# Grundlehren der mathematischen Wissenschaften 330

*A Series of Comprehensive Studies in Mathematics*

Boris S. Mordukhovich

# Variational Analysis and Generalized Differentiation I

## Basic Theory

Springer

Boris S. Mordukhovich
Department of Mathematics
Wayne State University
College of Science
Detroit, MI 48202-9861, U.S.A.
E-mail: *boris@math.wayne.edu*

Library of Congress Control Number: 2005932550

Mathematics Subject Classification (2000): 49J40, 49J50, 49J52, 49K24, 49K27, 49K40, 49N40, 58C06, 58C20, 58C25, 65K05, 65L12, 90C29, 90C31, 90C48, 93B35

ISSN 0072-7830
ISBN-10 3-540-25437-4 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-25437-9 Springer Berlin Heidelberg New York



Springer is a part of Springer Science+Business Media
springeronline.com

Printed in The Netherlands



Typesetting: by the author and TechBooks using a Springer LaTeX macro package
Cover design: *design & production* GmbH, Heidelberg

Printed on acid-free paper SPIN: 10922989 41/TechBooks 5 4 3 2 1 0

To Margaret, as always

# Preface

*Namely, because the shape of the whole universe is most perfect and, in fact, designed by the wisest creator, nothing in all of the world will occur in which no maximum or minimum rule is somehow shining forth.*

Leonhard Euler (1744)

We can treat this firm stand by Euler [411] ("...nihil omnino in mundo contingint, in quo non maximi minimive ratio quapiam eluceat") as the most fundamental principle of *Variational Analysis*. This principle justifies a variety of striking implementations of *optimization/variational* approaches to solving numerous problems in mathematics and applied sciences that may not be of a variational nature. Remember that optimization has been a major motivation and driving force for developing differential and integral calculus. Indeed, the *very concept of derivative* introduced by Fermat via the tangent slope to the graph of a function was motivated by solving an optimization problem; it led to what is now called the *Fermat stationary principle*. Besides applications to optimization, the latter principle plays a crucial role in proving the most important calculus results including the mean value theorem, the implicit and inverse function theorems, etc. The same line of development can be seen in the infinite-dimensional setting, where the Brachistochrone was the first problem not only of the calculus of variations but of all functional analysis inspiring, in particular, a variety of concepts and techniques in infinite-dimensional differentiation and related areas.

Modern variational analysis can be viewed as an outgrowth of the calculus of variations and mathematical programming, where the focus is on optimization of functions relative to various constraints and on sensitivity/stability of optimization-related problems with respect to perturbations. Classical notions of variations such as moving away from a given point or curve no longer play a critical role, while concepts of problem *approximations* and/or *perturbations* become crucial.

One of the most characteristic features of modern variational analysis is the intrinsic presence of *nonsmoothness*, i.e., the necessity to deal with nondifferentiable functions, sets with nonsmooth boundaries, and set-valued mappings. Nonsmoothness naturally enters not only through initial data of optimization-related problems (particularly those with inequality and geometric constraints) but largely via *variational principles* and other optimization, approximation, and perturbation techniques applied to problems with even smooth data. In fact, many fundamental objects frequently appearing in the framework of variational analysis (e.g., the distance function, value functions in optimization and control problems, maximum and minimum functions, solution maps to perturbed constraint and variational systems, etc.) are inevitably of nonsmooth and/or set-valued structures requiring the development of new forms of analysis that involve *generalized differentiation*.

It is important to emphasize that even the simplest and historically earliest problems of *optimal control* are *intrinsically nonsmooth*, in contrast to the classical calculus of variations. This is mainly due to pointwise constraints on control functions that often take only discrete values as in typical problems of automatic control, a primary motivation for developing optimal control theory. Optimal control has always been a major source of inspiration as well as a fruitful territory for applications of advanced methods of variational analysis and generalized differentiation.

Key issues of variational analysis in finite-dimensional spaces have been addressed in the book "Variational Analysis" by Rockafellar and Wets [1165]. The development and applications of variational analysis in infinite dimensions require certain concepts and tools that cannot be found in the finite-dimensional theory. The *primary goals* of this book are to present basic concepts and principles of variational analysis unified in finite-dimensional and infinite-dimensional space settings, to develop a comprehensive generalized differential theory at the same level of perfection in both finite and infinite dimensions, and to provide valuable applications of variational theory to broad classes of problems in constrained optimization and equilibrium, sensitivity and stability analysis, control theory for ordinary, functional-differential and partial differential equations, and also to selected problems in mechanics and economic modeling.

Generalized differentiation lies at the heart of variational analysis and its applications. We systematically develop a *geometric dual-space approach* to generalized differentiation theory revolving around the *extremal principle*, which can be viewed as a local *variational* counterpart of the classical convex separation in nonconvex settings. This principle allows us to deal with *nonconvex* derivative-like constructions for sets (normal cones), set-valued mappings (coderivatives), and extended-real-valued functions (subdifferentials). These constructions are defined directly in dual spaces and, being nonconvex-valued, cannot be generated by any derivative-like constructions in primal spaces (like tangent cones and directional derivatives). Nevertheless, our basic nonconvex constructions enjoy comprehensive calculi, which happen to be significantly better than those available for their primal and/or convex-valued counterparts. Thus passing to *dual spaces*, we are able to achieve more beauty and harmony in comparison with primal world objects. In some sense, the dual viewpoint does indeed allow us to meet the perfection requirement in the fundamental statement by Euler quoted above.

Observe to this end that dual objects (multipliers, adjoint arcs, shadow prices, etc.) have always been at the center of variational theory and applications used, in particular, for formulating principal optimality conditions in the calculus of variations, mathematical programming, optimal control, and economic modeling. The usage of variations of optimal solutions in primal spaces can be considered just as a convenient tool for deriving necessary optimality conditions. There are no essential restrictions in such a "primal" approach in smooth and convex frameworks, since primal and dual derivative-like constructions are equivalent for these classical settings. It is not the case any more in the framework of modern variational analysis, where even *nonconvex primal space* local approximations (e.g., tangent cones) inevitably yield, *under duality, convex sets* of normals and subgradients. This convexity of dual objects leads to significant restrictions for the theory and applications. Moreover, there are many situations particularly identified in this book, where primal space approximations simply cannot be used for variational analysis, while the employment of dual space constructions provides comprehensive results. Nevertheless, tangentially generated/primal space constructions play an important role in some other aspects of variational analysis, especially in finite-dimensional spaces, where they recover in duality the nonconvex sets of our basic normals and subgradients at the point in question by *passing to the limit* from points nearby; see, for instance, the afore-mentioned book by Rockafellar and Wets [1165]

Among the abundant bibliography of this book, we refer the reader to the monographs by Aubin and Frankowska [54], Bardi and Capuzzo Dolcetta [85], Beer [92], Bonnans and Shapiro [133], Clarke [255], Clarke, Ledyaev, Stern and Wolenski [265], Facchinei and Pang [424], Klatte and Kummer [686], Vinter [1289], and to the comments given after each chapter for significant aspects of variational analysis and impressive applications of this rapidly growing area that are not considered in the book. We especially emphasize the concurrent and complementing monograph "Techniques of Variational Analysis" by Borwein and Zhu [164], which provides a nice introduction to some fundamental techniques of modern variational analysis covering important theoretical aspects and applications not included in this book.

The book presented to the reader's attention is self-contained and mostly collects results that have not been published in the monographical literature. It is split into two volumes and consists of eight chapters divided into sections and subsections. Extensive comments (that play a special role in this book discussing basic ideas, history, motivations, various interrelations, choice of terminology and notation, open problems, etc.) are given for each chapter. We present and discuss numerous references to the vast literature on many aspects of variational analysis (considered and not considered in the book) including early contributions and very recent developments. Although there are no formal exercises, the extensive remarks and examples provide grist for further thought and development. Proofs of the major results are complete, while there is plenty of room for furnishing details, considering special cases, and deriving generalizations for which guidelines are often given.

*Volume I* "Basic Theory" consists of four chapters mostly devoted to basic constructions of generalized differentiation, fundamental extremal and variational principles, comprehensive generalized differential calculus, and complete dual characterizations of fundamental properties in nonlinear study related to Lipschitzian stability and metric regularity with their applications to sensitivity analysis of constraint and variational systems.

*Chapter 1* concerns the generalized differential theory in arbitrary *Banach spaces*. Our basic normals, subgradients, and coderivatives are directly defined in dual spaces via *sequential* weak$^*$ limits involving more primitive $\varepsilon$-normals and $\varepsilon$-subgradients of the Fréchet type. We show that these constructions have a variety of nice properties in the general Banach spaces setting, where the usage of $\varepsilon$-enlargements is crucial. Most such properties (including first-order and second-order calculus rules, efficient representations, variational descriptions, subgradient calculations for distance functions, necessary coderivative conditions for Lipschitzian stability and metric regularity, etc.) are collected in this chapter. Here we also define and start studying the so-called *sequential normal compactness* (SNC) properties of sets, set-valued mappings, and extended-real-valued functions that automatically hold in finite dimensions while being one of the most essential ingredients of variational analysis and its applications in infinite-dimensional spaces.

*Chapter 2* contains a detailed study of the *extremal principle* in variational analysis, which is the main single tool of this book. First we give a direct variational proof of the extremal principle in finite-dimensional spaces based on a smoothing penalization procedure via the method of *metric approximations*. Then we proceed by infinite-dimensional variational techniques in Banach spaces with a Fréchet smooth norm and finally, by separable reduction, in the larger class of *Asplund spaces*. The latter class is well-investigated in the geometric theory of Banach spaces and contains, in particular, every reflexive space and every space with a separable dual. Asplund spaces play a prominent role in the theory and applications of variational analysis developed in this book. In Chap. 2 we also establish relationships between the (geometric) extremal principle and (analytic) variational principles in both conventional and enhanced forms. The results obtained are applied to the derivation of novel variational characterizations of Asplund spaces and useful representations of the basic generalized differential constructions in the Asplund space setting similar to those in finite dimensions. Finally, in this chapter we discuss abstract versions of the extremal principle formulated in terms of axiomatically defined normal and subdifferential structures on appropriate Banach spaces and also overview in more detail some specific constructions.

*Chapter 3* is a cornerstone of the generalized differential theory developed in this book. It contains comprehensive *calculus rules* for basic normals, subgradients, and coderivatives in the framework of Asplund spaces. We pay most of our attention to *pointbased* rules via the limiting constructions *at* the points in question, for both assumptions and conclusions, having in mind that pointbased results indeed happen to be of crucial importance for applications. A number of the results presented in this chapter seem to be new even in the finite-dimensional setting, while overall we achieve the same level of perfection and generality in Asplund spaces as in finite dimensions. The main issue that distinguishes the finite-dimensional and infinite-dimensional settings is the necessity to invoke *sufficient amounts of compactness* in infinite dimensions that are not needed at all in finite-dimensional spaces. The required compactness is provided by the afore-mentioned SNC properties, which are included in the assumptions of calculus rules and call for their own calculus ensuring the preservation of SNC properties under various operations on sets and mappings. The absence of such a *SNC calculus* was a crucial obstacle for many successful applications of generalized differentiation in infinite-dimensional spaces to a range of infinite-dimensions problems including those in optimization, stability, and optimal control given in this book. Chapter 3 contains a broad spectrum of the SNC calculus results that are decisive for subsequent applications.

*Chapter 4* is devoted to a thorough study of Lipschitzian, metric regularity, and linear openness/covering properties of set-valued mappings, and to their applications to sensitivity analysis of parametric constraint and variational systems. First we show, based on variational principles and the generalized differentiation theory developed above, that the necessary coderivative conditions for these fundamental properties derived in Chap. 1 in arbitrary Banach spaces happen to be *complete characterizations* of these properties in the Asplund space setting. Moreover, the employed variational approach allows us to obtain verifiable formulas for computing the *exact bounds* of the corresponding moduli. Then we present detailed applications of these results, supported by generalized differential and SNC calculi, to sensitivity and stability analysis of parametric constraint and variational systems governed by perturbed sets of feasible and optimal solutions in problems of optimization and equilibria, implicit multifunctions, complementarity conditions, variational and hemivariational inequalities as well as to some mechanical systems.

*Volume II* "Applications" also consists of four chapters mostly devoted to applications of basic principles in variational analysis and the developed generalized differential calculus to various topics in constrained optimization and equilibria, optimal control of ordinary and distributed-parameter systems, and models of welfare economics.

*Chapter 5* concerns constrained optimization and equilibrium problems with possibly nonsmooth data. Advanced methods of variational analysis based on extremal/variational principles and generalized differentiation happen to be very useful for the study of constrained problems even with smooth initial data, since nonsmoothness naturally appears while applying penalization, approximation, and perturbation techniques. Our primary goal is to derive necessary optimality and suboptimality conditions for various constrained problems in both finite-dimensional and infinite-dimensional settings. Note that conditions of the latter – *suboptimality* – type, somehow underestimated in optimization theory, don't assume the existence of optimal solutions (which is especially significant in infinite dimensions) ensuring that "almost" optimal solutions "almost" satisfy necessary conditions for optimality. Besides considering problems with constraints of conventional types, we pay serious attention to rather new classes of problems, labeled as *mathematical problems with equilibrium constraints* (MPECs) and *equilibrium problems with equilibrium constraints* (EPECs), which are intrinsically nonsmooth while admitting a thorough analysis by using generalized differentiation. Finally, certain concepts of *linear subextremality* and *linear suboptimality* are formulated in such a way that the necessary optimality conditions derived above for conventional notions are seen to be *necessary and sufficient* in the new setting.

In *Chapter 6* we start studying problems of *dynamic optimization* and *optimal control* that, as mentioned, have been among the primary motivations for developing new forms of variational analysis. This chapter deals mostly with optimal control problems governed by *ordinary* dynamic systems whose state space may be infinite-dimensional. The main attention in the first part of the chapter is paid to the Bolza-type problem for evolution systems governed by constrained *differential inclusions*. Such models cover more conventional control systems governed by parameterized evolution equations with control regions generally dependent on state variables. The latter don't allow us to use control variations for deriving necessary optimality conditions. We develop the *method of discrete approximations*, which is certainly of numerical interest, while it is mainly used in this book as a direct vehicle to derive optimality conditions for continuous-time systems by passing to the limit from their discrete-time counterparts. In this way we obtain, strongly based on the generalized differential and SNC calculi, necessary optimality conditions in the extended Euler-Lagrange form for nonconvex differential inclusions in infinite dimensions expressed via our basic generalized differential constructions.

The second part of Chap. 6 deals with constrained optimal control systems governed by ordinary evolution equations of *smooth dynamics* in arbitrary Banach spaces. Such problems have essential specific features in comparison with the differential inclusion model considered above, and the results obtained (as well as the methods employed) in the two parts of this chapter are generally independent. Another major theme explored here concerns *stability* of the maximum principle under discrete approximations of nonconvex control systems. We establish rather surprising results on the *approximate maximum principle* for discrete approximations that shed new light upon both qualitative and quantitative relationships between continuous-time and discrete-time systems of optimal control.

In *Chapter 7* we continue the study of optimal control problems by applications of advanced methods of variational analysis, now considering systems with *distributed parameters*. First we examine a general class of *hereditary systems* whose dynamic constraints are described by both delay-differential inclusions and linear algebraic equations. On one hand, this is an interesting and not well-investigated class of control systems, which can be treated as a special type of variational problems for *neutral functional-differential inclusions* containing time delays not only in state but also in velocity variables. On the other hand, this class is related to differential-algebraic systems with a linear link between "slow" and "fast" variables. Employing the method of discrete approximations and the basic tools of generalized differentiation, we establish a strong variational convergence/stability of discrete approximations and derive extended optimality conditions for continuous-time systems in both Euler-Lagrange and Hamiltonian forms.

The rest of Chap. 7 is devoted to optimal control problems governed by *partial differential equations* with *pointwise* control and state constraints. We pay our primary attention to evolution systems described by *parabolic* and *hyperbolic* equations with controls functions acting in the Dirichlet and Neumann boundary conditions. It happens that such *boundary control* problems are the most challenging and the least investigated in PDE optimal control theory, especially in the presence of pointwise state constraints. Employing approximation and perturbation methods of modern variational analysis, we justify variational convergence and derive necessary optimality conditions for various control problems for such PDE systems including *minimax* control under *uncertain disturbances*.

The concluding *Chapter 8* is on applications of variational analysis to *economic modeling*. The major topic here is *welfare economics*, in the general nonconvex setting with infinite-dimensional commodity spaces. This important class of competitive equilibrium models has drawn much attention of economists and mathematicians, especially in recent years when nonconvexity has become a crucial issue for practical applications. We show that the methods of variational analysis developed in this book, particularly the extremal principle, provide adequate tools to study Pareto optimal allocations and associated price equilibria in such models. The tools of variational analysis and generalized differentiation allow us to obtain extended nonconvex versions of the so-called "second fundamental theorem of welfare economics" describing marginal equilibrium prices in terms of minimal collections of generalized normals to nonconvex sets. In particular, our approach and variational descriptions of generalized normals offer new economic interpretations of market equilibria via "nonlinear marginal prices" whose role in nonconvex models is similar to the one played by conventional linear prices in convex models of the Arrow-Debreu type.

The book includes a Glossary of Notation, common for both volumes, and an extensive Subject Index compiled separately for each volume. Using the Subject Index, the reader can easily find not only the page, where some notion and/or notation is introduced, but also various places providing more discussions and significant applications for the object in question.

Furthermore, it seems to be reasonable to title all the statements of the book (definitions, theorems, lemmas, propositions, corollaries, examples, and remarks) that are numbered in sequence within a chapter; thus, in Chap. 5 for instance, Example 5.3.3 precedes Theorem 5.3.4, which is followed by Corollary 5.3.5. For the reader's convenience, all these statements and numerated comments are indicated in the List of Statements presented at the end of each volume. It is worth mentioning that the list of acronyms is included (in alphabetic order) in the Subject Index and that the common principle adopted for the book notation is to use lower case Greek characters for numbers and (extended) real-valued functions, to use lower case Latin characters for vectors and single-valued mappings, and to use Greek and Latin upper case characters for sets and set-valued mappings.

Our notation and terminology are generally consistent with those in Rockafellar and Wets [1165]. Note that we try to distinguish everywhere the notions defined *at* the point and *around* the point in question. The latter indicates *robustness/stability* with respect to perturbations, which is critical for most of the major results developed in the book.

The book is accompanied by the abundant bibliography (with English sources if available), common for both volumes, which reflects a variety of topics and contributions of many researchers. The references included in the bibliography are discussed, at various degrees, mostly in the extensive commentaries to each chapter. The reader can find further information in the given references, directed by the author's comments.

We address this book mainly to researchers and graduate students in mathematical sciences; first of all to those interested in nonlinear analysis, optimization, equilibria, control theory, functional analysis, ordinary and partial differential equations, functional-differential equations, continuum mechanics, and mathematical economics. We also envision that the book will be useful to a broad range of researchers, practitioners, and graduate students involved in the study and applications of variational methods in operations research, statistics, mechanics, engineering, economics, and other applied sciences.

Parts of the book have been used by the author in teaching graduate classes on variational analysis, optimization, and optimal control at Wayne State University. Basic material has also been incorporated into many lectures and tutorials given by the author at various schools and scientific meetings during the recent years.

## Acknowledgments

My first gratitude go to Terry Rockafellar who has encouraged me over the years to write such a book and who has advised and supported me at all the stages of this project.

Special thanks are addressed to Rafail Gabasov, my doctoral thesis adviser, from whom I learned optimal control and much more; to Alec Ioffe, Boris Polyak, and Vladimir Tikhomirov who recognized and strongly supported my first efforts in nonsmooth analysis and optimization; to Sasha Kruger, my first graduate student and collaborator in the beginning of our exciting journey to generalized differentiation; to Jon Borwein and Marián Fabian from whom I learned deep functional analysis and the beauty of Asplund spaces; to Ali Khan whose stimulating work and enthusiasm have encouraged my study of economic modeling; to Jiři Outrata who has motivated and influenced my growing interest in equilibrium problems and mechanics and who has intensely promoted the implementation of the basic generalized differential constructions of this book in various areas of optimization theory and applications; and to Jean-Pierre Raymond from whom I have greatly benefited on modern theory of partial differential equations.

During the work on this book, I have had the pleasure of discussing its various aspects and results with many colleagues and friends. Besides the individuals mentioned above, I'm particularly indebted to Zvi Artstein, Jim Burke, Tzanko Donchev, Asen Dontchev, Joydeep Dutta, Andrew Eberhard, Ivar Ekeland, Hector Fattorini, René Henrion, Jean-Baptiste Hiriart-Urruty, Alejandro Jofré, Abderrahim Jourani, Michal Kočvara, Irena Lasiecka, Claude Lemaréchal, Adam Levy, Adrian Lewis, Kazik Malanowski, Michael Overton, Jong-Shi Pang, Teemu Pennanen, Steve Robinson, Alex Rubinov, Andrzej Świech, Michel Théra, Lionel Thibault, Jay Treiman, Hector Sussmann, Roberto Triggiani, Richard Vinter, Nguyen Dong Yen, George Yin, Jack Warga, Roger Wets, and Jim Zhu for valuable suggestions and fruitful conversations throughout the years of the fulfillment of this project.

The continuous support of my research by the National Science Foundation is gratefully acknowledged.

As mentioned above, the material of this book has been used over the years for teaching advanced classes on variational analysis and optimization attended mostly by my doctoral students and collaborators. I highly appreciate their contributions, which particularly allowed me to improve my lecture notes and book manuscript. Especially valuable help was provided by Glenn Malcolm, Nguyen Mau Nam, Yongheng Shao, Ilya Shvartsman, and Bingwu Wang. Useful feedback and text corrections came also from Truong Bao, Wondi Geremew, Pankaj Gupta, Aychi Habte, Kahina Sid Idris, Dong Wang, Lianwen Wang, and Kaixia Zhang.

I'm very grateful to the nice people in Springer for their strong support during the preparation and publishing this book. My special thanks go to Catriona Byrne, Executive Editor in Mathematics, to Achi Dosajh, Senior Editor in Applied Mathematics, to Stefanie Zoeller, Assistant Editor in Mathematics, and to Frank Holzwarth from the Computer Science Editorial Department.

I thank my younger daughter Irina for her interest in my book and for her endless patience and tolerance in answering my numerous question on English. I would also like to thank my poodle Wuffy for his sharing with me the long days of work on this book. Above all, I don't have enough words to thank my wife Margaret for her sharing with me everything, starting with our high school years in Minsk.

Ann Arbor, Michigan — *Boris Mordukhovich*
August 2005

# Contents

**Volume I Basic Theory**

# Volume I

## Basic Theory

# 1

# Generalized Differentiation in Banach Spaces

In this chapter we define and study basic concepts of *generalized differentiation* that lies at the heart of variational analysis and its applications considered in the book. Most properties presented in this chapter hold in *arbitrary* Banach spaces (some of them don't require completeness or even a normed structure, as one can see from the proofs). Developing a *geometric dual-space approach* to generalized differentiation, we start with *normals* to sets (Sect. 1.1), then proceed to *coderivatives* of set-valued mappings (Sect. 1.2), and then to *subdifferentials* of extended-real-valued functions (Sect. 1.3).

Unless otherwise stated, *all the spaces in question are Banach* whose norms are always denoted by $\|\cdot\|$. Given a space $X$, we denote by $I\!B_X$ its closed unit ball and by $X^*$ its dual space equipped with the weak$^*$ topology $w^*$, where $\langle\cdot,\cdot\rangle$ means the canonical pairing. If there is no confusion, $I\!B$ and $I\!B^*$ stand for the closed unit balls of the space and dual space in question, while $S$ and $S^*$ are usually stand for the corresponding unit spheres ; also $B_r(x) := x + rI\!B$ with $r > 0$. The symbol $^*$ is used everywhere to indicate relations to *dual* spaces (dual elements, adjoint operators, etc.)

In what follows we often deal with set-valued mappings (multifunctions) $F\colon X \rightrightarrows X^*$ between a Banach space and its dual, for which the notation

$$\begin{aligned}\mathop{\rm Lim\,sup}_{x\to\bar{x}} F(x) := \Big\{x^* \in X^* \;\Big|\; \exists \text{ sequences } \ x_k \to \bar{x} \ \text{ and } \ x_k^* \stackrel{w^*}{\to} x^* \\ \text{with } \ x_k^* \in F(x_k) \ \text{ for all } \ k \in I\!N\Big\}\end{aligned} \tag{1.1}$$

signifies the *sequential Painlevé-Kuratowski upper/outer limit* with respect to the norm topology of $X$ and the weak$^*$ topology of $X^*$. Note that the symbol $:=$ means "equal by definition" and that $I\!N := \{1, 2, \ldots\}$ denotes the set of all natural numbers.

The linear combination of the two subsets $\Omega_1$ and $\Omega_2$ of $X$ is defined by

$$\alpha_1\Omega_1 + \alpha_2\Omega_2 := \big\{\alpha_1 x_1 + \alpha_2 x_2 \;\big|\; x_1 \in \Omega_1,\ x_2 \in \Omega_2\big\}$$

with real numbers $\alpha_1, \alpha_2 \in I\!R := (-\infty, \infty)$, where we use the convention that $\Omega + \emptyset = \emptyset$, $\alpha\emptyset = \emptyset$ if $\alpha \in I\!R \setminus \{0\}$, and $\alpha\emptyset = \{0\}$ if $\alpha = 0$. Dealing with empty sets, we let $\inf \emptyset := \infty$, $\sup \emptyset := -\infty$, and $\|\emptyset\| := \infty$.

## 1.1 Generalized Normals to Nonconvex Sets

Throughout this section, $\Omega$ is a nonempty subset of a real Banach space $X$. Such a set is called *proper* if $\Omega \neq X$. In what follows the expressions

$$\mathrm{cl}\,\Omega,\quad \mathrm{co}\,\Omega,\quad \mathrm{clco}\,\Omega,\quad \mathrm{bd}\,\Omega,\quad \mathrm{int}\,\Omega$$

stand for the standard notions of *closure*, *convex hull* , *closed convex hull*, *boundary*, and *interior* of $\Omega$, respectively. The *conic hull* of $\Omega$ is

$$\mathrm{cone}\;\Omega := \big\{\alpha x \in X \big|\; \alpha \geq 0,\; x \in \Omega\big\}\;.$$

The symbol $\mathrm{cl}^*$ signifies the *weak* topological closure* of a set in a dual space.

### 1.1.1 Basic Definitions and Some Properties

We begin the generalized differentiation theory with constructing generalized normals to arbitrary sets. To describe basic normals to a set $\Omega$ at a given point $\bar{x}$, we use a two-stage procedure: first define more primitive $\varepsilon$-normals (prenormals) to $\Omega$ at points $x$ close to $\bar{x}$ and then pass to the sequential limit (1.1) as $x \to \bar{x}$ and $\varepsilon \downarrow 0$. Throughout the book we use the notation

$$x \xrightarrow{\Omega} \bar{x} \quad\Longleftrightarrow\quad x \to \bar{x}\;\text{ with }\; x \in \Omega\;.$$

**Definition 1.1 (generalized normals).** *Let $\Omega$ be a nonempty subset of $X$.*

**(i)** *Given $x \in \Omega$ and $\varepsilon \geq 0$, define the* SET OF $\varepsilon$-NORMALS *to $\Omega$ at $x$ by*

$$\widehat{N}_\varepsilon(x;\Omega) := \Big\{x^* \in X^* \;\Big|\; \limsup_{u \xrightarrow{\Omega} x} \frac{\langle x^*, u - x\rangle}{\|u - x\|} \leq \varepsilon\Big\}\;. \tag{1.2}$$

*When $\varepsilon = 0$, elements of (1.2) are called* FRÉCHET NORMALS *and their collection, denoted by $\widehat{N}(x;\Omega)$, is the* PRENORMAL CONE *to $\Omega$ at $x$. If $x \notin \Omega$, we put $\widehat{N}_\varepsilon(x;\Omega) := \emptyset$ for all $\varepsilon \geq 0$.*

**(ii)** *Let $\bar{x} \in \Omega$. Then $x^* \in X^*$ is a* BASIC/LIMITING NORMAL *to $\Omega$ at $\bar{x}$ if there are sequences $\varepsilon_k \downarrow 0$, $x_k \xrightarrow{\Omega} \bar{x}$, and $x_k^* \xrightarrow{w^*} x^*$ such that $x_k^* \in \widehat{N}_{\varepsilon_k}(x_k;\Omega)$ for all $k \in I\!N$. The collection of such normals*

$$N(\bar{x};\Omega) := \mathop{\mathrm{Lim\,sup}}_{\substack{x \to \bar{x}\\ \varepsilon \downarrow 0}} \widehat{N}_\varepsilon(x;\Omega) \tag{1.3}$$

*is the (basic, limiting)* NORMAL CONE *to $\Omega$ at $\bar{x}$. Put $N(\bar{x};\Omega) := \emptyset$ for $\bar{x} \notin \Omega$.*

It easily follows from the definitions that

$$\widehat{N}_\varepsilon(\bar{x};\Omega) = \widehat{N}_\varepsilon(\bar{x};\operatorname{cl}\Omega) \ \text{ and } \ N(\bar{x};\Omega) \subset N(\bar{x};\operatorname{cl}\Omega)$$

for every $\Omega \subset X$, $\bar{x} \in \Omega$, and $\varepsilon \geq 0$. Observe that both the prenormal cone $\widehat{N}(\cdot;\Omega)$ and the normal cone $N(\cdot;\Omega)$ are *invariant* with respect to equivalent norms on $X$ while the $\varepsilon$-normal sets $\widehat{N}_\varepsilon(\cdot;\Omega)$ depend on a given norm $\|\cdot\|$ if $\varepsilon > 0$. Note also that for each $\varepsilon \geq 0$ the sets (1.2) are obviously *convex and closed* in the norm topology of $X^*$; hence they are *weak* closed* in $X^*$ when $X$ is *reflexive*.

In contrast to (1.2), the basic normal cone (1.3) may be *nonconvex* in very simple situations as for $\Omega := \big\{(x_1,x_2) \in I\!\!R^2 |\ x_2 \geq -|x_1|\big\}$, where

$$N((0,0);\Omega) = \big\{(v,v)\ \big|\ v \leq 0\big\} \cup \big\{(v,-v)\ \big|\ v \geq 0\big\} \tag{1.4}$$

while $\widehat{N}((0,0);\Omega) = \{0\}$. This shows that $N(\bar{x};\Omega)$ *cannot be dual/polar* to any (even nonconvex) *tangential approximation* of $\Omega$ at $\bar{x}$ in the primal space $X$, since *polarity always implies convexity*; cf. Subsect. 1.1.2.

One can easily observe the following *monotonicity* properties of the $\varepsilon$-normal sets (1.2) with respect to $\varepsilon$ as well as with respect to the set order:

$$\widehat{N}_\varepsilon(\bar{x};\Omega) \subset \widehat{N}_{\tilde{\varepsilon}}(\bar{x};\Omega) \ \text{ if } \ 0 \leq \varepsilon \leq \tilde{\varepsilon}\ ,$$

$$\widehat{N}_\varepsilon(\bar{x};\Omega) \subset \widehat{N}_\varepsilon(\bar{x};\widetilde{\Omega}) \ \text{ if } \ \bar{x} \in \widetilde{\Omega} \subset \Omega \ \text{ and } \ \varepsilon \geq 0\ . \tag{1.5}$$

In particular, the decreasing property (1.5) holds for the prenormal cone $\widehat{N}(\bar{x};\cdot)$. Note however that neither (1.5) nor the opposite inclusion is valid for the basic normal cone (1.3). To illustrate this, we consider the two sets

$$\Omega := \big\{(x_1,x_2) \in I\!\!R^2\ \big|\ x_2 \geq -|x_1|\big\} \ \text{ and } \ \widetilde{\Omega} := \big\{(x_1,x_2) \in I\!\!R^2\ \big|\ x_1 \leq x_2\big\}$$

with $\bar{x} = (0,0) \in \widetilde{\Omega} \subset \Omega$. Then

$$N(\bar{x};\widetilde{\Omega}) = \big\{(v,-v)\ \big|\ v \geq 0\big\} \subset N(\bar{x};\Omega)\ ,$$

where the latter cone is computed in (1.4). Furthermore, taking $\Omega$ as above and $\widetilde{\Omega} := \big\{(x_1,x_2) \in I\!\!R^2 |\ x_2 \geq 0\big\} \subset \Omega$, we have

$$N(\bar{x};\Omega) \cap N(\bar{x};\widetilde{\Omega}) = \{(0,0)\}\ ,$$

which excludes any monotonicity relations.

The next property for representing normals to set products is common for both prenormal and normal cones.

**Proposition 1.2 (normals to Cartesian products).** *Consider an arbitrary point $\bar{x} = (\bar{x}_1, \bar{x}_2) \in \Omega_1 \times \Omega_2 \subset X_1 \times X_2$. Then*

$$\widehat{N}(\bar{x}; \Omega_1 \times \Omega_2) = \widehat{N}(\bar{x}_1; \Omega_1) \times \widehat{N}(\bar{x}_2; \Omega_2) ,$$

$$N(\bar{x}; \Omega_1 \times \Omega_2) = N(\bar{x}_1; \Omega_1) \times N(\bar{x}_2; \Omega_2) .$$

**Proof.** Since both prenormal and normal cones do not depend on equivalent norms on $X_1$ and $X_2$, we can fix any norms on these spaces and define a norm on the product $X_1 \times X_2$ by

$$\|(x_1, x_2)\| := \|x_1\| + \|x_2\| .$$

Given arbitrary $\varepsilon \geq 0$ and $x = (x_1, x_2) \in \Omega := \Omega_1 \times \Omega_2$, we easily check that

$$\widehat{N}_\varepsilon(x_1; \Omega_1) \times \widehat{N}_\varepsilon(x_2; \Omega_2) \subset \widehat{N}_{2\varepsilon}(x; \Omega) \subset \widehat{N}_{2\varepsilon}(x_1; \Omega_1) \times \widehat{N}_{2\varepsilon}(x_2; \Omega_2) ,$$

which implies both product formulas in the proposition. △

The prenormal cone $\widehat{N}(\cdot; \Omega)$ is obviously the smallest set among all the sets $\widehat{N}_\varepsilon(\cdot; \Omega)$. It follows from (1.2) that

$$\widehat{N}_\varepsilon(\bar{x}; \Omega) \supset \widehat{N}(\bar{x}; \Omega) + \varepsilon I\!B^*$$

for every $\varepsilon \geq 0$ and an arbitrary set $\Omega$. If $\Omega$ is convex, then this inclusion holds as *equality* due to the following representation of $\varepsilon$-normals.

**Proposition 1.3 ($\varepsilon$-normals to convex sets).** *Let $\Omega$ be convex. Then*

$$\widehat{N}_\varepsilon(\bar{x}; \Omega) = \big\{ x^* \in X^* \;\big|\; \langle x^*, x - \bar{x} \rangle \leq \varepsilon \|x - \bar{x}\| \;\; \text{whenever} \;\; x \in \Omega \big\}$$

*for any $\varepsilon \geq 0$ and $\bar{x} \in \Omega$. In particular, $\widehat{N}(\bar{x}; \Omega)$ agrees with the normal cone of convex analysis.*

**Proof.** Note that the inclusion "$\supset$" in the above formula obviously holds for an arbitrary set $\Omega$. Let us justify the opposite inclusion when $\Omega$ is convex. Consider any $x^* \in \widehat{N}_\varepsilon(\bar{x}; \Omega)$ and fix $x \in \Omega$. Then we have

$$x_\alpha := \bar{x} + \alpha(x - \bar{x}) \in \Omega \;\; \text{for all} \;\; 0 \leq \alpha \leq 1$$

due to the convexity of $\Omega$. Moreover, $x_\alpha \to \bar{x}$ as $\alpha \downarrow 0$. Taking an arbitrary $\gamma > 0$, we easily conclude from (1.2) that

$$\langle x^*, x_\alpha - \bar{x} \rangle \leq (\varepsilon + \gamma)\|x_\alpha - \bar{x}\| \;\; \text{for small} \;\; \alpha > 0 ,$$

which completes the proof. △

It follows from Definition 1.1 that

$$\widehat{N}(\bar{x};\Omega) \subset N(\bar{x};\Omega) \quad \text{for any} \quad \Omega \subset X \quad \text{and} \quad \bar{x} \in \Omega \,. \tag{1.6}$$

This inclusion may be *strict* even for simple sets as the one in (1.4), where $\widehat{N}(\bar{x};\Omega) = \{0\}$ for $\bar{x} = 0 \in I\!R^2$. The equality in (1.6) singles out a class of sets that have certain "regular" behavior around $\bar{x}$ and unify good properties of both prenormal and normal cones at $\bar{x}$.

**Definition 1.4 (normal regularity of sets).** *A set $\Omega \subset X$ is (normally)* REGULAR *at $\bar{x} \in \Omega$ if*

$$N(\bar{x};\Omega) = \widehat{N}(\bar{x};\Omega) \,.$$

An important example of set regularity is given by sets $\Omega$ *locally convex* around $\bar{x}$, i.e., for which there is a neighborhood $U \subset X$ of $\bar{x}$ such that $\Omega \cap U$ is convex.

**Proposition 1.5 (regularity of locally convex sets).** *Let $U$ be a neighborhood of $\bar{x} \in \Omega \subset X$ such that the set $\Omega \cap U$ is convex. Then $\Omega$ is regular at $\bar{x}$ with*

$$N(\bar{x};\Omega) = \big\{x^* \in X^* \;\big|\; \langle x^*, x - \bar{x}\rangle \le 0 \;\; \textit{for all} \;\; x \in \Omega \cap U\big\} \,.$$

**Proof.** The inclusion "$\supset$" follows from (1.6) and Proposition 1.3. To prove the opposite inclusion, we take any $x^* \in N(\bar{x};\Omega)$ and find the corresponding sequences of $(\varepsilon_k, x_k, x_k^*)$ from Definition 1.1(ii). Thus $x_k \in U$ for all $k \in I\!N$ sufficiently large. Then Proposition 1.3 ensures that, for such $k$,

$$\langle x_k^*, x - x_k\rangle \le \varepsilon_k \|x - x_k\| \quad \text{for all} \quad x \in \Omega \cap U \,.$$

Passing there to the limit as $k \to \infty$, we finish the proof. △

Further results and discussions on normal regularity of sets and related notions of regularity for functions and set-valued mappings will be presented later in this chapter and mainly in Chap. 3, where they are incorporated into *calculus rules*. We'll show that regularity is preserved under major calculus operations and ensure *equalities* in calculus rules for basic normal and subdifferential constructions. On the other hand, such regularity may fail in many situations important for the theory and applications. In particular, it *never holds* for sets in finite-dimensional spaces related to *graphs of nonsmooth locally Lipschitzian mappings*; see Theorem 1.46 below. However, the basic normal cone and associated subdifferentials and coderivatives enjoy desired properties in general "irregular" settings, in contrast to the prenormal cone $\widehat{N}(\bar{x};\Omega)$ and its counterparts for functions and mappings.

Next we establish two special representations of the basic normal cone to closed subsets of the finite-dimensional space $X = I\!R^n$. Since all the norms in finite dimensions are equivalent, we always select the *Euclidean norm*

$$\|x\| := \sqrt{x_1^2 + \ldots + x_n^2}$$

on $I\!R^n$, unless otherwise stated. In this case $X^* = X = I\!R^n$.

Given a nonempty set $\Omega \subset I\!R^n$, consider the associated *distance function*

$$\operatorname{dist}(x;\Omega) := \inf_{u\in\Omega} \|x-u\| \tag{1.7}$$

and define the *Euclidean projector* of $x$ to $\Omega$ by

$$\Pi(x;\Omega) := \big\{w \in \Omega \;\big|\; \|x-w\| = \operatorname{dist}(x;\Omega)\big\} .$$

If $\Omega$ is closed, the set $\Pi(x;\Omega)$ is nonempty for every $x \in I\!R^n$. The following theorem describes the basic normal cone to subsets $\Omega \subset I\!R^n$ that are *locally closed* around $\bar{x}$. The latter means that there is a neighborhood $U$ of $\bar{x}$ for which $\Omega \cap U$ is closed.

**Theorem 1.6 (basic normals in finite dimensions).** *Let $\Omega \subset I\!R^n$ be locally closed around $\bar{x} \in \Omega$. Then the following representations hold:*

$$N(\bar{x};\Omega) = \operatorname*{Lim\,sup}_{x\to\bar{x}} \widehat{N}(x;\Omega) , \tag{1.8}$$

$$N(\bar{x};\Omega) = \operatorname*{Lim\,sup}_{x\to\bar{x}} \big[\operatorname{cone}(x - \Pi(x;\Omega))\big] . \tag{1.9}$$

**Proof.** First we prove (1.8), which means that one can equivalently put $\varepsilon = 0$ in definition (1.3) of basic normals to locally closed sets in finite-dimensions. The inclusion "$\supset$" in (1.8) is obvious; let us justify the opposite inclusion.

Fix $x^* \in N(\bar{x};\Omega)$ and find, by Definition 1.1(ii), sequences $\varepsilon_k \downarrow 0$, $x_k \to \bar{x}$, and $x_k^* \to x^*$ such that $x_k \in \Omega$ and $x_k^* \in \widehat{N}_{\varepsilon_k}(x_k;\Omega)$ for all $k \in I\!N$. Taking into account that $X = X^* = I\!R^n$ and that $\Omega$ is locally closed around $\bar{x}$, for each $k = 1, 2, \ldots$ we form $x_k + \alpha x_k^*$ with some parameter $\alpha > 0$ and select $w_k \in \Pi(x_k + \alpha x_k^*;\Omega)$ from the Euclidean projector. Due to the choice of $w_k$ one has the inequality

$$\|x_k + \alpha x_k^* - w_k\|^2 \le \alpha^2 \|x_k^*\|^2$$

and, since the norm is Euclidean,

$$\|x_k + \alpha x_k^* - w_k\|^2 = \|x_k - w_k\|^2 + 2\alpha\langle x_k^*, x_k - w_k\rangle + \alpha^2\|x_k^*\|^2 .$$

This implies the estimate

$$\|x_k - w_k\|^2 \le 2\alpha\langle x_k^*, w_k - x_k\rangle \;\text{ for any }\; \alpha > 0 . \tag{1.10}$$

Using the convergence $w_k \to x_k$ as $\alpha \downarrow 0$ and the definition of the $\varepsilon_k$-normals $x_k^* \in \widehat{N}_{\varepsilon_k}(x_k;\Omega)$, we find a sequence of positive numbers $\alpha = \alpha_k$ along which

$$\langle x_k^*, w_k - x_k\rangle \le 2\varepsilon_k \|w_k - x_k\| \;\text{ for every }\; k \in I\!N .$$

This gives $\|x_k - w_k\| \le 4\alpha_k\varepsilon_k$ due to (1.10); hence $w_k \to \bar{x}$ as $k \to \infty$. Moreover, letting

$$w_k^* := x_k^* + \frac{1}{\alpha_k}(x_k - w_k) ,$$

we get $\|w_k^* - x_k^*\| \le 4\varepsilon_k$ and $w_k^* \to x^*$ as $k \to \infty$.

To justify (1.8), it remains to show that $w_k^* \in \widehat{N}(w_k; \Omega)$ for all $k$. Indeed, for every fixed $x \in \Omega$ we get

$$0 \le \|x_k + \alpha_k x_k^* - x\|^2 - \|x_k + \alpha_k x_k^* - w_k\|^2$$

$$= \langle \alpha_k x_k^* + x_k - x, \alpha_k x_k^* + x_k - w_k \rangle + \langle \alpha_k x_k^* + x_k - x, w_k - x \rangle$$

$$- \langle \alpha_k x_k^* + x_k - w_k, x - w_k \rangle - \langle \alpha_k x_k^* + x_k - w_k, \alpha_k x_k^* + x_k - x \rangle$$

$$= -2\alpha_k \langle w_k^*, x - w_k \rangle + \|x - w_k\|^2 ,$$

since the norm is Euclidean. The latter implies the estimate

$$\langle w_k^*, x - w_k \rangle \le \frac{1}{2\alpha_k}\|x - w_k\|^2 \quad \text{for all} \quad x \in \Omega ,$$

which obviously ensures that $w_k^* \in \widehat{N}(w_k; \Omega)$ by Definition 1.1(i). Thus we arrive at the first representation (1.8) of the basic normal cone.

To justify the second representation (1.9), it is sufficient to show that

$$\mathop{\mathrm{Lim\,sup}}_{x \to \bar{x}} \widehat{N}(x; \Omega) = \mathop{\mathrm{Lim\,sup}}_{x \to \bar{x}} \big[\mathrm{cone}(x - \Pi(x; \Omega))\big] .$$

Let us first prove the inclusion

$$\widehat{N}(x; \Omega) \subset \mathop{\mathrm{Lim\,sup}}_{u \to x} \big[\mathrm{cone}(u - \Pi(u; \Omega))\big] \quad \text{for any} \quad x \in \Omega . \tag{1.11}$$

Given $x \in \Omega$ and $x^* \in \widehat{N}(x; \Omega)$, we put $x_k := x + \frac{1}{k}x^*$ and pick some $w_k \in \Pi(x_k; \Omega)$ for each $k \in I\!N$. The latter is clearly equivalent to

$$0 \le \|x_k - v\|^2 - \|x_k - w_k\|^2 = \langle x_k - v, x_k - w_k \rangle$$

$$+ \langle x_k - v, w_k - v \rangle - \langle x_k - w_k, v - w_k \rangle - \langle x_k - w_k, x_k - v \rangle$$

$$= -2\langle x_k - w_k, v - w_k \rangle + \|v - w_k\|^2 \quad \text{for all} \quad v \in \Omega ,$$

which *characterizes the Euclidean projector*: $w_k \in \Pi(x_k; \Omega)$ if and only if

$$\langle x_k - w_k, v - w_k \rangle \le \frac{1}{2}\|v - w_k\|^2 \quad \text{for all} \quad v \in \Omega .$$

Letting $v = x$ and using the definition of $x_k$, we get

$$\|x - w_k\|^2 + \tfrac{1}{k}\langle x^*, x - w_k\rangle \le \tfrac{1}{2}\|x - w_k\|^2 .$$

Since $x^* \in \widehat{N}(x;\Omega)$, the latter inequality gives

$$k\|x - w_k\| \le \frac{2\langle x^*, w_k - x\rangle}{\|x - w_k\|} \to 0 \ \text{ as } \ k \to \infty$$

and therefore

$$k(x_k - w_k) = x^* + k(x - w_k) \to x^* \ \text{ as } \ k \to \infty .$$

Thus we have (1.11) that implies the inclusion "$\subset$" in (1.9) by taking the Painlevé-Kuratowski upper limit as $x \to \bar{x}$ and using (1.8).

It remains to prove the opposite inclusion in (1.9). To furnish this, let us consider the *inverse* Euclidean projector

$$\Pi^{-1}(x;\Omega) := \big\{ z \in X \ \big|\ x \in \Pi(z;\Omega) \big\}$$

to $\Omega$ at $x \in \Omega$. It follows from the above characterization of the Euclidean projector and the definition of $\widehat{N}(x;\Omega)$ that

$$\text{cone}\big[\Pi^{-1}(x;\Omega) - x\big] \subset \widehat{N}(x;\Omega) \ \text{ for any } \ x \in \Omega ,$$

which implies the inclusion "$\supset$" in (1.9) by taking the Painlevé-Kuratowski upper limit as $x \xrightarrow{\Omega} \bar{x}$ and using (1.8). △

Note that, although the proof of representation (1.8) essentially employs properties of the Euclidean norm , the representation itself doesn't depend on a specific norm on $I\!R^n$ all of which are equivalent. In Chap. 2 we show, using variational arguments, that this representation of the basic normal cone holds in any *Asplund space*, i.e., in a Banach space where every convex continuous function is generically Fréchet differentiable (in particular, in any reflexive space). In fact, (1.8) is a characterization of Asplund spaces. Note however that $\varepsilon > 0$ *cannot be removed* from the definition of basic normals and the corresponding subdifferential and coderivative constructions without loss of important properties in the general Banach space setting; see below, in particular, the next subsection. Moreover, we'll see that *stability* with respect to $\varepsilon$-enlargements plays an essential role in the proof of some principal results in Asplund spaces and even in finite-dimensions.

On the contrary, representation (1.9) heavily depends on the Euclidean norm on $I\!R^n$ and is not valid even for convex sets if a norm in non-Euclidean. For example, we have

$$N((0,0);\Omega) = \big\{(0,v) \ \big|\ v \le 0\big\} \ \text{ for } \ \Omega = \big\{x = (x_1,x_2) \in I\!R^2 \,|\, x_2 \ge 0\big\} ,$$

while the cone on the right-hand side of (1.9) equals to $\big\{(v_1,v_2) |\, v_2 + |v_1| \le 0\big\}$ when the norm is given by $\|x\| := \max\big\{|x_1|,|x_2|\big\}$.

We are not going to consider here special properties of the basic normal cone in finite-dimensional spaces referring the reader to the books by Mordukhovich [901] and Rockafellar and Wets [1165]. Let us just mention that this cone enjoys the following *robustness property*

$$N(\bar{x};\Omega) = \operatorname*{Lim\,sup}_{x\to\bar{x}} N(x;\Omega) \quad \text{for all} \quad \bar{x}\in\Omega\ ,$$

which can be easily obtained via the standard diagonal process in finite dimensions. For closed sets $\Omega \subset I\!\!R^n$ this means that the *graph* of the set-valued mapping $N(\cdot;\Omega)$ is *closed*, which obviously implies that the values $N(x;\Omega)$ are closed for all $x \in \Omega$.

It happens that these properties don't hold in infinite dimensions, even in the case of the simplest *Hilbert space* of sequences $X = X^* = \ell^2$. The reason is that the basic normal cone is defined in terms of *sequential* limits but the weak$^*$ topology of $X^*$ is not sequential, so the weak$^*$ sequential closure of a set may not be weak$^*$ sequentially closed. The following example, which is due to Fitzpatrick (1994, personal communication; see also [144]), shows that values of the basic normal cone may *not* be even *norm closed* in $X^*$, hence neither weak$^*$ closed nor weak$^*$ sequentially closed in the dual space.

**Example 1.7 (nonclosedness of the basic normal cone in $\ell^2$).** *There are a closed subset $\Omega$ of the Hilbert space $\ell^2$ and a boundary point $\bar{x} \in \Omega$ such that $N(\bar{x};\Omega)$ is not norm closed in $\ell^2$.*

**Proof.** Consider a complete orthonormal basis $\{e_1, e_2, \ldots\}$ in the Hilbert space $\ell^2$ and form a nonconvex subset of $\ell^2$ by

$$\Omega := \big\{s(e_1 - je_j) + t(je_1 - e_m)\big|\ m > j > 1,\ \ s,t \ge 0\big\} \cup \big\{te_1\big|\ t \ge 0\big\}\ ,$$

which is obviously a cone. We can check that $\Omega$ is closed in $\ell^2$. Let us show that the basic normal cone $N(0;\Omega)$ is not closed in the norm topology of $\ell^2$. This follows from:

(i) $e_1^* + \frac{1}{j}e_j^* \in N(0;\Omega)$ for all $j = 2, 3, \ldots$,

(ii) $e_1^* + \frac{1}{j}e_j^* \to e_1^*$ as $j \to \infty$,

(iii) $e_1^* \notin N(0;\Omega)$,

where $e_j^*$ are linear functionals generated by $e_j$. To justify (i), we define $e_{jm}^* := e_1^* + \frac{1}{j}e_j^* + je_m^*$ for $1 < j < m$ and observe that $e_{jm}^* \in \widehat{N}(\frac{1}{m}(je_1 - e_m);\Omega)$. For each $j$ we have $\frac{1}{m}(je_1 - e_m) \to 0$ and $e_{jm}^* \xrightarrow{w} e_1^* + \frac{1}{j}e_j^*$ as $m \to \infty$, which gives (i). It is easy to check (ii), and so it remains to verify (iii).

Suppose that (iii) doesn't hold, i.e., $e_1^* \in N(0;\Omega)$. Then, by the definition of basic normals with $w^* = w$ (the weak convergence in $X^* = \ell^2$), there are sequences $x_k \xrightarrow{\Omega} 0$, $\varepsilon_k \downarrow 0$, and $x_k^* \xrightarrow{w} e_1^*$ such that $x_k^* \in \widehat{N}_{\varepsilon_k}(x_k;\Omega)$ for all $k \in I\!\!N$. Assume that some of $x_k$ are of the form $x_k = t_k e_1$ with $t_k \ge 0$. Putting $u := x_k + re_1$ with $r > 0$, we get