\leqslant c - \int_{t_1}^{t_2} ||\varphi'(\sigma(\tau))||_*^2 d\tau \text{ (see Definition 5.1.1)} \\
&\leqslant c - \frac{1}{2}\int_{t_1}^{t_2} ||V(\sigma(\tau))||^2 d\tau \text{ (see Definition 5.1.1)} \\
&\leqslant c - \frac{\vartheta}{2}\int_{t_1}^{t_2} ||V(\sigma(\tau))|| d\tau \text{ (see (6.70))} \\
&= c - \frac{\vartheta}{2}\int_{t_1}^{t_2} ||\sigma'(\tau)|| d\tau \text{ (see (6.69))} \\
&\leqslant c - \frac{\vartheta}{2}\left\| \int_{t_1}^{t_2} \sigma'(\tau)d\tau \right\| \\
&= c - \frac{\vartheta}{2}||\sigma(t_2) - \sigma(t_1)|| \\
&\leqslant c - \frac{\vartheta}{2}\left[||\sigma(t_2) - u|| - ||\sigma(t_1) - u||\right] \text{ (by the triangle inequality)} \\
&= c - \frac{\vartheta}{2}\frac{r}{2} = c - \frac{\vartheta r}{4}.
\end{aligned}$$

We finish the proof by taking $\epsilon = \frac{\vartheta r}{4}$. □

This lemma leads to some other useful observations concerning the pseudogradient flow.

Lemma 6.3.3 *If X is a Banach space, $\varphi \in C^1(X)$, $u \in K_\varphi$ is isolated and φ satisfies the PS-condition over a closed ball $\bar{B}_r(u)$, then there exist $\epsilon > 0$ and $E \subseteq X$ such that*

(a) E is a closed neighborhood of u;
(b) E is positively invariant for the pseudogradient flow $\sigma(\cdot)$;
(c) $\varphi^{-1}([c-\epsilon, c+\epsilon]) \cap E$ is complete, where $c = \varphi(u)$;
(d) the PS-condition is satisfied over $\varphi^{-1}([c-\epsilon, c+\epsilon]) \cap E$.

Proof Let $C = \bar{B}_r(u)$ and let $\epsilon > 0$ and D, a neighborhood of u, be as postulated by Lemma 6.3.2. Consider the set

$$F = \{\sigma(t, v) : v \in D, 0 \leqslant t < \eta_+(v)\}$$

where $\sigma(\cdot, v)$ is the pseudogradient flow emanating from $v \in D$ and $[0, \eta_+(v))$ is the maximal interval of existence of the flow. We set

$$E = \overline{F}.$$

Evidently, E is a closed neighborhood of u which is positively invariant for the flow. So, we have proved (a) and (b). From Lemma 6.3.2 we have

$$\begin{aligned} &\varphi^{-1}([c-\epsilon, c+\epsilon]) \cap F \subseteq \bar{B}_r(u) \\ \Rightarrow\ &\varphi^{-1}([c-\epsilon, c+\epsilon]) \cap E \subseteq \bar{B}_r(u). \end{aligned}$$

So, $\varphi^{-1}([c-\epsilon, c+\epsilon]) \cap E$ is complete (being closed). This proves (c) and because by hypothesis $\varphi|_{\bar{B}_r(u)}$ satisfies the PS-condition (see Definition 6.3.1), we conclude that (d) holds. □

In what follows, for $\varphi \in C^1(X)$, $c \in \mathbb{R}$ and $A \subseteq X$, we set

$$A^c = A \cap \varphi^c.$$

Theorem 6.3.4 *If X is a separable reflexive Banach space, $\varphi, \psi \in C^1(X)$, $u \in X$, there exists an $r > 0$ such that $\bar{B}_r(u) \cap K_\varphi = \bar{B}_r(u) \cap K_\psi = \{u\}$ and both φ and ψ satisfy the PS-condition on $\bar{B}_r(u)$, then there exists a $\delta > 0$ depending only on φ such that*

$$\sup_{v \in X} ||\varphi - \psi||_{C^1(X)} \leqslant \delta \Rightarrow \text{rank}\, C_k(\varphi, u) = \text{rank}\, C_k(\psi, u)\ \textit{for all}\ k \in \mathbb{N}_0.$$

Proof Let $\epsilon > 0$ and $E \subseteq X$ be as postulated by Lemma 6.3.3. From Definition 6.2.18 and Theorem 6.2.20, we have

$$\begin{aligned} \text{rank}\, C_k(\varphi, u) = M_k(\varphi^{c+\epsilon} \cap E, \varphi^{c-\epsilon} \cap E) = B_k(\varphi^{c+\epsilon} \cap E, \varphi^{c-\epsilon} \cap E) \\ \text{for all } k \in \mathbb{N}_0,\ \text{where } c = \varphi(u). \end{aligned} \tag{6.72}$$

Let $\rho > 0$ be such that

$$\bar{B}_{2\rho}(u) \subseteq \varphi^{-1}([c - \frac{\epsilon}{3}, c + \frac{\epsilon}{3}]) \cap E. \tag{6.73}$$

Since by hypothesis φ satisfies the PS-condition on $\bar{B}_r(u)$, we have

$$m = \inf\{||\varphi'(v)||_* : \rho/2 \leqslant ||v - u|| \leqslant \rho\} > 0. \tag{6.74}$$

Choose $h \in C^1(X)$ such that

$$h|_{\bar{B}_{\rho/2}(u)} = 1,\ h|_{X \setminus B_\rho(u)} = 0,\ 0 \leqslant h \leqslant 1,\ \eta = \sup_{u \in X} ||h'(u)||_* < \infty. \tag{6.75}$$

This is a smooth bump function which exists since X is a separable reflexive Banach space. Let $\delta = \min\left\{\frac{\epsilon}{3}, \frac{n}{2(1+\eta)}\right\}$. We introduce the function $\hat{\psi} \in C^1(X)$ defined by

$$\hat{\psi}(v) = \varphi(v) + h(v)(\varphi(v) - \psi(v)) \text{ for all } v \in X. \tag{6.76}$$

Since $||\varphi - \psi||_{C^1(X)} \leqslant \delta$, we have

$$\begin{aligned} ||\hat{\psi}'(v)||_* &\geqslant ||\varphi'(v)||_* - h(v)||\varphi'(v) - \psi'(v)||_* - ||\psi'(u)||_*|\varphi(v) - \psi(v)| \\ &\geqslant m - (1+\eta)\delta \\ &\geqslant \frac{\delta}{2} \text{ for all } v \in X,\ \rho/2 \leqslant ||v-u|| \leqslant \rho \text{ see (6.74), (6.75).} \end{aligned} \tag{6.77}$$

Also, from (6.75) we have

$$|\hat{\psi}(v) - \varphi(v)| \leqslant h(v)|\varphi(v) - \psi(v)| \leqslant \delta \leqslant \frac{\epsilon}{3}. \tag{6.78}$$

From (6.75) and (6.76) we see that

$$\begin{aligned} &\hat{\psi}(v) = \varphi(v) \text{ for all } ||v-u|| \geqslant \rho \\ \Rightarrow\ &\hat{\psi}^{c\pm\epsilon} = \varphi^{c\pm\epsilon} \text{ (see (6.73), (6.78)).} \end{aligned}$$

Therefore $\hat{\psi}^{-1}([c-\epsilon, c+\epsilon]) \cap E = \varphi^{-1}([c-\epsilon, c+\epsilon]) \cap E$ is complete. From (6.77) it is clear that $\hat{\psi}$ satisfies the PS-condition over $\hat{\psi}^{-1}([c-\epsilon, c+\epsilon]) \cap E$. Moreover, $\bar{B}_\rho(u) \subseteq \operatorname{int} E$, so that E is positively invariant for the pseudogradient flow $\hat{\sigma}(t)$ corresponding to the functional $\hat{\psi} \in C^1(X)$. From (6.76) it is clear that $K_{\hat{\psi}} = \{u\}$. Then we have

$$\begin{aligned} \operatorname{rank} C_k(\hat{\varphi}, u) &= M_k(\hat{\psi}^{c+\epsilon} \cap E, \hat{\psi}^{c-\epsilon} \cap E) \\ &= M_k(\varphi^{c+\epsilon} \cap E, \varphi^{c-\epsilon} \cap E) \\ &= B_k(\varphi^{c+\epsilon} \cap E, \varphi^{c-\epsilon} \cap E) \text{ for all } k \in \mathbb{N}_0 \text{ (see (6.72)).} \end{aligned} \tag{6.79}$$

But from (6.75), (6.76) and the local character of the critical groups, we have

$$C_k(\hat{\psi}, u) = C_k(\psi, u) \text{ for all } k \in \mathbb{N}_0. \tag{6.80}$$

Then from (6.72), (6.79), (6.80) we conclude that

$$\operatorname{rank} C_k(\varphi, u) = \operatorname{rank} C_k(\psi, u) \text{ for all } k \in \mathbb{N}_0.$$

The proof is now complete. □

Remark 6.3.5 If we choose a field $\mathbb{F}$ for the coefficients of the homology groups in the definition of the critical groups of φ (see Definition 6.2.1), then we know that the critical groups are in fact $\mathbb{F}$-vector spaces (for example, we can take $\mathbb{F} = \mathbb{R}$, see Remark 6.2.2). Hence we avoid torsion phenomena and in the above theorem, we can say that $C_k(\varphi, u) = C_k(\psi, u)$ for all $k \in \mathbb{N}_0$.

As a consequence of Theorem 6.3.4, we see that the critical groups are invariant under homotopies which preserve the isolation of the critical point.

Theorem 6.3.6 *If X is a separable reflexive Banach space, $\{h_t\}_{t\in[0,1]} \subseteq C^1(X)$, $u \in X$ there exists $r > 0$ such that $\bar{B}_r(u) \cap K_{h_t} = \{u\}$ for all $t \in [0, 1]$, all $\{h_t\}_{t\in[0,1]}$ satisfy the PS-condition on $\bar{B}_r(u)$ and $t \to h_t$ is continuous from $[0, 1]$ into $C^1(X)$, then for all $k \in \mathbb{N}_0$ $C_k(h_t, u)$ is independent of $t \in [0, 1]$.*

Proof This follows at once from Theorem 6.3.4 and the compactness of $[0, 1]$. $\square$

Remark 6.3.7 Under the hypotheses of Theorem 6.3.6, if u is a local minimizer of h_0, then it is also a local minimizer for all h_t, $t \in (0, 1]$. In fact the above result remains valid if X is only a Banach space (see Theorem 6.3.8 below).

Theorem 6.3.8 *If X is a Banach space, $\{h_t\}_{t\in[0,1]} \subseteq C^1(X)$, each h_t satisfies the PS-condition, $a, b : [0, 1] \to \mathbb{R}$ are continuous functions such that $a(t) < b(t)$ for all $t \in [0, 1]$, both $a(t), b(t)$ are regular values of h_t, $t \in [0, 1]$, and $t \to h_t$ is continuous from $[0, 1]$ into $C^1(X)$, then for all $k \in \mathbb{N}_0$, $H_k\left(h_t^{b(t)}, h_t^{a(t)}\right)$ is independent of $t \in [0, 1]$.*

Proof To simplify an already cumbersome notation, for every $t_0, t_1 \in [0, 1]$ instead of

$$h_{t_i}, a(t_i), b(t_i) \text{ and } h_i^{-1}(a(t_i), b(t_i)) \cap K_{h_{t_i}}$$

we write

$$h_i, a_i, b_i \text{ and } K_i \text{ for } i = 0, 1.$$

Suppose that $|t_1 - t_0|$ is small. Since by hypothesis h_0 satisfies the PS-condition, we can find $c < d$

$$h_0(K_0) \subseteq (c, d) \subseteq [c, d] \subseteq (a_0, b_0) \cap (a_1, b_1).$$

The continuity of h_0 implies that we can find $\delta > 0$ such that

$$h_0((K_0)_\delta) \subseteq (c, d),$$

where $(K_0)_\delta = \{u \in X : d(u, K_0) < \delta\}$ (the δ-neighborhood of the set K_0). The PS-condition implies that there exists an $\epsilon = \epsilon(\delta) > 0$ such that

$$||h_0'(u)||_* \geqslant \epsilon \text{ for all } u \in h_0^{-1}([a_0, b_0])\backslash(K_0)_\delta.$$

Since $|t_1 - t_0|$ is small we have

$$\begin{aligned}&K_1 \subseteq (K_0)_\delta, h_1((K_0)_\delta) \subseteq (c, d), h_1^{-1}([c, d]) \subseteq h_0^{-1}(a_0, b_0)\\ \Rightarrow\ & h_i(K_j) \subseteq (c, d) \text{ for } i, j \in \{0, 1\}.\end{aligned}$$

We can construct a pseudogradient vector field for h, which coincides with a pseudogradient vector field for h_0 on $\left(h_1^{b_1} \cap h_0^{b_0}\right) \setminus (K_0)_\delta$. Then according to Corollary 5.3.13, we have that

$$\begin{aligned}&\left(h_0^d \cap h_1^d, h_0^c \cap h_1^c\right) \text{ is a strong deformation retract of } \left(h_1^{b_1}, h_0^c \cap h_1^c\right),\\ &h_0^c \cap h_1^c \text{ is a strong deformation retract of } h_1^c.\end{aligned}$$

In a similar fashion, we also show that

$$\begin{aligned}&\left(h_0^d \cap h_1^d, h_0^c \cap h_1^c\right) \text{ is a strong deformation retract of } \left(h_0^{b_0}, h_0^c \cap h_1^c\right),\\ &h_0^c \cap h_1^c \text{ is a strong deformation retract of } h_0^c.\end{aligned}$$

So, using Proposition 6.1.18, we have

$$H_k(h_0^{b_0}, h_0^c) = H_k(h_1^{b_1}, h_1^c) \text{ for all } k \in \mathbb{N}_0. \tag{6.81}$$

On the other hand, again from Corollary 5.3.13 we have

$$h_0^{a_0} \text{ is a strong deformation retract of } h_0^c, \tag{6.82}$$
$$h_1^{a_1} \text{ is a strong deformation retract of } h_1^c. \tag{6.83}$$

From (6.81), (6.82), (6.83) and Corollary 6.1.24, we infer that

$$H_k(h_0^{b_0}, h_0^{a_0}) = H_k(h_1^{b_1}, h_1^{a_1}) \text{ for all } k \in \mathbb{N}_0.$$

Finally the conclusion of the theorem follows from the compactness of $[0, 1]$. $\square$

Remark 6.3.9 A careful reading of this proof reveals that we may have $b(t) = +\infty$ for all $t \in [0, 1]$ (hence $h_t^{b(t)} = X$ for all $t \in [0, 1]$).

6.4 Extended Gromoll–Meyer Theory

In this section we present the Gromoll–Meyer theory of dynamically isolated critical sets, which is useful in dealing with resonant elliptic problems.

Definition 6.4.1 Let (X, d) be a metric space. A "flow" on X is a continuous map $\sigma : \mathbb{R} \times X \to X$ such that

(a) $\sigma(0, u) = u$ for all $u \in X$;

(b) $\sigma(t_1, \sigma(t_2, u)) = \sigma(t_1 + t_2, u)$ for all $t_1, t_2 \in \mathbb{R}$, all $u \in X$ (the group property).

Remark 6.4.2 It is an easy consequence of the group property that, for each $t \in \mathbb{R}$, $\sigma(t, \cdot)$ is a homeomorphism of X onto X (that is, a bicontinuous bijection).

Given a flow $\sigma(t, u)$, a set $C \subseteq X$ and $b > 0$, we introduce the following sets:

$$C^b = \bigcup_{|t| \leqslant b} \sigma(t, C) \text{ where } \sigma(t, C) = \{v = \sigma(t, u), u \in C\},$$

$$C^\infty = \bigcup_{t \in \mathbb{R}} \sigma(t, C), \; C_+^\infty = \bigcup_{t \geqslant 0} \sigma(t, C).$$

Also, given $t_1 \leqslant 0 \leqslant t_2$ and $0 \leqslant t_0 \leqslant +\infty$, we define

$$G_{t_1}^{t_2}(C) = \{u \in \bar{C} : \sigma([t_1, t_2], u) \subseteq \bar{C}\} = \bigcap_{t_1 \leqslant t \leqslant t_2} \sigma(t, \bar{C}),$$

$$G^{t_0}(C) = G_{-t_0}^{t_0}(C) = \bigcap_{|t| \leqslant t_0} \sigma(t, \bar{C}),$$

$$I(C) = G^\infty(C) = \bigcap_{t \in \mathbb{R}} \sigma(t, \bar{C}) \text{ (that is, } t_0 = +\infty),$$

$$\Gamma^b(C) = \{u \in G^b(C) : \sigma([0, b], u) \cap \partial C \neq \emptyset\}.$$

From these definitions, we easily deduce the following lemma.

Lemma 6.4.3 *(a)* $G^t(C) = G^t(\bar{C})$ *for all* $t \geqslant 0$.
(b) $G^{t_1}(C) \subseteq G^{t_2}(C)$ *if* $t_2 \leqslant t_1$.
(c) $G^t(C_1) \subseteq G^t(C_2)$ *for all* $t \geqslant 0$ *if* $C_1 \subseteq C_2$.
(d) $G^t(C)$ *is closed in* X *for every* $t \geqslant 0$.
(e) $G^{t_1+t_2}(C) = G^{t_2}(G^{t_1}(C))$ *for all* $t_1, t_2 \geqslant 0$.
(f) If $G^t(C) \subseteq \operatorname{int} C$, *then* $G^{2t}(C) \subseteq \operatorname{int} G^t(C)$.

Proof Only part (f) is not obvious. Suppose that the implication is not true. So, we have $G^{2t}(C) \cap \partial(G^t(C)) \neq \emptyset$. Let $\hat{u} \in G^{2t}(C) \cap \partial(G^t(C))$. So,

$$\text{there exists } u_n \to \hat{u} \text{ with } \sigma([-t, t], u_n),\; n \in \mathbb{N},\; \text{not a subset of } \bar{C}.$$

Hence we can find a sequence $\{t_n\}_{n \geqslant 1} \subseteq [-t, t]$ such that

$$\sigma(t_n, u_n) \notin \bar{C} \text{ for all } n \in \mathbb{N}.$$

By passing to a subsequence if necessary, we may assume that $t_n \to \hat{t} \in [-t, t]$. Then $\sigma(\hat{t}, \hat{u}) \in \partial C$. But $\hat{u} \in G^{2t}(C)$ and so $\sigma([-2t, 2t], \hat{u}) \subseteq \bar{C}$, which implies that $\sigma(\hat{t}, \hat{u}) \in G^t(\bar{C}) = G^t(C) \subseteq \operatorname{int} C$, a contradiction. ☐

We introduce the following family of closed sets

$$\Sigma = \Sigma(\sigma) = \{C \subseteq X : C \text{ is closed and there exists a } t > 0 \text{ such that } G^t(C) \subseteq \operatorname{int} C\}.$$

Definition 6.4.4 Let (X, d) be a metric space and $\sigma(t, u)$ a flow on it.

(a) A set $C \subseteq X$ is said to be "σ-invariant" if for all $u \in C$ and for all $t \in \mathbb{R}$, we have $\sigma(t, u) \in C$.
(b) A σ-invariant set C is said to be "isolated" if there is a neighborhood U of C such that

$$U \in \Sigma \text{ and } I(U) = C.$$

In this case, U is called an "isolating neighborhood of C".

Proposition 6.4.5 *If $C \subseteq X$ is σ-invariant, U is a compact neighborhood of C and $C = I(U) \subseteq \operatorname{int} C$, then $U \in \Sigma$.*

Proof Arguing indirectly, suppose that for every $n \in \mathbb{N}$, we have

$$G^n(U) \text{ is not a subset of int } C.$$

So, we can find $u_n \in G^n(U) \backslash \operatorname{int} U$. We have

$$\sigma([-n, n], u_n) \subseteq U \text{ but } u_n \notin \operatorname{int} U \text{ for all } n \in \mathbb{N}.$$

Due to the compactness of U, we may assume that

$$u_n \to u \in I(U) \text{ but } u \in \operatorname{int} U,$$

a contradiction. □

Proposition 6.4.6 *If $K \in \Sigma$, then there exists a $b > 0$ such that $G^b(K) \in \Sigma$ and for all $t \in \mathbb{R}$, $\sigma(t, K) \in \Sigma$.*

Proof Both sets $G^b(K)$ and $\sigma(t, K)$ are closed (see Remark 6.4.2).

Since $K \in \Sigma$, we can find $b > 0$ such that $G^b(K) \subseteq \operatorname{int} K$. Using Lemma 6.4.3 we have

$$\begin{aligned} & G^b(G^b(K)) = G^{2b}(K) \subseteq \operatorname{int} G^b(K) \\ \Rightarrow\ & G^b(K) \in \Sigma. \end{aligned}$$

Also, we have

$$G^b(\sigma(t, K)) = \bigcap_{|s| \leqslant b} \sigma(t + s, K) = \sigma(t, G^b(K)) \subseteq \sigma(t, \operatorname{int} K) = \operatorname{int} \sigma(t, K)$$
$$\Rightarrow \sigma(t, K) \in \Sigma \text{ for all } t \in \mathbb{R}.$$

The proof is now complete. □

Proposition 6.4.7 *If $K \in \Sigma$, and $b > 0$, then $\Gamma^b(K)$ is closed and $\Gamma^b(K) \subseteq \partial G^b(K)$.*

Proof Let $\{u_n\}_{n\geqslant 1} \subseteq \Gamma^b(K)$ and suppose that $u_n \to u$. Then we can find $t_n \in [0, b]$ such that $\sigma(t_n, u_n) \in \partial K$ for all $n \in \mathbb{N}$. We may assume that $t_n \to t \in [0, b]$. We have

$$\begin{aligned} &\sigma(t, u) \in \partial K \\ \Rightarrow\ &u \in \Gamma^b(K) \text{ and so } \Gamma^b(K) \text{ is closed.} \end{aligned}$$

Next, let $u \in \Gamma^b(K)$. We can find $t \in [0, b]$ and $v_n \in X \backslash K,\ n \in \mathbb{N}$, such that

$$v_n \to \sigma(t, u).$$

Let $y_n = \sigma(-t, v_n),\ n \in \mathbb{N}$. Then from the semigroup property (see Definition 6.4.1), we have

$$y_n \to u \text{ with } y_n \notin G^b(K) \text{ for all } n \in \mathbb{N}.$$

Since the set $G^b(K)$ is closed, we have $u \notin \operatorname{int} G^b(K)$ and so we conclude that $u \in \partial G^b(K)$. Therefore

$$\Gamma^b(K) \subseteq \partial G^b(K).$$

The proof is now complete. □

We take the next definition from the theory of dynamical systems.

Definition 6.4.8 Let (X, d) be a metric space and $\sigma(t, x)$ a flow on it. For every $u \in X$ the set

$$\omega(u) = \bigcap_{t>0} \overline{\sigma([t, +\infty), u)}$$

is called the "ω-limit set of u". The set

$$\omega^*(u) = \bigcap_{t>0} \overline{\sigma((-\infty, -t], u)}$$

is called the "ω^*-limit set of u". Also, given a set $S \subseteq X$, the "σ-invariant hull of S" is defined to be the set

$$[S] = \{u \in X :\ \omega(u) \cup \omega^*(u) \subseteq S\}.$$

Remark 6.4.9 Note that $\omega(u) = \omega(\sigma(t, u))$ and $\omega^*(u) = \omega^*(\sigma(t, u))$ for all $t \in \mathbb{R}$. So, it is clear that $[S]$ is σ-invariant and is the minimal σ-invariant set containing S. If S is σ-invariant, then $S \subseteq [S]$. Finally, these limit sets are described equivalently by

$$\omega(u) = \{v \in X : v = \lim_{n\to\infty} \sigma(t_n, u) \text{ for some sequence } t_n \to +\infty\},$$
$$\omega^*(u) = \{y \in X : y = \lim_{n\to\infty} \sigma(t_n, u) \text{ for some sequence } t_n \to -\infty\}.$$

Both sets are closed and σ-invariant.

Next we present two important flows with useful invariant sets.

Example 6.4.10 (a) Let X be a Banach space and $g : X \to X$ a compact map. Let $\varphi = \mathrm{id}_X - g$ and consider the abstract Cauchy problem

$$\left\{\begin{array}{l} \sigma'(t,u) = \varphi(\sigma(t,u)),\ (t,u) \in \mathbb{R} \times X, \\ \sigma(0,u) = u \end{array}\right\}.$$

Then $\sigma(t,u)$ is a flow on X and any subset of the fixed point set of g is σ-invariant.

(b) Let X be a Banach space and suppose $\varphi \in C^1(X)$ satisfies the PS-condition. Consider a pseudogradient vector field $V(\cdot)$ of φ. Let $g(u) = \min\{d(u, K_\varphi), 1\}$ and consider the abstract Cauchy problem.

$$\left\{\begin{array}{l} \sigma'(t,u) = -g(\sigma(t,u))\frac{V(\sigma(t,u))}{||V(t,u)||},\ (t,u) \in \mathbb{R} \times (X \backslash K_\varphi), \\ \sigma(0,u) = u. \end{array}\right\} \tag{6.84}$$

Then $\sigma(t,u)$ is a flow on X and any subset of K_φ is σ-invariant.

Motivated by Example 6.4.10(b), we make the following definition.

Definition 6.4.11 A triple (X, φ, σ) is a "pseudogradient flow", if X is a Banach space, $\varphi \in C^1(X)$ and satisfies the PS-condition and $\sigma(t,u)$ is the flow generated by (6.84).

Proposition 6.4.12 *If (X, φ, σ) is a pseudogradient flow, then for any $u \in X$, the sets $\omega(u)$, $\omega^*(u)$ are compact and*

$$\omega(u) \subseteq K_\varphi^c,\ \omega^*(u) \subseteq K_\varphi^{c^*},$$

for some critical values $c, c^* \in \mathbb{R}$.

Proof We prove the statement for $\omega(u)$, the proof for $\omega^*(u)$ being similar.

We claim that $\omega(u)$ is located in one level set of φ, that is,

$$\omega(u) \subseteq \varphi^{-1}(c) \text{ for some } c \in \mathbb{R}.$$

To see this, we argue by contradiction. So, suppose we could find $\{t_n\}_{n\geqslant 1}$ and $\{s_n\}_{n\geqslant 1}$ in $(0, +\infty)$ such that $t_n, s_n \uparrow +\infty$ and

$$\sigma(t_n, u) \to v,\ \sigma(s_n, u) \to \hat{v},\ \varphi(v) < \varphi(\hat{v}).$$

Without any loss of generality we may assume that $t_n < s_n$ for all $n \in \mathbb{N}$. Recalling that the pseudogradient flow is φ-decreasing (see the proof of Theorem 5.3.7), we have

$$\varphi(\sigma(s_n, u)) \leqslant \varphi(\sigma(t_n, u)) \text{ for all } n \in \mathbb{N}$$
$$\Rightarrow \varphi(\hat{v}) \leqslant \varphi(v), \text{ a contradiction.}$$

Next we show that $\omega(u) \subseteq K_\varphi$ and this, combined with the first part of the proof, implies that $\omega(u) \subseteq K_\varphi^c$. Again we argue by contradiction. So, suppose we can find $h \in \omega(u) \setminus K_\varphi$. Then choose regular values $a < b$ such that $\varphi(u), \varphi(h) \in (a, b)$. The set $K_\varphi^{[a,b]} = K_\varphi \cap \varphi^{-1}([a, b])$ is compact. So, we can find $r > 0$ such that

$$B_r(h) \cap K_\varphi^{[a,b]} = \emptyset.$$

Since $h \in \omega(u)$, from Remark 6.4.9, we know that there exists a sequence $t_n \to +\infty$ such that $u_n = \sigma(t_n, h) \in B_r(h)$ for all $n \in \mathbb{N}$.

There exists a sequence $s_n \to +\infty$ such that $y_n = \sigma(s_n, h) \in \partial(K_\varphi^{[a,b]})_r$, where

$$(K_\varphi^{[a,b]})_r = \{x \in X : d(x, K_\varphi^{[a,b]}) < r\}$$

(the r-neighborhood of $K_\varphi^{[a,b]}$). Indeed, if this is not true, then we can find $d > 0$ such that

$$\sigma([d, +\infty), h) \cap (K_\varphi^{[a,b]})_r = \emptyset.$$

Recall that φ satisfies the PS-condition. So, we can find $\eta > 0$ such that

$$||\varphi'(v)||_* \geqslant \eta \text{ for all } v \in \varphi^{-1}([a, b]) \setminus (K_\varphi^{[a,b]})_r.$$

Then

$$\varphi(h) = \lim_{n\to\infty} \varphi(y_n) \leqslant \liminf_{t\to\infty} \varphi(\sigma(t, u)) \leqslant a,$$

a contradiction to the choice of a.

Now choose $\tau_n \to +\infty$ with $s_n < \tau_n$ for all $n \in \mathbb{N}$ such that

$$\hat{y}_n = \sigma(\tau_n, u) \in B_r(h), \sigma([s_n, \tau_n], u) \cap (K_\varphi^{[a,b]})_r = \emptyset \text{ for all } n \in \mathbb{N}.$$

From the mean value theorem, we have

$$\varphi(y_n) - \varphi(\hat{y}_n) \geqslant \eta||y_n - \hat{y}_n|| \geqslant \eta d(B_r(h), (K_\varphi^{[a,b]})_r) \text{ for all } n \in \mathbb{N},$$

which is impossible since $\omega(u)$ is located in one level set of φ. □

We introduce a notion concerning a flow $\sigma(t, u)$, which is critical in our analysis.

Definition 6.4.13 Let (X, d) be a metric space and $\sigma(t, u)$ a flow on X. A set $D \subseteq X$ is said to have the "mean value property" (MVP for short) for the flow σ if for all $u \in X$ and all $t_1 < t_2$

$$\sigma(t_k, u) \in D \text{ for } k \in \{1, 2\} \text{ implies } \sigma([t_1, t_2], u) \subseteq D.$$

Using this definition, we have the following auxiliary result.

Proposition 6.4.14 *If* (X, φ, σ) *is a pseudogradient flow, S a critical set of* φ *(that is,* $S \subseteq K_\varphi$*) and D is a closed neighborhood of S with the MVP such that*

$$D \cap K_\varphi = S,$$

then

(a) $I(D) = [S] \subseteq \operatorname{int} D$;
(b) for any $t_1 < 0 < t_2$*, the set* $G_{t_1}^{t_2}(D)$ *is a closed neighborhood of* $[S]$ *with the MVP.*

Proof (a) We first show that

$$[S] \subseteq I(D). \tag{6.85}$$

To this end let $u \in [S]$. Then by Definition 6.4.8, we have

$$\omega(u) \cup \omega^*(u) \subseteq S.$$

So, according to Remark 6.4.9 we can find $t_n^+ \to +\infty$ and $t_n^- \to -\infty$ as $n \to \infty$ such that

$$\sigma(t_n^\pm, u) \in D \text{ for all } n \in \mathbb{N}.$$

Because D has the MVP, it follows that

$$\begin{aligned} &\sigma([t_n^-, t_n^+], u) \subseteq D \text{ for all } n \in \mathbb{N} \\ \Rightarrow\ &\sigma(t, u) \in D \text{ for all } t \in \mathbb{R} \\ \Rightarrow\ &u \in I(D) \text{ (from the Definition of } I(D)). \end{aligned}$$

Next we show that

$$I(D) \subseteq [S]. \tag{6.86}$$

Let $u \in I(D)$. Then $\sigma(t, u) \in D$ for all $t \in \mathbb{R}$. Recall that $D \subseteq X$ is closed. So, from Remark 6.4.9 if follows that

$$\omega(u) \cup \omega^*(u) \subseteq D. \tag{6.87}$$

Also, from Proposition 6.4.12, we have

$$\omega(u) \cup \omega^*(u) \subseteq K_\varphi. \tag{6.88}$$

From (6.87) and (6.88), we have

$$\omega(u) \cup \omega^*(u) \subseteq D \cap K_\varphi = S$$
$$\Rightarrow u \in [S] \text{ (see Definition 6.4.8).}$$

So, we have proved (6.86). From (6.85) and (6.86), we conclude that

$$I(D) = [S].$$

Next, we show that $[S] \subseteq \operatorname{int} D$. So, let $u \in [S]$. There exist $t^- < 0 < t^+$ and neighborhoods $U^\pm$ of $h^\pm = \sigma(t^\pm, u)$ such that $U^\pm \subseteq D$. Let

$$V^\pm = \sigma(t^\pm, U^\pm) \text{ and } V = V^+ \cap V^-.$$

Evidently, V is a neighborhood of u such that

$$\sigma(t^\pm, V) \subseteq U^\pm \subseteq D.$$

The MVP of D implies that $V \subseteq D$. Therefore we conclude that

$$[S] \subseteq \operatorname{int} D.$$

(b) From its definition it is clear that the set $G_{t_1}^{t_2}(D)$ is closed and has the MVP. We have

$$[S] = I(D) \subseteq G_{t_1}^{t_2}(D) \text{ (see part(a)).}$$

We need to show that $[S] \subseteq \operatorname{int} G_{t_1}^{t_2}(D)$. Arguing by contradiction, suppose that we can find $u \in [S] \cap \partial(G_{t_1}^{t_2}(D))$. So, we can find a sequence $\{u_n\}_{n \geqslant 1}$ such that

$$u_n \to u \text{ and } u_n \notin G_{t_1}^{t_2}(D) \text{ for all } n \in \mathbb{N}.$$

Therefore we can find $t_n \in [t_1, t_2]$ such that $\sigma(t_n, u_n) \notin D$ for all $n \in \mathbb{N}$. We may assume that $t_n \to t$. Then $\sigma(t_n, u_n) \to \sigma(t, u)$ (see Definition 6.4.1) and $\sigma(t, u) \notin \operatorname{int} D$. But $u \in [S]$ and so from Definition 6.4.8 (see also Remark 6.4.9), we have

$$\sigma(t, u) \in [S] \subseteq \operatorname{int} D \text{ (see part (a)),}$$

a contradiction. $\square$

Proposition 6.4.15 *If (X, φ, σ) is a pseudogradient flow, S is a critical set of φ (that is, $S \subseteq K_\varphi$) and D is a closed neighborhood of S with the MVP such that*

$$D \cap K_\varphi = S \text{ and } D \subseteq \varphi^{-1}([a, b])$$

with $a, b \in \mathbb{R}$ being regular values of φ, then for any neighborhood U of S, we can find $c > 0$ such that

$$G^c(D) = \bigcap_{|t|\leqslant c} \sigma(t, D) \subseteq \operatorname{int} U.$$

Proof We will show the contrapositive. Namely, we will show that if $u \notin \operatorname{int} U$, then we can find $c > 0$ and $t \in [-c, c]$ such that $\sigma(t, u) \notin D$.

Since φ satisfies the PS-condition, we can find $\eta > 0$ such that

$$d(u, K_\varphi) \geqslant \eta \text{ and } ||\varphi'(u)||_* \geqslant \eta \text{ for all } u \in D\backslash\operatorname{int} U. \tag{6.89}$$

Choose $c > \frac{1}{\eta_2}(b - a)$. We consider three distinct cases:

(1) $u \notin D$: Let $t = 0$. We have $\sigma(0, u) = u \notin D$.
(2) $u \in D\backslash(\operatorname{int} U)^\infty$ (recall $(\operatorname{int} U)^\infty = \bigcup_{t\in\mathbb{R}} \sigma(t, U)$): Suppose that for $c > 0$

$$\begin{aligned} &\sigma([-c, c], u) \subseteq D \\ \Rightarrow\ &\sigma([-c, c], u) \subseteq D\backslash(\operatorname{int} U)^\infty \\ \Rightarrow\ &\varphi(\sigma(-c, u)) - \varphi(\sigma(c, u)) \geqslant 2\eta^2 c > b - a \text{ (see (6.89))}, \end{aligned}$$

which contradicts the hypothesis that $D \subseteq \varphi^{-1}([a, b])$.
(3) $u \in [(\operatorname{int} U)^\infty\backslash\operatorname{int} U] \cap D = [(\operatorname{int} U)^\infty \cap D]\backslash\operatorname{int} U$: Then

$$(3)_i\ u \in \bigcup_{t>0} \sigma(t, \operatorname{int} U) \text{ or } (3)_{ii}\ u \in \bigcup_{t<0} \sigma(t, \operatorname{int} U).$$

If $(3)_i$ holds, then we can find $t_1 \leqslant 0 \leqslant t_2$ such that

$$\sigma([t_1, t_2], u) \subseteq [(\operatorname{int} U)^\infty \cap D]\backslash\operatorname{int} U \text{ and } \sigma(t_1 - \epsilon, u) \in U, \sigma(t_2 + \epsilon, u) \notin D,$$

for $\epsilon > 0$ small. Then

$$\begin{aligned} &b - a \geqslant \eta^2(t_2 - t_1) \\ \Rightarrow\ &t_2 < c. \end{aligned}$$

If $(3)_{ii}$ holds, then in a similar fashion we show that $t_1 > -c$.
So, for both $(3)_i$ and $(3)_{ii}$ we have

$$\sigma([-c, c], u) \cap D^c \neq \emptyset.$$

The proof is now complete. □

Now we are ready for our first theorem.

Theorem 6.4.16 *If (X, φ, σ) is a pseudogradient flow, S is a critical set of φ, that is, $S \subseteq K_\varphi$ and D is a closed neighborhood of S with the MVP such that*

$$D \cap K_\varphi = S \text{ and } D \subseteq \varphi^{-1}([a, b]),$$

then $[S]$ is an isolated σ-invariant set and any closed neighborhood U of $[S]$ with $U \subseteq D$ is an isolating neighborhood of $[S]$.

Proof From Proposition 6.4.14, we have

$$[S] = I(D) \subseteq \operatorname{int} D. \tag{6.90}$$

Let U be a closed neighborhood of $[S]$ such that $U \subseteq D$. Then

$$\begin{aligned} &[S] = I([S]) \subseteq I(U) \subseteq I(D) = [S] \text{ (see (6.90))} \\ \Rightarrow\ & I(U) = [S]. \end{aligned}$$

By definition

$$I(U) \subseteq G_{t_1}^{t_2}(U) \text{ for any } t_1 < 0 < t_2.$$

So, we can use Proposition 6.4.15 and find $c > 0$ such that

$$\begin{aligned} &G^c(U) \subseteq G^c(D) \subseteq \operatorname{int} U \\ \Rightarrow\ & U \text{ is an isolating neighborhood of } [S] \text{ (see Definition 6.4.4(b))}. \end{aligned}$$

The proof is now complete. ☐

We introduce the fundamental notion of "dynamically isolated critical set".

Definition 6.4.17 Let (X, φ, σ) be a pseudogradient flow and S a critical set of φ (that is, $S \subseteq K_\varphi$). We say that S is a "dynamically isolated critical set" if there exists a closed neighborhood D of S and regular values $a < b$ of φ such that

$$D \subseteq \varphi^{-1}([a, b]) \text{ and } \overline{D^\infty} \cap K_\varphi \cap \varphi^{-1}([a, b]) = S$$

(recall $D^\infty = \bigcup\limits_{t \in \mathbb{R}} \sigma(t, D)$). We say that (D, a, b) is an "isolating triplet" for S.

Remark 6.4.18 If C is an isolated critical value of φ (that is, there exists an $\epsilon > 0$ such that $[c - \epsilon, c + \epsilon]$ contains no critical values other than c), then K_φ^c is a dynamically isolated critical set. Similarly, if u_0 is an isolated critical point of φ, then the singleton $S = \{u_0\}$ is a dynamically isolated critical set. In particular, if u_0 is a nondegenerate critical point of $\varphi \in C^2(H)$ (H being a Hilbert space), then the singleton $S = \{u_0\}$ is a dynamically isolated critical set.

Lemma 6.4.19 *If (X, φ, σ) is a pseudogradient flow, S is a critical set of φ (that is, $S \subseteq K_\varphi$) and (D, a, b) is an isolating triplet for S, then there exists a $c > 0$ such that*

$$D^c \cap \varphi^{-1}([a, b]) = D^\infty \cap \varphi^{-1}([a, b]) = \overline{D^\infty} \cap \varphi^{-1}([a, b]);$$

moreover the set $D^\infty \cap \varphi^{-1}([a, b])$ is a closed neighborhood of both S and $[S]$, which has the MVP.

Proof Let $V = \overline{D^\infty} \cap \varphi^{-1}([a, b])$. We need to show that $V = D^c \cap \varphi^{-1}([a, b])$ for some $c > 0$. Since φ satisfies the PS-condition, we can find $\eta > 0$ such that

$$d(u, K_\varphi) \geqslant \eta \text{ and } ||\varphi'(u)||_* \geqslant \eta \text{ for all } u \in V \backslash D.$$

If $\sigma([0, t], u) \subseteq V \backslash D$, then

$$b - a \geqslant \varphi(u) - \varphi(\sigma(t, u)) \geqslant -\int_0^t \big\langle \varphi'(\sigma(s, u)), \sigma'(s, u)\big\rangle ds \geqslant \eta^2 t.$$

Choose $c > \frac{1}{\eta^2}(b - a)$. If the equality $V = D^c \cap \varphi^{-1}([a, b])$ fails, we can find

$$\begin{aligned} & h \in V \backslash D^c \\ \Rightarrow\ & h = \sigma(t, u) \text{ and } \sigma([0, t], u) \cap D = \emptyset \text{ for some } u \in D \text{ and some } t > c, \end{aligned}$$

a contradiction (see Definition 6.4.17). The last part of the lemma follows from Theorem 6.4.16. ☐

Combining Lemma 6.4.19 with Theorem 6.4.16, we obtain:

Theorem 6.4.20 *If (X, φ, σ) is a pseudogradient flow and S is a dynamically isolated critical set of φ, then $[S]$ is an isolated σ-invariant set and if (D, a, b) is an isolating triplet for S, then any closed neighborhood U of $[S]$ such that $U \subseteq D$ is an isolating neighborhood of $[S]$.*

Now we can extend the notion of critical groups from an isolated critical point (see Definition 6.2.1) to a dynamically isolated critical set.

Definition 6.4.21 Let (X, φ, σ) be a pseudogradient flow, S a dynamically isolated critical set and (D, a, b) an isolating triplet for S. The "critical groups" of S are defined by

$$C_k(\varphi, S) = H_k(\varphi^b \cap D_+^\infty, \varphi^a \cap D_+^\infty) \text{ for all } k \in \mathbb{N}_0.$$

For this definition to make sense, we need to show that it is independent of the choice of the isolating triplet (D, a, b) and the choice of the pseudogradient vector field for φ. This is done in the next proposition.

Proposition 6.4.22 *The definition of critical groups for a dynamically isolated critical set S (see Definition 6.4.21) is independent of the particular choice of the isolating triplet and of the pseudogradient vector field.*

Proof First we assume that in the isolating triple (D, a, b), the neighborhood D is fixed and the regular values a, b vary. Then the invariance of $C_k(\varphi, S),\ k \in \mathbb{N}_0$, is a consequence of Corollary 5.3.13.

Next suppose that (D, a, b) and (D_*, a, b) are two isolating triplets for S such that

$$S \subseteq \operatorname{int} D_* \subseteq D_* \subseteq \operatorname{int} D \subseteq D. \tag{6.91}$$

We need to show that

$$H_k(\varphi^b \cap D_+^\infty, \varphi^a \cap D_+^\infty) = H_k(\varphi^b \cap D_{*,+}^\infty, \varphi^a \cap D_{*,+}^\infty) \text{ for all } k \in \mathbb{N}_0. \tag{6.92}$$

Claim 1: There exists a $\delta > 0$ such that $d(\partial D_{*,+}^\infty, [S]) \geqslant \delta$.
If Claim 1 is not true, then we can find a sequence $\{u_n\}_{n\geqslant 1}$ such that

$$u_n \in \partial D_{*,+}^\infty\ n \in \mathbb{N} \text{ and } u_n \to u \in [S] \text{ as } n \to \infty.$$

So, we can find $v_n \in \partial D_*$ and $t_n \in [0, c]$, $n \in \mathbb{N}$, such that

$$u_n = \sigma(t_n, v_n) \to u.$$

By passing to a suitable subsequence if necessary, we may assume that

$$v_n = \sigma(-t_n, u_n) \to \sigma(-t, u) = v.$$

The set ∂D_* is closed, hence $v \in \partial D_*$. But recall that $u \in [S]$. Hence $v \in [S]$, a contradiction [see (6.91)]. This proves Claim 1.

Choose $a_* < a$ so that (D_*, b, a_*) is still an isolating triple for S (recall a is a regular value). We set

$$A = D_+^\infty \cap \varphi^{-1}(a_*),\ A_* = D_{*,+}^\infty \cap \varphi^{-1}(a_*) \text{ and } E = \{u \in D_+^\infty : \omega(u) \cap S^c \neq \emptyset\}.$$

For each $u \in E$, there exists a unique $h \in A$ and a unique $t \in \mathbb{R}$ such that

$$h = \sigma(t, u). \tag{6.93}$$

Let $p : E \to A$ and $q : E \to \mathbb{R}$ be the maps defined by

$$p(u) = h \text{ and } q(u) = t \text{ (see (6.93))}.$$

We set

$$C = A \backslash p(E).$$

Claim 2: $d(C, p(\partial D_+^\infty)) > 0$.
If Claim 2 is not true, then we can find a sequence $\{h_n\}_{n\geqslant 1}$ such that

$$h_n \in p(\partial B_+^\infty) \text{ and } h_n \to h \in C.$$

So, we can find $v_n \in \partial D$ and $t_n \in [0, c]$ such that $v_n = \sigma(-t_n, h_n)$ for all $n \in \mathbb{N}$. We may assume that $t_n \to t$ and so we have

$$\begin{aligned} v_n = \sigma(-t_n, h_n) \to \sigma(-t, h) = v \\ \Rightarrow v \in \partial D \text{ (recall that } \partial D \text{ is closed).} \end{aligned}$$

But $h \in C$ and so we have a contradiction. This proves Claim 2.

Let $E_* = \{u \in D_{*,+}^\infty : \omega(u) \cap S^c \neq \emptyset\}$. From Claim 2 we have

$$r_0 = d(C, p(\partial B_{*,+}^\infty)) > 0.$$

Consider the parametric family of continuous functions $\gamma_\tau : \mathbb{R}_+ \to \mathbb{R}$, $\tau \in [0, 1]$, defined by

$$\gamma_\tau(t) = \begin{cases} 1 - \frac{\tau}{r_0} t & \text{if } 0 \leqslant t \leqslant r_0 \\ 1 - \tau & \text{if } r_0 < t. \end{cases}$$

Then we introduce the deformation $\xi : [0, 1] \times (\varphi^b \cap D_+^\infty) \to \varphi^b \cap D_+^\infty$ defined by

$$\xi(\tau, u) = \begin{cases} \sigma(-\gamma_\tau(d(h, C)s), h) & \text{if } u \in E \cap \varphi^{-1}([a, b]) \\ u & \text{if } u \in (\varphi^b \cap D_+^\infty)\backslash E, \end{cases}$$

where $h = p(u)$ and $s = q(u)$. Then $\xi(0, u) = u$ and

$$\xi(1, u) = \begin{cases} h & \text{if } u \notin D_{*,+}^\infty \cap \varphi^{-1}([a_*, b]) \\ \sigma(-(1 - d(h, C)\frac{1}{r_0})s, h) & \text{if } u \in E_* \cap \varphi^{-1}([a_*, b]) \\ u & \text{if } u \in (\varphi^b \cap D_+^\infty)\backslash E. \end{cases}$$

Let $L_1 = \xi(1, \varphi^b \cap D_+^*)$ and $L_2 = L_1 \backslash (A_+^\infty \backslash A_1)$. Using the properties of relative singular homology groups (see Sect. 6.1), we have

$$\begin{aligned} & H_k(\varphi^b \cap D_+^\infty, \varphi^a \cap D_+^\infty) \\ = & H_k(L_1, \xi(1, \varphi^a \cap D_+^\infty)) \text{ (by the deformation invariance)} \\ = & H_k(L_2, \xi(1, \varphi^a \cap D_+^\infty)) \text{ (by excision)} \\ = & H_k(L_2, \xi(1, \varphi^{a_*} \cap D_{*,+}^\infty)) \text{ (by the deformation invariance)} \\ = & H_k(\varphi^b \cap D_{*,+}^\infty, \varphi^{a_*} \cap D_{*,+}^\infty) \text{ (by the deformation invariance)} \\ = & H_k(\varphi^b \cap D_{*,+}^\infty, \varphi^a \cap D_{*,+}^\infty) \text{ (by Corollary 5.3.13) for all } k \in \mathbb{N}_0. \end{aligned}$$

In a similar fashion, we also show the invariance with respect to the pseudogradient vector field. $\square$

To describe the topological properties of a dynamically isolated critical set, we will need the following notion.

Definition 6.4.23 Let (X, φ, σ) be a pseudogradient flow and S a critical set of φ (that is, $S \subseteq K_\varphi$). A pair of sets (W, W_-) is said to be a "Gromoll–Meyer pair" (a GM-pair for short) associated with the pseudogradient flow if the following conditions hold:

(a) W is a closed neighborhood of S with the MVP such that

$$W \cap K_\varphi = S \text{ and } W \cap \varphi^a = \emptyset \text{ for some } a \in \mathbb{R}.$$

(b) W_- is an exit set of W, that is, for every $u \in W$ and $t_1 > 0$ such that $\sigma(t_1, u) \notin W$, we can find $\hat{t} \in [0, t_1)$ for which we have

$$\sigma([0, \hat{t}], u) \subseteq W \text{ and } \sigma(\hat{t}, u) \in W_-.$$

(c) W_- is closed and is the union of a finite number of submanifolds which are transversal to the flow σ.

Example 6.4.24 In this example, we construct a GM-pair for an isolated critical point. So, let $X = H$ be a Hilbert space and suppose $\varphi \in C^1(H)$ satisfies the PS-condition. For simplicity we assume that the critical point $u_0 = 0$ and that $\varphi(0) = 0$. Let $\epsilon, \delta > 0$ such that

$$0 \text{ is the only critical value in } [-\epsilon, \epsilon] \text{ and } \bar{B}_\delta(0) \cap K_\varphi = \{0\}.$$

The PS-condition implies that

$$\eta = \inf\{||\nabla\varphi(u)|| : u \in \bar{B}_\delta(0)\backslash\bar{B}_{\delta/2}(0)\} > 0.$$

Let $\lambda \in \left(0, \frac{2\delta}{\eta}\right)$ and consider the functional

$$\psi(u) = \varphi(u) + \lambda||u||^2.$$

We choose $\gamma, \mu > 0$ in such a way that if $W = \psi^\mu \cap \varphi^{-1}([-\gamma, \gamma])$ and $W_- = W \cap \varphi^{-1}(-\gamma)$, then the following conditions hold

$$0 < \gamma < \min\{\epsilon, \frac{3\delta^2\lambda}{8}\} \text{ and } \delta^2\lambda/4 + \gamma < \mu < \delta^2\lambda - \gamma,$$

$$\bar{B}_{\delta/2}(0) \cap \varphi^{-1}([-\gamma, \gamma]) \subseteq W \subseteq \bar{B}_\delta(0) \cap \varphi^{-1}([-\epsilon, \epsilon]), \tag{6.94}$$

$$\varphi^{-1}([-\gamma, \gamma]) \cap \psi^{-1}(\mu) \subseteq \bar{B}_\delta(0)\backslash\bar{B}_{\delta/2}(0), \tag{6.95}$$

$$(\nabla\varphi(u), \nabla\psi(u))_H > 0 \text{ for all } u \in \bar{B}_\delta(0)\backslash B_{\delta/2}(0). \tag{6.96}$$

We claim that (W, W_-) is a GM-pair.

First we show that W has the MVP. There is no loss of generality if we assume that $\sigma(0,u), \sigma(t,u) \in W$ $u \in W$. Let $t_0 = \sup\{s \in [0,t] : \sigma(s',u) \in W \text{ for all } u \leqslant s' \leqslant s\}$. If $t_0 < t$, then $\sigma(t_0,u) \notin B_{\delta/2}(0)$. But we have

$$\gamma \geqslant \varphi(\sigma(0,u)) \geqslant \varphi(\sigma(t_0,u)) \geqslant \varphi(\sigma(t,u)) \geqslant -\gamma \tag{6.97}$$

(recall that σ is φ-decreasing),

$$(\psi \circ \sigma(\cdot,u))'(t_0) = -(\nabla\psi(s), \nabla\varphi(s))_H|_{\sigma(t_0,u)} < 0 \text{ see (6.94), (6.96).} \tag{6.98}$$

From (6.97) and (6.98) we have a contradiction to the maximality of t_0.

Let $\tilde{W}_- = \{u \in W : \sigma(t,u) \notin W \text{ for all } t > 0\}$. Evidently,

$$W_- \subseteq \tilde{W}_-. \tag{6.99}$$

By definition, $\tilde{W}_- \subseteq \partial W = W_- \cup (\varphi^{-1}(\gamma) \cap \text{int}\psi^\mu) \cup (\psi^{-1}(\mu) \cap (W \setminus W_-))$. If $u \in \varphi^{-1}(\gamma) \cap \text{int}\,\psi^\mu$, then $u \notin W_-$. If $u \in \psi^{-1}(\mu) \cap (W \setminus W_-)$, then from (6.95) and (6.96) we have

$$(\psi \circ \sigma(\cdot,u))'(0) < 0 \text{ and } \varphi(u) > -\gamma.$$

So, we can find $\tau > 0$ such that

$$\begin{aligned} &\psi(\sigma(\tau,u)) \leqslant \mu \text{ and } |\varphi(\sigma(\tau,u))| \leqslant \gamma \\ \Rightarrow\ & u \notin \tilde{W}_- \\ \Rightarrow\ & \tilde{W}_- \subseteq W_- \\ \Rightarrow\ & W_- = \tilde{W}_- \text{ (see (6.99)).} \end{aligned}$$

From Definition 6.4.23 it follows that (W, W_-) is a GM-pair.

We can extend this example from an isolated critical point to a dynamically isolated critical set.

Theorem 6.4.25 *If (X,φ,σ) is a pseudogradient flow, S is a dynamically isolated critical set and (D,a,b) is an isolating triplet for S, then for any neighborhood U of $[S]$ such that*

$$U \subseteq D^\infty \cap \varphi^{-1}([a,b]),$$

there exists a GM-pair (W, W_-) for S such that $W \subseteq U$.

Proof Let $a < a_0 < \min[\varphi(u) : u \in S]$. From Lemma 6.4.19 we know that $D^\infty \cap \varphi^{-1}([a_0,b])$ is a closed neighborhood of $[S]$ which has the MVP. From Proposition 6.4.15, we know that there exists a $c > 0$ such that

$$W = G^c(D^\infty \cap \varphi^{-1}([a_0,b])) \subseteq \text{int}\,U.$$

Proposition 6.4.14 implies that

$$W \text{ is a closed neighborhood of } [S] \text{ which has the MVP.}$$

Moreover, we have

$$W \cap K_\varphi = S \text{ and } W \cap \varphi^a = \emptyset.$$

We look for an exit set E of W. Let $L_{a_0} = D^\infty \cap \varphi^{-1}(a_0)$. This is a submanifold of $\varphi^{-1}(a_0)$.

Since W is a neighborhood of S, we have

$$(D^\infty \cap \varphi^{-1}([a_0, b]) \setminus W) \cap K_\varphi = \emptyset.$$

So, for all $u \in E$, we can find $t > 0$ such that

$$\begin{aligned}
& h = \sigma(t, u) \in L_{a_0} \\
\Rightarrow\ & t = -c \text{ (recall the definition of } W) \\
\Rightarrow\ & E = \sigma(-c, L_{a_0}) \text{ is a submanifold which is transversal to } \sigma \\
\Rightarrow\ & (W, W_-) = (W, E) \text{ is a GM-pair for } S.
\end{aligned}$$

The proof is now complete. □

Theorem 6.4.26 *If (X, φ, σ) is a subgradient flow and S is a dynamically isolated critical set of φ, then for any GM-pair for S we have*

$$C_k(\varphi, S) = H_k(W, W_-) \textit{ for all } k \in \mathbb{N}_0.$$

Proof Let (D, a, b) be an isolating triplet for S. Using Theorem 6.4.25, we replace D by W. First we show that

$$H_k(\varphi^b \cap W_+^\infty, \varphi^a \cap W_+^\infty) = H_k(W_+^\infty, (W_-)_+^\infty) \text{ for all } k \in \mathbb{N}_0 \tag{6.100}$$

(recall that $W_+^\infty = \bigcup\limits_{t \geqslant 0} \sigma(t, W)$ and $(W_-)_+^\infty = \bigcup\limits_{t \geqslant 0} \sigma(t, W_-)$).

We introduce two deformation retractions

$$\begin{aligned}
d_1 &: [0, 1] \times (W_-)_+^\infty \to \varphi^a \cap (W_-)_+^\infty, \\
d_2 &: [0, 1] \times W_+^\infty \to \varphi^b \cap W_+^\infty,
\end{aligned}$$

defined as follows. Let $\vartheta_1 : (W_-)_+^\infty \to \mathbb{R}$ be the first hitting time from $(W_-)_+^\infty$ at the level set $\varphi^{-1}(a)$. So, we have

$$\begin{aligned}
& \sigma(\vartheta_1(u), u) \in \varphi^{-1}(a) \text{ for all } u \in (W_-)_+^\infty \setminus \varphi^a, \\
& \vartheta_1(u) = 0 \text{ for all } u \in \varphi^a \cap (W_-)_+^\infty.
\end{aligned}$$

The transversality of σ to $\varphi^{-1}(a)$ implies the continuity of ϑ_1.

Similarly, let $\vartheta_2 : W_+^\infty \to \mathbb{R}$ be the first hitting time from W_+^∞ at the level set $\varphi^{-1}(b)$. So, as before we have

$$\sigma(\vartheta_2(u), u) \in \varphi^{-1}(b) \text{ for all } u \in W_+^\infty \setminus \varphi^b,$$
$$\vartheta_2(u) = 0 \text{ for all } u \in \varphi^b \cap W_+^\infty.$$

For the same reason ϑ_2 is continuous.

We set

$$d_1(s, u) = \sigma(s\vartheta_1(u), u) \text{ for all } (s, u) \in [0, 1] \times (W_-)_+^\infty,$$
$$d_2(s, u) = \sigma(s\vartheta_2(u), u) \text{ for all } (s, u) \in [0, 1] \times W_+^\infty.$$

Since $W \cap \varphi^a = \emptyset$ (see Definition 6.4.23), we have

$$\varphi^a \cap (W_-)_+^\infty = \varphi^a \cap W_+^\infty.$$

Then using the deformation retractions d_1 and d_2, we have that

$$(W_+^\infty, (W_-)^\infty) \text{ and } (\varphi^b \cap W_+^\infty, \varphi^a \cap W_+^\infty)$$

are homotopy equivalent. Hence by Proposition 6.1.14 we have (6.100).

Next we show that

$$H_k(W_+^\infty, (W_-)_+^\infty) = H_k(W, W_-) \text{ for all } k \in \mathbb{N}_0. \tag{6.101}$$

Let $\delta > 0$ and set $W_\delta = \bigcup\limits_{t>} \sigma(t, W_-)$. We consider $\vartheta : W_+^\infty \to \mathbb{R}$, the first hitting time at the set W_-. We have

$$\sigma(-\vartheta(u), u) \in W_- \text{ for all } u \in W_+^\infty,$$
$$\vartheta(u) = 0 \text{ for all } u \in W_+^\infty \setminus (W_-)_+^\infty.$$

Recall that the flow σ is transversal to W_- [see Definition 6.4.23(c)]. So $\vartheta(\cdot)$ is continuous. Also, we have

$$W_\delta = \{u \in W^\infty : \vartheta(u) > \delta\}$$

and so W_δ is relatively open in W^∞. We have

$$\overline{W_\delta}^{W^\infty} = \{u \in W^\infty : \vartheta(u) \geqslant \delta\} \subseteq \{u \in W^\infty : \vartheta(u) > 0\} = \operatorname{int}(W_-)_+^\infty.$$

The excision property of singular homology implies that

$$H_k(W_+^\infty, (W_-)_+^\infty) = H_k(W_+^\infty \backslash W_\delta, (W_-)_+^\infty \backslash W_\delta) \text{ for all } k \in \mathbb{N}_0. \tag{6.102}$$

Let $d_3 : [0, 1] \times (W^\infty \backslash W_\delta) \to W$ and $d_4 : [0, 1] \times ((W_-)_+^\infty \backslash W_\delta) \to W$ be the deformations defined by reversing the flow, that is,

$$d_3(t, u) = \sigma(-t\vartheta(u), u) \text{ for all } (t, u) \in [0, 1] \times (W^\infty \backslash W_\delta),$$
$$d_4(t, u) = \sigma(-t\vartheta(u), u) \text{ for all } (t, u) \in [0, 1] \times ((W_-)_+^\infty \backslash W_\delta).$$

These are strong deformation retractions. So, we have

$$H_k(W_+^\infty \backslash W_\delta, (W_-)_+^\infty \backslash W_\delta) = H_k(W, W_-) \text{ for all } k \in \mathbb{N}_0. \tag{6.103}$$

From (6.100), (6.101), (6.102), (6.103) and invoking Proposition 6.4.22, we conclude that

$$C_k(\varphi, S) = H_k(W, W_-) \text{ for all } k \in \mathbb{N}_0.$$

The proof is now complete. □

We want to study the stability of the critical groups for dynamically isolated critical sets under perturbations of the flow (therefore under changes in the pseudogradient vector field).

So, let $(S, \varphi, \sigma_\lambda)$ $\lambda \in [0, 1]$, be a family of pseudogradient flows. We impose the following "uniform continuity condition".

(UC) : "For every $\epsilon > 0$ and $b > 0$, there exists a $\delta(\epsilon, b) > 0$ such that

$$||u - v|| + |t - s| + |\lambda - \eta| < \delta \text{ and } |t|, |s| \leqslant b$$

imply

$$||\sigma_\lambda(t, u) - \sigma_\eta(s, v)|| < \epsilon.$$

The following proposition is an immediate consequence of this uniformity condition.

Proposition 6.4.27 *If* $(X, \varphi, \sigma_\lambda)$, $\lambda \in [0, 1]$, *is a family of pseudogradient flows which satisfies the* (UC), *S is a dynamically isolated critical set of the flow* σ_{λ_0} *and* (D, a, b) *is an isolating triplet for S, then there exists a* $\delta > 0$ *such that*

$$|\lambda - \lambda_0| < \delta \Rightarrow (D, a, b) \textit{ is also an isolating triple for } S \textit{ for the flow } \sigma_\lambda.$$

Remark 6.4.28 In fact the isolating neighborhood of *S* (see Definition 6.4.4(b)) is also stable under small changes of the parameter $\lambda \in [0, 1]$.

Next we examine the effect on GM-pairs when we perturb the functional φ.

Theorem 6.4.29 *If* (X, φ, σ) *is a pseudogradient flow,* S_φ *is critical set of* φ *and* (W, W_-) *is a GM-pair for* S_φ, *then there exists an* $\epsilon = \epsilon(\varphi, W) > 0$ *such that for all* $\psi \in C^1(X)$ *with*

$$||\psi - \varphi||_{C^1(W)} < \epsilon$$

(W, W_-) is still a GM-pair for any critical set S_ψ of ψ such that

$$W \cap K_\psi = S_\psi.$$

Proof Let U be a neighborhood of S_φ such that $U \subseteq \bar{U} \subseteq \operatorname{int} W$. Let V be the pseudogradient vector field associated with the flow (X, φ, σ). Since φ satisfies the PS-condition (see Definition 6.4.11), we can find an open neighborhood U_0 of S_φ such that

$$U_0 \subseteq \bar{U}_0 \subseteq U \text{ and } \eta = \inf\{||\varphi'(u)||_* : u \in W \backslash U_0\} > 0.$$

Let $\epsilon \in (0, \frac{\eta}{6})$ and let $\psi \in C^1(X)$ be such that

$$||\psi - \varphi||_{C^1(W)} < \epsilon.$$

Evidently, $S_\psi \subseteq U_0$.

Consider a pseudogradient vector field $\hat{V}$ for ψ such that

$$||\hat{V}(u) - V(u)|| < \epsilon \text{ for all } u \in W.$$

Let $\vartheta : X \to \mathbb{R}$ be a locally Lipschitz function such that

$$0 \leqslant \vartheta \leqslant 1 \text{ and } \vartheta(u) = \begin{cases} 1 \text{ if } u \in \bar{U}_0 \\ 0 \text{ if } u \in X \backslash U. \end{cases}$$

Let $V_0(u) = \frac{5}{4}\left[(1 - \vartheta(u))V(u) + \vartheta(u)\hat{V}(u)\right]$ for all $u \in X$. For $u \in X \backslash U_0$ we have

$$\begin{aligned} ||V_0(u)|| &\leqslant \frac{5}{4}\left[||\hat{V}(u)|| + \epsilon\right] \leqslant \frac{13}{4}||\psi'(u)||_* \\ \text{and } \langle \psi'(u), V_0(u) \rangle &\geqslant \frac{5}{4}\left[||\psi'(u)||_*^2 - \epsilon||\psi'(u)||_*\right] \\ &\geqslant \frac{5}{4}\left[||\psi'(u)||_*^2 - \frac{1}{5}||\psi'(u)||_*^2\right] \\ &= ||\psi'(u)||_*^2. \end{aligned}$$

If $u \in U_0$, then $V_0(u) = \hat{V}(u)$. Therefore V_0 is also a pseudogradient vector field for ψ.

Note that $V_0(u) = \frac{5}{4}V(u)$ for all $u \in X \backslash U$. Therefore the flow σ_0 corresponding to V_0 remains the same as the flow σ. In particular, they are the same on W_-. Also, we can easily check that W satisfies the MVP for the flow σ_0. Therefore (W, W_-) remains a GM-pair for S_ψ. $\square$

Definition 6.4.30 Let (X, φ, σ) be a pseudogradient flow and S a critical set of φ (that is, $S \subseteq K_\varphi$). A subset $A \subseteq [S]$ is called an "attractor" in $[S]$ if there is a neighborhood U of A such that $\omega(U \cap [S]) = A$. The dual "repeller" of A in $[S]$ is defined by $A^* = \{u \in [S] : \omega(u) \cap A = \emptyset\}$. The pair (A, A^*) is said to be an "attractor-repeller pair". An ordered collection $\{M_k\}_{k=1}^n$ of σ-invariant subsets $M_k \subseteq [S]$ is said to be a "Morse decomposition" of $[S]$ if there is an increasing family of attractors

$$\emptyset = A_0 \subseteq A_1 \subseteq \ldots \subseteq A_n = [S]$$

such that $M_k = A_k \cap A_{k-1}^*$ for all $k \in \{1, \ldots, n\}$.

Remark 6.4.31 An attractor-repeller pair (A, A^*) of $[S]$ is a Morse decomposition with $A_0 = \emptyset$, $A_1 = A$, $A_2 = [S]$. More generally, suppose that $a, b \in \mathbb{R}$ are regular values of φ and assume that $\varphi^{-1}([a, b]) \cap K_\varphi = \{u_k\}_{k=1}^n$ with $\varphi(u_k) \leqslant \varphi(u_{k+1})$ for all $k \in \{1, \ldots, n-1\}$. Then $\{\{u_k\}\}_{k=1}^n$ is a Morse decomposition of $[S] = I(\varphi^{-1}([a, b]))$.

Then we can have an extension of the Morse relation from Theorem 6.2.20. For a proof of this result, we refer to Chang [118] (Sect. 5.5).

Theorem 6.4.32 *If (X, φ, σ) is a pseudogradient flow, S is a critical set of φ (that is, $S \subseteq K_\varphi$), $\{M_i\}_{i=1}^n$ is a Morse decomposition of $[S]$ and (W, W_-) is a GM-pair for $[S]$, then $\sum_{k \in \mathbb{N}_0} \left(\sum_{i=1}^n \operatorname{rank} H_k(W_i, W_{i-1}) t^k \right) = \sum_{k \in \mathbb{N}_0} \operatorname{rank} H_k(W, W_-) t^k + (1 + t) Q(t)$, where (W_i, W_{i-1}) is the GM-pair for M_i, $i \in \{1, \ldots, n\}$ and $Q(t)$ is a formal series with nonnegative integer coefficients.*

Remark 6.4.33 Suppose that $[S] = \{u_i\}_{i=1}^n$. From Remark 6.4.31 we know that $\{u_i\}_{i=1}^n$ is a Morse decomposition of $[S]$. We set

$$M_k = \sum_{i=1}^n \operatorname{rank} C_k(\varphi, u_i), \ \beta_k = \operatorname{rank} H_k(W, W_-) \text{ for all } k \in \mathbb{N}_0.$$

We assume that they are all finite and that the series that we are about to formally introduce converge. We have

$$\sum_{i=0}^k (-1)^{k-i} \beta_i \leqslant \sum_{i=0}^k (-1)^{k-i} M_i \text{ for all } k \in \mathbb{N}_0,$$
$$\sum_{k \in \mathbb{N}_0} (-1)^k \beta_k = \sum_{k \in \mathbb{N}_0} (-1)^k M_k.$$

These are the extended Morse relations.

6.5 Local Extrema and Critical Points of Mountain Pass Type

The main idea of Morse theory is that different critical points of a functional $\varphi \in C^1(X)$ can be distinguished by the topological structure of their neighborhoods in the sublevel sets of φ. In fact such topological information can also be extracted from the minimax characterization of the corresponding critical values.

We start with some easy observations concerning local extrema.

In Proposition 6.2.3, we saw that if u_0 is a local minimum of φ, then

$$C_k(\varphi, u_0) = \delta_{k,0}\mathbb{Z} \text{ for all } k \in \mathbb{N}_0.$$

In the next proposition we complete this result.

Proposition 6.5.1 *If X is a reflexive Banach space, $\varphi \in C^1(X)$, φ satisfies the C-condition, $u_0 \in K_\varphi$ is isolated and $c_0 = \varphi(u_0)$ is isolated in $\varphi(K_\varphi)$, then the following statements are equivalent:*

(a) u_0 is local minimizer of φ;
(b) $C_k(\varphi, u_0) = \delta_{k,0}\mathbb{Z}$ for all $k \in \mathbb{N}_0$;
(c) $C_0(\varphi, u_0) \neq 0$.

Proof $(a) \Rightarrow (b)$: This implication is Proposition 6.2.3.

$(b) \Rightarrow (c)$: Obvious.

$(c) \Rightarrow (a)$: Arguing by contradiction suppose that u_0 is not a local minimizer of φ. By Lemma 6.2.35, we can find $a, b \in \mathbb{R}$ such that

$$a < c_0 < b \text{ and } K_\varphi \cap \varphi^{-1}([a, b]) = \{u_0\}.$$

Then Definition 6.2.1, Theorem 5.3.12 and Corollary 6.1.24 imply that

$$C_0(\varphi, u_0) = H_0(\varphi^b, \varphi^a) = H_0(\varphi^b, \varphi^{c_0}\setminus\{u_0\}).$$

Let $h : [0, 1] \times \varphi^b \to \varphi^b$ be the deformation into φ^{c_0} provided by Theorem 5.3.12 (the second deformation theorem). Then for any $u \in \varphi^b$, $h(\cdot, u)$ is a path in φ^b which connects u to $b(1, u) \in \varphi^{c_0}$.

Next note that we can find a small $r > 0$ such that $\varphi(y) < b$ for all $y \in B_r(u_0) = \{v \in X : ||v - u_0|| < r\}$. Since by hypothesis u_0 is not a local minimizer of φ, we can find $\hat{u} \in B_r(u_0)$ such that $\varphi(\hat{u}) < c_0$. Then $\gamma(t) = (1 - t)u_0 + t\hat{u}$, $t \in [0, 1]$ is a path connecting u_0 and $\hat{u}$ and staying in φ^b.

So, we have seen that every element $u \in \varphi^b$ can be connected to an element of $\varphi^{c_0}\setminus\{u_0\}$ by a path staying in φ^b. Then according to Remark 6.1.50 we have $C_0(\varphi, u_0) = H_0\left(\varphi^b, \varphi^{c_0}\setminus\ \{u_0\}\right) = 0$. So, u_0 is a local minimizer of φ. $\square$

Combining Proposition 6.5.1 and Lemma 6.2.38, we obtain:

Proposition 6.5.2 *If $\varphi \in C^2(\mathbb{R}^N)$ and $u_0 \in K_\varphi$ is isolated, then the following statements are equivalent:*

(a) u_0 is a local maximizer of φ;
(b) $C_k(\varphi, u_0) = \delta_{k,0}\mathbb{Z}$ for all $k \in \mathbb{N}_0$;
(c) $C_N(\varphi, u_0) \neq 0$.

In fact a similar result holds for C^2-functions on $\mathbb{R}^N$ and local maximizers (see Mawhin and Willem [293, p. 193]).

Proposition 6.5.3 *If $\varphi \in C^2(\mathbb{R}^N)$ and $u_0 \in K_\varphi$ is isolated, then the following statements are equivalent:*

(a) u_0 is local minimizer of φ;
(b) $C_k(\varphi, u_0) = \delta_{k,N}\mathbb{Z}$ for all $k \in \mathbb{N}_0$;
(c) $C_n(\varphi, u_0) \neq 0$.

Remark 6.5.4 So, if $\varphi \in C^2(\mathbb{R}^N)$ and $u \in K_\varphi$ is isolated and it is neither a local minimizer or a local maximizer, then $C_0(\varphi, u_0) = C_N(\varphi, u_0) = 0$.

Also, as a consequence of Lemma 6.2.38, we have:

Proposition 6.5.5 *If $\varphi \in C^2(\mathbb{R}^N)$ and $u_0 \in K_\varphi$ is isolated, then* rank $C_k(\varphi, u_0)$ *is finite for all $k \in \mathbb{N}_0$ and $C_k(\varphi, u_0) = 0$ for all $k \notin \{0, 1, \dots, N\}$.*

Next we recall a notion from Sect. 5.7 (see Definition 5.7.2):

Definition 6.5.6 Let X be a Banach space, $\varphi \in C^1(X)$ and $u \in K_\varphi$. We say that u is of "mountain pass type" if for any open neighborhood U of u, the set $\{v \in U : \varphi(v) < \varphi(u)\}$ is nonempty and not path connected.

Remark 6.5.7 In Theorem 5.7.7, we established that if $\varphi \in C^1(X)$ satisfies the C-condition and the mountain pass geometry and K_φ is discrete, then we can find $u \in K_\varphi^c$ which is of mountain pass type (recall $c = \inf_{\gamma\in\Gamma} \max_{0\leqslant t\leqslant 1} \varphi(\gamma(t))$, see Theorem 5.4.6).

In the next theorem we establish a useful property of the critical groups of a $u \in K_\varphi$ which is of mountain pass type.

Theorem 6.5.8 *If X is a reflexive Banach space, $\varphi \in C^1(X)$, $u_0 \in K_\varphi$ is isolated, $c_0 = \varphi(u_0)$ is isolated in $\varphi(K_\varphi)$ and u_0 of mountain pass type, then $C_1(\varphi, u_0) \neq 0$.*

Proof Let $\psi \in C^1(X)$ as postulated by Lemma 6.2.35. We know that $\varphi \leqslant \psi$ and $\varphi|_U = \psi|_U$ with U some open neighborhood of u_0.

Claim 1. *u_0 is a critical point of mountain pass type for the functional ψ.*

From Lemma 6.2.35 we know that $K_\varphi = K_\psi$ and so $u_0 \in K_\psi$. Let V be an open neighborhood of u_0 and define

$$\hat{U} = \{u \in V : \psi(u) < c_0\} \cup (U \cap V).$$

Then we have

$$\{u \in V : \psi(u) < c_0\} = \{u \in \hat{U} : \varphi(u) < c_0\}. \tag{6.104}$$

By hypothesis, $u_0 \in K_\varphi$ is of mountain pass type. So, from (6.104) and Definition 6.5.6 we have that

$$\begin{aligned} &\{u \in V : \psi(u) < c_0\} \text{ is nonempty and not path connected,} \\ \Rightarrow\ & u \text{ is of mountain pass type for } \psi. \end{aligned}$$

This proves Claim 1.

Using Lemma 6.2.35(d) and Claim 1, without any loss of generality we may assume that there are $a, b \in \mathbb{R}$ such that

$$a < c_0 < b \text{ and } K_\varphi \cap \varphi^{-1}([a, b]) = \{u_0\}.$$

Let C be the connected component of $U = \varphi^{-1}((a, b))$ which contains $u_0 \in K_\varphi$. Then C is open, path-connected, and contains u_0 and $K_\varphi \cap C = \{u_0\}$. Therefore from Definition 6.2.1, we have

$$C_1(\varphi, u_0) = H_1(C \cap \varphi^{c_0}, C \cap \varphi^{c_0} \backslash \{u_0\}). \tag{6.105}$$

From the second deformation theorem (see Theorem 5.3.12), we can find a deformation $\hat{h} : [0, 1] \times \varphi^b \to \varphi^b$ of φ^b into φ^{c_0} with the properties provided by Theorem 5.3.12. We have

$$\begin{aligned} &\hat{h}([0, 1] \times V) = V \\ \Rightarrow\ &\hat{h}([0, 1] \times C) = C \\ &\text{(recall that } \hat{h}([0, 1] \times C) \text{ is connected and contains C).} \end{aligned}$$

So, it follows that $\hat{h} : [0, 1] \times C \to C$ is a deformation into $C \cap \varphi^{c_0}$ and this means that $C \cap \varphi^{c_0}$ is a strong deformation retract of C. Then from (6.105) and Corollary 6.1.24(b), we have

$$C_1(\varphi, u_0) = H_1(C, C \cap \varphi^{c_0} \backslash \{u_0\}).$$

Using Axiom 4 in Definition 6.1.12 we have the exact sequence

$$C_1(\varphi, u_0) \xrightarrow{\partial} H_0(C \cap \varphi^{c_0} \backslash \{u_0\}) \to H_0(C) = \mathbb{Z} \text{ (see Remark 6.1.22).} \tag{6.106}$$

Let $C_0 = \{u \in C : \varphi(u) < c_0\}$, $d \in (a, c_0)$ and $\hat{h} : [0, 1] \times (\varphi^{c_0} \backslash \{u_0\}) \to \varphi^{c_0} \backslash \{u_0\}$ be the deformation into φ^d provided by the second deformation theorem (see Theorem 5.3.12).

Claim 2. $\tilde{h}([0,1] \times (C \cap \varphi^{c_0} \backslash \{u_0\})) \subseteq C \cap \varphi^{c_0} \backslash \{u_0\}$ *and* $\tilde{h}([0,1] \times C_0) \subseteq C_0$.

Take $u \in C \cap \varphi^{c_0} \backslash \{u_0\}$. Then $\tilde{h}([0,1] \times \{u\}) \subseteq V$, it is connected and intersects C (since it contains u). So, from the definition of C, we have

$$\begin{aligned} & \tilde{h}([0,1] \times \{u\}) \subseteq C \\ \Rightarrow\ & \tilde{h}([0,1] \times (C \cap \varphi^{c_0} \backslash \{u_0\})) \subseteq C \cap \varphi^{c_0} \backslash \{u_0\}. \end{aligned}$$

If $u \in C_0$, then $\varphi(\tilde{h}(t,u)) \leqslant \varphi(u) < c_0$ for all $t \in [0,1]$ (see Theorem 5.3.12). This proves Claim 2.

Then according to Claim 2, using $\tilde{h}(t,u)$ we have deformations of

$$C \cap \varphi^{c_0} \backslash \{u_0\} \text{ and } C_0 \text{ into } C \cap \varphi^d.$$

It follows that

$$\begin{aligned} & C \cap \varphi^d \text{ is a strong deformation retract of both } C \cap \varphi^c \backslash \{u_0\} \text{ and } C_0 \\ & \qquad\qquad\qquad\qquad\qquad\qquad \text{(see Theorem 5.3.12)}, \\ \Rightarrow\ & H_0(C \cap \varphi^c \backslash \{u_0\}) = H_0(C_0). \end{aligned} \tag{6.107}$$

Then from (6.106) and (6.107), we have

$$C_1(\varphi, u_0) \xrightarrow{\partial} H_0(C_0) \xrightarrow{\gamma} \mathbb{Z}. \tag{6.108}$$

Since u_0 is of mountain pass type, the set C_0 is nonempty and not path connected. Therefore

$$1 < \operatorname{rank} H_0(C_0).$$

Hence the homomorphism γ in (6.108) cannot be injective and so by the exactness of (6.108) it follows that $C_1(\varphi, u_0) \neq 0$. $\square$

In the context of Hilbert spaces and of C^2-functionals, we can improve Remark 6.5.7.

Proposition 6.5.9 *If H is a Hilbert space, $\varphi \in C^2(H)$, $u_0 \in K_\varphi$ is isolated with finite Morse index $m_0 = m(u_0)$ and finite nullity $\nu_0 = \nu(u_0) = \dim \ker \varphi''(u_0)$, when $m_0 = 0$ we have $\nu_0 \in \{0, 1\}$ and $C_1(\varphi, u_0) \neq 0$, then $C_k(\varphi, u_0) = \delta_{k,1}\mathbb{Z}$ for all $k \in \mathbb{N}_0$.*

Proof Let $\hat{\varphi} \in C^2(W)$ be as postulated by Proposition 6.2.9, with $W \subseteq \ker \varphi''(u_0)$ a neighborhood of the origin. Then from Theorem 6.2.13 (the shifting theorem), we have

$$C_k(\varphi, u_0) = C_{k-m_0}(\hat{\varphi}, 0) \text{ for all } k \in \mathbb{N}_0. \tag{6.109}$$

Since by hypothesis $C_1(\varphi, u_0) \neq 0$, it follows that $m_0 \in \{0, 1\}$.

Case 1. $m_0 = 1$.

Then from (6.109) we have $C_0(\hat{\varphi}, 0) \neq 0$ and so Proposition 6.5.2 implies that

$$C_k(\varphi, u_0) = \delta_{k,0}\mathbb{Z} \text{ for all } k \in \mathbb{N}_0.$$

Case 2. $m_0 = 0$.
Then from (6.109) and the hypothesis we have

$$C_1(\hat{\varphi}, 0) \neq 0.$$

By hypothesis $\dim\ker \varphi''(u_0) \leqslant 1$. Then from Proposition 6.2.5 we have

$$\begin{aligned} & C_k(\hat{\varphi}, 0) = \delta_{k,1}\mathbb{Z} \text{ for all } k \in \mathbb{N}_0 \\ \Rightarrow\ & C_k(\varphi, 0) = \delta_{k,1}\mathbb{Z} \text{ for all } k \in \mathbb{N}_0. \end{aligned}$$

The proof is now complete. □

Remark 6.5.10 The hypotheses of the above proposition imply that $\varphi''(u_0)$ is a Fredholm operator (see Definition 6.2.7).

Finally, combining Theorem 6.5.8 and Proposition 6.5.9, we can state the following theorem.

Theorem 6.5.11 *If H is a Hilbert space, $\varphi \in C^2(H)$, φ satisfies the C-condition, $u_0 \in K_\varphi$ is isolated and so is $c_0 = \varphi(u_0)$ in $\varphi(K_\varphi)$, the Morse index $m_0 = m(u_0)$ and the nullity $\nu_0 = \nu(u_0) = \dim\ker \varphi''(u_0)$ are finite and $m_0 = 0$ implies $\nu_0 \in \{0, 1\}$ and u_0 is of mountain pass type, then $C_k(\varphi, u_0) = \delta_{k,1}\mathbb{Z}$ for all $k \in \mathbb{N}_0$.*

6.6 Computation of Critical Groups

In this section we use Morse theory to compute the critical groups at certain particular critical points.

The notion of linking sets introduced in Definition 5.4.1 played a central role in the minimax theory of the critical values of a functional $\varphi \in C^1(X)$. Next, we introduce an analogous notion, which will help us to produce pairs of sublevel sets with nontrivial homology groups. From Proposition 6.2.15 we know that such pairs lead to a critical level between them.

Definition 6.6.1 Let X be a Banach space and E_0, E, D nonempty subsets of X such that

$$E_0 \subseteq E \text{ and } E_0 \cap D = \emptyset.$$

We say that the pair $\{E_0, E\}$ "homologically links" D in dimension m if the homomorphism $i_* : H_m(E, E_0) \to H_m(X, X \setminus D)$ induced by the corresponding inclusion of (E, E_0) into $(X, X \setminus D)$ is nontrivial.

Remark 6.6.2 (a) In the literature, the notion of linking introduced in Definition 5.4.1 is often called "homotopical linking" in order to distinguish it from the above concept of linking, called "homological linking".

(b) For every $m \in \mathbb{N}_0$ and $* \in E_0$, we have the following commutative diagram of homomorphisms

$$\begin{array}{ccc} H_m(E, E_0) & \xrightarrow{\partial_1} & H_{m-1}(E_0, *) \\ \Big\downarrow{\scriptstyle i_*} & & \Big\downarrow{\scriptstyle j_*} \\ H_m(X, X\backslash D) & \xrightarrow{\partial_2} & H_{m-1}(X\backslash D, *) \end{array}$$

with j_* being the homomorphism induced by the corresponding inclusion map. Suppose that E is contractible. Using the long exact sequence of Proposition 6.1.14 we see that the boundary homomorphisms ∂_1, ∂_2 are isomorphisms. So, it follows that

$$\begin{gathered} \text{“}\{E_0, E\} \text{ homologically links } D \text{ in dimension } m \\ \text{if and only if} \\ \text{the homomorphism } j_* \text{ is nontrivial''.} \end{gathered} \tag{6.110}$$

In Example 5.4.3, we introduced some triplets of sets $\{E_0, E, D\}$ which arise in the main minimax theorems and which are homotopically linking in the sense of Definition 5.4.1. In the sequel we show that these triplets are also homologically linking.

Proposition 6.6.3 *If X is a Banach space, $u_0 \in X$, U is a bounded open neighborhood of u_0, $u_1 \notin \bar{U}$, $E_0 = \{u_0, u_1\}$, $E = \{tu_0 + (1-t)u_1 : t \in [0, 1]\}$ and $D = \partial U$, then the pair $\{E_0, E\}$ homologically links D in dimension 1.*

Proof Let $j : (E_0, \{u_1\}) \to (X, \backslash D, \{u_1\})$ be the inclusion map and consider the map $r : (X\backslash D, \{u_1\}) \to (E_0, \{u_1\})$ defined by

$$r(u) = \begin{cases} u_0 \text{ if } u \in U \\ u_1 \text{ if } u \in X\backslash\bar{U} \end{cases} \quad \text{for all } u \in X\backslash D.$$

Then $r \circ j = \mathrm{id}_{(E_0,\{u_1\})}$ (here by $\mathrm{id}_{(E_0,\{u_1\})}$ we denote the identity map seen as a map of pairs (see Definition 6.1.1(b))). Hence $j_* : H_0(E_0, \{u_1\}) \to H_0(X\backslash D, \{u_1\})$ is injective. From Example 6.1.34(b), we have

$$\begin{aligned} & H_0(E_0, \{u_1\}) = \mathbb{Z} \\ \Rightarrow\ & j_* \text{ is nontrivial} \\ \Rightarrow\ & \{E_0, E\} \text{ homologically links } D \text{ in dimension 1 (see (6.110)).} \end{aligned}$$

The proof is now complete. ∎

Recall that for $\rho > 0$, we have

$$\partial B_\rho(0) = \{u \in X : ||u|| = \rho\} \text{ and } \bar{B}_\rho(0) = \{u \in X : ||u|| \leqslant \rho\}.$$

Proposition 6.6.4 *If X is a Banach space, $X = Y \oplus V$ with $dimY < \infty$, $E_0 = \partial B_\rho(0) \cap Y$, $E = \bar{B}_\rho(0) \cap Y$ and $D = V$, then the pair (E_0, E) homologically links D in dimension $d = \dim Y$.*

Proof From the proof of Proposition 6.2.30, we know that

$$E_0 \text{ is a strong deformation retract of } X\backslash D = X\backslash V.$$

So, we have

$$H_k(X\backslash D, E_0) = 0 \text{ for all } k \in \mathbb{N}_0 \text{ (see Proposition 6.1.15).}$$

Let $* \in E_0$ and consider the triple $\{*\} \subseteq E_0 \subseteq X\backslash D$. Using the long exact sequence from Proposition 6.1.29, we have that $j_* : H_{d-1}(E_0, *) = \mathbb{Z} \to H_{d-1}(X\backslash D, *)$ is an isomorphism, thus nontrivial. Again from (6.110) we conclude that the pair $\{E_0, E\}$ homologically links D in dimension $d = \dim Y$. □

In a similar fashion, we also establish the following propositions.

Proposition 6.6.5 *If X is a Banach space, $X = Y \oplus V$ with $\dim,\ Y < +\infty$, $v_0 \in V$ with $||v_0|| = 1$ $0 < \rho < r_1,\ 0 < r_2$,*

$$E_0 = \{y + \lambda v_0 : y \in Y, (0 < \lambda < r_1, ||y|| = r_2) \text{ or } (\lambda \in (0, r_1), ||y|| \leqslant r_2)\},$$
$$E = \{y + \lambda v_0 : y \in Y, 0 \leqslant \lambda \leqslant r_1, ||y|| \leqslant r_2\},$$
$$D = \partial B_\rho(0) \cap V,$$

then the pair $\{E_0, E\}$ homologically links D in dimension $d = \dim Y + 1$.

Homological linking is invariant under homeomorphisms.

Proposition 6.6.6 *If X is a Banach space, the pair $\{E_0, E\}$ homologically links D in dimension m and $h : X \to X$ is a homeomorphism, then the pair $(h(E), h(E_0))$ homologically links $h(D)$ in dimension m.*

Proof Just consider the following commutative diagram of homomorphisms

$$\begin{tikzcd} H_m(E, E_0) \arrow[r, "i_*"] \arrow[d, "h_*"] & H_m(X, X\backslash D) \arrow[d, "h_*"] \\ H_m(h(E), h(E_0)) \arrow[r, "j_*"] & H_m(X, X\backslash h(D)) \end{tikzcd}$$

The proof is now complete. □

Homological linking implies homotopical linking.

Proposition 6.6.7 *If X is a Banach and the pair (E_0, E) homotopically links D in dimension d, then the pair (E_0, E) homologically links D (that is, in the sense of Definition 5.4.1).*

Proof Since by Definition 6.6.1, i_* is nontrivial and the homology class $[\mathrm{id}_{(E,E_0)}]$ generates $H_m(E, E_0)$, we have that $i_*([\mathrm{id}_{(E,E_0)}]) \neq 0$ in $H_m(X, X\backslash D)$. So, there is no relative singular homology $(m+1)$-chain of $(X, X\backslash D)$ with boundary $\mathrm{id}_{(E,E_0)}$. Therefore there is no map $\gamma \in C(X, X\backslash D)$ such that $\gamma|_{E_0} = \mathrm{id}|_{E_0}$. □

Next we present a useful consequence of the notion of homological linking.

Proposition 6.6.8 *If X is a Banach space, the pair $\{E_0, E\}$ homologically links D in dimension m, $\varphi \in C^1(X)$ and $a < b \leqslant +\infty$ are such that*

$$\varphi|_{E_0} \leqslant a < \varphi|_D \text{ and } \sup_E \varphi \leqslant b,$$

then

(a) $H_m(\varphi^b, \varphi^a) \neq \emptyset$;
(b) *if in addition φ satisfies the C-condition, $a, b \notin \varphi(K_\varphi)$ and $K_\varphi \cap \varphi^{-1}((a,b))$ is finite then there exists a $u \in K_\varphi \cap \varphi^{-1}((a,b))$ such that $C_m(\varphi, u) \neq 0$.*

Proof (a) Consider the following inclusion maps of pairs of spaces

$$(E, E_0) \xrightarrow{j} (\varphi^b, \varphi^a) \xrightarrow{e} (X, X\backslash D). \tag{6.111}$$

We have $i_* = e_* \circ j_* : H_m(E, E_0) \to H_m(X, X\backslash D)$ and by hypothesis it is nontrivial. Therefore $j_* \neq 0$, $e_* \neq 0$. From (6.111) it follows that $H_m(\varphi^b, \varphi^a) \neq 0$.

(b) Follows from Theorem 6.2.20(b). □

Corollary 6.6.9 *If X is a Banach space, $\varphi \in C^1(X)$, φ satisfies the C-condition, K_φ is finite, $u_0, u_1 \in X$, $0 < \rho < ||u_1 - u_0||$ and*

$$c = \max\{\varphi(u_0), \varphi(u_1)\} < \inf\{\varphi(u) : ||u - u_0|| = \rho\} = d,$$

then there exists a $u \in K_\varphi$ with $d \leqslant \varphi(u)$ and $C_1(\varphi, u) \neq 0$.

Proof Choose $a \in (c, d)$ so that $[a, b)$ contains no critical values of φ. Let

$$E_0 = \{u_0, u_1\},\ E = \{tu_0 + (1-t)u : 0 \leqslant t \leqslant 1\} \text{ and } D = \partial B_\rho(0).$$

From Proposition 6.6.3 we know that the pair $\{E, E_0\}$ homologically links D in dimension 1. Using Proposition 6.6.8(b) (with $b = +\infty$), we can find

$$u \in K_\varphi, \varphi(u) > a \text{ and } C_1(\varphi, u) \neq 0.$$

Since $[a, b)$ contains no critical values of φ, we must have $\varphi(u) \geqslant d$. □

Remark 6.6.10 This corollary is essentially Theorem 6.5.8.

Proposition 6.6.11 *If X is a Banach space, $\varphi \in C^1(X)$, the pair $\{E, E_0\}$ homologically links D in dimension m and*

$$\varphi|_{E_0} \leqslant a < \varphi|_{X \setminus D} \tag{6.112}$$

then $H_m(X, \varphi^a) \neq 0$.

Proof We consider the following commutative diagram of homomorphisms

$$\begin{tikzcd} H_m(E, E_0) \arrow[r] \arrow[dr, "i_*"'] & H_m(X, \varphi^a) \arrow[d] & \text{(see (6.112))} \\ & H_m(X, X \setminus D) & \end{tikzcd}$$

By hypothesis, $i_* \neq 0$ (see Definition 6.6.1). It follows that

$$E_0 \text{ is a strong deformation retract of } X \setminus D.$$

Hence we have that i_* is an isomorphism and so

$$\operatorname{rank} i_* = \operatorname{rank} H_m(X, X \setminus D) = 1.$$

The proof is now complete. $\square$

Corollary 6.6.12 *If X is a Banach space, $\varphi \in C^1(X)$, φ satisfies the C-condition, the pair $\{E, E_0\}$ homologically links D in dimension m and*

$$\varphi|_{E_0} \leqslant a < \varphi|_{X \setminus D}, \ a < \inf \varphi(K_\varphi),$$

then $C_m(\varphi, \infty) \neq 0$.

In Definition 5.4.14 we introduced the notion of local linking, which is important in many variational problems. In Remark 5.4.15 we observed that this condition implies that $u = 0$ is a critical point of the functional. So, we would like to compute its critical groups. We will do this as a consequence of our analysis of a more general notion called "homological local linking".

Definition 6.6.13 Let X be a Banach space, $\varphi \in C^1(X)$, $\varphi(0) = 0$ and $0 \in K_\varphi$ be isolated. Let $m, n \in \mathbb{N}$. We say that φ has a "local (m, n)-linking" near the origin if there is a neighborhood U of the origin and nonempty sets $E_0 \subseteq E \subseteq U$, $D \subseteq X$ such that $E_0 \cap D = \emptyset$ and

(a) $\varphi^0 \cap U \cap K_\varphi = \{0\}$;
(b) $\operatorname{rank} i_* - \operatorname{rank} \operatorname{im} j_* \geqslant n$, where $i_* : H_{m-1}(E_0) \to H_{m-1}(X \setminus D)$ and $j_* : H_{m-1}(E_0) \to H_{m-1}(E)$ are the homomorphisms induced by the inclusions $i : E_0 \to X \setminus D$ and $j : E_0 \to E$;

(c) $\varphi|_E \leqslant 0 < \varphi|_{U \cap D \setminus \{0\}}$.

We want to know how this notion is related to the concept of local linking (see Definition 5.4.14). To do this, we will need the following deformation lemma.

Lemma 6.6.14 *If X is a Banach space, $X = Y \oplus V$ with $d = \dim Y < +\infty$, $\varphi \in C^1(X)$ has a local linking at 0 (with respect to the pair (Y, V)), that is, there exists an $r > 0$ such that*

$$\varphi(u) \leqslant 0 \text{ if } u \in Y, ||u|| \leqslant r,$$
$$\varphi(u) \geqslant 0 \text{ if } u \in V, ||u|| \leqslant r,$$

$0 \in K_\varphi$ is isolated and one of the following conditions holds

(i) 0 is a strict local minimizer of $\varphi|_V$; or
(ii) $X = H =$ a Hilbert space and φ' is Lipschitz near 0,

then there exist $\rho \in (0, r)$ and a homeomorphism $h : X \to X$, $h(0) = 0$ such that

(a) $h(\bar{B}_\rho(0)) \subseteq B_r(0)$;
(b) $h(u) = u$ for all $u \in Y \cap \bar{B}_\rho(0)$;
(c) $\varphi(u) > 0$ for all $u \in h(V \cap \bar{B}_\rho(0)),\ u \neq 0$.

Proof Suppose that hypothesis (i) holds. Then we can take $\rho \in (0, r)$ such that

$$0 < \varphi(u) \text{ for all } u \in V \cap \bar{B}_\rho(0), u \neq 0 \text{ and } h = \mathrm{id}_X$$

So, suppose that hypothesis (ii) holds. Let $0 < \rho_1 < \rho_2 < r$ be such that

$$K_\varphi \cap B_{\rho_1}(0) = \{0\} \text{ and } \varphi'|_{B_{\rho_2}(0)} \text{ is Lipschitz.}$$

Take $\rho \in (0, \rho_1)$. Then the sets $\bar{B}_\rho(0)$ and $H \setminus B_{\rho_1}(0)$ are closed and disjoint. Consider the function $f : H \to [0, 1]$ defined by

$$f(u) = \frac{d(u, H \setminus \bar{B}_{\rho_1}(0))}{d(u, \bar{B}_\rho(0)) + d(u, H \setminus \bar{B}_{\rho_1}(0))} \text{ for all } u \in H.$$

Evidently, $f(\cdot)$ is locally Lipschitz and $f|_{\bar{B}_\rho(0)} = 1,\ f|_{H \setminus B_{\rho_1}(0)} = 0$. Let $p_V : H \to V$ the projection operator onto V and consider the map $\xi : H \to H$ defined by

$$\xi(u) = f(u)||p_V(u)||\,\nabla\varphi(u) \text{ for all } u \in H. \tag{6.113}$$

Clearly, $\xi(\cdot)$ is Lipschitz and bounded. We consider the abstract Cauchy problem defined by

$$\frac{d\sigma(t)}{dt} = \xi(\sigma(t)) \text{ for } t \geqslant 0,\ \xi(t_0) = u\ (t_0 \geqslant 0). \tag{6.114}$$

From Proposition 5.3.5 we know that problem (6.114) admits a unique global solution $\sigma(t_0, u) : [0, +\infty) \to H$. Consider the maps $h, l : H \to H$ defined by

$$h(u) = \sigma(0, u)(1) \text{ and } l(u) = \sigma(1, u)(0).$$

The continuous dependence of the flow on the initial condition (see Proposition 5.3.5) implies that $h(\cdot)$ and $l(\cdot)$ are both continuous. We have

$$\begin{aligned} & h \circ l = l \circ h = \mathrm{id}_H \\ \Rightarrow\ & h \text{ is a homeomorphism.} \end{aligned}$$

Note that $h(0) = 0$. Also, we have:

(a) If $u \in H \setminus B_{\rho_1}(0)$, then $\xi(u) = 0$ and so $h(u) = u$. Therefore $h(H \setminus \bar{B}_{\rho_1}(0)) = H \setminus \bar{B}_{\rho_1}(0)$. This mean that $h(\bar{B}_\rho(0)) \subseteq h(B_{\rho_1}(0)) \subseteq B_{\rho_1}(0) \subseteq B_r(0)$. This proves part (a) of the lemma.
(b) If $u \in Y$, then $\xi(u) = 0$ (see (6.113)) and so $h(u) = u$, which proves part (b) of the lemma.
(c) If $u \in V \cap \bar{B}_\rho(0)$, then

$$\varphi(h(u)) = \varphi(u) + \int_0^1 \vartheta(t)dt,$$

where $\vartheta(t) = f(\sigma(0, u)(t))||p_V \sigma(0, u)(t)||\ ||\varphi'(\sigma(0, u)(t))||^2$. Evidently, $\vartheta \geqslant 0$ and since $f(u) = 1,\ p_V(u) = u, u \notin K_\varphi$, we have

$$\begin{aligned} & \vartheta(0) = ||u||\ ||\nabla\varphi(u)||^2 > 0 \\ \Rightarrow\ & \varphi(h(u)) > \varphi(u) \geqslant 0. \end{aligned}$$

This proves part (c) of the lemma and completes the proof. □

Using this lemma, we obtain a precise relation between the notions of local linking and of homological local linking.

Proposition 6.6.15 *If X is a Banach space, $X = Y \oplus V$ with $d = \dim Y < \infty$, $\varphi \in C^1(X)$ has a local linking set at 0 (with respect to the pair (Y, V)), that is, there exists an $r > 0$ such that*

$$\begin{aligned} & \varphi(u) \leqslant 0 \textit{ if } u \in Y, ||u|| \leqslant r, \\ & \varphi(u) \geqslant 0 \textit{ if } u \in V, ||u|| \leqslant r, \end{aligned}$$

$0 \in K_\varphi$ is isolated and one of the following conditions holds

(i) 0 is a strict local minimizer of $\varphi|_V$; or
(ii) $X = H =$ a Hilbert space and φ' is Lipschitz near 0,

then φ has local $(d, 1)$-linking at 0.

Proof By taking $r > 0$ even smaller if necessary, we may assume that $K_\varphi \cap B_r(0) = \{0\}$. Using Lemma 6.6.14, we can find $\rho \in (0, r)$ and a homeomorphism $h : X \to X$ which have properties (a), (b), (c) from Lemma 6.6.14. We set

$$U = h(\bar{B}_\rho(0))\ E_0 = Y \cap \partial B_\rho(0),\ \ E = Y \cap \bar{B}_\rho(0) \text{ and } D = h(V).$$

Then conditions (a) and (c) in Definition 6.6.13 follow from the above choices, the local linking property and part (c) of Lemma 6.6.14. So, we need to verify property (b) in Definition 6.6.13.

From the proof of Proposition 6.2.30, we know that

$$\begin{aligned} &E_0 \text{ is a strong deformation retract of } X \backslash D = h(X \backslash V) \\ \Rightarrow\ & i_* : H_{d-1}(E_0) \to H_{d-1}(X \backslash D) \text{ is a bijection.} \end{aligned} \tag{6.115}$$

Also, from Example 6.1.34(b), we have

$$\text{rank im } i_* = \text{rank } H_{d-1}(E) = \begin{cases} 2 \text{ if } d = 1 \\ 1 \text{ if } d \geqslant 2. \end{cases} \tag{6.116}$$

Since E is contractible, using Proposition 6.1.29, we see that

$$H_{d-1}(E, E_0) = H_{d-2}(E_0, *) = 0\ (* \in E_0).$$

Axiom 4 in Definition 6.1.12 implies that

$$j_* : H_{d-1}(E_0) \to H_{d-1}(E) \text{ is surjective.}$$

Therefore

$$\text{rank im } j_* = \text{rank } H_{d-1}(E) = \begin{cases} 1 \text{ if } d = 1 \\ 0 \text{ if } d \geqslant 2. \end{cases} \tag{6.117}$$

From (6.116) and (6.117) we see that

$$\text{rank im } i_* - \text{rank im } j_* = 1.$$

So, property (b) in Definition 6.6.13 is satisfied. This completes the proof of the proposition. ☐

Remark 6.6.16 Homological local linking in general does not imply local linking. Consider the function $\varphi : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$\varphi(u) = x^3 - 3xy^2 \text{ for all } u = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2.$$

Then φ has a local $(1, 2)$-linking near 0, but it does not have a local linking at 0.

Now we estimate the critical groups of $\varphi \in C^1(X)$, when it exhibits homological linking at 0.

Theorem 6.6.17 *If X is a Banach space, $\varphi \in C^1(X)$, $0 \in K_\varphi$ is isolated and φ has a local (m, n)-linking near the origin, then* $\operatorname{rank} C_m(\varphi, 0) \geqslant n$.

Proof Let U, E_0, E and D be as postulated by Definition 6.6.13. From Definition 6.2.1 we have

$$C_m(\varphi, 0) = H_m(\varphi^0 \cap U, \varphi^0 \cap U \backslash \{0\}).$$

Consider the exact chain

$$C_m(\varphi, 0) \xrightarrow{\partial_*} H_{m-1}(\varphi^0 \cap U \backslash \{0\}) \xrightarrow{e_*} H_{m-1}(\varphi^0 \cap U), \tag{6.118}$$

with e_* being the homomorphism induced by the inclusion $e : \varphi^0 \cap U \backslash \{0\} \to \varphi^0 \cap U$. From (6.118) and the rank formula, we have

$$\operatorname{rank} \ker e_* = \operatorname{rank} \operatorname{im} \partial_* \leqslant \operatorname{rank} C_m(\varphi, 0). \tag{6.119}$$

Using Definition 6.6.13, we have the following commutative diagram

$$\begin{tikzcd}
H_{m-1}(X \backslash D) \arrow[r, "(i_*)"] \arrow[dr, "l_*"'] & H_{m-1}(E_0) \arrow[r, "j_*"] \arrow[d, "\eta_*"] & H_{m-1}(E) \arrow[d, "\partial_*"] \\
& H_{m-1}(\varphi^0 \cap U \backslash \{0\}) \arrow[r, "e_*"] & H_{m-1}(\varphi^0 \cap U)
\end{tikzcd}$$

Here l_*, η_*, ∂_* are the homomorphisms induced by the corresponding inclusion maps. Then we have

$$\operatorname{rank} \operatorname{im} i_* = \operatorname{rank} \operatorname{im} l_* \leqslant \operatorname{rank} \operatorname{im} \eta_*, \tag{6.120}$$

$$\begin{aligned} &\operatorname{rank} \operatorname{im} \eta_* - \operatorname{rank} \ker e_* \leqslant \operatorname{rank} \operatorname{im} \eta_* - \operatorname{rank} \ker e_*|_{\operatorname{im} \eta_*} = \\ &\qquad \operatorname{rank} \operatorname{im} (e_* \circ \eta_*) \leqslant \operatorname{rank} \operatorname{im} j_* \quad \text{(use the rank formula)}. \end{aligned} \tag{6.121}$$

From (6.119), (6.120), (6.121) it follows that

$$n \leqslant \operatorname{rank} \operatorname{im} i_* - \operatorname{rank} \operatorname{im} j_* \leqslant \dim \ker e_* \leqslant \operatorname{rank} C_m(\varphi, 0).$$

The proof is now complete. $\square$

Corollary 6.6.18 *If X is a Banach space, $X = Y \oplus V$ with $d = \dim Y < \infty$, $\varphi \in C^1(X)$, φ has a local linking at 0, $0 \in K_\varphi$ is isolated and one of the conditions (i) and (ii) from Proposition 6.6.15 holds, then $C_d(\varphi, 0) \neq 0$.*

We can improve this corollary by restricting ourselves to Hilbert spaces and to C^2-functionals.

Proposition 6.6.19 *If H is a Hilbert space, $H = \bar{H} \oplus \hat{H}$ with $d = \dim \bar{H}$, $\varphi \in C^2(H)$, φ has a local linking at 0 (with respect to the pair $(\bar{H}, \hat{H})$), $0 \in K_\varphi$ is isolated with Morse index m_0 and nullity ν_0 (that is, $\dim \ker \varphi''(0) = \nu_0$ and $\varphi''(0)$ is Fredholm), then $C_k(\varphi, 0) = \begin{cases} \delta_{k,m_0}\mathbb{Z} \text{ if } d = m_0 \\ \delta_{k,m_0}\mathbb{Z} \text{ if } d = m_0 + \nu_0 \end{cases}$ for all $k \in \mathbb{N}$.*

Proof From the shifting theorem (see Theorem 6.2.13), we have

$$C_k(\varphi, 0) = C_{k-m_0}(\hat{\varphi}, 0) \text{ for all } k \in \mathbb{N}_0,$$

where $\hat{\varphi}$ is as in Proposition 6.2.9. From Corollary 6.6.18, we have

$$C_d(\varphi, 0) \neq 0.$$

If $d = m_0$, then 0 is a local minimizer of $\hat{\varphi}$ and so

$$C_k(\varphi, 0) = C_{k-m_0}(\hat{\varphi}, 0) = \delta_{k-m_0,0}\mathbb{Z} = \delta_{k,m_0}\mathbb{Z} \text{ for all } k \in \mathbb{N}_0.$$

If $d = m_0 + \nu_0$, then 0 is a local maximizer of $\hat{\varphi}$ and so

$$C_k(\varphi, 0) = C_{k,m_0}(\hat{\varphi}, 0) = \delta_{k-m_0,\nu_0}\mathbb{Z} = \delta_{k,m_0+\nu_0}\mathbb{Z} \text{ for all } k \in \mathbb{N}_0.$$

The proof is now complete. □

The next result is a useful tool in the computation of critical groups at infinity.

Proposition 6.6.20 *If X is a Banach space, $(t, u) \to h_t(u)$ is a function in $C^1([0, 1] \times X)$, the maps $u \to (h_t)'(u)$ and $t \to \partial_t h_t(u)$ are both locally Lipschitz, h_0 and h_1 satisfy the C-condition,*

$$|\partial_t h_t(u)| \leqslant c_0(||u||^q + ||u||^p) \text{ for all } u \in X$$

with $c_0 > 0$, $1 < q < p < \infty$ and there exist $\gamma_0 > 0$ and $\delta_0 > 0$ such that

$$h_t(u) \leqslant \gamma_0 \Rightarrow (1 + ||u||)||(h_t)'(u)||_* \geqslant \delta_0(||u||^q + ||u||^p) \text{ for all } t \in [0, 1],$$

then $C_k(h_0, \infty) = C_k(h_1, \infty)$ for all $k \in \mathbb{N}_0$.

Proof Since by hypothesis $(t, u) \to h_t(u)$ belongs to the space $C^1([0, 1] \times X)$, it admits a pseudogradient vector field $\hat{v}_t(u)$ (see Theorem 5.1.4). Moreover, from the construction of the pseudogradient vector field (see the proof of Theorem 5.1.4), we have

$$\hat{v}_t(u) = (\partial_t h_t(u), v_t(u)),$$

with $(t, u) \to v_t(u)$ locally Lipschitz and for all $t \in [0, 1]$, $v_t(\cdot)$ is the pseudogradient vector field corresponding to $h_t(\dot{})$. So, for all $t \in [0, 1]$ and all $u \in X$, we have

$$||(h_t)'(u)||_*^2 \leqslant \left\langle (h_t)'(u), v_t(u) \right\rangle \text{ and } ||v_t(u)|| \leqslant 2||(h_t)'(u)||_*. \tag{6.122}$$

Given $t \in [0, 1]$, we consider the map $w_t : X \to X$ defined by

$$w_t(u) = -\frac{|\partial_t h_t(u)|}{||(h_t)'(u)||_*} v_t(u) \text{ for all } u \in X.$$

Clearly, this is a well-defined vector field and $(t, u) \to w_t(u)$ is locally Lipschitz. Let $\gamma \leqslant \gamma_0$ be such that

$$h_0^\gamma \neq 0 \text{ or } h_t^\gamma \neq 0.$$

If no such $\gamma \leqslant \gamma_0$ can be found, then $C_k(h_0, \infty) = C_k(h_1, \infty) = \delta_{k,0}\mathbb{Z}$ for all $k \in \mathbb{N}_0$.

So, to fix things we assume that $h_0^\gamma \neq 0$. Let $y \in h_0^\gamma$ and consider the following abstract Cauchy problem

$$\frac{d\sigma}{dt} = w_t(\sigma) \text{ on } [0, 1],\ \sigma(0) = y. \tag{6.123}$$

This Cauchy problem admits a unique local flow (see Proposition 5.3.4) denoted by $\sigma(t, y)$. In the sequel, for notational simplicity, from σ we drop the initial condition y. We have

$$\begin{aligned}
\frac{d}{dt} h_t(\sigma) &= \left\langle (h_t)'(\sigma), \frac{d\sigma}{dt} \right\rangle + \partial_t h_t(\sigma) \text{ (by the chain rule)} \\
&= \left\langle (h_t)'(\sigma), \frac{-|\partial_t h_t(\sigma)|}{||(h_t)'(\sigma)||_*^2} v_t(\sigma) \right\rangle + \partial_t h_t(\sigma) \text{ (see (6.123))} \\
&\leqslant -|\partial_t h_t(\sigma)| + \partial_t h_t(\sigma) \text{ (see (6.122))} \\
&\leqslant 0 \\
\Rightarrow \quad & t \to h_t(\sigma) \text{ is nonincreasing.}
\end{aligned}$$

Hence for small $t \geqslant 0$, we have

$$\begin{aligned}
& h_t(\sigma(t)) \leqslant h_0(\sigma(0)) = h_0(y) \leqslant \gamma \leqslant \gamma_0 \\
\Rightarrow\ & (1 + ||\sigma(t)||)||(h_t)'(\sigma(t))||_* \geqslant \delta_0(||\sigma(t)||^q + ||\sigma(t)||^p).
\end{aligned} \tag{6.124}$$

Then

$$\begin{aligned}
|w_t(\sigma(t))| &\leqslant \frac{|\partial_t h_t(\sigma(t))|}{||(h_t)'(\sigma(t))||_*^2}||v_t(\sigma(t))|| \\
&\leqslant \frac{c_0(||\sigma(t)||^q + ||\sigma(t)||^p)}{||(h_t)'(\sigma(t))||_*^2}\, 2||(h_t)'(\sigma(t))||_* \text{ (see (6.122))} \\
&\leqslant \frac{c_0(||\sigma(t)||^q + ||\sigma(t)||^p)}{\delta_0(||\sigma(t)||^q + ||\sigma(t)||^p)}\,(1+||\sigma(t)||) \text{ (see (6.124))} \\
&= \frac{c_0}{\delta_0}(1+||\sigma(t)||) \text{ for all small } t \in [0,1]
\end{aligned}$$

$\Rightarrow$ the flow $\sigma(\cdot)$ is global on $[0, 1]$.

We have that $\sigma(t, \cdot)$ is a homeomorphism of h_0^γ onto a subset D_0 of h_1^γ. Reversing the time $(t \to 1-t)$ and using the corresponding flow $\sigma_*(\cdot, v)(v \in h_1^\gamma)$, we have that h_1^γ is homeomorphic to a subset D_1 of h_0^γ.

We set

$$\eta(t, y) = \sigma_*(t, \sigma(1, y)) \text{ for all } (t, y) \in [0,1] \times h_0^\gamma.$$

Then we have

$$\eta(0, \cdot) \text{ is homotopy equivalent to } \mathrm{id}|_{D_0}(\cdot), \tag{6.125}$$

$$\eta(1, \cdot) = (\sigma_*)_1 \circ \sigma_1(\cdot). \tag{6.126}$$

Similarly, if

$$\eta_*(t, v) = \sigma(t, \sigma_*(1, v)) \text{ for all } (t, v) \in [0,1] \times h_1^\gamma,$$

then

$$\eta(0, \cdot) \text{ is homotopy equivalent to } \mathrm{id}|_{D_1}(\cdot), \tag{6.127}$$

$$\eta(1, \cdot) = \sigma_1 \circ (\sigma_*)_1(\cdot). \tag{6.128}$$

Recall that D_0 and h_0^γ are homeomorphic. Similarly D_1 and h_1^γ are homeomorphic too. These facts together with (6.125), (6.126), (6.127), (6.128) imply that

$$\begin{aligned}
& h_0^\gamma \text{ and } h_1^\gamma \text{ are homotopy equivalent} \\
\Rightarrow\ & H_k(X, h_0^\gamma) = H_k(X, h_1^\gamma) \text{ for all } k \in \mathbb{N}_0 \text{ (see Proposition 6.1.14)} \\
\Rightarrow\ & C_k(h_0, \infty) = C_k(h_1, \infty) \text{ for all } k \in \mathbb{N}_0.
\end{aligned}$$

The proof is now complete. □

From Definition 6.2.1 it is clear that if $\varphi \in C^1(X)$ and $u \in K_\varphi$ is isolated, then the critical groups of φ at u depend only on the values of φ near u. Now suppose that $\Omega \subseteq \mathbb{R}^N$ is a bounded domain with a C^2-boundary $\partial\Omega$ and let $f : \Omega \times \mathbb{R} \to \mathbb{R}$

be a Carathéodory function, that is, for all $x \in \mathbb{R}$ $z \to f(z,x)$ is measurable and for almost all $z \in \Omega$, $x \to f(z,x)$ is continuous. We assume that

$$|f(z,x)| \leqslant a(z)(1+|x|^{r-1}) \text{ for almost all } z \in \Omega \text{ and all } x \in \mathbb{R}, \tag{6.129}$$

with $a \in L^\infty(\Omega)_+$, $2 \leqslant r < 2^*$ (recall that $2^* = \begin{cases} \frac{2N}{N-2} & \text{if } N \geqslant 3, \\ +\infty & \text{if } N = 1,2 \end{cases}$, the Sobolev critical exponent, see Definition 1.9.1). We set $F(z,x) = \int_0^x f(z,s)ds$ and consider the C^1-functional $\varphi : H_0^1(\Omega) \to \mathbb{R}$ defined by

$$\varphi(u) = \frac{1}{2}||Du||_2^2 - \int_\Omega F(z,u(z))dz \text{ for all } u \in H_0^1(\Omega).$$

We recall that for $N \geqslant 2$, the Sobolev space is not embedded into $L^\infty(\Omega)$ (see Theorem 1.7.4, the Rellich–Kondrachov embedding theorem). So, if $u_0 \in K_\varphi$ is isolated, then a priori it seems that the critical groups of φ depend on values of $f(z,\cdot)$ far away from $u_0(z)$. We will show that in fact this is not true, establishing in an emphatic way the local character of critical groups.

First we show that without any loss of generality, we may assume that $u_0 = 0$. Indeed, let

$$\begin{aligned} \hat{\varphi}(u) &= \varphi(u+u_0) - \varphi(u_0) \\ &= \frac{1}{2}||Du||_2^2 + \int_\Omega (Du, Du_0)_{\mathbb{R}^N} dz - \int_\Omega (F(z,u+u_0) - F(z,u_0))dz \\ &= \frac{1}{2}||Du||_2^2 - \int_\Omega [F(z,u+u_0) - F(z,u_0) - f(z,u_0)u]dz \text{ (since } u_0 \in K_\varphi). \end{aligned} \tag{6.130}$$

We set

$$g(z,x) = f(z,x+u_0(z)) - f(z,u_0(z)) \text{ for all } (z,x) \in \Omega \times \mathbb{R}.$$

Evidently, this is a Carathéodory function. Also, since $u_0 \in K_\varphi$, by the standard regularity theory for semilinear elliptic problems, we have $u_0 \in L^\infty(\Omega)$. So, it follows that $g(z,\cdot)$ has the same polynomial growth as $f(z,\cdot)$ (see (6.129)). We set $G(z,x) = \int_0^x g(z,s)ds$. Then

$$G(z,u(z)) = F(z,(u+u_0)(z)) - F(z,u_0(z)) - f(z,u_0(z))u(z).$$

Therefore we can write (6.130) as follows:

$$\hat{\varphi}(u) = \frac{1}{2}||Du||_2^2 - \int_\Omega G(z,u(z))dz \text{ for all } u \in H_0^1(\Omega).$$

Evidently, $u = 0 \in K_{\hat{\varphi}}$ and it is isolated.

So, we have seen that without any loss of generality, we may assume that $u_0 = 0$.

Also, let $\delta > 0$ and consider a function $\xi \in C^1(\mathbb{R})$ defined by

$$\xi(x) = \begin{cases} -\delta \text{ if } x \leqslant -\delta \\ x \quad \text{if } -\frac{\delta}{2} \leqslant x \leqslant \frac{\delta}{2} \\ \delta \quad \text{if } \delta \leqslant x. \end{cases}$$

Let $\psi : H_0^1(\Omega) \to \mathbb{R}$ be the C^1-functional defined by

$$\psi(u) = \frac{1}{2}||Du||_2^2 - \int_\Omega F(z, \xi(u(z)))dz \text{ for all } u \in H_0^1(\Omega).$$

In the next lemma, we compare the critical groups of φ and ψ.

Lemma 6.6.21 *If* $0 \in K_\varphi$ *is isolated, then 0 is an isolated critical point of* ψ *too and we have*

$$C_k(\varphi, 0) = C_k(\psi, 0) \textit{ for all } k \in \mathbb{N}_0.$$

Proof We consider the following family of functions $h_t(u)$ defined on $[0, 1] \times H_0^1(\Omega)$

$$h_t(u) = \frac{1}{2}||Du||_2^2 - \int_\Omega F(z, (1-t)u(z) + t\xi(u(z)))dz \text{ for all } t \in [0, 1] \text{ and all } u \in H_0^1(\Omega).$$

Evidently, $h_0(u) = \varphi(u)$ and $h_1(u) = \psi(u)$ for all $u \in H_0^1(\Omega)$.

We will show that $0 \in K_{h_t}$ is isolated uniformly in $t \in [0, 1]$. Arguing by contradiction, suppose we can find $\{t_n\}_{n\geqslant 1} \subseteq [0, 1]$ and $\{u_n\}_{n\geqslant 1} \subseteq H_0^1(\Omega)$ such that

$$t_n \to t \text{ in } [0, 1],\ u_n \to 0 \text{ in } H^1(\Omega) \text{ and } h'_{t_n}(u_n) = 0 \text{ for all } n \in \mathbb{N}. \tag{6.131}$$

From (6.131) we have for all $n \in \mathbb{N}$

$$-\Delta u_n(z) = f(z, (1-t_n)u_n(z) + t_n\xi(u_n(z))) \tag{6.132}$$
$$\text{for almost all } z \in \Omega,\ u_n|_{\partial\Omega} = 0.$$

From (6.132) and the regularity theory for semilinear elliptic equations (the Calderon–Zygmund estimates), we can find $\alpha \in (0, 1)$ and $M > 0$ such that

$$u_n \in C_0^{1,\alpha}(\overline{\Omega}) \text{ and } ||u_n||_{C_0^{1,\alpha}(\overline{\Omega})} \leqslant M \text{ for all } n \in \mathbb{N}. \tag{6.133}$$

Exploiting the compact embedding of $C_0^{1,\alpha}(\overline{\Omega})$ into $C_0^1(\overline{\Omega})$, from (6.131) and (6.133) we have

$$u_n \to 0 \text{ in } C^1(\overline{\Omega}) \text{ as } n \to \infty. \tag{6.134}$$

So, we can find $n_0 \in \mathbb{N}$ such that

$$|u_n(z)| \leqslant \delta/2 \text{ for all } n \geqslant n_0 \text{ and all } z \in \overline{\Omega}$$
$$\Rightarrow \varphi'(u_n) = 0 \text{ for all } n \geqslant n_0,$$

a contradiction to our hypothesis that $0 \in K_\varphi$ is isolated (see (6.134)).

Having the isolation of the critical point $u = 0$ for the family $\{h_t(\cdot)\}_{t\in[0,1]}$ we conclude that

$$0 \in K_\varphi \text{ is isolated and } C_k(\varphi, 0) = C_k(\psi, 0) \text{ for all } k \in \mathbb{N}_0 \text{ (see Theorem 6.3.6).}$$

The proof is now complete. □

This lemma leads to the following theorem, stressing the really local character of critical groups. Its proof is immediate from Lemma 6.6.21 and the previous observations.

So, let $f, g : \Omega \times \mathbb{R} \to \mathbb{R}$ be two Carathéodory functions satisfying

$$|f(z, x)|, |g(z, x)| \leqslant a(z)(1 + |x|^{r-1}) \text{ for almost all } z \in \Omega \text{ and all } x \in \mathbb{R},$$

with $a \in L^\infty(\Omega)$, $2, \leqslant r < 2^*$. We set

$$F(z, x) = \int_0^x f(z, s)ds \text{ and } G(z, x) = \int_0^x g(z, s)ds$$

and consider the C^1-functionals $\varphi, \psi : H_0^1(\Omega) \to \mathbb{R}$ defined by

$$\begin{aligned} \varphi(u) &= \frac{1}{2}||Du||_2^2 - \int_\Omega F(z, u(z))dz, \\ \psi(u) &= \frac{1}{2}||Du||_2^2 - \int_\Omega G(z, u(z))dz \text{ for all } u \in H_0^1(\Omega). \end{aligned}$$

Theorem 6.6.22 *If* $0 \in K_\varphi$ *is isolated and there exists a* $\delta > 0$ *such that*

$$f(z, x + u_0(z)) = g(z, x + u_0(z)) \textit{ for almost all } z \in \Omega \textit{ and all } |x| \leqslant \delta,$$

then $0 \in K_\psi$ *is isolated too and* $C_k(\varphi, u_0) = C_k(\psi, u_0)$ *for all* $k \in \mathbb{N}_0$.

Remark 6.6.23 Evidently, this theorem is also valid for problems other than the Dirichlet problem, such as the Neumann or Robin problems. Then $H_0^1(\Omega)$ is replaced by $H^1(\Omega)$.

From Palais [325] (Theorem 16), we have the following result.

Theorem 6.6.24 *If* V_1, V_2 *are two paracompact locally convex topological vector spaces,* $j : V_1 \to V_2$ *is a continuous, linear map,* $j(V_1)$ *is dense in* V_2*,* $W \subseteq V_2$ *is open and* $U = j^{-1}(W)$*, then* $\hat{j} = j|_U : U \to W$ *is a homotopy equivalence.*

Remark 6.6.25 Using this general result, we see that if X is a Banach space which is embedded continuously and densely into a Hilbert space H, then for any pair of open sets (D, E) in H, we have

$$H_k(D, E) = H_k(D \cap X, E \cap X) \text{ for all } k \in \mathbb{N}_0.$$

This equality leads to the following useful result.

Theorem 6.6.26 *If H is a Hilbert space, X is a Banach space which is continuously and densely embedded in H, $\varphi \in C^2(H)$ and $u \in K_\varphi^c$ is isolated, then $C_k(\varphi, u) = C_k(\varphi|_X, u)$ for all $k \in \mathbb{N}_0$.*

6.7 Existence and Multiplicity of Critical Points

In this section we use critical groups to establish the existence and multiplicity of critical points.

We start with some auxiliary results related to the so-called "Lyapunov–Schmidt reduction method". With this method the initial infinite-dimensional problem is reduced to a finite-dimensional one which is easier to deal with. In Volume 2 we will see that this method, under some reasonable hypotheses on the data of the problem, is very effective in dealing with resonant equations.

The setting is the following. We have H, a separable Hilbert space with H^* its topological dual, and $\langle \cdot, \cdot \rangle$, the duality brackets for the pair (H^*, H). We assume that H admits the following orthogonal direct sum decomposition

$$H = Y \oplus V \text{ with } \dim Y < +\infty.$$

So, every $u \in H$ admits a unique decomposition

$$u = y + v \text{ with } y \in Y, v \in V.$$

Proposition 6.7.1 *If $\varphi \in C^1(H)$, φ is sequentially weakly lower semicontinuous and*

$$\hat{c}||v - v'||^2 \leqslant \left\langle \varphi'(y + v) - \varphi'(y + v'), v - v' \right\rangle$$

for all $y \in Y$, all $v, v' \in V$ and some $\hat{c} > 0$, then there exists a continuous map $\vartheta : Y \to V$ such that

$$\varphi(y + \vartheta(y)) = \inf\{\varphi(y + v) : v \in V\} \text{ for all } y \in Y.$$

Proof Fix $y \in Y$ and consider the C^1-functional $\varphi_y : H \to \mathbb{R}$ defined by

$$\varphi_y(u) = \varphi(y + u) \text{ for all } u \in H.$$

Let $i_V : V \to H$ denote the inclusion map and consider the map $\hat{\varphi}_y : V \to H$ defined by

$$\hat{\varphi}_y = \varphi_y \circ i_V.$$

Evidently, $\hat{\varphi}_y \in C^1(V, H)$ and from the chain rule we have

$$\hat{\varphi}'_y = p_{V^*} \circ \varphi'_y, \tag{6.135}$$

with p_{V^*} being the orthogonal projection of H^* onto V^* (recall that $H^* = Y^* \oplus V^*$). In what follows, by $\langle \cdot, \cdot \rangle_V$ we denote the duality brackets for the pair (V^*, V). For $v, v' \in V$ we have

$$\begin{aligned}
&\left\langle \hat{\varphi}'_y(v) - \hat{\varphi}'_y(v'), v - v' \right\rangle_V \\
&= \left\langle \varphi'_y(v) - \varphi'_y(v'), v - v' \right\rangle \text{ (see (6.135))} \\
&= \left\langle \varphi'(y + v) - \varphi'(y + v'), v - v' \right\rangle \\
&\geqslant \hat{c}||v - v'||^2 \text{ (by hypothesis)} && (6.136) \\
&\Rightarrow \hat{\varphi}'_y \text{ is strongly monotone, hence } \hat{\varphi}_y \text{ is strictly convex.} && (6.137)
\end{aligned}$$

For all $v \in V$, we have

$$\begin{aligned}
&\left\langle \hat{\varphi}'_y(v), v \right\rangle_V = \left\langle \hat{\varphi}'_y(v) - \hat{\varphi}'_y(0), v \right\rangle_V + \left\langle \hat{\varphi}'_y(0), v \right\rangle_V \geqslant \\
&\hat{c}||v||^2 - c_1||v|| \text{ for some } c_1 > 0 && (6.138) \\
&\Rightarrow \hat{\varphi}'_y \text{ is coercive.}
\end{aligned}$$

Also, note that $\hat{\varphi}'_y : V \to V^*$ is monotone and continuous, hence by Proposition 2.6.12 $\hat{\varphi}'_y(\cdot)$ is maximal monotone. Therefore $\hat{\varphi}'_y$ is maximal monotone and coercive, so it is surjective (see Theorem 2.8.6). Hence, we can find $v_0 \in V$ such that

$$\hat{\varphi}'_y(v_0) = 0. \tag{6.139}$$

From (6.136) it is clear that $v_0 \in V$ is unique and is the unique minimizer of the strictly convex functional $\hat{\varphi}_y = \varphi_y|_V$ (see (6.137)). This means that we can define the map $\vartheta : Y \to V$ by setting $\vartheta(y) = v_0$. Then we have

$$p_{V^*}\varphi(y + \vartheta(u)) = 0 \text{ and } \varphi(y + \vartheta(y)) = \inf\{\varphi(y + v) : v \in V\} \text{ for all } y \in Y \tag{6.140}$$
$$\text{(see (6.135) and (6.139)).}$$

Next we show the continuity of the map $\vartheta : Y \to V$. So, let $y_n \to y$ in Y. For every $n \in \mathbb{N}$, we have

$$\begin{aligned}
0 &= \left\langle \hat{\varphi}'_{y_n}(\vartheta(y_n)), \vartheta(y_n) \right\rangle_V \text{ (see (6.139))} \\
&\geqslant \hat{c}||\vartheta(y_n)||^2 - c_1||\vartheta(y_n)|| \text{ (see (6.138))} \\
&\Rightarrow \{\vartheta(y_n)\}_{n \geqslant 1} \subseteq V \text{ is bounded.}
\end{aligned}$$

Passing to a suitable subsequence if necessary, we may assume that

$$\vartheta(y_n) \xrightarrow{w} \hat{v} \text{ in } H,\ \hat{v} \in V. \tag{6.141}$$

Since by hypothesis $\varphi(\cdot)$ is sequentially weakly lower semicontinuous, we have

$$\varphi(y+\hat{v}) \leqslant \liminf_{n\to\infty} \varphi(y_n+\vartheta(y_n)) \text{ (see (6.141))}. \tag{6.142}$$

From (6.140) we know that

$$\begin{aligned}
&\varphi(y_n+\vartheta(y_n)) \leqslant \varphi(y_n+v) \text{ for all } n \in \mathbb{N} \text{ and all } v \in V\\
\Rightarrow\ &\limsup_{n\to\infty} \varphi(y_n+\vartheta(y_n)) \leqslant \varphi(y+v) \text{ (recall that } y_n \to y \text{ in } Y)\\
\Rightarrow\ &\varphi(y+\hat{v}) \leqslant \varphi(y+v) \text{ for all } v \in V \text{ (see (6.142))}\\
\Rightarrow\ &\hat{v} = \vartheta(y).
\end{aligned}$$

So, by the Urysohn criterion for the initial sequence $\{\vartheta(y_n)\}_{n\geqslant 1} \subseteq V$ we have

$$\begin{aligned}
&\vartheta(y_n) \to \vartheta(y) \text{ as } n \to \infty\\
\Rightarrow\ &\vartheta(\cdot) \text{ is continuous.}
\end{aligned}$$

Moreover, from (6.140) we have

$$\varphi(y+\vartheta(y)) = \inf\{\varphi(y+v) : v \in V\}.$$

The proof is now complete. □

We set

$$\varphi_0(y) = \varphi(y+\vartheta(y)) \text{ for all } y \in Y. \tag{6.143}$$

From Proposition 6.7.1 it is clear that $\varphi_0 : Y \to \mathbb{R}$ is continuous. In fact, we can say more.

Proposition 6.7.2 *If $\varphi \in C^1(H)$, φ is sequentially weakly lower semicontinuous,*

$$\left\langle \varphi'(y+v) - \varphi'(y+v'), v - v' \right\rangle \geqslant \hat{c}||v-v'||^2$$

for all $y \in Y$, all $v, v' \in V$, some $\hat{c} > 0$ and $\varphi_0 : Y \to \mathbb{R}$ is given by (6.143), then $\varphi_0 \in C^1(Y)$.

Proof Let $y, h \in Y$ and $t > 0$. From (6.143) and Proposition 6.7.1, we have

$$\begin{aligned}
&\frac{1}{t}\left[\varphi_0(y+th) - \varphi_0(y))\right]\\
\leqslant\ &\frac{1}{t}\left[\varphi(y+th+\vartheta(y)) - \varphi(y+\vartheta(y))\right]\\
\Rightarrow\ &\limsup_{t\to 0} \frac{1}{t}\left[\varphi_0(y+th) - \varphi_0(y)\right] \leqslant \left\langle \varphi'(y+\vartheta(y)), h\right\rangle.
\end{aligned} \tag{6.144}$$

Also, we have

$$\begin{aligned}
&\frac{1}{t}\left[\varphi_0(y+th)-\varphi_0(y)\right] \\
&\geqslant \frac{1}{t}\left[\varphi(y+th+\vartheta(y+th))-\varphi(y+\vartheta(y+th))\right] \\
&\Rightarrow \liminf_{t\to 0} \frac{1}{t}\left[\varphi_0(y+th)-\varphi_0(y)\right] \geqslant \left\langle \varphi'(y+\vartheta(y)), h\right\rangle \qquad (6.145) \\
&\quad \text{(since } \varphi \in C^1(H) \text{ and } \vartheta(\cdot) \text{ is continuous, see Proposition 6.7.1).}
\end{aligned}$$

Let $\langle\cdot,\cdot\rangle_Y$ denote the duality brackets for the pair (Y^*, Y). From (6.144) and (6.145) we have

$$\left\langle (\varphi_0)'_+(y), h\right\rangle = \left\langle \varphi'(y+\vartheta(y)), h\right\rangle \text{ for all } y, h \in Y. \qquad (6.146)$$

Similarly if $t < 0$, then

$$\left\langle (\varphi_0)'_-(y), -h\right\rangle = \left\langle \varphi'(y+\vartheta(y)), -h\right\rangle \text{ for all } y, h \in Y. \qquad (6.147)$$

From (6.146) and (6.147) we conclude that

$$\varphi_0 \in C^1(Y) \text{ and } \varphi_0'(y) = \varphi'(y+\vartheta(y)) \text{ for all } y \in Y.$$

The proof is now complete. □

Proposition 6.7.3 *If $\varphi \in C^1(H)$, φ is sequentially weakly lower semicontinuous,*

$$\left\langle \varphi'(y+v)-\varphi'(y+v'), v-v'\right\rangle \geqslant \hat{c}||v-v'||^2$$

for all $y \in Y$, all $v, v' \in V$, some $\hat{c} > 0$ and $\varphi_0 : Y \to \mathbb{R}$ is given by (6.143), then $y \in K_{\varphi_0}$ if and only if $y+\vartheta(y) \in K_\varphi$.

Proof $\Leftarrow$: This is immediate from (6.135) and (6.140).
$\Rightarrow$: Suppose that $y \in K_{\varphi_0}$. Then

$$0 = \varphi_0'(y) = p_{V^*}\varphi'(y+\vartheta(y)) \text{ (see (6.135), (6.143)).} \qquad (6.148)$$

Since $H^* = Y^* \oplus V^*$, it follows that

$$\begin{aligned}
&\varphi'(y+\vartheta(y)) \in Y^* \\
&\left\langle \varphi'(y+\vartheta(y)), h\right\rangle_Y = 0 \text{ for all } h \in Y \text{ (see (6.148))} \\
\Rightarrow\ &\varphi'(y+\vartheta(y)) = 0 \\
\Rightarrow\ &y+\vartheta(y) \in K_\varphi.
\end{aligned}$$

The proof is now complete. □

Remark 6.7.4 So $y \in K_{\varphi_0}$ is isolated if and only if $y + \vartheta(y) \in K_\varphi$ is isolated.

Based on this remark, one may ask what the relation is between the critical groups

$$C_k(\varphi_0, y) \text{ and } C_k(\varphi, y + \vartheta(y)) \text{ for all } k \in \mathbb{N}_0.$$

The next theorem answers this equation and shows the effectiveness of the Lyapunov–Schmidt reduction method.

Theorem 6.7.5 *If $\varphi \in C^1(H)$, φ is sequentially weakly lower semicontinuous,*

$$\left\langle \varphi'(y+v) - \varphi'(y+v'), v - v' \right\rangle \geqslant ||v - v'||^2$$

for all $y \in Y$, all $v, v' \in V$, some $\hat{c} > 0$, $\varphi_0 : Y \to \mathbb{R}$ is given by (6.143) and $\hat{y} \in K_{\varphi_0}$ is isolated, then $C_k(\varphi_0, \hat{y}) = C_k(\varphi, \hat{y} + \vartheta(\hat{y}))$ for all $k \in \mathbb{N}_0$.

Proof Recall that $\hat{y} + \vartheta(\hat{y}) \in K_\varphi$ is isolated (see Proposition 6.7.3 and Remark 6.7.4).

Let $c = \varphi_0(y) = \varphi(y + \vartheta(y))$ (see (6.143)) and

$$D = \{(y, \vartheta(y)) \in Y \times V : y \in \varphi_0^c\}.$$

We set $\hat{u} = \hat{y} + \vartheta(\hat{y}) \in H$ and consider the maps

$$\xi : \varphi^c \to D \text{ and } \eta : \varphi_0^c \to D$$

defined by

$$\begin{aligned} &\xi(y, v) = (y, \vartheta(y)) \text{ (with } y + v \in \varphi^c), \\ &\eta(y) = (y, \vartheta(y)) \text{ (that is } \xi(y, v) = \eta(y)). \end{aligned}$$

We have

$$(\varphi^c, \varphi^c \setminus \{\hat{u}\}) \xrightarrow{\xi} (D, D \setminus \{\hat{u}\}) \text{ and } (\varphi_0^c, \varphi_0^c \setminus \{\hat{y}\}) \xrightarrow{\eta} (D, D \setminus \{\hat{u}\}). \tag{6.149}$$

Note that η is a homeomorphism and $\eta^{-1}(y, \vartheta(y)) = y$ for all $(y, \vartheta(y)) \in D$. Recall that $v \to \varphi(y + v)$ is strictly convex (see the proof of Proposition 6.7.1). So, identifying $u = y + v \in H$ with $y \in Y, v \in V$ (uniquely) with the pair (y, v), we have

$$\begin{aligned} &\varphi(y, (1-t)v + tv') \leqslant (1-t)\varphi(y, v) + t\varphi(y, v') \\ &\text{for all } t \in [0, 1], \text{ all } y \in Y, \text{ and all } v, v' \in V. \end{aligned}$$

Therefore we can define the homotopy $e : ([0, 1] \times \varphi^c, [0, 1] \times (\varphi^c \setminus \{\hat{u}\})) \to (\varphi^c, \varphi^c \setminus \{\hat{u}\})$ by setting

$$e(t, (y, v)) = (y, (1-t)v + t\vartheta(y)).$$

Let $i : (D, D\backslash\{\hat{u}\}) \to (\varphi^c, \varphi^c\backslash\{\hat{u}\})$ be the inclusion map. Using the homotopy we can easily see that

$$\xi \circ i = \mathrm{id}_{(D, D\backslash\{\hat{u}\})} \text{ and } i \circ \xi \simeq \mathrm{id}_{(\varphi^c, \varphi^c\backslash\{\hat{u}\})}.$$

So, the pairs $(D, D\backslash\{\hat{u}\})$ and $(\varphi^c, \varphi^c\backslash\{\hat{u}\})$ are homotopy equivalent. Then from Proposition 6.1.14 we have

$$\begin{aligned} & H_k(\varphi_0^c, \varphi_0^c\backslash\{\hat{u}\}) = H_k(D, D\backslash\{\hat{u}\}) \\ \Rightarrow\ & C_k(\varphi, y + \vartheta(y)) = C_k(\varphi_0, y) \text{ for all } k \in \mathbb{N}_0. \end{aligned}$$

The proof is now complete. □

Suppose that $\varphi \in C^1(H)$ satisfies the assumptions of the above theorem. Let $\nabla\varphi$ denote the gradient of φ, that is,

$$(\nabla\varphi(u), h)_H = \left\langle \varphi'(u), h \right\rangle \text{ for all } u, h \in H,$$

with $(\cdot, \cdot)_H$ denoting the inner product of H. Suppose that

$$\nabla\varphi = I - K \text{ with } K \in \mathscr{L}_c(H).$$

Therefore the Leray–Schauder index $i_{LS}(\nabla\varphi, \hat{y} + \vartheta(\hat{y}))$ (see Definition 6.2.43) can be defined. Also, since Y is finite-dimensional, the Brouwer index $i(\varphi_0, \hat{y}, c)$ (see Definition 3.8.1) is also defined. We expect the two to be related. Indeed using Theorem 6.7.5, we have the following result.

Corollary 6.7.6 *If everything is as above with $\hat{y} \in K_{\varphi_0}$ isolated, then $i_{LS}(\nabla\varphi, \hat{y} + \vartheta(\hat{y})) = i(\varphi_0, \hat{y}, c)$.*

Proof Using Proposition 6.2.44 we have

$$\begin{aligned} i(\varphi_0, \hat{y}, c) &= \sum_{k \geqslant 0} (-1)^k \operatorname{rank} C_k(\varphi_0, \hat{y}) \\ &= \sum_{k \geqslant 0} (-1)^k \operatorname{rank} C_k(\varphi, \hat{y} + \vartheta(\hat{y})) \text{ (see Theorem 6.7.5)} \\ &= i_{LS}(\nabla\varphi, \hat{y} + \vartheta(\hat{y})) \text{ (see Proposition 6.2.44)}. \end{aligned}$$

The proof is now complete. □

The next existence result is in the spirit of Proposition 6.2.42.

Proposition 6.7.7 *If X is a Banach space, $\varphi \in C^1(X)$, φ satisfies the C-condition, $u \in X$, $a, b, c \in \mathbb{R}$ with $a < c < b$, K_φ is finite, $K_\varphi^c = \{u\}$, $a, b \notin \varphi(K_\varphi)$ and*

$C_k(\varphi, u) \neq 0,\ H_k(\varphi^b, \varphi^a) = 0$ for some $k \in \mathbb{N}_0$, then we can find $u_0 \in K_\varphi$ such that

$$\begin{aligned} &a < \varphi(u_0) < c \textit{ and } C_{k-1}(\varphi, u_0) \neq 0 \textit{ or} \\ &c < \varphi(u_0) < b \textit{ and } C_{k+1}(\varphi, u_0) \neq 0. \end{aligned}$$

Proof Choose $\epsilon > 0$ small so that

$$K_\varphi \cap \varphi^{-1}([c-\epsilon, c+\epsilon]) = \{u\}$$

and $a < c - \epsilon < c + \epsilon < b$.

From Proposition 6.2.16, we have

$$\begin{aligned} &H_k(\varphi^{c+\epsilon}, \varphi^{c-\epsilon}) = C_k(\varphi, u) \neq 0 \\ \text{and} \quad &H_k(\varphi^b, \varphi^a) = 0 \text{ (by hypothesis)}. \end{aligned}$$

We consider the sets $\varphi^a \subseteq \varphi^{c-\epsilon} \subseteq \varphi^{c+\epsilon} \subseteq \varphi^b$ and use Proposition 6.1.37. We obtain $H_{k-1}(\varphi^{c-\epsilon}, \varphi^a) \neq 0$ or $H_{k+1}(\varphi^b, \varphi^{c+\epsilon}) \neq 0$. Then Proposition 6.2.15 implies that we can find $u_0 \in K_\varphi$ such that

$$\begin{aligned} &a < \varphi(u_0) < c - \epsilon \text{ and } C_{k-1}(\varphi, u_0) \neq 0 \text{ or} \\ &c + \epsilon < \varphi(u_0) < b \text{ and } C_{k+1}(\varphi, u_0) \neq 0. \end{aligned}$$

The proof is now complete. □

Corollary 6.7.8 *If X is a Banach space, $\varphi \in C^1(X)$, φ satisfies the C-condition, K_φ is finite, $K_\varphi^c = \{u\}$ and $C_k(\varphi, u) \neq 0,\ C_k(\varphi, \infty) = 0$ for some $k \in \mathbb{N}_0$, then we can find $u_0 \in K_\varphi$ such that*

$$\begin{aligned} &\varphi(u_0) < \varphi(u) = c \textit{ and } C_{k-1}(\varphi, u_0) \neq 0 \textit{ or} \\ &c = \varphi(u) < \varphi(u_0) \textit{ and } C_{k+1}(\varphi, u_0) \neq 0. \end{aligned}$$

Proof Since K_φ is finite, we can find $a, b \in \mathbb{R}$ such that

$$a < \inf \varphi(K_\varphi) < \sup \varphi(K_\varphi) < b.$$

Invoking Proposition 6.2.28(a), we have

$$H_k(\varphi^b, \varphi^a) = C_k(\varphi, \infty) = 0 \text{ (by hypothesis)}.$$

So, we can apply Proposition 6.7.7 and find $u_0 \in K_\varphi$ such that

$$\begin{aligned} &\varphi(u_0) < c = \varphi(u) \text{ and } C_{k-1}(\varphi, u_0) \neq 0 \text{ or} \\ &\varphi(u) = c < \varphi(u_0) \text{ and } C_{k+1}(\varphi, u_0) \neq 0. \end{aligned}$$

The proof is now complete. □

Next, we present some multiplicity results.

Proposition 6.7.9 *If X is a Banach space, $\varphi \in C^1(X)$, φ is bounded below and satisfies the C-condition (or equivalently the PS-condition, see Proposition 5.1.14), $u \in K_\varphi$ is isolated and not a global minimizer of φ and for some $m \in \mathbb{N}_0$, $C_m(\varphi, u) \neq 0$, then K_φ has at least three elements.*

Proof Since φ is bounded below and satisfies the C-condition, by Proposition 5.1.8 we can find $u_0 \in K_\varphi$ which is a global minimizer of φ. Since by hypothesis $u \in K_\varphi$ is not a global minimizer, we must have $u \neq u_0$ and $\varphi(u_0) < \varphi(u)$. Suppose that $K_\varphi = \{u, u_0\}$ and choose $a, b \in \mathbb{R}$ such that

$$\varphi(u_0) < a < \varphi(u) < b.$$

Then Corollary 5.3.13 implies that φ^b is a strong deformation retract, while Theorem 5.3.12 (the second deformation theorem) says that $\{u_0\}$ is a strong deformation retract of φ^a. So, we have

$$H_k(\varphi^b, \{u_0\}) = H_k(X, \{u_0\}) \text{ for all } k \in \mathbb{N}_0 \text{ (see Corollary 6.1.24(b))}, \tag{6.150}$$

$$H_k(X, u_0) = 0 \text{ for all } k \in \mathbb{N}_0 \tag{6.151}$$

(since X is a contractible, see Proposition 6.1.30),

$$H_k(\varphi^a, \{u_0\}) = 0 \text{ for all } k \in \mathbb{N}_0 \text{ (see Proposition 6.1.15)}. \tag{6.152}$$

Using the log exact sequence from Proposition 6.1.29 and (6.150), (6.151), (6.152), we infer that

$$\begin{aligned} & H_k(\varphi^b, \varphi^a) = 0 \text{ for all } k \in \mathbb{N}_0 \\ \Rightarrow\ & C_k(\varphi, u) = 0 \text{ for all } k \in \mathbb{N}_0 \text{ (see Proposition 6.2.16)}, \end{aligned}$$

a contradiction to the hypothesis $C_m(\varphi, u) \neq 0$. So, K_φ has a third element $\hat{u}$. The proof is now complete. □

Corollary 6.7.10 *If X is a Banach space, $\varphi \in C^1(X)$ is bounded below and satisfies the C-condition, φ has a local (m, n)-linking at 0 with $m, n \in \mathbb{N}$ and 0 is not a global minimizer of φ, then φ has at least three critical points.*

Proof From Theorem 6.6.17 we know that $C_m(\varphi, 0) \neq 0$. So, we can use Proposition 6.7.9 and conclude that K_φ has at least three elements. □

Proposition 6.7.11 *If X is a Banach space, $\varphi \in C^1(X)$ is bounded below and satisfies the C-condition, and $C_m(\varphi, 0) \neq 0$ for some $m \in \mathbb{N}$, then*

(a) φ has a nontrivial critical point;
(b) if $m \geqslant 2$, then φ has at least two nontrivial critical points.

Proof (*a*) Since φ is bounded below and satisfies the C-condition, from Proposition 5.1.8, φ has a global minimizer $u \in K_\varphi$. Then Proposition 6.2.3 implies

$$C_k(\varphi, u) = \delta_{k,0}\mathbb{Z} \text{ for all } k \in \mathbb{N}_0. \tag{6.153}$$

By hypothesis we have

$$C_m(\varphi, 0) \neq 0 \text{ for some } m \in \mathbb{N}. \tag{6.154}$$

Comparing (6.153) and (6.154) we conclude that $u \in K_\varphi$ is nontrivial.
(*b*) From Proposition 6.2.24 we have

$$C_k(\varphi, \infty) = \delta_{k,0}\mathbb{Z} \text{ for all } k \in \mathbb{N}_0. \tag{6.155}$$

From (6.154), (6.155) and Proposition 6.2.42, we know that we can find $u_0 \in K_\varphi$ such that

$$C_{m-1}(\varphi, u_0) \neq 0 \text{ or } C_{m+1}(\varphi, u_0) \neq 0. \tag{6.156}$$

Since $1 \leqslant m - 1$ (recall $m \geqslant 2$), from (6.153) and (6.156) it follows that $u_0 \neq u$. □

6.8 Remarks

6.1: The material on Algebraic Topology is standard and can be found in the books of Dold [148], Eilenberg and Steenrod [156], Hatcher [203], Maunder [292] and Spanier [390]. The axiomatic treatment (see Definition 6.1.12) was first introduced by Eilenberg and Steenrod [156]. The term "homology group" is due to Vietoris [410]. Singular homology was introduced by Eilenberg. Preceding that we had simplicial homology, which is the result of the work of many mathematicians, including Betti and Poincaré. The systematic introduction of group theoretic methods occurred in the 1920s through the works of Alexander, Hopf and Lefschetz, who developed simplicial homology. Cohomology theories can be axiomatized in the same way as homology theories. The formalism of category theory is helpful in this respect.

6.2: Critical groups provide a powerful tool to distinguish between critical points and to produce additional critical points of a given functional. For this reason they are important in the study of nonlinear boundary value problems. Let X be a Hausdorff topological space, $\varphi : X \to \mathbb{R}$ a continuous function, $K \subseteq X$ closed and $c, d \in \mathbb{R}$, $c \leqslant d$. As before we set

$$\begin{aligned} \varphi^c &= \{u \in X : \varphi(u) \leqslant c\}, \\ K_c &= \{u \in K : \varphi(u) = c\}, \\ \varphi_c^d &= \{u \in X : c \leqslant \varphi(u) \leqslant d\}. \end{aligned}$$

In the case when X is a Banach space and φ a C^1-functional, K is the critical set of φ (the set K_φ in the notation of Sect. 6.2).

Definition 6.8.1 If $D \subseteq \varphi^{-1}(c)$, the critical groups for the pair (φ, D) are defined by

$$C_k(\varphi, D) = H_k(\varphi^c, \varphi^c \setminus D) \text{ for all } k \in \mathbb{N}_0.$$

For $u \in \varphi^{-1}(c)$, we set

$$C_k(\varphi, u) = C_k(\varphi, \{u\}) = H_k(\varphi^c, \varphi^c \setminus \{u\}) \text{ for all } k \in \mathbb{N}_0.$$

Remark 6.8.2 By excision, the critical groups depend only on the restriction of φ on an arbitrary neighborhood U of u. This way, in the differentiable setting we recover Definition 6.2.1, which stresses the local character of the theory.

As an easy application of Proposition 6.1.23, we get the following result.

Proposition 6.8.3 *If $D_1, D_2 \subseteq \varphi^{-1}(C)$ are disjoint closed sets then $C_k(\varphi, D_1 \cup D_2) = C_k(\varphi, D_1) \oplus C_k(\varphi, D_2)$ for all $k \in \mathbb{N}_0$.*

The Morse lemma (see Proposition 5.4.19) was extended to the degenerate case by Hofer [209], under the assumption that $\varphi''(u_0)$ is of the form of a compact perturbation of the identity. The more general form included here (see Proposition 6.2.9) is due to Mawhin and Willem [293, p. 185]. The shifting theorem (see Theorem 6.2.13) is due to Gromoll and Meyer [198]. The Morse relation (see Theorem 6.2.20) can be found in the important paper of Marino and Prodi [288] on perturbation methods in Morse theory. In the same work of Marino–Prodi we can find Lemmata 6.2.35, 6.2.37 and 6.2.38. Critical groups at infinity were introduced by Bartsch and Li [38] as the appropriate tool for the global theory of critical points. Condition (A_∞) is a slightly more general version of the one employed by Bartsch and Li [38]. In Bartsch and Li [38] it is assumed that $\varphi''(u) \to 0$ as $||u|| \to \infty$. However, a careful reading of their proof reveals that it is enough to assume that $\psi'(u) = 0(||u||)$ (see condition (A_∞)). This generalization was first used by Su and Zhao [394]. Proposition 6.2.44 is due to Rothe [361].

The next proposition gives a more convenient version of the Shifting Theorem (see Theorem 6.2.13).

Proposition 6.8.4 *If H is a Hilbert space, $\varphi \in C^2(H)$ and $u \in K_\varphi$ is isolated with finite Morse index m and nullity ν, then one of the following holds:*

(a) $C_k(\varphi, u) = 0$ for all $k \leqslant m$ and all $k \geqslant m + \nu$;
(b) $C_k(\varphi, u) = \delta_{k,m}\mathbb{Z}$ for all $k \in \mathbb{N}_0$;
(c) $C_k(\varphi, u) = \delta_{k,m+\nu}\mathbb{Z}$ for all $k \in \mathbb{N}_0$.

Remark 6.8.5 In fact, the result is also true for nontrivial critical points of C^{2-0} functionals (that is, C^1-functionals φ whose derivative $\varphi'(\cdot)$ is locally Lipschitz, alternatively the notation $C^{1,1}$ is also used). This extension was proved by Li et al. [268]. In the same paper, it is also proved that Proposition 6.2.44 is in fact true for $\varphi \in C^1(H)$.

6.3: The invariance of the critical groups with respect to small $C^1(X)$-perturbations (see Theorem 6.3.4) and with respect to homotopies which preserve the isolation of the critical point (see Theorem 6.3.6) are two very useful tools for computing critical groups in concrete situations. The results can be found in Chang and Ghoussoub [120] and Corvellec and Hantoute [128]. In the latter, the setting is more general (continuous functions on metric spaces using the notions of weak slope and lower critical point). The result on the continuity of critical groups with respect to the C^1-topology can also be found in Chang [118, p. 336] and in Mawhin and Willem [293, p. 196] (for C^2-functionals on Hilbert spaces). Similarly, the homotopy invariance of critical groups can also be found in Chang [118, p. 53]. Theorem 6.3.8 is due to Chang [118, p. 334].

6.4: The critical group theory can be extended to critical subsets. More precisely, we saw that an isolated critical point is replaced by a dynamically isolated critical set (see Definition 6.4.17) and for such sets we have an analogous critical group theory. In the presentation of this theory we follow Chang [118] (see also Chang and Ghoussoub [120]), who developed an extension of the Gromoll–Meyer theory (see Gromoll and Meyer [198]). This extended theory is also related to the Conley index theory of isolated invariant sets for gradient flows.

6.5: Recall that the notion of a critical point of mountain pass type (see Definition 6.5.6) is due to Hofer [210, 211]. Theorem 6.5.8 is usually proved using the second deformation theorem (see Theorem 5.3.12) and arguments based on the existence of neighborhoods which are stable with respect to the pseudogradient flow. Here, we follow a different approach based on Lemma 6.2.35. This allows us to slightly weaken the hypotheses in Theorem 6.5.8 and assume that the critical value $\varphi(u_0)$ is isolated in $\varphi(K_\varphi)$. Theorem 6.5.11 can also be found in Chang [118, p. 91] and in Mawhin and Willem [293, p. 195], under a little more restrictive conditions (see also Bartsch [37]).

6.6: The notion of homological linking (see Definition 6.6.1) goes back to the work of Liu [276], who also proved the nontriviality of the first critical group $C_1(\varphi, u_0)$ for a critical point $u_0 \in K_\varphi$ produced by an application of the mountain pass theorem (see also Corollary 6.6.9). The notion of local (m, n)-linking (see Definition 6.6.13) is due to Perera [335], who also proved Theorem 6.6.17. Corollary 6.6.18 is due to Liu [276]. Proposition 6.6.20 can be found in Papageorgiou and Rădulescu [330]. Theorem 6.6.22, stressing the local character of critical groups, is essentially due to Degiovanni et al. [141], who employed an approach using the truncation function $\xi(\cdot)$ (see Lemma 6.6.21). Theorem 6.6.26 can also be found in Chang [118, p. 14] and in Bartsch [36].

6.7: The "reduction method" for elliptic problems was developed by Amann [12, 13] and Castro and Lazer [109]. Theorem 6.7.5 is due to Liu and Li [278] and Liu [276]. Multiplicity results for critical points using critical groups can also be found in Motreanu et al. [309, 310] and in Perera et al. [336].

References

1. G. Acosta, R.G. Durán, An optimal Poincaré inequality in L^1 for convex domains. Proc. Am. Math. Soc. **132**, 195–202 (2004)
2. R.A. Adams, *Sobolev Spaces* (Academic Press, New York, 1975)
3. R.A. Adams, J. Fournier, *Sobolev Spaces*, 2nd edn. (Academic Press (Elsevier), Amsterdam, 2003)
4. D. Adams, L. Hedberg, *Function Spaces and Potential Theory* (Springer, Berlin, 1996)
5. S. Aizicovici, N.S. Papageorgiou, V. Staicu, *Degree Theory for Operators of Monotone Type and Nonlinear Elliptic Equations with Inequality Constraints*, Memoirs American Mathematical Society, vol. 196, No. 915 (2008)
6. S. Aizicovici, N.S. Papageorgiou, V. Staicu, The spectrum and an index formula for the Neumann p-Laplacian and multiple solutions for problems with a crossing nonlinearity. Discret. Contin. Dyn. Syst. **25**, 431–456 (2009)
7. N.I. Akhiezer, I.M. Glazman, *Theory of Linear Operators in Hilbert Space*, vol. I and II (Frederick Ungar Publishing Company, New York, 1961 and 1963)
8. V. Alexeev, V. Tikhomirov, S. Fomin, *Commande Optimale* (Mir, Moscou, 1982)
9. C. Aliprantis, K. Border, *Infinite Dimensional Analysis. A Hitchhiker's Guide* (Springer, Berlin, 1994)
10. D. Alspach, A fixed point free nonexpansive map. Proc. Am. Math. Soc. **82**, 423–424 (1981)
11. M. Altman, A fixed point theorem for completely continuous operators in Banach spaces. Bull. Acad. Polon. Sci. **3**, 409–413 (1955)
12. H. Amann, Fixed point equations and nonlinear eigenvalue problems in ordered Banach spaces. SIAM Rev. **18**, 620–709 (1976)
13. H. Amann, *Ordered structures and fixed points*, in Atti 2 Sem. Anal. Funz. Appl., Univ. Cosenza, Italy (1977), pp. 50
14. H. Amann, A note on degree theory for gradient mappings. Proc. Am. Math. Soc. **85**, 591–595 (1982)
15. H. Amann, S. Weiss, On the uniqueness of the topological degree. Math. Z. **130**, 39–54 (1973)
16. A. Ambrosetti, A. Malchiodi, *Nonlinear Analysis and Semilinear Elliptic Problems* (Cambridge University Press, Cambridge, UK, 2007)
17. A. Ambrosetti, P. Rabinowitz, Dual variational methods in critical point theory and applications. J. Funct. Anal. **14**, 349–381 (1973)
18. E. Asplund, Positivity of duality mappings. Bull. Am. Math. Soc. **73**, 200–203 (1967)
19. E. Asplund, Averaged norms. Isr. J. Math. **5**, 227–233 (1967)

N. S. Papageorgiou et al., *Nonlinear Analysis–Theory and Methods*,
Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-030-03430-6

20. H. Attouch, G. Buttazzo, G. Michaille, *Variational Analysis in Sobolev and BV Spaces. Applications to PDEs and Optimization*, MPS/SIAM Series on Optimization, 6. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Programming Society (MPS), Philadelphia, PA, 2006
21. J.P. Aubin, I. Ekeland, *Applied Nonlinear Analysis* (Wiley, New York, 1984)
22. J.P. Aubin, H. Frankowska, *Set-Valued Analysis* (Birkhäuser, Boston, 1990)
23. T. Aubin, Problèmes isopérimétriques et espaces de Sobolev. J. Differ. Geom. **11**, 573–598 (1976)
24. R. Aumann, *Measurable utility and the measurable choice theorem* (Aix-en-Provence, Actes Colloq. Internat., 1969), pp. 15–26
25. R. Baire, Sur les fonctions de variables réelles. Ann. Mat. Pura Appl. **3**, 1–123 (1899)
26. R. Baire, Sur l'origine de la notion de semi-continuité. Bull. Soc. Math. Fr. **55**, 141–142 (1927)
27. E.J. Balder, Necessary and sufficient conditions for L_1-strong-weak lower semicontinuity of integral functionals. Nonlinear Anal. **11**, 1399–1404 (1987)
28. E.J. Balder, Lectures on Young measure theory and its applications in economics, Workshop on Measure Theory and Real Analysis (Italian) (Grado, 1997). Rend. Istit. Mat. Univ. Trieste **31**(suppl. 1), 1–69 (2000)
29. J.M. Ball, K.-W. Zhang, Lower semicontinuity of multiple integrals and the biting lemma. Proc. R. Soc. Edinb. Sect. A **114**, 367–379 (1990)
30. S. Banach, Sur les opérations dans les ensembles abstraits et leur application aux équations integrales. Fundam. Math. **3**, 133–181 (1922)
31. S. Banach, Théorie des Opérations Linéaires, Monografje Matematyczne, Warsaw, 1932 (English Translation: Theory of Linear Operators, North Holland, Amsterdam, 1987)
32. V. Barbu, *Nonlinear Semigroups and Differential Equations in Banach Spaces* (Noordhoff, Leiden, The Netherlands, 1975)
33. V. Barbu, T. Precupanu, *Convexity and Optimization in Banach Spaces* (Reidel Publishing Company, Dordrecht, 1986)
34. P. Bartolo, V. Benci, D. Fortunato, Abstract critical point theorems and applications to some nonlinear problems with "strong" resonance at infinity. Nonlinear Anal. **7**, 981–1012 (1983)
35. T. Bartsch, Infinitely many solutions of a symmetric Dirichlet problem. Nonlinear Anal. **20**, 1205–1216 (1993)
36. T. Bartsch, *Topological Methods for Variational Problems with Symmetries*, Lecture Notes in Mathematics, vol. 1560 (Springer, Berlin, 1993)
37. T. Bartsch, Critical point theory on partially ordered Hilbert spaces. J. Funct. Anal. **186**, 117–152 (2001)
38. T. Bartsch, S. Li, Critical point theory for asymptotically quadratic functionals and applications to problems with resonance. Nonlinear Anal. **28**, 419–441 (1997)
39. H. Bauschke, P. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces* (Springer, New York, 2011)
40. M. Bebendorf, A note on the Poincaré inequality for convex domains. Z. Anal. Anwendungen **22**, 751–756 (2003)
41. R. Bellman, *Introduction to Matrix Analysis* (McGraw-Hill Company, New York, 1970)
42. V. Benci, A new approach to the Morse-Conley theory and some applications. Ann. Mat. Pura Appl. **158**, 231–305 (1991)
43. V. Benci, P. Rabinowitz, Critical point theorems for indefinite functionals. Invent. Math. **52**, 241–273 (1979)
44. M. Berger, *Nonlinearity and Functional Analysis* (Academic Press, New York, 1977)
45. J. Berkovits, A note on the imbedding theorem of Browder and Ton. Proc. Am. Math. Soc. **131**, 2963–2966 (2003)
46. J. Berkovits, V. Mustonen, On the topological degree for mappings of monotone type. Nonlinear Anal. **10**, 1373–1383 (1986)
47. J. Berkovits, V. Mustonen, *Nonlinear mappings of monotone type: classification and degree theory*, No. 2/1988 in Mathematics, Faculty of Science, University of Oulu, Finland, 1988

48. H. Berliocchi, J.-M. Lasry, Intégrandes normales et mesures paramétrées en calcul des variations. Bull. Soc. Math. Fr. **101**, 129–184 (1973)
49. C. Bessaga, A. Pelczynski, *Topics in Infinite Dimensional Topology* (Polish Scientific Publishers, Warsaw, 1975)
50. A. Beurling, A. Livingston, A theorem on duality mappings in Banach spaces. Ark. Mat. **4**, 405–411 (1962)
51. A. Bielecki, Une remarque sur la méthode de Banach-Caccioppoli-Tikhonov dans la théorie des équations différentielles ordinaires. Bull. Acad. Polon. Sci. **4**, 261–264 (1956)
52. H.F. Bohnenblust, S. Karlin, On a theorem of Ville, *Contributions to the Theory of Games*, Annals of Mathematics Studies, No. 24 (Princeton University Press, Princeton, 1950), pp. 155–160
53. P. Bohl, Über die Bewegung eines mechnischen Systems in der Nähe einer Gleichgewichtslage. J. Reine Angew. Math. **127**, 179–276 (1904)
54. K.C. Border, *Fixed Point Theorems with Applications to Economics and Game Theory* (Cambridge University Press, Cambridge, 1985)
55. E. Borel, *Leçons sur la théorie des fonctions* (Gauthier-Villars, Paris, 1928)
56. K. Borsuk, Sur les rétractes. Fund. Math. **17**, 152–170 (1931)
57. K. Borsuk, Sur un espace de transformations continues et ses applications topologiques. Monatsh. Math. Phys. **38**, 381–386 (1931)
58. K. Borsuk, Drei Sätze uber die n-dimensionale euklidische Sphäre. Fund. Math. **20**, 177–190 (1933)
59. N. Bourbaki, *Elements of Mathematics–General Topology, Part 1* (Addison-Wesley Publishing Company, Reading, Mass, 1966)
60. A. Bressan, G. Colombo, Extensions and selections of maps with decomposable values. Studia Math. **90**, 69–86 (1988)
61. G.E. Bredon, Some examples for the fixed point property. Pacific J. Math. **38**, 571–575 (1971)
62. H. Brezis, Équations et inéquations non linéaires dans les espaces vectoriels en dualité. Ann. Inst. Fourier (Grenoble) **18**, 115–175 (1968)
63. H. Brezis, Monotonicity methods in Hilbert spaces and some applications to nonlinear differential equations, in *Contributions to Nonlinear Functional Analysis*, ed. by E. Zarantonello (Academic Press, New York, 1971), pp. 101–156
64. H. Brezis, *Opérateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert* (North Holland, London, 1973)
65. H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations* (Springer, New York, 2011)
66. H. Brezis, F. Browder, A general principle on ordered sets in nonlinear functional analysis. Adv. Math. **21**, 355–364 (1976)
67. H. Brezis, M. Crandall, A. Pazy, Perturbations of maximal monotone sets. Commun. Pure Appl. Math. **23**, 123–144 (1970)
68. H. Brezis, L. Nirenberg, Characterization of the ranges of some nonlinear operators and applications to boundary value problems. Ann. Scuola Norm. Sup. Pisa Cl. Sci. Ser. IV **5**, 225–326 (1978)
69. H. Brezis, L. Nirenberg, Remarks on finding critical points. Comm. Pure Appl. Math. **44**, 939–963 (1991)
70. M.S. Brodskii, D.P. Milman, On the center of a convex set. Dokl. Akad. Nauk SSSR **59**, 837–840 (1948)
71. A. Brondsted, Conjugate convex functions in topological function spaces. Mat. Fys. Medd. Danske Vid. Selsk. **34**, 1–26 (1964)
72. A. Brondsted, R.T. Rockafellar, On the subdifferentiability of convex functions. Proc. Am. Math. Soc. **16**, 605–611 (1965)
73. L.E.J. Brouwer, On continuous one-to-one transformations of surfaces into themselves. Proc. Kon Nederl. Akad. Wetensch. **11**, 788–798 (1909)
74. L.E.J. Brouwer, Beweis der Invarianz der Dimensionenzahl. Math. Ann. **70**, 161–165 (1911)
75. L.E.J. Brouwer, Zur Invarianz des n-dimensionalen Gebiets. Math. Ann. **72**, 55–56 (1912)

76. L.E.J. Brouwer, Über Abbildung von Mannigfaltigkeiten. Math. Ann. **71**, 598 (1912)
77. F. Browder, On a generalization of the Schauder fixed point theorem. Duke Math. J. **26**, 291–304 (1959)
78. F. Browder, Nonlinear elliptic boundary value problems. Bull. Am. Math. Soc. **69**, 862–874 (1963)
79. F. Browder, Continuity properties of monotone nonlinear operators in banach spaces. Bull. Am. Math. Soc. **70**, 551–553 (1964)
80. F. Browder, Infinite dimensional manifolds and nonlinear elliptic eigenvalue problems. Ann. Math. **82**, 459–477 (1965)
81. F. Browder, Multi-valued monotone nonlinear mappings and duality mappings in Banach spaces. Trans. Am. Math. Soc. **118**, 338–351 (1965)
82. F. Browder, On a theorem of Beurling and Livingston. Canad. J. Math. **17**, 617–622 (1965)
83. F. Browder, Fixed point theorems for noncompact mappings in Hilbert spaces. Proc. Nat. Acad. Sci. USA **43**, 1272–1276 (1965)
84. F. Browder, Nonlinear maximal monotone operators in Banach spaces. Math. Ann. **175**, 89–113 (1968)
85. F. Browder, Nonlinear eigenvalue problems and Galerkin approximations. Bull. Am. Math. Soc. **74**, 651–656 (1968)
86. F. Browder, Nonlinear operators and nonlinear equations of evolution in Banach spaces, Nonlinear Functional Analysis, in *Proceedings of Symposia in Pure Mathematics, vol. XVIII, Part 2, Chicago, Ill, 1968* (American Mathematical Society, Providence, R.I, 1976), pp. 1–308
87. F. Browder, Fixed point theory and nonlinear problems. Bull. Am. Math. Soc. (N.S.) **9**, 1–39 (1983)
88. F. Browder, Degree of mapping for nonlinear mappings of monotone type. Proc. Nat. Acad. Sci. USA **80**, 1771–1773 (1983)
89. F. Browder, The degree of mapping and its generalizations. Contemp. Math. **21**, 1–39 (1983)
90. F. Browder, Degree theory for nonlinear mappings, in *Nonlinear Functional Analysis and Applications, Part 1*, Proceedings of Symposia in Pure Mathematics, ed. by F. Browder (American Mathematical Society, Providence, R.I, 1986), pp. 203–226
91. F. Browder, P. Hess, Nonlinear mappings of monotone type in Banach spaces. J. Funct. Anal. **11**, 251–294 (1972)
92. F. Browder, W. Petryshyn, The topological degree and Galerkin approximation for noncompact operators in Banach spaces. Bull. Am. Math. Soc. **74**, 641–646 (1968)
93. F. Browder, Approximation methods and the generalized topological degree for mappings in Banach spaces. J. Funct. Anal. **3**, 217–245 (1969)
94. F. Browder, B.A. Ton, Nonlinear functional equations in Banach spaces and elliptic superregularization. Math. Z. **105**, 177–195 (1968)
95. R. Brown, An elementary proof of the uniqueness of the fixed point index. Pacific J. Math. **35**, 548–558 (1970)
96. V.I. Burenkov, *Sobolev Spaces on Domains*, Teubner-Texte zur Mathematik [Teubner Texts in Mathematics], vol. 137 (Stuttgart, 1998)
97. G. Buttazzo, *Semicontinuity, Relaxation and Integral Representation in the Calculus of Variations* (Longman Scientific and Technical, London, 1989)
98. L. Čaklović, S. Li, M. Willem, A note on Palais-Smale condition and coercivity. Differ. Integr. Equ. **3**, 799–800 (1990)
99. A. Calderon, Lebesgue spaces and differentiable functions and distributions, in *Proceedings of the Symposia Pure Mathematics*, vol. 4 (American Mathematical Society, Providence, R.I., 1961), pp. 33–49
100. B. Calvert, J.R.L. Webb, An existence theorem for quasimonotone operators. Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. Ser. VIII **50**, 690–696 (1971)
101. J. Caristi, Fixed point theorems for mappings satisfying inwardness conditions. Trans. Am. Math. Soc. **215**, 241–251 (1976)
102. S. Carl, S. Heikkila, *Fixed Point Theory in Ordered Sets and Applications* (Springer, Berlin, 2011)

103. S. Carl, V.K. Le, D. Motreanu, *Nonsmooth Variational Problems and Their Inequalities* (Springer, New York, 2007)
104. H. Cartan, *Calcul Différentiel* (Hermann, Paris, 1967)
105. E. Casas, L.A. Fernández, A Green's formula for quasilinear elliptic operators. J. Math. Anal. Appl. **142**, 62–73 (1989)
106. C. Castaing, Sur les multi-applications mesurables. Rev. Française Informat. Recherche Opérat. **1**, 91–126 (1967)
107. C. Castaing, P. Raynaud de Fitte, M. Valadier, *Young Measures on Topological Spaces. With Applications in Control Theory and Probability Theory*, Mathematics and its Applications, vol. 571 (Kluwer Academic Publishers, Dordrecht, 2004)
108. C. Castaing, M. Valadier, *Convex Analysis and Measurable Multifunctions*, Lecture Notes in Mathematics, vol. 580 (Springer, Berlin, 1977)
109. A. Castro, A.C. Lazer, Critical point theory and the number of solutions of a nonlinear Dirichlet problem. Ann. Mat. Pura Appl. **120**, 113–137 (1979)
110. A. Cauchy, *Oeuvres Complètes*, vol. XI (Gauthier-Villars, Paris, 1958), pp. 399–465
111. A. Cellina, Approximation of set-valued functions and fixed point theorems. Ann. Mat. Pura Appl. **82**, 17–24 (1969)
112. A. Cellina, The role of approximation in the theory of multi-valued mappings, in *Differential Games and Related Topics*, ed. by H.W. Kuhn, G.P. Szegö (North Holland Publishing Company, Amsterdam, 1971), pp. 209–220
113. A. Cellina, G. Colombo, A. Fonda, Approximate selections and fixed points for upper semicontinuous maps with decomposable values. Proc. Am. Math. Soc. **98**, 663–666 (1986)
114. A. Cellina, A. Lasota, A new approach to the definition of topological degree for multivalued mappings. Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Math. Natur. **47**, 154–160 (1969)
115. G. Cerami, Un criterio di esistenza per i punti critici su varieta illimitate, Istit. Lombardo Accad. Sci. Lett. Rend. A **112**, 332–336 (1978)
116. L. Cesari, *Optimization Theory and Applications* (Springer, New York, 1983)
117. K.C. Chang, Solutions of asymptotically linear operator equations via Morse theory. Commun. Pure Appl. Math. **34**, 693–712 (1981)
118. K.C. Chang, *Infinite Dimensional Morse Theory and Multiple Solution Problems* (Birkhäuser, Boston, 1993)
119. K.C. Chang, *Methods in Nonlinear Analysis* (Springer, Berlin, 2005)
120. K.C. Chang, N. Ghoussoub, The Conley index and the critical groups via an extension of Gromoll-Meyer theory. Topol. Methods Nonlinear Anal. **7**, 77–93 (1996)
121. G. Choquet, Theory of capacities. Ann. Inst. Fourier, Grenoble **5**, 131–295 (1955)
122. G. Choquet, in *Lectures on Analysis. Vol. I: Integration and Topological Vector Spaces*, ed. by J. Marsden, T. Lance, S. Gelbart (W.A. Benjamin, Inc., New York, Amsterdam, 1969)
123. S.N. Chow, J.K. Hale, *Methods of Bifurcation Theory* (Springer, New York, 1982)
124. S.K. Chua, R.L. Wheeden, Estimates of best constants for weighted Poincaré inequalities on convex domains. Proc. Lond. Math. Soc. **93**, 197–226 (2006)
125. I. Cioranescu, *Geometry of Banach Spaces, Duality Mappings and Nonlinear Problems* (Kluwer Academic Publishers, Boston, 1990)
126. D.C. Clark, A variant of the Lusternik-Schnirelmann theory. Indiana Univ. Math. J. **22**, 65–74 (1972/1973)
127. C.V. Coffman, A minimum-maximum principle for a class of nonlinear integral equations. J. Anal. Math. **22**, 391–419 (1969)
128. J.-N. Corvellec, A. Hantoute, Homotopical stability of isolated critical points of continuous functionals. Set-Valued Anal. **10**, 143–162 (2002)
129. D. Costa, *An Invitation to Variational Methods in Differential Equations* (Birkhäuser, Boston, 2007)
130. D. Costa, E.A. Silva, The Palais-Smale condition versus coercivity. Nonlinear Anal. **16**, 371–381 (1991)
131. J. Cronin, *Fixed Points and Topological Degree in Nonlinear Analysis*, Mathematical Surveys, No. 11 (American Mathematical Society, Providence, R.I., 1964)

132. M. Cuesta, Minimax theorems on C^1-manifolds via Ekeland variational principle. Abstr. Appl. Anal. **13**, 757–768 (2003)
133. B. Dacorogna, *Direct Methods in the Calculus of Variations* (Springer, New York, 1989)
134. G. Dal Maso, *An Introduction to* Γ*-Convergence* (Birkhäuser, Boston, 1993)
135. J. Daneš, A geometric theorem useful in nonlinear functional analysis. Boll. Un. Mat. Ital. **6**, 369–375 (1972)
136. G. Darbo, Punti uniti in trasformazioni a codominio non compatto. Rend. Sem. Mat. Univ. Padova **24**, 84–92 (1955)
137. R. Dautray, J.-L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology. Vol. 2. Functional and Variational Methods* (Springer, Berlin, 1988)
138. M. Day, *Normed Linear Spaces-Selected Topics*, Lecture Notes in Mathematics, vol. 486 (Springer, Berlin, 1975)
139. F.S. De Blasi, J. Myjak, A remark on the definition of topological degree for set-valued mappings. J. Math. Anal. Appl. **92**, 445–451 (1983)
140. H. Debrunner, P. Flor, Ein Erweiterungssatz für monotone Mengen. Arch. Math. **15**, 445–447 (1964)
141. M. Degiovanni, S. Lancelotti, K. Perera, Nontrivial solutions of p-superlinear p-Laplacian problems via a cohomological splitting. Commun. Contemp. Math. **12**, 475–486 (2010)
142. K. Deimling, *Nonlinear Functional Analysis* (Springer, Berlin, 1985)
143. Z. Denkowski, S. Migorski, N.S. Papageorgiou, *An Introduction to Nonlinear Analysis: Theory* (Kluwer Academic Publishers, Boston, 2003)
144. J. Deny, J.-L. Lions, Les espaces du type Beppo-Levi. Ann. Inst. Fourier (Grenoble) **5**, 305–370 (1955)
145. R. DiPerna, Measure-valued solutions to conservation laws. Arch. Rat. Mech. Anal. **88**, 223–270 (1985)
146. R. DiPerna, A. Majda, Oscillations and concentrations in weak solutions of the incompressible fluid equations. Commun. Math. Phys. **108**, 667–689 (1987)
147. J. Diestel, *Geometry of Banach Spaces*, Lecture Notes in Mathematics, vol. 485 (Springer, Berlin-New York, 1975)
148. A. Dold, *Lectures on Algebraic Topology*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 200 (Springer, Berlin-New York, 1980)
149. J. Dugundji, An extension of Tietze's theorem. Pacific J. Math. **1**, 353–367 (1951)
150. J. Dugundji, *Topology* (Allyn and Bacon, Boston, 1966)
151. N. Dunford, J. Schwartz, *Linear Operators I* (Wiley, New York, 1958)
152. E. Dyer, A fixed point theorem. Proc. Am. Math. Soc. **7**, 662–672 (1956)
153. M. Edelstein, The construction of an asymptotic center with a fixed-point property. Bull. Am. Math. Soc. **78**, 206–208 (1972)
154. R.E. Edwards, *Functional Analysis: Theory and Applications* (Holt, Rinehart and Winston, New York, 1965)
155. S. Eilenberg, D. Montgomery, Fixed point theorems for multivalues transformations. Am. J. Math. **58**, 214–222 (1946)
156. S. Eilenberg, N. Steenrood, *Foundations of Algebraic Topology* (Princeton University Press, Princeton, New Jersey, 1952)
157. I. Ekeland, On the variational principle. J. Math. Anal. Appl. **47**, 324–353 (1974)
158. I. Ekeland, Nonconvex minimization problems. Bull. Am. Math. Soc. (NS) **1**, 443–474 (1979)
159. I. Ekeland, The ε-variational principle revisited, in *Methods of Nonconvex Analysis*, Lecture Notes in Mathematics, vol. 1446, ed. by A. Cellina (Springer, Berlin, 1989), pp. 1–15
160. I. Ekeland, *Convexity Methods in Hamiltonian Mechanics* (Springer, Berlin, 1990)
161. I. Ekeland, R. Temam, *Convex Analysis and Variational Problems* (North-Holland, Amsterdam, 1976)
162. P. Enflo, A counterexample to the approximation problem in Banach spaces. Acta Math. **130**, 309–317 (1973)

163. L. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics, vol. 19 (American Mathematical Society, Providence, R.I., 1998)
164. L. Evans, R. Gariepy, *Measure Theory and Fine Properties of Functions* (CRC Press, Boca Raton, FL, 1992)
165. E.R. Fadell, P. Rabinowitz, Generalized cohomological index theories for Lie groups actions with applications to bifurcation questions for Hamiltonian systems. Invent. Math. **45**, 139–174 (1978)
166. K. Fan, A generalization of Tychonoff's fixed point theorem. Math. Ann. **142**, 305–310 (1961)
167. W. Fenchel, On conjugate convex functions. Canad. J. Math. **1**, 73–77 (1949)
168. G. de Figueiredo, *Lectures on the Ekeland Variational Principle with Applications and Detours*, Tata Institute of Fundamental Research, Bombay (Springer, Berlin, 1989)
169. G. de Figueiredo, S. Solimini, A variational approach to superlinear elliptic problems. Commun. Partial Differ. Equ. **9**, 699–717 (1984)
170. M. Filippakis, A. Kristaly, N. Papageorgiou, Existence of five nonzero solutions with exact sign for a p-Laplacian equation. Discrete Contin. Dyn. Syst. **24**, 405–440 (2009)
171. K. Floret, *Weakly Compact Sets*, Lecture Notes in Mathematics, vol. 801 (Springer, Berlin, 1980)
172. I. Fonseca, W. Gangbo, *Degree Theory in Analysis and Applications* (Clarendon Press, Oxford, UK, 1995)
173. K. Friedrichs, Spektraltheorie halbbeschränkter Operatoren und Anwendung auf die Spektralzerlegung von Differentialoperatoren. Math. Ann. **109**, 465–487 and 685–713 (1934)
174. A. Fryszkowski, Continuous selections for a class of nonconvex multivalued maps. Studia Math. **75**, 163–174 (1983)
175. A. Fryszkowski, *Fixed Point Theory for Decomposable Sets* (Kluwer Academic Publishers, New York, 2004)
176. L. Führer, Ein elementarer analytischer Beweis zur Eindeutigkeit des Abbildungsgrades im $\mathbb{R}^n$. Math. Nachr. **54**, 259–267 (1972)
177. M. Furi, A. Vignoli, On a property of the unit sphere in a linear normed space. Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. **18**, 333–334 (1970)
178. E. Gagliardo, Proprietà di alcune classi di funzioni in più variabili. Ric. Mat. **7**, 102–137 (1958)
179. R.V. Gamkrelidze, *Principles of Optimal Control Theory, Mathematical Concepts and Methods in Science and Engineering*, vol. 7 (Plenum Press, New York-London, 1978)
180. L. Gasinski, N.S. Papageorgiou, Solutions and multiple solutions for quasilinear hemivariational inequalities at resonance. Proc. R. Soc. Edinb. Sect. A **131**, 1091–1111 (2001)
181. L. Gasinski, N.S. Papageorgiou, *Nonsmooth Critical Point Theory and Nonlinear Boundary Value Problems* (Chapman and Hall/CRC Press, Boca Raton, FL, 2005)
182. L. Gasinski, N.S. Papageorgiou, *Nonlinear Analysis* (Chapman and Hall/CRC, Boca Raton, 2006)
183. L. Gasinski, N.S. Papageorgiou, Nodal and multiple constant sign solutions for resonant p-Laplacian equations with a nonsmooth potential. Nonlinear Anal. **71**, 5747–5772 (2009)
184. N. Ghoussoub, *Duality and Perturbation Methods in Critical Point Theory* (Cambridge University Press, Cambridge, UK, 1993)
185. N. Ghoussoub, D. Preiss, A general mountain pass principle for locating and classifying critical points. Ann. Inst. Henri Poincaré Anal. Non Linéaire **6**, 321–330 (1989)
186. M. Giaquinta, *Multiple Integrals in the Calculus of Variations and Nonlinear Elliptic Systems* (Princeton University Press, Princeton, 1983)
187. D. Gilbarg, N. Trudinger, *Elliptic Partial Differential Equations of Second Order* (Springer, Berlin, 1983)
188. J. Giles, *Convex Analysis with Applications in the Differentiation of Convex Functions*, Research Notes in Mathematics, vol. 58 (Pitman, Boston, Mass, 1982)
189. I.L. Glicksberg, A further generalization of the Kakutani fixed theorem, with application to Nash equilibrium points. Proc. Am. Math. Soc. **3**, 170–174 (1952)

190. K. Goebel, W.A. Kirk, *Topics in Metric Fixed Point Theory*, Cambridge Studies in Advanced Mathematics, vol. 28 (Cambridge University Press, Cambridge, 1990)
191. D. Göhde, Zum Prinzip der kontraktiven Abbildung. Math. Nachr. **30**, 251–258 (1965)
192. L.S. Goldenstein, I. Gohberg, A.S. Markus, Investigation of some properties of bounded linear operators in the connection with their g-norm (Russian). Uchen Zap. Kishinev Gos. Univ. **29**, 29–36 (1957)
193. M. Golomb, Zur Theorie der nichtlinearen Integralgleichungen, Integralgleichungssysteme und allgemeinen Funktionalgleichungen. Math. Z. **39**, 45–75 (1935)
194. L. Górniewicz, *Topological Fixed Point Theory of Multivalued Mappings*, Topological Fixed Point Theory and Its Applications, vol. 4 (Springer, Dordrecht, 2006)
195. A. Granas, Sur la notion du degré topologique pour une certaine classe de transformations multivalentes dans les espaces de Banach. Bull. Acad. Polon. Sci. Sér. Sci. Math. Astr. Phys. **7**, 191–194 (1959)
196. A. Granas, The theory of compact vector fields and some of its applications to topology of functional spaces. Rozpr. Mat. **30**, 93 (1962)
197. A. Granas, J. Dugundji, *Fixed Point Theory* (Springer, New York, 2003)
198. D. Gromoll, W. Meyer, On differentiable functions with isolated critical points. Topology **8**, 361–369 (1969)
199. D.J. Guo, V. Lakshmikantham, *Nonlinear Problems in Abstract Cones*, Notes and Reports in Mathematics in Science and Engineering (Academic Press, Boston, 1988)
200. J. Hadamard, Sur quelques applications de l'indice de Kronecker, Appendix in *Introduction à la théorie des fonctions d'une variable II*, 2nd edn. (1910)
201. B.R. Halpern, G.M. Bergman, A fixed point theorem for inward and outward maps. Trans. Am. Math. Soc. **130**, 353–358 (1968)
202. P. Hartman, G. Stampacchia, On some nonlinear elliptic differential-functional equations. Acta Math. **115**, 271–310 (1966)
203. A. Hatcher, *Algebraic Topology* (Cambridge University Press, Cambridge, 2002)
204. J. Heinonen, T. Kilpelainen, O. Martio, *Nonlinear Potential Theory of Degenerate Elliptic Equations* (Oxford Science Publications, Oxford, UK, 1993)
205. E. Heinz, An elementary analytic theory of the degree of mapping in n-dimensional space. J. Math. Mech. **8**, 231–247 (1959)
206. P. Hess, On nonlinear problems of monotone type with respect to two Banach spaces. J. Math. Pures Appl. **52**, 13–26 (1973)
207. M.R. Hestenes, Extension of the range of a differentiable function. Duke Math. J. **8**, 183–192 (1941)
208. D. Hilbert, *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen I–VI*, Nachr. Ges. Wiss. Göttingen Math.-Phys. Klasse; **I** (1904), 49–91; **II** (1905), 213–259; **III** (1905), 307–338; **IV** (1906), 157–227; **V** (1906), 439–480; **VI** (1910), 355–417; [published in book form by Teubner, Leipzig (1912); reprinted by Chelsea, New York (1953)]
209. H. Hofer, Variational and topological methods in partially ordered Hilbert spaces. Math. Ann. **261**, 493–514 (1982)
210. H. Hofer, A geometric description of the neighborhood of a critical point given by the mountain pass theorem. J. Lond. Math. Soc. **31**, 566–570 (1985)
211. H. Hofer, The topological degree at a critical point of mountain-pass type, Nonlinear functional analysis and its applications, Part 1 (Berkeley, Calif., 1983), pp. 501–509, *Proceedings of Symposia in Pure Mathematics*, vol. 45, Part 1, (American Mathematical Society, Providence, RI, 1986)
212. H. Hofer, A strong form of the mountain pass theorem and application, *Nonlinear diffusion equations and their equilibrium states I*, (Berkeley, CA, 1986), Mathematical Sciences Research Institute Publications, vol. 12 (Springer, New York, 1988), pp. 341–350
213. H. Hopf, Abbildungsklassen n-dimensionaler Mannigfaltigkeiten. Math. Ann. **96**, 209–224 (1927)
214. H. Hopf, Vektorfelder in n-dimensionalen Mannigfaltigkeiten. Math. Ann. **96**, 225–249 (1927)

215. H. Hopf, Zur Topologie der Abbildungen von Mannigfaltigkeiten. Math. Ann. **100**, 579–608 (1928)
216. S.T. Hu, *Theory of Retracts* (Wayne State University Press, Detroit, 1965)
217. S. Hu, N.S. Papageorgiou, Generalizations of Browder's degree theory. Trans. Am. Math. Soc. **347**, 233–259 (1995)
218. S. Hu, N.S. Papageorgiou, *Handbook of Multivalued Analysis. Volume I: Theory* (Kluwer Academic Publishers, Dordrecht, 1997)
219. S. Hu, N.S. Papageorgiou, *Handbook of Multivalued Analysis. Volume II: Applications* (Kluwer Academic Publishers, Dordrecht, 2000)
220. H. Hukuhara, Sur l'application semi-continue dont la valeur est un compact convexe. Funkcial. Ekvac. **10**, 43–66 (1967)
221. A. Ioffe, V. Tichomirov, *Theory of Extremal Problems* (North-Holland, Amsterdam, 1979)
222. V. Istrăţescu, *Fixed Point Theory: An Introduction* (Reidel Publishing Company, Dordrecht-Boston, 1981)
223. Y. Jabri, *The Mountain Pass Theorem. Variants, Generalizations and Some Applications* (Cambridge University Press, Cambridge, UK, 2003)
224. R. Kachurovski, On monotone operators and convex functionals (Russian). Uspekhi Mat. Nauk **15**, 213–215 (1960)
225. R. Kachurovski, Nonlinear monotone operators on Banach spaces. Russ. Math. Surv. **23**, 117–165 (1968)
226. S. Kakutani, A generalization of Brouwer's fixed point theorem. Duke Math. J. **8**, 457–459 (1941)
227. M. Kamenskii, V. Obukhovskii, P. Zecca, *Condensing Multivalued Maps and Semilinear Differential Inclusions in Banach Spaces*, De Gruyter Series in Nonlinear Analysis and Applications (Walter de Gruyter, Berlin, 2001)
228. R. Kannan, C.K. Krueger, *Advanced Analysis on the Real Line* (Springer, New York, 1996)
229. T. Kato, *Perturbation Theory for Linear Operators* (Springer, Berlin, 1976)
230. O. Kavian, *Introduction à la théorie des points critiques et applications aux problemes elliptiques*, Mathématiques & Applications, vol. 13 (Springer, Paris, 1993)
231. N. Kenmochi, Pseudo-monotone operators and nonlinear elliptic boundary value problems. J. Math. Soc. Jpn. **27**, 121–149 (1975)
232. B.T. Kien, M.M. Wong, N.C. Wong, On the degree theory for general mappings of monotone type. J. Math. Anal. Appl. **340**, 707–720 (2008)
233. D. Kinderlehrer, P. Pedregal, Weak convergence of integrands and the Young measure representation. SIAM J. Math. Anal. **23**, 1–19 (1992)
234. D. Kinderlehrer, G. Stampacchia, *An Introduction to Variational Inequalities and Their Applications* (Academic Press, New York-London, 1980)
235. W.A. Wirk, A fixed point theorem for mappings which do not increase distances. Am. Math. Mon. **72**, 1004–1006 (1965)
236. M. Kisielewicz, *Differential Inclusions and Optimal Control* (Kluwer Academic Publishers, Dordrecht, 1991)
237. V. Klee, An example related to the fixed-point property. Nieuw Arch. Wisk. **8**, 81–82 (1960)
238. E. Klein, A. Thompson, *Theory of Correspondences* (Wiley, New York, 1984)
239. B. Knaster, Un théorème sur les fonctions d'ensembles. Ann. Soc. Polon. Math. **6**, 133–134 (1928)
240. B. Knaster, C. Kuratowski, S. Mazurkiewicz, Ein Beweis des Fixpunktsatzes für n-dimensionale Simplexe. Fund. Math. **14**, 132–137 (1929)
241. J. Kobayashi, M. Otani, The principle of symmetric criticality for nondifferentiable mappings. J. Funct. Anal. **214**, 428–449 (2004)
242. J. Kobayashi, M. Otani, Topological degree for $(S)_+$-mappings with maximal monotone perturbations and its applications to variational inequalities. Nonlinear Anal. **59**, 147–172 (2004)
243. J. Kobayashi, M. Otani, Degree for subdifferential operators in Hilbert spaces. Adv. Math. Sci. Appl. **14**, 307–325 (2004)

244. J. Kobayashi, M. Otani, An index formula for the degree of $(S)_+$-mappings associated with one-dimensional p-Laplacian. Abstr. Appl. Anal. **11**, 981–995 (2004)
245. A.N. Kolmogorov, S.V. Fomin, *Elements of the Theory of Functions and Functional Analysis. Vol. 1. Metric and Normed Spaces* (Translated from the first Russian edition by Leo F. Boron). (Graylock Press, Rochester, N.Y., 1957)
246. V.I. Kondrachov, Certain properties of functions in the spaces L^p. Dokl. Akad. Nauk SSR **48**, 535–538 (1945)
247. N.C. Kourogenis, N.S. Papageorgiou, Nonsmooth critical point theory and nonlinear elliptic equations at resonance. J. Aust. Math. Soc. A **69**, 245–271 (2000)
248. N.C. Kourogenis, N.S. Papageorgiou, A weak nonsmooth Palais-Smale condition and coercivity. Rend. Circ. Mat. Palermo **49**, 521–526 (2000)
249. N.C. Kourogenis, N.S. Papageorgiou, Multiple solutions for discontinuous strongly resonant elliptic problems. J. Math. Soc. Jpn. **53**, 17–34 (2001)
250. M.A. Krasnoselskii, *Topological Methods in the Theory of Nonlinear Integral Equations* (Pergamon, Oxford, UK, 1965)
251. W. Krawcewicz, J. Wu, *Theory of Degrees with Applications to Bifurcations and Differential Equations*, Canadian Mathematical Society Series of Monographs and Advanced Texts (Wiley, New York, 1997)
252. M.G. Krein, Propriétés fondamentales des ensembles coniques normaux dans l'espace de Banach. Dokl. Akad. Nauk USSR **28**, 13–17 (1940)
253. K. Kuratowski, Sur les espaces complets. Fund. Math. **15**, 301–309 (1930)
254. K. Kuratowski, C. Ryll-Nardzewski, A general theorem on selectors. Bull. Acad. Polon. Sci. Ser. Sci. Math. Astronom. Phys. **13**, 397–403 (1965)
255. I. Kuzin, S. Pohozaev, *Entire Solutions of Semilinear Elliptic Equations* (Birkhäuser, Basel, 1997)
256. S. Kyritsi, N.S. Papageorgiou, Nonsmooth critical point theory on closed convex sets and nonlinear hemivariational inequalities. Nonlinear Anal. **61**, 373–403 (2005)
257. V. Lakshmikantham, S. Leela, *Nonlinear Differential Equations in Abstract Spaces, International Series in Nonlinear Mathematics: Theory, Methods and Applications*, vol. 2 (Pergamon Press, Oxford-New York, 1981)
258. S. Lang, *Real and Functional Analysis*, Graduate Texts in Mathematics, vol. 142 (Springer, New York, 1993)
259. J.-M. Lasry, R. Robert, Degré pour les fonctions multivoques et applications. C. R. Acad. Sci. Paris Sér. A-B **280**, 1435–1438 (1975)
260. P. Laurent, *Approximation et Optimisation*, Collection Enseignement des Sciences, No. 13, Hermann, Paris, 1972
261. P.D. Lax, A.N. Milgram, Parabolic equations, *Contributions to the Theory of Partial Differential Equations*, Annals of Mathematics Studies, No. 33 (Princeton University Press, Princeton, 1954), pp. 167–190
262. G. Leoni, *A First Course in Sobolev Spaces*, Graduate Studies in Mathematics, vol. 105 (American Mathematical Society, Providence, R.I., 2009)
263. J. Leray, Les problèmes non linéaires. L'Enseignement Mathématique, Revue Internationale **35**, 139–151 (1936)
264. J. Leray, Valeurs propres et vecteurs propres d'un endomorphisme complètement continu d'un espace vectoriel à voisinages convexes. Acta Sci. Math. Szeged **12**, 177–186 (1950)
265. J. Leray, La théorie des points fixes et ses applications en analyse, in *Proceedings of the International Congress of Mathematicians*, Cambridge, 1950, vol. 2, (American Mathematical Society, Providence, 1952), pp. 202–208
266. J. Leray, J. Schauder, Topologie et équations fonctionnelles. Ann. Sci. École Norm. Sup. **51**, 45–78 (1934)
267. B. Levi, Sul principio di Dirichlet. Rend. Circ. Mat. Palermo **22**, 293–359 (1906)
268. C. Li, S. Li, J. Liu, Splitting theorem, Poincaré-Hopf theorem and jumping nonlinear problems. J. Funct. Anal. **221**, 439–455 (2005)

269. S. Li, M. Willem, Applications of local linking to critical point theory. J. Math. Anal. Appl. **189**, 6–32 (1995)
270. L. Lichtenstein, Eine elementare Bemerkung zur reellen Analysis. Math. Z. **30**, 794–795 (1929)
271. T.-C. Lim, On fixed point stability for set-valued contractive mappings with applications to generalized differential equations. J. Math. Anal. Appl. **110**, 436–441 (1985)
272. J. Lindenstrauss, L. Tzafriri, *Classical Banach Spaces I* (Springer, Berlin-New York, 1977)
273. J.-L. Lions, Sur les espaces d'interpolation; dualité. Math. Scand. **9**, 147–177 (1961)
274. J.-L. Lions, Théorèmes de trace et d'interpolation, IV. Math. Ann. **151**, 42–56 (1963)
275. J.-L. Lions, *Quelques Méthodes de Résolution des Problèmes aux Limites Non Linéaires* (Gauthier-Villars, Paris, 1969)
276. J. Liu, The Morse index of a saddle point. Syst. Sci. Math. Sci. **2**, 32–39 (1989)
277. S. Liu, Remarks on multiple solutions for elliptic resonant problems. J. Math. Anal. Appl. **336**, 498–505 (2007)
278. J. Liu, S. Li, An existence theorem for multiple critical points and its applications. Kexue Tongbao (Chinese) **29**, 1025–1027 (1984)
279. S. Liu, S. Li, Critical groups at infinity, saddle point reduction and elliptic resonant problems. Comm. Contemp. Math. **5**, 761–773 (2003)
280. L. Ljusternik, Topologische Grundlangen der allgemeinen Eigenwertheorie. Monatsch. Math. Phys. **37**, 125–130 (1930)
281. L. Ljusternik, L. Schnirelmann, *Méthodes Topologiques dans les Problèmes Variationnels* (Hermann, Paris, 1934)
282. L. Ljusternik, L. Schnirelmann, Topological methods in variational problems and their application to the differential geometry of surfaces. Uspekhi Mat. Nauk **2**, 166–217 (1947). (Russian)
283. N.G. Lloyd, *Degree Theory*, Cambridge Tracts in Mathematics, No. 73 (Cambridge University Press, Cambridge-New York-Melbourne, 1978)
284. W. Lopez, An example in the fixed point theory of polyhedra. Bull. Am. Math. Soc. **73**, 922–924 (1967)
285. T.-W. Ma, Topological degrees of set-valued compact fields in locally convex spaces. Diss. Math. **92**, 43 (1972)
286. M. Marcus, V. Mizel, Absolute continuity on tracks and mappings of Sobolev spaces. Arch. Rat. Mech. Anal. **45**, 294–320 (1972)
287. M. Marcus, V. Mizel, Complete characterization of functions which act, via superposition, on Sobolev spaces. Trans. Am. Math. Soc. **251**, 187–218 (1979)
288. A. Marino, G. Prodi, Metodi perturbativi nella teoria di Morse. Boll. Unione Mat. Italiana **11**, 1–34 (1975)
289. J.T. Markin, A fixed point theorem for multivalued mappings. Bull. Am. Math. Soc. **74**, 639–640 (1968)
290. J.T. Markin, Continuous dependence of fixed point sets. Proc. Am. Math. Soc. **38**, 545–557 (1973)
291. R.H. Martin Jr., *Nonlinear Operators and Differential Equations in Banach Spaces*, Pure and Applied Mathematics (Wiley, New York-London-Sydney, 1976)
292. C.R.F. Maunder, *Algebraic Topology* (Dover Publications Inc., Mincola, NY, 1996)
293. J. Mawhin, M. Willem, *Critical Point Theory and Hamiltonian Systems*, Applied Mathematical Sciences, vol. 74 (Springer, New York, 1989)
294. V. Mazja, *Sobolev Spaces*, Springer Series in Soviet Mathematics (Springer, Berlin, 1985)
295. R. Megginson, *An Introduction to Banach Space Theory* (Springer, New York, 1998)
296. N. Meyers, J. Serrin, $H = W$. Proc. Nat. Acad. USA **51**, 1055–1056 (1964)
297. E. Michael, Topologies on spaces of subsets. Trans. Am. Math. Soc. **71**, 152–182 (1951)
298. E. Michael, Continuous selections I. Ann. Math. **63**, 361–382 (1956)
299. J. Milnor, *Topology from the differentiable viewpoint*, Based on notes by David W. Weaver (The University Press of Virginia, Charlottesville, 1965)
300. G. Minty, Monotone nonlinear operators in Hilbert spaces. Duke Math. J. **29**, 341–346 (1962)

301. G. Minty, On a monotonicity method for the solution of nonlinear equations in Banach spaces. Proc. Natl. Acad. Sci. USA **50**, 1038–1041 (1963)
302. H. Mönch, Boundary value problems for nonlinear ordinary differential equations of second order in Banach spaces. Nonlinear Anal. **4**, 985–999 (1980)
303. J.-J. Moreau, Proximité et dualité dans un espace hilbertien. Bull. Soc. Math. Fr. **93**, 273–299 (1965)
304. J.-J. Moreau, *Fonctionnelles Convexes*, Seminaire sur les équations aux dérivées partielles II, College de France, Paris, 1966–1967
305. C.B. Morrey, Functions of several variables and absolute continuity II. Duke Math. J. **6**, 187–215 (1940)
306. C.B. Morrey, *Multiple Integrals in the Calculus of Variations* (Springer, New York, 1966)
307. M. Morse, Relations between the critical points of a real function of n independent variables. Trans. Am. Math. Soc. **27**, 345–396 (1925)
308. M. Morse, *The Calculus of Variations in the Large*, vol. 18 (Colloquium Publications, American Mathematical Society, New York, 1934)
309. D. Motreanu, V. Motreanu, N.S. Papageorgiou, A degree theoretic approach for multiple solutions of constant sign for nonlinear elliptic equations. Manuscr. Math. **124**, 507–531 (2007)
310. D. Motreanu, V. Motreanu, N.S. Papageorgiou, *Topological and Variational Methods with Applications to Nonlinear Boundary Value Problems* (Springer, New York, 2014)
311. D. Motreanu, V. Rădulescu, *Variational and Non-Variational Methods in Nonlinear Analysis and Boundary Value Problems* (Kluwer Academic Publishers, Dordrecht, The Netherlands, 2003)
312. S.B. Nadler Jr., *The Fixed Point Property for Continua*, Aportaciones Matemáticas: Textos [Mathematical Contributions: Texts], vol. 30 (Sociedad Matemática Mexicana, México, 2005)
313. M. Nagumo, A theory of degree of mapping based on infinitesimal analysis. Am. J. Math. **73**, 485–496 (1951)
314. M. Nagumo, Degree of mapping in convex linear topological spaces. Am. J. Math. **73**, 497–511 (1951)
315. J.P. Natanson, *Theory of Functions of a Real Variable*, vol. I (Frederick Ungar Publishing Company, New York, 1961)
316. Z. Nehari, Characteristic values associated with a class of nonlinear second order differential equations. Acta Math. **105**, 141–175 (1961)
317. J. von Neumann, On rings of operators. Reduction theory. Ann. Math. **50**, 401–485 (1949)
318. L. Nirenberg, An extended interpolation inequality. Ann. Scuola Norm. Sup. Pisa **20**, 733–737 (1966)
319. L. Nirenberg, *Topics in Nonlinear Functional Analysis*, Lecture Notes (Courant Institute of Mathematical Sciences New York University, New York, 1974)
320. R.D. Nussbaum, The fixed point index for local condensing maps. Ann. Mat. Pura Appl. **89**, 217–258 (1971)
321. R.D. Nussbaum, Degree theory for local condensing maps. J. Math. Anal. Appl. **37**, 741–766 (1972)
322. J. Oinas, The degree theory and the index of a critical point for mappings of the type (S_+). *Acta Univ. Oulu. Ser. A Sci. Rerum Natur.***488**, 117 (2007)
323. B. O'Neil, Essential sets and fixed points. Am. J. Math. **75**, 497–509 (1953)
324. M. Otani, Nonmonotone perturbations for nonlinear evolution equations associated with subdifferential operators, periodic problems. J. Differ. Equ. **54**, 248–273 (1984)
325. R.S. Palais, Morse theory on Hilbert manifolds. Topology **2**, 299–340 (1963)
326. R.S. Palais, Ljusternik-Schnirelmann theory on Banach manifolds. Topology **5**, 115–132 (1966)
327. R.S. Palais, The principle of symmetric criticality. Commun. Math. Phys. **69**, 19–30 (1979)
328. R.S. Palais, S. Smale, A generalized Morse theory. Bull. Am. Math. Soc. **70**, 165–172 (1964)
329. N.S. Papageorgiou, S. Kyritsi, *Handbook of Applied Analysis* (Springer, New York, 2009)

330. N.S. Papageorgiou, V.D. Rădulescu, Resonant $(p, 2)$-equations with asymmetric reaction. Anal. Appl. (Singap.) **13**, 481–506 (2015)
331. D. Pascali, S. Sburlan, *Nonlinear Mappings of Monotone Type, Martinus Nijhoff Publishers* (The Hague; Sijthoff & Noordhoff International Publishers, Alphen aan den Rijn, 1978)
332. P. Pedregal, *Parametrized Measures and Variational Principles* (Birkhäuser, Basel, 1997)
333. J.-P. Penot, The drop theorem, the petal theorem and Ekeland's variational principle. Nonlinear Anal. **10**, 813–822 (1986)
334. A. Peressini, *Ordered Topological Vector Spaces* (Harper & Row Publishers, New York-London, 1967)
335. K. Perera, Homological local linking. Abstr. Appl. Anal. **3**, 181–189 (1998)
336. K. Perera, R. Agarwal, D. O'Regan, *Morse Theoretic Aspects of p-Laplacian Type Operators*, Mathematical Surveys and Monographs, vol. 161 (American Mathematical Society, Providence, RI, 2010)
337. W. Petryshyn, On the approximation-solvability of nonlinear equations. Math. Ann. **177**, 156–164 (1968)
338. W. Petryshyn, On the projectional solvability and the Fredholm alternative for equations involving linear A-proper operators. Arch. Rat. Mech. Anal **30**, 270–284 (1968)
339. W. Petryshyn, Nonlinear equations involving noncompact operators, in Nonlinear Functional Analysis, in *Proceedings of Symposia in Pure Mathematics*, vol. XVIII, Part 1, Chicago, ed. by F. Browder (American Mathematical Society, Providence, R.I., 1968), pp. 206–233
340. W. Petryshyn, On the approximation-solvability of equations involving A-proper and pseudo-A-proper mappings. Bull. Am. Math. Soc. **81**, 223–312 (1975)
341. R. Phelps, *Convex Functions, Monotone Operators and Differentiability*, Lecture Notes in Mathematics, vol. 1364 (Springer, Berlin, 1989)
342. E. Picard, *Traité d'Analyse*, Éditions Gauthier-Villars, 1905 (réédition Jacques Gabay, Sceaux, 1991)
343. P. Pucci, J. Serrin, A mountain pass theorem. J. Differ. Equ. **60**, 142–149 (1985)
344. P. Pucci, J. Serrin, The structure of the critical set in the mountain pass theorem. Trans. Am. Math. Soc. **299**, 115–132 (1987)
345. P. Rabinowitz, *Théorie du Degré Topologique et Applications à des Problèmes aux Limites Non Linéaires*, Lecture Notes (ed.) by H. Berestycki (Université Paris VI, Laboratoire d'Analyse Numérique, Paris, 1975)
346. P. Rabinowitz, Some minimax theorems and applications to nonlinear partial differential equations. Nonlinear Anal. (collection of papers in honor of Erich H. Rothe), 161–177
347. P. Rabinowitz, *Minimax methods in critical point theory with applications to differential equations*, CBMS Regional Conference Series in Mathematics, vol. 65. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1986
348. V.D. Rădulescu, Mountain pass theorems for nondifferentiable functions and applications. Proc. Jpn. Acad. Ser. Math. Sci. **69**(6), 193–198 (1993)
349. V.D. Rădulescu, Locally Lipschitz functionals with the strong Palais-Smale property. Rev. Roum. Math. Pures Appl. **40**, 355–372 (1995)
350. V.D. Rădulescu, A Lusternik-Schnirelmann type theorem for locally Lipschitz functionals with applications to multivalued periodic problems. Proc. Jpn. Acad. Ser. Math. Sci. **71**(7), 164–167 (1995)
351. V.D. Rădulescu, *Qualitative Analysis of Nonlinear Elliptic Partial Differential Equations: Monotonicity, Analytic, and Variational Methods*, Contemporary Mathematics and Its Applications, vol. 6 (Hindawi Publishing Corporation, New York, 2008)
352. M. Reed, B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis* (Academic Press, New York, 1972)
353. F. Rellich, Ein Satz über mittlere Konvergenz. Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. I 30–35 (1930)
354. F. Riesz, Über lineare Funktionalgleichungen. Acta Math. **41**, 71–98 (1918)

355. R.T. Rockafellar, Characterization of the subdifferential of convex functions. Pacific J. Math. **17**, 497–510 (1966)
356. R.T. Rockafellar, Duality and stability in extremum problems involving convex functions. Pacific J. Math. **21**, 200–205 (1967)
357. R.T. Rockafellar, Conjugates and Legendre transforms of convex functions. Canad. J. Math. **19**, 200–205 (1967)
358. R.T. Rockafellar, *Convex Analysis* (Princeton University Press, Princeton, N.J., 1970)
359. R.T. Rockafellar, On the maximal monotonicity of subdifferential mapping. Pacific J. Math. **33**, 209–216 (1970)
360. R.T. Rockafellar, *Conjugate Duality and Optimization*, CBMS, Regional Conference Series in Mathematics, vol. 16 (SIAM, Philadelphia PA, 1974)
361. F. Rothe, A relation between the type numbers of a critical point and the index of the corresponding field of gradient vectors. Math. Nachr. **4**, 12–27 (1950)
362. F. Rothe, Gradient mappings. Bull. Am. Math. Soc. **59**, 5–19 (1953)
363. F. Rothe, Morse theory in Hilbert space. Rocky Mt. J. Math. **3**, 251–274 (1973)
364. F. Rothe, *Introduction to Various Aspects of Degree Theory in Banach Spaces*, vol. 23 (American Mathematical Society, Providence, R.I., 1986)
365. T. Roubicek, *Relaxation in Optimization Theory and Variational Calculus* (Walter de Gruyter, Berlin, 1997)
366. W. Rudin, *Functional Analysis*, International Series in Pure and Applied Mathematics (McGraw-Hill Inc, New York, 1991)
367. B.N. Sadovskii, On a fixed point principle (Russian). Funkcional. Anal. i Prilozhen. **1**, 74–76 (1967)
368. B.N. Sadovskii, Limit compact and condensing operators. Russ. Math. Surv. **27**, 85–155 (1972)
369. M.-F. Saint-Beuve, On the extension of von Neumann-Aumann's theorem. J. Funct. Anal. **17**, 112–129 (1974)
370. A. Sard, The measure of critical values of differentiable maps. Bull. Am. Math. Soc. **48**, 883–890 (1942)
371. H. Schaefer, Über die Methode der a priori-Schranken. Math. Ann. **129**, 415–416 (1955)
372. J. Schauder, Zur theorie stetiger abbildungen in funktionalraumen. Math. Z. **26**, 417–431 (1927)
373. J. Schauder, Invarianz des gebietes in funktionalraumen. Studia Math. **1**, 123–139 (1929)
374. J. Schauder, Der fixpunktsatz in funktionalraumen. Studia Math. **2**, 171–180 (1930)
375. J. Schwartz, Generalizing the Ljusternik-Schnirelmann theory of critical points. Commun. Pure Appl. Math. **17**, 307–315 (1964)
376. J. Schwartz, *Nonlinear Functional Analysis* (Gordon and Breach Science Publishers, New York, 1969)
377. L. Schwartz, *Théorie des Distributions* (Hermann, Paris, 1966)
378. M. Schechter, New linking theorems. Rend. Sem. Mat. Univ. Padova **99**, 255–269 (1998)
379. M. Schechter, *Linking Methods in Critical Point Theory* (Birkhäuser, Boston, 2009)
380. R.T. Seeley, Extensions of C^∞-functions defined in a half-space. Proc. Am. Math. Soc. **15**, 625–626 (1964)
381. R.E. Showalter, *Hilbert Space Methods for Partial Differential Equations, Monographs and Studies in Mathematics*, vol. 1 (Pitman, London-San Francisco, 1977)
382. I.V. Skrypnik, *Nonlinear Higher Order Elliptic Equations (Russian)* (Izdat. Naukova Dumka, Kiev, 1973), p. 219
383. S. Smale, Morse theory and a nonlinear generalization of the Dirichlet problem. Ann. Math. **80**, 382–396 (1964)
384. D.R. Smart, *Fixed Point Theorems*, Cambridge Tracts in Mathematics, No. 66 (Cambridge University Press, London-New York, 1974)
385. G. Stampacchia, *Équations Elliptiques du Second Ordre à Coefficients Discontinus* (Presses University, Montreal, 1966)

386. G. Stampacchia, *Topological methods of investigation of general nonlinear elliptic boundary value problems* (Russian), Geometry and topology in global nonlinear problems, pp. 68–91, Novoe Global. Anal., Voronezh. Gos. Univ., Voronezh, 1984
387. G. Stampacchia, *Methods for Analysis of Nonlinear Elliptic Boundary Value Problems*, Translations of Mathematical Monographs, vol. 139 (American mathematical Society, Providence, RI, 1994)
388. S.L. Sobolev, Méthode nouvelle à résoudre le problème de Cauchy pour les équations linéaires hyperboliques normales. Mat. Sbornik **1**, 37–72 (1936)
389. S.L. Sobolev, On a theorem of functional analysis. Mat. Sbornik **4**, 471–496 (1938)
390. E.H. Spanier, *Algebraic Topology* (McGraw-Hill Book Company, New York-Toronto, 1966)
391. H. Steinlein, Ein Satz über den Leray-Schauderschen Abbildungsgrad. Math. Z. **126**, 176–208 (1972)
392. M. Struwe, On a critical point theory for minimal surfaces spanning a wire in $\mathbb{R}^n$. J. Reine Angew. Math. **349**, 1–23 (1984)
393. M. Struwe, *Variational Methods. Applications to Nonlinear Partial Differential Equations and Hamiltonian Systems*, vol. 34, 4th edn., Results in Mathematics and Related Areas, 3rd Series, A Series of Modern Surveys in Mathematics (Springer, Berlin, 2008)
394. J. Su, L. Zhao, An elliptic resonance problem with multiple solutions. J. Math. Anal. Appl. **319**, 604–616 (2006)
395. M.A. Sychev, A new approach to Young measure theory, relaxation and convergence in energy. Ann. Inst. H. Poincaré Anal. Non Linéaire **16**, 773–812 (1999)
396. S. Szarek, A Banach space without a basis which has the bounded approximation property. Acta. Math. **159**, 81–98 (1987)
397. A. Szulkin, Minimax principles for lower semicontinuous functions and applications to nonlinear boundary value problems. Ann. Inst. Henri Poincaré Anal. Non Linéaire **3**, 77–109 (1986)
398. A. Szulkin, Ljusternik-Schnirelmann theory on C^1-manifolds. Ann. Inst. Henri Poincaré Anal. Non Linéaire **5**, 119–139 (1988)
399. G. Talenti, Best constant in Sobolev inequality. Ann. Mat. Pura. Appl. **110**, 353–372 (1976)
400. A. Tarski, A lattice theoretical fixed point theorem and its applications. Pacific J. Math. **5**, 285–309 (1955)
401. L. Tartar, *An Introduction to Sobolev Spaces and Interpolation Spaces*, Lecture Notes of UMI, vol. 3 (Springer, Berlin, 2007)
402. A. Tarski, A lattice-theoretical fixpoint theorem and its applications. Pacific J. Math. **5**, 285–309 (1955)
403. V. Tichomirov, *Fundamental Principles of the Theory of Extremal Problems* (Wiley, New York, 1986)
404. L. Tonelli, *Fondamenti di Calcolo delle Variazioni* (Zanichelli, Bologna, 1921)
405. S. Troyanski, On locally uniformly convex and differentiable norms in certain nonseparable Banach spaces. Studia Math. **37**, 173–180 (1971)
406. N. Trudinger, On embeddings into Orhcz spaces and some applications. J. Math. Mech. **17**, 473–483 (1967)
407. A. Tychonoff, Ein fixpunktsatz. Math. Ann. **111**, 767–776 (1935)
408. M. Valadier, *Structure des atomes conditionnels d'un espace de probabilité*, Travaux Sém. Anal. Convexe **10** (1980), no. 1, exp. no. 1, 17 pp
409. M. Valadier, Young measures, in *Methods of Nonconvex Analysis*, Lecture Notes in Mathematics, vol. 1446, ed. by A. Cellina (Springer, Berlin, 1990), pp. 152–188
410. L. Vietoris, Über den höheren Zusammenhang kompakter Räume und eine Klasse von zusammenhangstreuen Abbildungen. Math. Ann. **97**, 454–472 (1927)
411. A. Visintin, Strong convergence results related to strict convexity. Commun. Partial Differ. Equ. **9**, 439–466 (1984)
412. Z.B. Wang, N.J. Huang, Degree theory for a generalized set-valued variational inequality with an application in Banach spaces. J. Glob. Optim. **49**, 343–357 (2011)

413. Z.B. Wang, N.J. Huang, The degree theory for set-valued compact perturbation of monotone-type mappings with an application. Appl. Anal. **92**, 616–635 (2013)
414. J. Warga, *Optimal Control of Differential and Functional Equations* (Academic Press, New York, 1972)
415. M. Willem, *Minimax Theorems* (Birkhäuser, Basel, 1996)
416. C.T. Yang, On theorems of Borsuk-Ulam, Kakutani-Yamabe-Yujobo and Dyson I. Ann. Math. **60**, 262–282 (1954)
417. V. Yankov, On the uniformization of A-sets. Doklady Akad. Nauk USSR **30**, 591–592 (1941)
418. K. Yosida, *Functional Analysis*, 5th edn. (Springer, Berlin, 1978)
419. L.C. Young, Generalized curves and existence of an attained absolute minimum in the calculus of variations. Comptes Rendus de la Société des Sciences et des Lettres de Varsovie cl. III **30**, 212–234 (1937)
420. L.C. Young, Generalized surfaces in the calculus of variations I. Ann. Math. **43**, 84–103 (1942)
421. L.C. Young, Generalized surfaces in the calculus of variations II. Ann. Math. **43**, 530–544 (1942)
422. L.C. Young, Surfaces paramétriques généralisées. Bull. Soc. Math. Fr. **79**, 59–84 (1951)
423. L.C. Young, On generalized surfaces of finite topological types. Mem. Am. Math. Soc. **17**, 63 (1955)
424. L.C. Young, *Lectures in the Calculus of Variations and Optimal Control Theory* (Saunders, New York, 1969)
425. E. Zeidler, *Nonlinear Functional Analysis and its Applications III* (Springer, New York, 1985)
426. E. Zeidler, *Nonlinear Functional Analysis and its Applications I* (Springer, New York, 1986)
427. E. Zeidler, *Nonlinear Functional Analysis and its Applications II/A and II/B* (Springer, Berlin, 1990)
428. C.K. Zhang, On Ekeland's variational principle and a minimax theorem. J. Math. Anal. Appl. **205**, 239–250 (1997)
429. W. Ziemer, *Weakly Differentiable Functions* (Springer, New York, 1989)
430. W. Zou, Variant fountain theorems and their applications. Manuscr. Math. **104**, 343–358 (2001)
431. W. Zou, M. Schechter, *Critical Point Theory and its Applications* (Springer, New York, 2006)

Index

N. S. Papageorgiou et al., *Nonlinear Analysis–Theory and Methods*,
Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-030-03430-6