

Springer Monographs in Mathematics

Alexander D. Ioffe

Variational Analysis of Regular Mappings

Theory and Applications

 Springer

Springer Monographs in Mathematics

Editors-in-chief

Isabelle Gallagher, Paris, France

Minhyong Kim, Oxford, UK

Series editors

Sheldon Axler, San Francisco, USA

Mark Braverman, Toronto, Canada

Maria Chudnovsky, Princeton, USA

Sinan C. Güntürk, New York, USA

Claude Le Bris, Marne la Vallée, France

Pascal Massart, Orsay, France

Alberto Pinto, Porto, Portugal

Gabriella Pinzari, Napoli, Italy

Ken Ribet, Berkeley, USA

René Schilling, Dresden, Germany

Panagiotis Souganidis, Chicago, USA

Endre Süli, Oxford, UK

Shmuel Weinberger, Chicago, USA

Boris Zilber, Oxford, UK

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>

Alexander D. Ioffe

Variational Analysis of Regular Mappings

Theory and Applications

Alexander D. Ioffe
Department of Mathematics
Technion – Israel Institute of Technology
Haifa
Israel

ISSN 1439-7382 ISSN 2196-9922 (electronic)
Springer Monographs in Mathematics
ISBN 978-3-319-64276-5 ISBN 978-3-319-64277-2 (eBook)
DOI 10.1007/978-3-319-64277-2

Library of Congress Control Number: 2017947747

Mathematics Subject Classification (2010): 49J53, 49K40, 47J07, 49K99

© Springer International Publishing AG 2017

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

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

To my family

Preface

It is ingrained in mathematical sciences that any real advance goes hand in hand with the invention of sharper tools and simpler methods which also assist in understanding earlier theories and cast aside earlier more complicated developments.

David Hilbert

In science everything should be made as simple as possible, but not simpler.

Albert Einstein

Metric regularity has emerged during last 2–3 decades as one of the central concepts of *variational analysis*. The roots of this concept go back to a circle of fundamental regularity ideas from classical analysis embodied in such results as the implicit function theorem, the Banach open mapping theorem, and theorems of Lyusternik and Graves, on the one hand, and Sard’s theorem and transversality theory, on the other hand.

Smoothness is the key property of the objects to which the classical results are applied. Variational analysis, on the other hand, appeals to objects that may lack this property: functions and maps that are nondifferentiable at points of interest, set-valued mappings, etc. Such phenomena naturally appear in optimization theory and elsewhere.¹

In traditional nonlinear analysis, regularity of a continuously differentiable mapping (e.g., from a normed space or a manifold to another space or manifold) at a certain point means that its derivative at the point is an operator onto. This property, translated through available analytic or topological means to corresponding local

¹Grothendieck mentions the “ubiquity of stratified structures in practically all domains of geometry” in his 1984 *Esquisse d’un Programme*, see [140].

properties of the mapping, plays a crucial role in the study of some basic problems of analysis such as the existence and behavior of solutions of a nonlinear equation $F(x) = y$. The most fundamental consequence of regularity of F at some $\bar{x}$ is that the equation has a solution for any y in a neighborhood of $\bar{y} = F(\bar{x})$ and moreover, the distance from the solution to $\bar{x}$ is controlled by $\|y - \bar{y}\|$.

Similar problems appear if, instead of an equation, we consider an inclusion

$$y \in F(x) \tag{1}$$

(with F a set-valued mapping this time) which, in essence, is the main object of study in variational analysis. The challenge here is evident: There is no clear way to approximate the mapping by simple objects, like linear operators in the classical case.

The key step in the answer to this challenge was connected with the understanding of the metric nature of some key phenomena that appear in the classical theory. This eventually led to the choice of the class of metric spaces as the main playground and subsequently to abandoning approximation as the primary tool of analysis in favor of a direct study of the phenomena as such. The “metric theory” offers a rich collection of results that, being fairly general and stated in a purely metric language, are easily adaptable to Banach and finite-dimensional settings (still the most important in applications) and to various classes of mappings with special structure.

Moreover, however surprising this may sound, the techniques coming from the metric theory may appear in certain circumstances more efficient, flexible, and easy to use and at the same time able to produce more precise results than the available Banach techniques (e.g., connected with generalized differentiation), especially in infinite-dimensional Banach spaces. Furthermore, it should be added that the central role played by distance estimates has determined a quantitative character of the theory (contrary to the predominantly qualitative character of the classical theory). Altogether, this opens the gates to a number of new applications, such as, say, metric fixed point theory, differential inclusions, all chapters of optimization theory, and numerical methods.

Our goal is to give a systematic account of the theory of metric regularity. The three principal themes that will be at the focus of our attention are as follows: regularity criteria (containing quantitative estimates for rates of regularity), the effect of perturbations of a mapping on its regularity properties, and the role of metric regularity in analysis and optimization. The structure of this book corresponds to the logical structure of the theory. We start with a thorough study of metric theory that lays a solid foundation for the subsequent study of metric regularity of mappings, first between Banach and then between finite-dimensional spaces. In the last two cases, special attention is paid to mappings with special structures (e.g., mappings with convex graphs, single-valued Lipschitz mappings, polyhedral and semi-algebraic mappings). We also consider a number of applications of the theory to concrete problems of analysis and optimization, including those mentioned in the previous paragraph.

But we begin, in Chap. 1, with a brief survey of the classical theory, providing complete proofs of most of the results. Hopefully, this will help to make the threads connecting the classical and modern theories more visible as far as both the basic ideas and the specific techniques are concerned.

The proper study of the theory of metric regularity starts in Chap. 2. It is concentrated on a direct analysis of the phenomena exhibited by the three equivalent regularity properties: openness at a linear rate, metric regularity proper, and the pseudo-Lipschitz property of the inverse mapping. The main results of the chapter are the regularity criteria and perturbation theorems describing the effect of Lipschitz perturbations of the mapping on the rates of regularity. Both will be systematically used in the sequel. Note that along with the typical local regularity “near a point of the graph” that dominates the research and publications, we thoroughly consider nonlocal metric regularity “on a fixed set,” which so far has attracted less attention. Meanwhile, it leads to important applications, especially connected with various existence problems.

The chapter also contains a section in which we introduce and study weaker regularity concepts such as subregularity, calmness, and controllability. These properties may not be stable under small perturbations of the mapping, and hence can hardly be used for practical computations, but nonetheless prove to be extremely useful in some problems of analysis, e.g., in subdifferential calculus and the theory of necessary optimality conditions.

In Chap. 3, we continue to study metric theory, this time its infinitesimal aspects, with the concept of slope at the center. The main results are infinitesimal analogues of the corresponding general results of Chap. 2, actually consequences of the latter. But they are equal in strength only under some restrictions on the class of possible range spaces. The restrictions are not particularly strong. Length spaces, for instance (that is, spaces in which the distance between points is defined by the length of curves joining the points), would work. We also consider in some detail the so-called nonlinear regularity models in which the basic estimates involve certain functions of distances that appear in the definitions of basic regularity properties, rather than distances themselves. The chapter concludes with a study of global regularity which in certain respects is closer to the local theory than to the regularity theory on fixed sets.

Chapter 4 is rather a service chapter providing a bridge between the metric and the Banach space theories. It contains necessary information about tangential set-valued approximations as well as the theory of subdifferentials, mainly relating to the five main types of subdifferentials: Fréchet, Dini–Hadamard, limiting Fréchet, G-subdifferential, and Clarke’s generalized gradients. All results are supplied with proofs. The latter makes the chapter, together with § 7.2, in which we consider applications of regularity theory to subdifferential calculus, a reasonably complete, albeit short introduction to the subdifferential theory of variational analysis in arbitrary Banach spaces, not covered, by the way, by the existing literature (except, to a certain extent, in the recent monograph by Penot [265]).

Regularity criteria for set-valued mappings between Banach spaces established in Chap. 5, either dual, using subdifferentials and coderivatives, or primal, using directional derivatives, tangent cones, and contingent derivatives, all follow from the slope-based criteria of Chap. 3 through a series of simple propositions connecting the values of slopes of certain distance functions, naturally connected with the mapping, on the one hand, and norms of suitable elements of subdifferentials, coderivatives, or tangent cones, on the other hand. The propositions also allow us to make a fairly detailed comparison between various Banach space criteria that results in the rather surprising conclusion that certain dual criteria are never worse than their primal counterparts. Another result to be mentioned is the separable reduction theorem, which says that in the Banach case metric regularity of a set-valued mapping near a point is fully determined by its restrictions to separable subspaces of the domain and range spaces. This is a substantial simplification from the theoretical viewpoint, in particular because in separable spaces subdifferential regularity criteria are much more convenient to work with, especially if the space is not reflexive.

In Chap. 6, we turn to the study of regularity properties of some special classes of mappings between Banach spaces. Information about the structure of a mapping may help to use more specialized techniques and obtain more precise results, e.g., better estimates for regularity rates. This is the case we are dealing with in the first three sections devoted, respectively, to error bounds, mappings with convex graphs, and single-valued locally Lipschitz mappings. In the last section, we briefly review implications of regularity for two types of mappings from a Banach space into its dual: monotone operators and subdifferentials of lower semi-continuous functions.

In Chap. 7, we consider a number of applications of regularity theory to analysis and optimization, mainly in infinite-dimensional Banach spaces. We begin with a discussion of possible extensions of the classical transversality concepts to settings of variational analysis. Applications to subdifferential calculus are considered next with fairly short proofs of the strongest available versions of calculus rules for practically all operations of interest in variational analysis. The metric qualification conditions in the statements of the rules are not just the most general. Remarkably (and contrary to popular qualification conditions involving subdifferentials), they are formulated in exactly the same way for all spaces, whether finite or infinite-dimensional, and for all functions, whether convex or not.

We then present a Banach space version of the implicit function theorem for set-valued mappings with special attention to generalized equations. The existence theorem for differential inclusions proved in the fourth section is the first application of the regularity-on-fixed-sets theory. Another application is considered in the seventh section, where we discuss connections between metric regularity and metric fixed point theory. The theorems proved in this section cover a number of well-known and recently established results. But the main innovation is the proofs, void of any iterations and fully based on regularity arguments. It seems that the proofs may substantially change the common perception of the relationship between metric regularity and metric fixed point theories.

The remaining two sections of Chap. 7, the fifth and the sixth, are devoted to necessary conditions in optimization problems. In the fifth section, we discuss two “nonvariational” approaches, both based on regularity theory, and, in particular, demonstrate one of them by giving a nontraditional proof of second-order optimality conditions in smooth optimization problems with equality and inequality constraints. In the sixth section, we give a new proof of Clarke’s necessary conditions for optimal control problems with differential inclusions, so far the strongest for problems of that sort.

The finite-dimensional theory is studied in Chap. 8. Naturally, all regularity criteria here are the best possible and give the exact values of regularity rates. Proofs of the criteria and results relating to stability analysis in finite-dimensional spaces easily follow from what we have already obtained in the previous chapters. We then pass to the study of two classes of finite-dimensional sets that often appear in practice and have many remarkable properties, namely polyhedral sets and their finite unions (called semi-linear sets), on the one hand, and semi-algebraic sets, on the other hand. (The first is, of course, a subclass of the second.) Locally, polyhedral sets have the structure of polyhedral cones, that is, convex hulls of finitely many directions, which tremendously simplifies working with them. The geometry of semi-algebraic sets is more complex. The principal structural property of a semi-algebraic set is that it admits Whitney stratification into a smooth manifold (a sort of stratification in which different strata meet each other in a certain regular way). This makes it possible to obtain a fairly strong version of Sard’s theorem (in which the exceptional set is not just of measure zero but of a smaller dimension) for semi-algebraic set-valued mappings. These structural properties make the regularity theory of semi-linear and semi-algebraic sets and mappings especially rich and interesting.

Finally, in Chap. 9, we apply the theory to a variety of finite-dimensional problems of analysis and optimization. The problems are not essentially connected and cross through a spectrum of disciplines that can be observed in the titles of the sections. In the first section, we offer a new treatment of the theory of variational inequalities over polyhedral sets, fully based on the regularity theory and elementary polyhedral geometry. Some very recent results emphasizing the role of transversality properties for linear convergence of the method of alternating projections for convex and nonconvex sets are presented in the second section. In the third section, we introduce and study a class of curves of “almost steepest descent” for lower semi-continuous functions. We prove the existence of such curves under some natural assumptions on the function and the possibility to obtain them as solutions of the anti-subgradient inclusion involving limiting subdifferentials or generalized gradients. Then, in the fourth section, we return to discussions on the connection between regularity properties of the subdifferential mapping and the characterization of minima of nonconvex functions, in particular tilt stability of the minima under linear perturbations of the function. Finally, in the fifth section, we apply the semi-algebraic Sard’s theorem and transversality theorem to prove the typically regular behavior of solutions of nonsmooth optimization problems with

semi-algebraic data and of equilibrium prices in (also nonsmooth) models of exchange economies in the spirit of the famous Debreu theorem.

We have substantially benefited from the existing monographs of Klatte and Kummer [200] and, especially, of Dontchev and Rockafellar [96], which have allowed us to mainly concentrate on those basic aspects of the theory that have not been touched upon in these monographs. This is first of all the bulk of the metric theory, including general regularity criteria and all local theory involving slopes, but also quite a bit of the infinite-dimensional Banach space theory, everything connected with semi-algebraic mappings and the majority of applications. On the other hand, there are a number of issues of fundamental importance that have been thoroughly studied in the two quoted monographs and that we address using very different approaches based on the theory developed in this book.

The first to be mentioned is the circle of problems associated with implicit functions. This is, of course, one of the principal themes of any regularity theory and by far the subject of main interest in [96]. We start with a version of the implicit function theorem stated in the most general situation of inclusion (1), where both the domain and the range spaces are metric, and follow the evolution of this result step-by-step as the assumptions on the environments and properties of the mapping change. The proofs at every step are surprisingly simple, and the main idea of the standard proof of the classical implicit function theorem (given in Chap. 1) works already at a very early stage, still in the fully metric setting. Another example is the theory of variational inequalities over polyhedral sets, which we have already mentioned, in many respects very different from the theories available in the existing literature.

Proofs have been the subject of special attention in the process of writing. I have already mentioned that the metric theory offers some new and efficient technical instruments that have been systematically used. To a large extent thanks to them, new, shorter, and simpler proofs have been given to quite a few known results, especially associated with applications. This does not change the fact that in variational analysis we have to deal with rather complicated objects and structural information is often helpful in pursuing simpler and more transparent proofs. This partly explains the close attention paid in this book to classes of objects with special structures. Fortunately, such objects seem to be rather typical in practice.

The book is essentially a research monograph whose aim is to present the state of the art of a fast developing and widely applicable theory. A certain level of advanced knowledge (e.g., in functional analysis and optimization) and mathematical maturity is desirable. But I believe the book will be accessible to a broad audience, including graduate students in mathematical departments and engineering departments with an advanced mathematical education (typical for computer science, electrical engineering, and industrial engineering/operation research departments in many universities).

We conclude with a few technical remarks about the organization of the book, terminology and notation.

How the Book is Organized

Every chapter begins with a short preface explaining the content of the chapter, the main results, techniques, and connections with other parts of the book. Whenever needed, we then add an introduction containing all necessary prerequisite, often with proofs, and the notation that appears in the chapter for the first time. At the end of every chapter (starting with Chap. 2), and in Chaps. 7 and 9 at the end of every section, we add bibliographic comments whose main purpose is not only to indicate the source of one result or another, or the relation of the results presented in the text with those in the literature, but also to give some information about the development of the ideas, the connections with some other related areas of analysis, open questions, etc. There are also many exercises scattered throughout the text.

Terminology and Notation

I have tried to avoid introducing new terminology and notation, unless there was a real necessity (very rarely). Concerning objects, properties, etc., for which there is more than one term often used in the literature, I have usually chosen one for systematic use but mentioned some others as well in definitions (usually in parentheses). The most essential notation is repeatedly reintroduced to free the reader from having to search for its meaning.

It is to be finally mentioned that the number of publications connected with metric regularity is enormous and continues to grow. So the bibliography presented in this book is definitely far from complete. In addition to publications most immediately connected with the results and proofs contained in this book, I have tried to mention works in which, to the best of my knowledge, ideas and results were originated or substantially improved or received new understanding in one way or another, plus of course available monographs, survey articles, and closely related publications from other areas. Needless to say, by doing this, despite all attempts, one cannot avoid being subjective. Therefore, I wish to apologize in advance for (hopefully not many) possible and inevitable omissions, misquotations, and plain mistakes.

Haifa, Israel
April 2017

Alexander D. Ioffe

Acknowledgements

My first thanks are to D. Drusvyatskiy, who is the author of the section on semi-algebraic geometry. A number of other parts of this book benefited from the many conversations I had with him and A. Lewis over years of cooperation. I am deeply thankful to A. Dontchev and J.-P. Penot for their many remarks and comments. I also wish to thankfully acknowledge helpful feedback I got from F. Aragon Artacho, M. Fabian, D. Klatte, A. Kruger, and R. Lucchetti at various stages of writing this book.

And a few words about something not directly connected with regularity theory. In 1976, my family and I applied for emigration to Israel but were not allowed to leave the Soviet Union until early 1988. For me, as well for other “refuseniks” involved in science, one of the immediate results was that any access to professional facilities was completely cut off: no seminars, no conferences, no publications.² My professional survival became possible thanks to the support and help of my friends and colleagues. V.M. Alekseev, S.S. Kutateladze, B.T. and G.M. Polyak, V.M. Tikhomirov, N.D. Vvedenskaya, unlike many others, continued to keep close contact with me (considered dangerous by many). J.P. Aubin, F.H. Clarke, R.T. Rockafellar, R. Vinter, and J. Warga sent me many materials, papers, and books, using various visitors to the USSR (my mail was basically cut off as well), and also monitored the publication process of the papers I managed to smuggle to the West with the same courageous visitors. This is impossible to forget.

²That time in the USSR you could not submit a paper for publication without a permission of the security department of the university, institute, etc., you were affiliated with.

Contents

1	The Classical Theory	1
1.1	Introduction	2
1.2	Regular Points of Smooth Maps: Theorems of Lyusternik and Graves	8
1.3	The Inverse and Implicit Function Theorems	12
1.4	Regular and Critical Values: Theorems of Sard and Smale	15
1.5	Manifolds	21
1.6	Transversality	23
1.7	The Morse Lemma	26
1.8	Exercises	28
2	Metric Theory: Phenomenology	33
2.1	Introduction	35
2.2	Regularity: Definitions and Equivalences	42
2.2.1	Local Regularity	42
2.2.2	General (Non-local) Case: Definitions and Discussion	48
2.2.3	Restricted Regularity	54
2.2.4	Regularity and Completeness	55
2.3	General Regularity Criteria	57
2.3.1	The Basic Lemma and Error Bounds	58
2.3.2	Main Regularity Criteria	60
2.3.3	A Fundamental Application – The Density Theorem	65
2.4	Related Concepts: Metric Subregularity, Calmness, Controllability, Linear Recession	66
2.5	Perturbations and Stability	71
2.6	Metric Implicit Function Theorems: Strong Regularity	76
2.7	Nonlinear Regularity Models	83
2.8	Supplement: Regularity of a Composition	87
2.9	Comments	90

3	Metric Theory: The Infinitesimal Viewpoint	99
3.1	Slopes and Metric Geometry	100
3.1.1	Curves and Lengths	101
3.1.2	Slope: An Infinitesimal Characterization of Length Spaces	102
3.1.3	Locally Coherent Spaces	105
3.2	Infinitesimal Regularity Criteria	107
3.2.1	Some Non-local Infinitesimal Results	108
3.2.2	Three Infinitesimal Local Regularity Criteria	110
3.2.3	Slope Criteria for Calmness and Subregularity	116
3.2.4	A Perturbation Theorem: Milyutin's Estimate Is Sharp	118
3.2.5	A Variational Approach	120
3.3	The Metric Implicit Function Theorem – Infinitesimal Version	122
3.4	Nonlinear Regularity Models – An Infinitesimal Viewpoint	126
3.4.1	Regularity of Order k	126
3.4.2	Error Bounds of Order k	129
3.4.3	The Variational Viewpoint	130
3.5	A Few Remarks Concerning Global Regularity	134
3.6	Comments	137
4	Subdifferentials: A Short Introduction	141
4.1	Introduction	142
4.2	Subdifferentials: General View	151
4.3	Subdifferentials of Fréchet and Hadamard. Tangent Cones	155
4.4	Subdifferentials of the Distance Function and Normal Cones	164
4.5	The Limiting Fréchet Subdifferential and Subdifferential Calculus in Asplund Spaces	171
4.6	Approximate G -subdifferential and Subdifferential Calculus in General Banach Spaces	174
4.6.1	G -subdifferential of a Lipschitz Function	177
4.6.2	General Case	182
4.6.3	The G -subdifferential on WCG Spaces	186
4.7	The Generalized Gradient	188
4.8	Comments	190
5	Banach Space Theory: Regularity Criteria	195
5.1	Homogeneous Mappings, Derivatives and Coderivatives	197
5.2	General Primal and Dual Regularity Criteria	202
5.2.1	The Case of a Single-Valued Mapping	203
5.2.2	Primal Estimates	205
5.2.3	Dual (Subdifferential or Coderivative) Estimates	209

5.2.4	Comparison of the Estimates	213
5.2.5	Exact Estimates	217
5.3	Point Criteria and Separable Reduction	222
5.3.1	Point Criteria	224
5.3.2	Separable Reduction	225
5.4	Linear Perturbations and Perfect Regularity	229
5.4.1	Radius of Regularity	229
5.4.2	Stability of Solutions with Respect to Linear Perturbations	232
5.4.3	A Counterexample: Failure of the Radius Formula in an Infinite-Dimensional Case	234
5.5	Comments	241
6	Banach Space Theory: Special Classes of Mappings	245
6.1	Error Bounds	246
6.1.1	Error Bounds for Convex Functions	247
6.1.2	General Results on Global Error Bounds	251
6.1.3	Comments	254
6.2	Mappings with Convex Graphs	254
6.2.1	Convex Processes	254
6.2.2	Mappings with Convex Graphs: The Robinson–Ursescu Theorem	259
6.2.3	Mappings with Convex Graphs: Regularity Rates	261
6.2.4	Linear Perturbations and Distance to Infeasibility	264
6.2.5	Systems of Convex Inequalities	270
6.2.6	Comments	272
6.3	Single-Valued Lipschitz Maps	273
6.3.1	The Case of an Asplund Space	274
6.3.2	Prederivatives and Directional Compactness	275
6.3.3	Scalarized Neighborhood Regularity Criteria	278
6.3.4	Semi-Fredholm Mappings and Point Criteria	282
6.3.5	Approximation by Sets of Linear Operators	288
6.3.6	Comments	292
6.4	Subdifferentials and Monotone Operators	293
6.4.1	Subdifferentials and Quadratic Growth	293
6.4.2	Enter Convexity and Monotonicity	296
6.4.3	Comments	297
7	Applications to Analysis and Optimization 1	299
7.1	Transversality	301
7.1.1	Transversality of Sets	301
7.1.2	Transversality of a Mapping to a Set in the Range Space	306
7.1.3	Comments	309

7.2	Subdifferential Calculus	310
7.3	Implicit Functions in Banach Spaces	319
7.3.1	General Results on Implicit Functions.	320
7.3.2	Implicit Function Theorem for Generalized Equations	323
7.3.3	Comments.	324
7.4	Differential Inclusions – Existence of Solutions	326
7.4.1	Local Existence Theorem	326
7.4.2	Existence on the Entire Interval	328
7.4.3	Dependence on Initial Conditions	330
7.4.4	Comments.	332
7.5	Necessary Conditions for Optimality	333
7.5.1	Non-controllability and Optimality	334
7.5.2	Exact Penalty	336
7.5.3	Second-Order Conditions	338
7.5.4	The Optimality Alternative	342
7.6	Optimal Control of Differential Inclusions	345
7.7	Fixed Points of Set-Valued Mappings	353
7.7.1	Existence of Fixed Points	353
7.7.2	Consequences for Stability and Globally Defined Mappings	359
7.7.3	The Two Maps Paradigm: Double Fixed Points and Related Problems	361
7.7.4	Comments.	362
8	Regularity in Finite-Dimensional Spaces	365
8.1	Regularity in IR^n	367
8.1.1	Regularity	367
8.1.2	Subregularity.	371
8.1.3	Transversality	374
8.1.4	Linear Perturbations and Stability	376
8.1.5	Local Versus Global Regularity	377
8.1.6	A Few Remarks Concerning Subdifferentials in IR^n	378
8.2	Polyhedral and Semi-linear Sets and Mappings	379
8.2.1	Face Geometry	380
8.2.2	Regularity Properties of Polyhedral and Semi-linear Mappings	385
8.2.3	Piecewise Affine Mappings.	390
8.3	Semi-algebraic Geometry (D. Drusvyatskiy).	394
8.3.1	Recognizing Semi-algebraicity	394
8.3.2	Basic Structure	396
8.3.3	The Size of Semi-algebraic Subdifferential Graphs	402
8.3.4	The Kurdyka–Łojasiewicz Inequality	407

8.4	A Semi-algebraic Sard Theorem and Generic Behavior	409
8.4.1	Sard's Theorem for Semi-algebraic Set-Valued Mappings	409
8.4.2	Generic Transversality	411
8.5	Comments	413
9	Applications to Analysis and Optimization 2	415
9.1	Variational Inequalities over Polyhedral Sets	417
9.1.1	The Main Result	418
9.1.2	Characterization Theorem	422
9.1.3	Regularity Rates	426
9.1.4	Application to KKT Systems	430
9.1.5	Comments	433
9.2	The Method of Alternating Projections	434
9.2.1	The Case of Convex Sets	434
9.2.2	The Case of Nonconvex Sets	437
9.2.3	Comments	442
9.3	Curves of Descent	443
9.3.1	Curves of Maximal and Near Maximal Slope	443
9.3.2	Existence Theorem	445
9.3.3	Representation by Subdifferential Dynamical Systems	449
9.3.4	Curves of Near Maximal Slope for Semi-algebraic Functions	451
9.3.5	Comments	454
9.4	Regularity Properties of Subdifferentials and Local Minima of Functions	454
9.4.1	The Case of a Semi-algebraic Function	454
9.4.2	First-Order Characterization of Second-Order Behavior	457
9.4.3	Strong Metric Regularity, Tilt Stability and Strong Optimality	461
9.4.4	Comments	465
9.5	Some Genericity Results	466
9.5.1	Tame Optimization	466
9.5.2	Debreu's Theorem on Regular Economies	472
9.5.3	Comments	476
	Notation	477
	Bibliography	481
	Index	493

Chapter 1

The Classical Theory

This chapter contains a brief account of the classical “smooth” regularity theory with two major blocks of results that lie at the heart of modern nonlinear analysis:

Lyusternik–Graves theorem $\Rightarrow$ Inverse Function & Implicit Function theorems and

Sard & Smale–Sard theorems $\Rightarrow$ Thom’s transversality theorem, the first explaining the virtues of the regularity properties and the second showing that they are typical (generic and full measure) for sufficiently smooth mappings.

The theory has a wide range of applications: in control theory, mathematical economics and, of course, in mathematics itself from optimization and numerical analysis to differential topology and dynamical systems. But the theory also has a beauty and perfection sufficient to make it a worthy object to study and develop.

Almost all main results and a number of accompanying results presented in this chapter are supplied with complete proofs. A few exceptions are those theorems that can be found in practically every book on functional or nonlinear analysis such as, for example, the Banach open mapping theorem. A partial exception, traditional for books on nonlinear analysis and differential topology, is Sard’s theorem, which is proved only for the C^∞ -case: the general proof (e.g. [1]) is much more complicated and not really needed for further discussions; the latter though equally applies to the simplified proof of the C^∞ -version of Sard’s theorem. But the elegance of the proof and the role of Sard’s theorem in analysis make the availability of the proof highly desirable.

We have tried to present the results in a form that would help to make the future passage to non-differentiable settings natural and understandable by emphasizing, whenever appropriate, along with the basic ideas and principles, the metric and quantitative aspects to which the classical theory has not paid much attention.

1.1 Introduction

Notation. Everywhere in this chapter X, Y, Z are Banach spaces over reals. As a rule we denote the norms in different spaces by the same symbol $\| \cdot \|$;

X^* – adjoint of X ;

$\langle x^*, x \rangle$ – the value of x^* on x (canonical bilinear form on $X^* \times X$);

$\mathbb{R}^n$ – the n -dimensional Euclidean space;

B – the closed unit ball in a Banach space (sometimes indicated by a subscript to avoid confusion, e.g. B_X is the unit ball in X and B_{X^*} is the unit ball in X^*);

$\overset{\circ}{B}$ – the open unit ball

$B(x, r)$ and $\overset{\circ}{B}(x, r)$ – closed and open balls of radius r and center at x ;

ker A – kernel of the (linear) operator A ;

$L^\perp = \{x^* \in X^* : \langle x^*, x \rangle = 0, \forall x \in L\}$ – annihilator of $L \subset X$;

Im A – image of the operator A ;

$d_Q(x)$ or $d(x, Q)$ – distance from x to Q ;

$d(Q, P) = \inf\{\|x - u\| : x \in Q, u \in P\}$ – distance between Q and P ;

$\text{ex}(Q, P) = \sup\{d(x, P) : x \in Q\}$ – excess of Q over P ;

$H(Q, P) = \max\{\text{ex}(Q, P), \text{ex}(P, Q)\}$ – Hausdorff distance between Q and P ;

I – the identity mapping (the space is sometimes indicated by a subscript, e.g. I_X);

$F|_Q$ – the restriction of a mapping F to the set Q ;

$\mathcal{L}(X, Y)$ – the space of linear bounded operators $X \rightarrow Y$ with the *operator norm*:

$$\|A\| = \sup_{\|x\|=1} \|Ax\|.$$

$X \times Y$ – Cartesian product of spaces;

$L \oplus M$ – direct sum of subspaces.

We use the standard conventions $d(x, \emptyset) = \infty$; $\inf \emptyset = \infty$; $\sup \emptyset = -\infty$ with one exception: when we deal with non-negative quantities we set $\sup \emptyset = 0$.

The Banach–Schauder open mapping theorem. Recall that a Banach space X is a *direct sum* of its closed subspaces L and M : $X = L \oplus M$ if the mapping $(x, y) \rightarrow x + y$ from $L \times M$ into X is a linear isomorphism, that is, its image is all of X and there is a K such that $\|x\| + \|y\| \leq K\|x + y\|$ for all $x \in L$ and $y \in M$. It is said that a closed subspace $L \subset X$ *splits* X if there is another closed subspace M such that $X = L \oplus M$. This is tantamount to the existence of a continuous linear projection $\pi : X \rightarrow L$, which is a continuous linear mapping $X \rightarrow X$ such that $\pi(X) = L$ and the restriction of π to L is the identity. Any finite-dimensional subspace and any closed subspace of finite codimension splits X . In a Hilbert space any closed subspace splits the space.

Theorem 1.1 (Banach–Schauder open mapping theorem). *Let $A : X \rightarrow Y$ be a linear bounded operator onto Y , that is, $A(X) = Y$. Then $0 \in \text{int } A(B)$.*

A close look at the proof of the theorem (available in most textbooks on functional analysis) allows us to deduce the following result.

Corollary 1.2. *If $A : X \rightarrow Y$ is a bounded linear operator and zero belongs to the interior of the closure of $A(B)$, then $0 \in \text{int } A(B)$.*

The corollary says that there is a $K > 0$ such that for any $y \in Y$ there is an $x \in X$ such that $A(x) = y$ and $\|x\| \leq K\|y\|$ (take as K the reciprocal of the radius of a ball in Y contained in the image of the unit ball in X under A). This, in turn, can be interpreted as an existence theorem for a *bounded right inverse* of a linear bounded operator onto Y , that is to say, the existence of a mapping R from Y into X such that $(A \circ R)(y) = y$ for all $y \in Y$ and $\|R(y)\| \leq K\|y\|$. It is to be emphasized that R is not uniquely defined unless A is one-to-one, and in general R cannot be a linear operator unless $\ker A$ splits X (see Proposition 1.9 and the subsequent remark).

As an immediate consequence of this interpretation we get another famous result of Banach. Recall that a bounded linear operator $A \in \mathcal{L}(X, Y)$ is *invertible*, or a *linear homeomorphism*, if it is one-to-one, maps X onto Y and the inverse mapping is also a bounded linear operator.

Corollary 1.3. *A linear bounded operator which is one-to-one and onto is a linear homeomorphism.*

There are other important consequences of the Banach–Schauder open mapping theorem.

Corollary 1.4 (triangle lemma). *Let X, Y and Z be Banach spaces, let $A : X \rightarrow Y$ be a linear bounded operator onto Y , and let $T : X \rightarrow Z$ be a bounded linear operator such that $\ker A \subset \ker T$. Then there is a bounded linear operator $S : Y \rightarrow Z$ such that $T = S \circ A$.*

Proof. By the assumption (as $\ker T \subset \ker A$) $T(x) = T(x')$ if $Ax = Ax'$. Thus the formula $S(y) = T(A^{-1}(y))$ defines a single-valued mapping from all of Y (as $\text{Im } A = Y$) into Z having the property that $S \circ A = T$. This mapping is linear because $A^{-1}(\lambda y) = \lambda A^{-1}(y)$ and $A^{-1}(y + y') = A^{-1}(y) + A^{-1}(y')$. Finally, by the Banach–Schauder theorem there is a $K > 0$ such that $Ax = y$ for some x with $\|x\| \leq K\|y\|$. This implies that $\|S(y)\| = \|Tx\| \leq K\|T\|\|y\|$, which means that S is a bounded linear operator. $\square$

Corollary 1.5 (annihilator of the kernel). *If A is a bounded linear operator from X onto Y , then*

$$(\ker A)^\perp = \text{Im } A^*.$$

Proof. If $x^* = A^*y^*$ and $x \in \ker A$, then $\langle x^*, x \rangle = \langle y^*, Ax \rangle = 0$. Conversely, let $x^* \in (\ker A)^\perp$. If we view x^* as a linear operator from X into $\mathbb{R}$, then A and x^* satisfy the conditions of the triangle lemma. Hence there is a $y^* \in Y^*$ such that $\langle y^*, Ax \rangle = \langle x^*, x \rangle$ for all x , that is, $x^* = A^*y^*$. $\square$

Definition 1.6 (*Banach constant*). Let $A : X \rightarrow Y$ be a bounded linear operator. The quantity

$$C(A) = \sup\{r \geq 0 : rB_Y \subset A(B_X)\} = \inf\{\|y\| : y \notin A(B_X)\}$$

will be called the *Banach constant* of A .

The Banach–Schauder theorem effectively says that *the Banach constant of a bounded linear operator onto a Banach space is positive*. The following proposition offers two more expressions for the Banach constant. Given a linear operator $A : X \rightarrow Y$, we set

$$\|A^{-1}\| = \sup_{\|y\| \leq 1} d(0, A^{-1}(y)) = \sup_{\|y\|=1} \inf\{\|x\| : Ax = y\}.$$

We shall see later in Chap. 5 that the quantity $\|A^{-1}\|$ is precisely the lower norm of the (generally) set-valued mapping A^{-1} . Of course, if A is a linear homeomorphism, this is the usual norm of the inverse operator.

Proposition 1.7 (calculation of $C(A)$). *For a bounded linear operator $A : X \rightarrow Y$*

$$C(A) = \inf_{\|y^*\|=1} \|A^*y^*\| = \|A^{-1}\|^{-1}. \quad (1.1.1)$$

In particular, if A is onto, then A^ is one-to-one and*

$$\|A^*y^*\| \geq C(A)\|y^*\|.$$

Proof. If $C(A) = 0$, then $\|A^*y^*\| \geq C(A)$ for any y^* . If $C(A) > 0$, take an $r < C(A)$. For any $y^* \in Y^*$ with $\|y^*\| = 1$

$$\|A^*y^*\| = \sup_{\|x\| \leq 1} \langle y^*, Ax \rangle \geq \sup\{\langle y^*, y \rangle : y \in rB_Y\},$$

so that (as r can be arbitrarily close to $C(A)$), $\inf_{\|y^*\|=1} \|A^*y^*\| \geq C(A)$.

Furthermore, if $C(A) > 0$, that is, $\text{int } A(B_X) \neq \emptyset$, then for any $y \notin A(B_X)$ there is a y^* , $\|y^*\| = 1$, separating y and $A(B_X)$, that is,

$$\|y\| \geq \langle y^*, y \rangle \geq \sup\{\langle y^*, v \rangle : v \in A(B_X)\} = \sup_{\|x\| \leq 1} \langle y^*, Ax \rangle = \|A^*y^*\|,$$

which proves the opposite inequality: $C(A) \geq \inf\{\|A^*y^*\| : \|y^*\| = 1\}$.

Finally, if $C(A) = 0$ then by Corollary 1.2 the closure of $A(B_X)$ cannot contain interior points and therefore there is an arbitrarily small y not belonging to the closure of $A(B_X)$, and the same separation argument shows that for any $\varepsilon > 0$ there is a y^* with $\|y^*\| = 1$ such that $\|A^*y^*\| \leq \varepsilon$. This completes the proof of the left equality in (1.1.1).

Let us prove that $C(A) = \|A^{-1}\|^{-1}$. If $C(A) = 0$, then by the Banach–Schauder theorem there is a $y \notin A(X)$, that is, such that $A^{-1}(y) = \emptyset$. According to our general convention $d(0, A^{-1}(y)) = \infty$ for such y , therefore $\|A^{-1}\| = \infty$, whence the equality. Assume now that $C(A) > 0$. By definition, for any positive $r < C(A)$ and any $y \in Y$ there is an x such that $y = Ax$ and $\|x\| \leq r^{-1}\|y\|$ and for any $r > C(A)$ there is a y for which such an x cannot be found. In other words,

$$C(A)^{-1} = \sup_{\|y\|=1} \inf\{\|x\| : y = A(x)\} = \sup_{\|y\|=1} d(0, A^{-1}(y)) = \|A^{-1}\|. \quad \square$$

Remark 1.8. (a) By analogy, we can introduce the *dual Banach constant* of A :

$$C^*(A) = \inf_{\|x\|=1} \|Ax\|.$$

Clearly, if $C^*(A) > 0$ then A is nonsingular, that is, $\ker A = \{0\}$. Another obvious relation is: $C^*(A^*) = C(A)$. We shall see in Chap. 5 that these trivial relations have fairly non-trivial analogues in variational analysis.

(b) If $X = \mathbb{R}^n$ and A is an operator from X into itself, then $C(A)$ is the minimal singular value of A . (Recall that singular values of a compact operator A are square roots of eigenvalues of AA^* .) This is an easy consequence of Proposition 1.7: consider the problem of minimizing $\|A^*x\|^2$ subject to $\|x\|^2 = 1$ and apply the Lagrange multiplier rule.

Proposition 1.9. *If $A \in \mathcal{L}(X, Y)$ is onto and $\ker A$ splits X , then there is a bounded linear operator $T : Y \rightarrow X$ which is a right inverse of A with $\|T\| \geq C(A)^{-1}$.*

Proof. Take a closed subspace $M \subset X$ which is complementary to $L = \ker A$. Then the restriction $A|_M$ of A to M is an invertible bounded linear operator onto Y . Set $T = (A|_M)^{-1}$ and consider it an operator into X . Clearly $(AT)y = y$ for all $y \in Y$, so T is a right inverse of A . On the other hand, as $C(A|_M) \leq C(A)$, then, as follows from Proposition 1.7, $\|T\| = C(A|_M)^{-1} \geq C(A)^{-1}$. $\square$

Remark 1.10. If $\ker A$ does not split X , we cannot state the existence of a linear bounded right inverse. However, we can be sure that *there is a continuous bounded right inverse if $C(A) > 0$* . This follows from the continuous selection theorem of Michael (see e.g. [12], p. 82)¹ because $A^{-1}(y)$ is a closed subspace and, as follows from the Banach–Schauder theorem, $\mathbf{H}(A^{-1}(y), A^{-1}(y')) \leq C(A)^{-1}\|y - y'\|$.

Fredholm operators. A linear bounded operator $A : X \rightarrow Y$ is called a *Fredholm operator* if its kernel has finite dimension and its image is a closed subspace of finite codimension in Y . The integer

¹Michael's theorem states that a lower semicontinuous set-valued mapping from a metric space into a Banach space whose values are closed convex sets admits a continuous selection.

$$\text{ind } A = \dim(\ker A) - \text{codim}(\text{Im } A)$$

is called the *index* of A . The collection of Fredholm operators from X into Y is an open set in the space $\mathcal{L}(X, Y)$ and the index is constant on each connected component (Gohberg–Krein [137]).

If both X and Y are finite-dimensional, then any linear mapping $X \rightarrow Y$ is obviously Fredholm and the index of any of them is equal to $\dim X - \dim Y$.

Derivatives. Everywhere in the book, when we speak about a (continuous) mapping from X into Y , we mean a mapping F defined on an (open) subset of X , the domain of F . So let F be a mapping from a neighborhood of $\bar{x} \in X$ into another Banach space Y . A bounded linear operator $A : X \rightarrow Y$ is the (Fréchet) *derivative* of F at $\bar{x}$ if

$$\|F(x) - F(\bar{x}) - A(x - \bar{x})\| = o(\|x - \bar{x}\|).$$

The derivative is usually denoted $F'(\bar{x})$ (sometimes $DF(x)$). If the derivative exists (it is unique in this case), then F is said to be (Fréchet) *differentiable* at $\bar{x}$. F is called *strictly differentiable* at $\bar{x}$ if it is differentiable at $\bar{x}$ and there is a neighborhood of $\bar{x}$ such that for all x in the neighborhood

$$F(x + h) - F(x) - F'(\bar{x})h = r(x, h)\|h\|,$$

where $\|r(x, h)\| \rightarrow 0$ when $x \rightarrow \bar{x}$ and $h \rightarrow 0$. This essentially says that the mapping $h \rightarrow F(x + h) - F(x) - F'(\bar{x})h$ satisfies the Lipschitz condition on small neighborhoods of $\bar{x}$ with Lipschitz constants going to zero as the neighborhoods shrink to $\bar{x}$.

It is said that F is *continuously* (Fréchet) differentiable at $\bar{x}$ if it is differentiable at every point in a neighborhood of $\bar{x}$ and the mapping $x \rightarrow F'(x)$ is continuous from the norm topology of X into the uniform operator topology in the space of operators, that is, if for any x in the neighborhood $\|F'(u) - F'(x)\| \rightarrow 0$ if $\|u - x\| \rightarrow 0$. Continuous differentiability implies strict differentiability.

A mapping is *twice differentiable* at x if it is Fréchet differentiable and there is a bounded bilinear mapping $B(x, u)$ from $X \times X$ into Y such that

$$f(x + h) = F(x) + F'(x)h + \frac{1}{2}B(h, h) + r(h),$$

where $\|r(h)\| = o(\|h\|^2)$. Among bilinear operators satisfying this equality (if such operators exist), there is precisely one which is symmetric (that is, $B(x, u) = B(u, x)$ for all x, u). The linear symmetric operator from X into X^* associated with the latter (that is, such that $\langle Ax, u \rangle = B(x, u)$), is called the *second (Fréchet) derivative* of F at x and is denoted $F''(x)$. F is said to be *twice continuously differentiable* at x if it is continuously differentiable at x , twice differentiable in a neighborhood of x and F'' is a continuous mapping from a neighborhood of x into the space of bilinear operators $X \times X \rightarrow Y$. Note that if F is twice continuously differentiable, the second derivative

at any point is the Fréchet derivative of the first derivative mapping $x \mapsto F'(x)$. The definition of derivatives of higher orders follows this model in an obvious way. It is said that F is of class C^k if it is defined on an open set and is k times continuously differentiable at every point of its domain. A mapping F from an open set of X into Y is a C^k -diffeomorphism if it is a homeomorphism and both F and F^{-1} are k times continuously Fréchet differentiable on U and $F(U)$, respectively.

Suppose now we have a mapping F from the product of two Banach spaces, e.g. X and Z into Y . Then by F_x and F_z we shall denote the *partial derivatives* of F with respect to the corresponding argument. If the partial derivatives exist and are continuous on an open set, then the mapping is continuously differentiable on this set, in which case we have the standard relation

$$F'(x, z)(h, v) = F_x(x, z)h + F_z(x, z)v.$$

Baire category and residual sets. A set in a topological space is called *residual* or a set of *second Baire category* if it contains a countable intersection of dense open sets. A complement of a residual set, that is, a set which is contained in a countable union of closed nowhere dense sets, is called a set of the *first Baire category*.

The Baire Category Theorem states that a *residual set in a complete metric space is dense* (hence nonempty). A property is called *generic* if it holds on a residual set.

A countable intersection of residual sets is again a residual set. Hence, if we have a sequence (P_n) of generic properties, then the property $\bigwedge P_n$ (meaning that all P_n are simultaneously satisfied) is also generic.

Manifolds. Many of the results that will be discussed in this chapter and in the book in general can be extended to functions and mappings on general smooth manifolds. But we will usually work with functions and maps on Banach spaces and occasionally on submanifolds of Banach spaces. The latter form a convenient class of sets in which a metric structure, which plays a key role in future discussions, is well coordinated with the structure of a differentiable manifold.

Namely, we shall call a subset M of a Banach space X a C^k -submanifold if for any $x \in M$ there are a closed subspace L_x of X , open neighborhoods $U_x \subset M$ of x and $V_x \subset L_x$ of zero and a C^k -diffeomorphism φ_x of V_x onto U_x such that $\varphi_x(0) = x$.

This definition is somewhat different from the standard definition involving local charts. It has been chosen to facilitate using the induced metric structures on submanifolds along with differential structures. To emphasize the difference we shall call the triple (L_x, V_x, φ_x) a *local parameterization* of M at x . It can be shown, however, (see Exercise 11) that this definition implies the existence of local charts if all spaces L_x split X (cf. Aubin–Ekeland [13]). In variational analysis such a property may not always be natural. But if it is satisfied, the submanifold receives another and often much more convenient representation.

Proposition 1.11. *Let $M \subset X$ be a C^k -manifold. Let (L, V, φ) be a local parameterization of M at a certain $\bar{x} \in M$. If L splits X , then there is a neighborhood U of $\bar{x}$ and a C^k mapping $\psi(x)$ from U into X such that $M \cap U = \psi^{-1}(0) = \{x \in U : \psi(x) = 0\}$.*

Proof. Let E be a closed subspace of X which is complementary to L , and let π be the projection to E parallel to L , that is, $\pi(x) = 0$ for $x \in L$. Set $\psi(x) = \pi \circ \varphi^{-1}$. $\square$

If L_x can be chosen the same for all $x \in M$, we shall say that M is *modeled on L* . In this case we shall mean by a local parameterization the pair (V_x, φ_x) . If the model space L has dimension (codimension) r , then we shall say that M is a submanifold of dimension (codimension) r .

The subspace $T_x M = \varphi'_x(0)(L_x)$ (clearly isomorphic to L_x) is called the *tangent space* to M at x . The elements of the tangent space are called *tangent vectors* (to M at x). It is an easy exercise to show that the tangent space does not depend on the choice of a local parameterization. Note also that the tangent space is a closed subspace as so is L_x . The annihilator of the tangent space, that is, the subspace $N_x M = (T_x M)^\perp$ is the *normal space* to M at x and its elements are *normal vectors* or *normals* to M at x .

Everywhere in the book the word “manifold” means “submanifold of a Banach space” with the induced metric. It is also convenient to consider the empty set as a C^∞ -manifold.

1.2 Regular Points of Smooth Maps: Theorems of Lyusternik and Graves

Theorem 1.12 (Graves [138]). *Let F be a continuous mapping from a neighborhood of $\bar{x} \in X$ into Y . Suppose that there are a linear bounded operator $A : X \rightarrow Y$ and positive numbers $\delta > 0$, $\gamma > 0$, $\varepsilon > 0$ such that $C(A) > \delta + \gamma$ and*

$$\|F(x') - F(x) - A(x' - x)\| < \delta \|x' - x\|, \quad (1.2.1)$$

whenever x and x' belong to the open ε -ball around $\bar{x}$. Then

$$B(F(\bar{x}), \gamma t) \subset F(B(\bar{x}, t))$$

for all $t \in (0, \varepsilon)$.

Proof. We can assume for convenience that $F(\bar{x}) = 0$. Take $K > 0$ such that $K C(A) > 1 > K(\delta + \gamma)$, and let $\|y\| < \gamma t$ for some $t < \varepsilon$. Set $x_0 = \bar{x}$, $y_0 = y$ and define recursively first x_n and then y_n as follows:

$$\begin{aligned} y_{n-1} &= A(x_n - x_{n-1}), \quad \|x_n - x_{n-1}\| \leq K \|y_{n-1}\|, \\ y_n &= A(x_n - x_{n-1}) - (F(x_n) - F(x_{n-1})). \end{aligned} \quad (1.2.2)$$

Subtracting the second equality from the first, we get $y_{n-1} - y_n = F(x_n) - F(x_{n-1})$, so that

$$y - y_n = \sum_{i=1}^n (y_{i-1} - y_i) = \sum_{i=1}^n (F(x_i) - F(x_{i-1})) = F(x_n).$$

Thus if the process can be extended infinitely and x_n converges to some x , we would have $F(x) = y$.

To see that this is indeed the case and to find how far x is from $\bar{x}$, we notice that $\|y_n\| \leq \delta \|x_n - x_{n-1}\|$ by (1.2.1) and the second equation in (1.2.2). Together with the inequality in (1.2.2), this gives $\|y_n\| \leq K\delta \|y_{n-1}\|$, so that

$$\|x_n - x_{n-1}\| \leq (K\delta)^{n-1} K \|y\|, \quad \|y_n\| \leq (K\delta)^n \|y\|, \quad (1.2.3)$$

and therefore

$$\|x_n - \bar{x}\| \leq \frac{K}{1 - K\delta} \|y\| \leq \gamma^{-1} \|y\| < t.$$

This means that the process can be continued and by (1.2.3) (x_n) is a Cauchy sequence. Finally, if $x = \lim x_n$, then $\|x - \bar{x}\| \leq t$, as claimed. $\square$

Corollary 1.13 (Graves' theorem at nearby points). *Under the assumptions of the theorem, the inclusion*

$$B(F(x), \gamma t) \subset F(B(x, t))$$

holds whenever $\|x - \bar{x}\| + t < \varepsilon$, $t > 0$.

Proof. Replace $\bar{x}$ by x and ε by $\varepsilon - \|x\|$. $\square$

The following result is also an easy corollary of Graves' theorem.

Theorem 1.14 (Banach constant under perturbation). *For any two bounded linear operators $A, T : X \rightarrow Y$*

$$|C(A + T) - C(A)| \leq \|T\|.$$

If, moreover, A is invertible and $C(A)^{-1} \|T\| < 1$, then $A + T$ is also invertible and $\|(A + T)^{-1}\| \leq (C(A) - \|T\|)^{-1}$.

Proof. To prove the first statement, we only have to show that $C(A + T) \geq C(A) - \|T\|$ (because $A = (A + T) + (-T)$). The inequality is trivial if $C(A) \leq \|T\|$. If $\|T\| < C(A)$, set $F(x) = A(x) + T(x)$, $\delta = \|T\|$, $\bar{x} = 0$. Then by Theorem 1.12 $(C(A) - \|T\|)\overset{\circ}{B} \subset (A + T)(B)$, from which the inequality easily follows.

If A is also invertible, then by Proposition 1.7

$$\|A(x)\| \geq \|A^{-1}\|^{-1} \|x\| = C(A) \|x\|$$

for all x . Therefore

$$\|(A + T)(x)\| \geq \|A(x)\| - \|T(x)\| \geq (C(A) - \|T\|) \|x\|,$$

which means that $A + T$ is one-to-one. But it is also onto since under the assumption, $C(A + T) > 0$, as we have seen. Hence $A + T$ is invertible and the last inequality is precisely the announced inequality for the norm of $(A + T)^{-1}$. $\square$

Theorem 1.15 ([189]). Under the assumptions of Graves' theorem, the inequality

$$d(x, F^{-1}(y)) \leq \frac{1}{\gamma} \|y - F(x)\|$$

holds for all (x, y) satisfying $\|x - \bar{x}\| + \gamma^{-1} \|y - F(x)\| < \varepsilon$.

Proof. Set $t = \gamma^{-1} \|y - F(x)\|$. Then $\|x - \bar{x}\| + t < \varepsilon$ and $y \in B(F(x), \gamma t)$. By Corollary 1.13 there is a $u \in X$ such that $F(u) = y$ and $\|x - u\| \leq \gamma^{-1} t$. Thus

$$d(x, F^{-1}(y)) \leq \|x - u\| \leq \gamma^{-1} \|y - F(x)\|.$$

$\square$

Up to now we have not needed any differentiability of the maps. It is time to return to the classical situation of continuously differentiable mappings and to see what can be obtained from the results just proved in this case.

Definition 1.16 (*regular point*). Let F be a continuously differentiable mapping from an open set $U \subset X$ into Y . A vector $x \in U$ is called a *regular point* of F if $F'(x)$ maps X onto the whole of Y : $\text{Im } F'(x) = Y$. If $\text{Im } F'(x) \neq Y$, we say that x is a *singular point* of F .

The first result we are going to state follows immediately from Theorem 1.14 and the Lyusternik–Graves theorem.

Theorem 1.17 (stability of the regularity property). *The set of regular points of a C^1 -mapping is open. Hence the set of singular points is closed in the domain of the mapping. Moreover, let P be a topological space of parameters, and let $F(x, p)$ be continuous and have a continuous (jointly in (x, p)) Fréchet derivative with respect to x . Assume that for certain $\bar{p} \in P$ the mapping $F(\cdot, \bar{p})$ is regular at $\bar{x}$. Then there are neighborhoods $U \subset X$ and $Q \subset P$ of $\bar{p}$ such that any $x \in U$ is a regular point of $F(\cdot, p)$ for every $p \in Q$.*

This result introduces a property that can be characterized as a stability of the regularity property under small variations of the mapping and the point. This is an important fact since (as we have already mentioned) only those phenomena of the real world that do not disappear if the environment changes slightly can be observed. We shall see, however, in the course of further discussions that this is in a sense the weakest stability property and that regularity of mappings display much more robust behavior in various situations.

Definition 1.18 (*tangent cone*). A vector $h \in X$ is called a *tangent* to a set $S \subset X$ at $x \in S$ if $d(x + th, S) = o(t)$ when $t \rightarrow +0$. The collection of all tangent vectors to S at x (which is always a cone containing zero) is denoted by $T(S, x)$. If this set is a subspace, it is called the *tangent space* to S at x .

The last definition applies to any set, not necessarily to submanifolds. The different notation (though standard in variational analysis) emphasizes the difference with the definition of a tangent space to a submanifold given in §1. Note also that the collection of tangent vectors to an arbitrary set need not be a subspace. It can be shown, however, that for submanifolds both definitions are equivalent.

Theorem 1.19 (Lyusternik [235]). *Suppose that F is strictly differentiable and regular at $\bar{x}$. Then the collection of vectors tangent to the level set $M = \{x : F(x) = F(\bar{x})\}$ at $\bar{x}$ is a subspace coinciding with $\ker F'(\bar{x})$.*

Proof. Clearly $T(M, \bar{x}) \subset \ker F'(\bar{x})$. Indeed, if h is a tangent vector to M at x , then for any t there is an $h(t)$ such that $F(x + th(t)) \in M$ and $\|h - h(t)\| = o(t)$. This means that

$$F'(x)h = \lim_{t \rightarrow 0} F'(\bar{x})h(t) = \lim_{t \rightarrow 0} t^{-1}(F(\bar{x} + th(t)) - F(\bar{x}) + o(\|th(t)\|)) = 0.$$

This implication holds without any assumption on the range of $F'(\bar{x})$.

On the other hand, if we set $A = F'(\bar{x})$, then F and A satisfy the assumptions of Theorem 1.12 with any positive δ . Therefore, given an $h \in \ker A$, for any small δ (say, all $\delta < C(A)/2$) by Theorem 1.15

$$d(\bar{x} + th, M) \leq \frac{C(A)}{2} \|F(\bar{x}) - F(\bar{x} + th)\| = o(t).$$

□

Lyusternik's theorem was proved in 1934 and Graves' theorem in 1950. The methods they used in their proofs were very similar (and our proof of Graves' theorem follows the same lines). For that reason the following statement, which is somewhat weaker than Graves' theorem and somewhat stronger than the Lyusternik's theorem, is usually called the Lyusternik–Graves theorem.

Theorem 1.20 (Lyusternik–Graves theorem). *Assume that $F : X \rightarrow Y$ is strictly differentiable and regular at $\bar{x}$. Then for any positive $r < C(F'(\bar{x}))$, there is an $\varepsilon > 0$ such that*

$$B(F(x), rt) \subset F(B(x, t)),$$

whenever $\|x - \bar{x}\| < \varepsilon$, $0 \leq t < \varepsilon$.

We shall see later (see Corollary 2.80 in the next chapter) that the converse is also true. This fact will appear as a consequence of a deep theorem of Milyutin to be proved in the next chapter.