Already published

1 W.M.L. Holcombe *Algebraic automata theory*
2 K. Petersen *Ergodic theory*
3 P.T. Johnstone *Stone spaces*
4 W.H. Schikhof *Ultrametric calculus*
5 J.-P. Kahane *Some random series of functions, 2nd edition*
6 H. Cohn *Introduction to the construction of class fields*
7 J. Lambek & P.J. Scott *Introduction to higher-order categorical logic*
8 H. Matsumura *Commutative ring theory*
9 C.B. Thomas *Characteristic classes and the cohomology of finite groups*
10 M. Aschbacher *Finite group theory*
11 J.L. Alperin *Local representation theory*
12 P. Koosis *The logarithmic integral I*
13 A. Pietsch *Eigenvalues and s-numbers*
14 S.J. Patterson *An introduction to the theory of the Riemann zeta-function*
15 H.J. Baues *Algebraic homotopy*
16 V.S. Varadarajan *Introduction to harmonic analysis on semisimple Lie groups*
17 W. Dicks & M. Dunwoody *Groups acting on graphs*
18 L.J. Corwin & F.P. Greenleaf *Representations of nilpotent Lie groups and their applications*
19 R. Fritsch & R. Piccinini *Cellular structures in topology*
20 H Klingen *Introductory lectures on Siegel modular forms*
22 M.J. Collins *Representations and characters of finite groups*
24 H. Kunita *Stochastic flows and stochastic differential equations*
25 P. Wojtaszczyk *Banach spaces for analysts*
26 J.E. Gilbert & M.A.M. Murray *Clifford algebras and Dirac operators in harmonic analysis*
27 A. Frohlich & M.J. Taylor *Algebraic number theory*
28 K. Goebel & W.A. Kirk *Topics in metric fixed point theory*
29 J.F. Humphreys *Reflection groups and Coxeter groups*
30 D.J. Benson *Representations and cohomology I*
31 D.J. Benson *Representations and cohomology II*
33 C. Soulé et al *Lectures on Arakelov geometry*

A Primer of Nonlinear Analysis

Antonio Ambrosetti
Scuola Normale Superiore, Pisa

Giovanni Prodi
Department of Mathematics, University of Pisa

CAMBRIDGE UNIVERSITY PRESS

Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
40 West 20th Street, New York, NY 10011-4211, USA
10 Stamford Road, Oakleigh, Victoria 3166, Australia

First published 1993

Printed in Great Britain at the University Press, Cambridge

Library of Congress cataloguing in publication data available
A catalogue record for this book is available from the British Library

ISBN 0 521 37390 5 hardback

Contents

Preface

In the last few decades, once linear functional analysis was quite widely and thoroughly estabilished, the interest of scientists in Nonlinear Analysis has been increasing a lot. On the one hand the treatments of various classical problems have been unified; on the other, theories specifically nonlinear, of great significance and applicability, have come out.

This book provides an introduction to basic aspects of Nonlinear Analysis, namely those based on differential calculus in Banach spaces. The matter is expressed in a geometric style, in the sense that the results obtained are often a transposition to infinite dimensions of events which are intuitive in $\mathbb{R}^2$ or $\mathbb{R}^3$. Indeed, this was a primary characteristic of the works of Pincherle, Volterra and Fréchet.

The topics treated can be divided into two main parts and are preceded by a short chapter in which some introductory material is recalled, and also the main notation fixed.

In the first part, differential calculus in Banach spaces is discussed, together with local and global inversion theorems.

The second part deals with bifurcation theory which in spite of its elementary character is, perhaps, one of the most powerful tools used in Nonlinear Analysis. Our attention is here devoted almost entirely to the case of simple eigenvalues, but an accurate analysis of hypotheses is made, in order to include, for example, also the celebrated Hopf theorem.

A specific feature of Nonlinear Analysis is that the theoretical setting is strictly linked to applications, especially those related to differential equations, where the power of nonlinear methods is expressed in a more striking way. Moreover, a relevant fact to be emphasized is that problems that are often considered of formidable difficulty, once they are framed in an appropriate functional setting, may be faced and solved quite easily.

It is, indeed, this aspect, peculiar to Nonlinear Analysis, that has driven us to leave considerable space to applications to differential equations, including various important classical problems such as Bénard Problem, the problem of water waves, the restricted three-body problem and some others. Thus, in addition to more elementary examples and applications that usefully accomplish theoretical results, in separate paragraphs and/or chapters, we deal with those problems which require more care both in formulation and in resolution.

Tools, still of remarkable importance, such as the theory of Leray–Schauder topological degree, or the critical point theory, which would require wider theoretical background and more subtle arguments, are left out in this treatise.

The book in its outlines is self-contained for a reader who, besides infinitesimal calculus, is acquainted with fundamental results of Linear Functional Analysis such as the Hahn–Banach Theorem, the "Closed Graph" Theorem and the Fredholm Alternative Principle. Only some of the problems dealing with partial differential equations require a certain knowledge of Sobolev spaces and therefore, in just a few cases, we refer to results contained in original papers.

This volume is partially based on an earlier booklet, published in Italian by the Scuola Normale Superiore di Pisa in the series "Quaderni". The authors wish to thank the Scuola Normale Superiore for the encouragement.

0

Preliminaries and notation

This chapter contains the notation and some preliminary tools used throughout the book. Almost always, the results are not quoted in the most general form, but in a way appropriate to our purposes; nevertheless some of them are actually slightly more general than we strictly need. For more detail we refer to any book of (linear) Functional Analysis (for example [**Y**] or [**Br**] for topics reviewed in sections 0.1–0.4; to [**Br**], [**KFS**], [**GT**] for 0.5–0.6).

0.1 Some notation and definitions

$\mathbb{R}^n$ will denote the n-dimensional Euclidian space with scalar product $x \cdot y$ and norm given by $|x|^2 = x \cdot x$.

$X, Y, Z, \ldots$ denote (real) Banach spaces with norm $\|.\|_X$, $\|.\|_Y$, etc., respectively (the subscript will be omitted if no possible confusion arises). $B(x^*, r)$ denotes the ball $\{x \in X : \|x - x^*\| < r\}$ and $B(r)$ stands for $B(0, r)$.

If X^* is the topological dual of X the symbol $\langle ., . \rangle$ will indicate the duality pairing between X and X^*.

Let $\{x_n\}$ be a sequence in X. We say that x_n converges (strongly) to $x \in X$, written as $x_n \to x$, if $\|x_n - x\| \to 0$ as $n \to \infty$; we say that x_n converges weakly to x, written as $x_n \rightharpoonup x$, if $\langle \psi, x_n - x \rangle \to 0$ as $n \to \infty$ for all $\psi \in X^*$.

Let X be a Banach space and let V be a closed subspace of X. A *topological complement* of V in X is a *closed* subspace W of X such that $V \cap W = \{0\}$ and $X = V \oplus W$; $V \oplus W$ is called a *splitting* of X.

Recall also that, associated with such a splitting, there are (continuous) projections P and Q onto V and W respectively.

0.2 Continuous mappings

We will deal with continuous maps $f : U \to Y$, where U is an open subset of X. Continuity means that $f(x_n) \to f(x)$ (strongly) for any sequence x_n strongly convergent to $x \in X$. The set of all continuous $f : U \to Y$ will be denoted by $C(U, Y)$.

0.3 Integration

For continuous maps $f : [a, b] \to Y$ the definition of the Cauchy integral is given as in the elementary case, as the (strong) limit of the finite sums $\Sigma f(\xi_i)(t_i - t_{i-1})$ (with obvious meaning).

From

$$\|\Sigma_i f(\xi_i)(t_i - t_{i-1})\| \leq \Sigma_i \|f(\xi_i)(t_i - t_{i-1})\| \leq \Sigma_i \|f(\xi_i)\|(t_i - t_{i-1})$$

there follows immediately the inequality

$$\left\| \int_a^b f(t)\mathrm{d}t \right\| \leq \int_a^b \|f(t)\|\mathrm{d}t.$$

0.4 Linear continuous maps

The space of linear continuous maps $A : X \to Y$ will be denoted by $L(X, Y)$. The range of A, $R(A)$, is the linear space $\{A(x) : x \in X\}$. Sometimes, when $Y = X$, we will use the notation $L(X)$ instead of $L(X, X)$. Equipped with the norm

$$\|A\| = \sup\{\|A(x)\| : \|x\| \leq 1\},$$

$L(X, Y)$ is a Banach space. The identity map in $L(X)$ will be denoted by I_X.

Hereafter, for linear maps, the notation Ax or $A[x]$ may replace $A(x)$.

An *eigenvalue* of $A \in L(X)$ is a $\mu \in \mathbb{C}$ such that the equation $Ax = \mu x$ has solutions $x \neq 0$. Any such solution is an *eigenvector* associated to μ and Ker $(\mu I - A)$ is the *eigenspace* associated to μ. We will be mainly interested in the case when $A \in L(X)$ is compact, namely when A is completely continuous, if this is the case, the following results hold.

Theorem 0.1 **(Fredholm Alternative)** *Let $A \in L(X)$ be compact and $\mu \neq 0$. Then*

(i) Ker $(\mu I - A) = \{0\}$ *if and only if* $R(\mu I - A) = X$,

(ii) $R(\mu I - A) = [\text{Ker } (\mu I - A^*)]^\perp = \{u \in X : (\psi, u) = 0$ *for all* $\psi \in$ Ker $(\mu I - A^*)\}$.

Moreover one has the following

Theorem 0.2 *Let $A \in L(X)$ be compact and $\mu \neq 0$. Then*

(i) Ker $(\mu I - A)$ *is finite-dimensional and* Range$(\mu I - A)$ *is closed,*

(ii) *the sequence* Ker $((\mu I - A)^n)$ $(n \in \mathbb{N})$ *is increasing, that is* Ker $((\mu I - A)^m) \subset$ Ker $((\mu I - A)^{m+1})$ *for all* $m \geq 1$,

(iii) *there exists a finite* $p \in \mathbb{N}$ *such that* Ker $((\mu I - A)^p)$ =Ker $((\mu I - A)^q)$ *if and only if* $q \geq p$.

The (algebraic) multiplicity of μ is the dimension of the linear subspace

$$\cup_{n \in \mathbb{N}} \mathrm{Ker}((\mu I - A)^n) = \mathrm{Ker}((\mu I - A)^p).$$

It is worth pointing out that the *algebraic* multiplicity of μ is, in general, different from the *geometric* multiplicity, defined as the dimension of Ker $(\mu I - A)$ (algebraic and geometric multiplicity coincide for self-adjoint operators on Hilbert spaces). Hereafter, by the multiplicity of an eigenvalue $\mu \neq 0$ of a completely continuous $A \in L(X)$ we will always mean the algebraic multiplicity. An eigenvalue will be said to be *simple* if its multiplicity is 1.

0.5 Function spaces

Let Ω be an open subset of $\mathbb{R}^n$ with boundary $\partial\Omega$ and closure $\overline{\Omega}$.

We will use standard notation for spaces of continuous or differentiable real-valued functions $C^k(\overline{\Omega})$ $(k \geq 0)$, for Lebesgue spaces $L^p(\Omega)$ $(1 \leq p < \infty)$ or $L^\infty(\Omega)$. In some cases we will write $C(\overline{\Omega})$ instead of $C^0(\overline{\Omega})$. The spaces above are Banach spaces under the norms defined, respectively, by

$$\|u\|_C = \sup\{|u(x)| : x \in \overline{\Omega}\},$$

$$\|u\|_{C^k} = \sum_{0 \leq |\beta| \leq k} \|\mathrm{D}^\beta u\|_C \quad (\beta \text{ is a multi index}),$$

$$\|u\|_{L^p} = \left[\int_\Omega |u|^p \right]^{1/p}$$

(the symbol $\mathrm{d}x$ will be omitted whenever there is no ambiguity)

$$\|u\|_{L^\infty} = \text{ess sup}\{|u(x)| : x \in \Omega\}.$$

For $k \geq 0$ and $0 < \alpha \leq 1$, $C^{k,\alpha}(\overline{\Omega})$ denotes the space of Hölder functions with exponent α, namely the $u \in C^k(\overline{\Omega})$ such that, for all multi-index $\beta, |\beta| = k$,

$$\sup \left\{ \frac{|D^\beta u(x) - D^\beta u(y)|}{|x-y|^\alpha} : x, y \in \Omega, x \neq y \right\} < \infty.$$

For $k = 0$ and $\alpha = 1$, $C^{0,1}(\overline{\Omega})$ is nothing but the space of Lipschitz-continuous functions on $\overline{\Omega}$.
Equipped with the norm

$$\|u\|_{C^{k,\alpha}} = \|u\|_{C^k} + $$
$$+ \sup \left\{ \frac{|D^\beta u(x) - D^\beta u(y)|}{|x-y|^\alpha} : x, y, \in \Omega, x \neq y,\ |\beta| = k \right\},$$

$C^{k,\alpha}(\overline{\Omega})$ is a Banach space.
In some cases, we shall also work with Sobolev spaces $H^{k,p}(\Omega)$ ($k \geq 1, p \in [1, \infty)$) equipped with the norm

$$\|u\|_{H^{k,p}} = \sum_{0 \leq |\beta| \leq k} \|D^\beta u\|_{L^p}.$$

The notation H^k will stand for $H^{k,2}$ while $H_0^k(\Omega)$ will denote the closure of $C_0^\infty(\Omega)$, the space of C^∞ functions with compact support in Ω, under the norm $\|u\|_{H^{k,2}}$. Among others, let us recall the following result.

Theorem 0.3 (Poincaré Inequality) *Let Ω be bounded. Then there exists a constant $c = c(\Omega)$ such that*

$$\int_\Omega |u|^2 \leq c \int_\Omega |\nabla u|^2 \text{ for all } u \in H_0^1(\Omega).$$

As a consequence, $\|\nabla u\|_{L^2}$ is a norm in $H_0^1(\Omega)$ equivalent to $\|u\|_{H^{1,2}}$.

In addition to the Poincaré Inequality one has that the *embedding* of $H_0^1(\Omega)$ in $L^2(\Omega)$ is compact (Rellich's Theorem). Let us recall that X is embedded in Y, $X \hookrightarrow Y$, if $X \subset Y$ and the inclusion $\imath : X \to Y$ is continuous. If $X \hookrightarrow Y$ then $\exists\, c > 0$ such that

$$\|u\|_Y \leq c\|u\|_X, \text{ for all } u \in X.$$

If the inclusion $\imath : X \to Y$ is *compact* we will write $X \hookrightarrow\hookrightarrow Y$.

The following result is a particular case of the "Sobolev Embedding Theorems".

Theorem 0.4 *Suppose that Ω is bounded open set in $\mathbb{R}^n$, with boundary $\partial\Omega$ of class $C^{0,1}$, and let $k \geq 1$ and $1 \leq p \leq \infty$.*

(i) *If $kp < n$, then $H^{k,p}(\Omega) \hookrightarrow L^q(\Omega)$, for all $1 \leq q \leq np/(n-kp)$.*
(ii) *If $kp = n$, then $H^{k,p}(\Omega) \hookrightarrow L^q(\Omega)$, for all $q \in [1, \infty)$.*

(iii) *If $kp > n$, then $H^{k,p}(\Omega) \hookrightarrow C^{0,\alpha}(\overline{\Omega})$, where $\alpha = k - n/p$ if $k - n/p < 1$; $\alpha \in [0,1)$ is arbitrary if $k - n/p = 1$ and $p > 1$; $\alpha = 1$ if $k - n/p > 1$.*

In addition, there result the following.

(i′) *If $kp < n$, then $H^{k,p}(\Omega) \hookrightarrow\hookrightarrow L^q(\Omega)$, for all $1 \leq q < np/(n-kp)$.*
(ii′) *If $kp = n$, then $H^{k,p}(\Omega) \hookrightarrow\hookrightarrow L^q(\Omega)$, for all $q \in [1, \infty)$.*
(iii′) *If $kp > n$, then $H^{k,p}(\Omega) \hookrightarrow\hookrightarrow C(\overline{\Omega})$.*

0.6 Elliptic boundary value problems

Let Ω be a bounded domain (i.e. open connected) in $\mathbb{R}^n$ with smooth boundary $\partial\Omega$ (this will always be understood hereafter) and let $\mathcal{L}$ denote the differential operator

$$\mathcal{L} = \sum_{1 \leq i,j \leq n} \frac{\partial}{\partial x_i} \left(a_{ij}(x) \frac{\partial}{\partial x_j} \right) \tag{0.1}$$

where

$$a_{ij} = a_{ji} \in C^\infty(\overline{\Omega}). \tag{0.2}$$

$\mathcal{L}$ is (uniformly) elliptic if there exists $\alpha > 0$ such that

$$\sum_{1 \leq i,j \leq n} a_{ij}(x)\xi_i\xi_j \geq \alpha|\xi|^2, \text{ for all } x \in \Omega \text{ and } \xi \in \mathbb{R}^n \tag{0.3}$$

Throughout the book, any elliptic operator will be an *elliptic operator with smooth coefficients*, namely an $\mathcal{L}$ of the form (0.1) and such that (0.2)–(0.3) hold.

Consider the Dirichlet Boundary Value Problem (b.v.p. for short)

$$\left. \begin{aligned} -\mathcal{L}u &= h(x) \text{ in } \Omega, \\ u &= 0 \text{ on } \partial\Omega, \end{aligned} \right\} \tag{0.4}$$

where h is given a function on Ω.
Let $h \in L^2(\Omega)$; a *weak solution* of (0.4) is a $u \in H_0^1(\Omega)$ such that

$$\sum_{1 \leq i,j \leq n} \int_\Omega a_{ij} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} = \int_\Omega hv, \text{ for all } v \in C_0^\infty(\Omega).$$

If u is a weak solution of (0.4) and $u \in C^2(\Omega)$, then u is a classical solution.

Theorem 0.5 *Suppose $\mathcal{L}$ is an elliptic operator. Then the following results hold.*

(i) *Let $h \in L^p(\Omega)$, $2 \leq p < \infty$. Then (0.4) has a unique (weak) solution $u \in H_0^1(\Omega) \cap H^{2,p}(\Omega)$ and the following estimate holds:*

$$\|u\|_{H^{2,p}} \leq c\,\|h\|_{L^p}.$$

(ii) *If $h \in L^\infty(\Omega)$ then $u \in C^{1,\alpha}(\overline{\Omega})$ for any $0 < \alpha < 1$ and*
$$\|u\|_{C^{1,\alpha}} \le c\|h\|_{L^\infty}.$$
(iii) *If $h \in C^{0,\alpha}(\overline{\Omega})$ then $u \in C^{2,\alpha}(\overline{\Omega})$ is a classical solution of (0.4) and*
$$\|u\|_{C^{2,\alpha}} \le c\|h\|_{C^{0,\alpha}}.$$
In the above c stands for a positive constant, depending on Ω.

As a consequence of the preceding results, we can define an operator $K : L^2(\Omega) \to L^2(\Omega)$ (the *Green* operator of $-\mathcal{L}$ with zero Dirichlet boundary conditions) setting $Ku = v$ if and only if $-\mathcal{L}v = u$, $v \in H_0^1(\Omega)$. From the Rellich Theorem it follows immediately that K is compact.

Given a function $m \in L^\infty(\Omega)$, let us consider the linear eigenvalue problem
$$\left.\begin{aligned} -\mathcal{L}u &= \lambda m u \text{ in } \Omega,\\ u &= 0 \text{ on } \partial\Omega. \end{aligned}\right\} \tag{0.5}$$
An eigenvalue of (0.5) is a λ such that (0.5) has a solution $u \ne 0$. Any $\phi \ne 0$ satisfying (0.5) is an eigenfunction associated to the eigenvalue λ. If we set $\mu = 1/\lambda$ and $K_m(u) = K(mu)$, problem (0.5) is equivalent to $\mu u = K_m u$. The eigenvalues λ_k of (0.5) correspond, through $\mu_k = 1/\lambda_k$ to the eigenvalues of K_m. The multiplicity of λ_k is the multiplicity of μ_k. In some cases we will write $\lambda_k(m)$ or $\lambda_k(\Omega)$ to highlight the dependence of the eigenvalues of (0.5) on m or Ω.

Theorem 0.6 *Let $m \in L^\infty(\Omega)$, $m \ge 0$ and $m(x) > 0$ in a set of positive measure.*

(i) *Problem (0.5) has a sequence*
$$0 < \lambda_1(m) < \lambda_2(m) \le \ldots \le \lambda_k(m) \le \ldots$$
of eigenvalues such that $\lambda_k(m) \to +\infty$ as $k \to \infty$. The first eigenvalue $\lambda_1(m)$ is simple and the corresponding eigenfunctions do not change sign in Ω. We will let denote ϕ_1, (sometimes only ϕ) the eigenfunction such that (a) $\phi > 0$ in Ω and (b) $\int_\Omega \phi^2 = 1$.
We will also let ϕ_k denote the eigenfunctions corresponding to λ_k normalized by
$$\int_\Omega \phi_h\phi_k = \delta_{hk} = \begin{cases} 1 & \text{if } h = k,\\ 0 & \text{if } h \ne k. \end{cases}$$
When $m \equiv 1$ we will simply write λ_k instead of $\lambda_k(1)$.

(ii) (Comparison property) *If $m \le M$ in Ω then $\lambda_k(m) \ge \lambda_k(M)$; if $m < M$ in a subset of positive measure then $\lambda_k(m) > \lambda_k(M)$. In particular, if $m < \lambda_k$ (resp. $> \lambda_k$) then $\lambda_k(m) > 1$ (resp. < 1).*

(iii) (Variational characterization) *There results*
$$\lambda_k(m) = \max\left\{\int_\Omega mv^2 : v \in H_0^1(\Omega),\ \int_\Omega \sum a_{ij}\frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_j} = 1\ ,\ \int_\Omega v\phi_i = 0, \text{ for all } i = 1,\ldots,k-1\right\}.$$

(iv) (Continuity property) *$\lambda_1(m)$ depends continuously on m in the $L^{n/2}(\Omega)$ topology.*

(v) *Let Ω' be a bounded domain, such that $\Omega' \subset \Omega$. Then $\lambda_k(\Omega') \ge \lambda_k(\Omega)$ for all $k \ge 1$.*

Consider the non-homogenous b.v.p.
$$\left.\begin{aligned} -\mathcal{L}u &= \lambda m u + h \text{ in } \Omega,\\ u &= 0 \text{ on } \partial\Omega, \end{aligned}\right\} \tag{0.6}$$
with, say, $h \in L^2(\Omega)$.

From the Fredholm Alternative Theorem 0.1 we get the following.

Theorem 0.7

(i) *If λ is not an eigenvalue of (0.5), then (0.6) has a unique solution for all $h \in L^2(\Omega)$;*

(ii) *if λ is an eigenvalue of (0.5), then (0.6) has a solution if and only if $\int_\Omega h\phi_k = 0$ for any k such that $\lambda = \lambda_k$.*

According to Theorem 0.4 (iii) all the preceding discussion can be carried over taking $X = C^{2,\alpha}(\overline{\Omega})$, $h \in C^{0,\alpha}(\overline{\Omega})$ and m smooth.

The arguments above apply to Sturm–Liouville Problems
$$\left.\begin{aligned} -\frac{d}{dx}\left(\alpha\frac{d}{dx}u\right) + \beta u &= h(x) \quad (0 < x < \pi),\\ a_0u(0) + b_0u'(0) &= a_1u(\pi) + b_1u'(\pi) = 0, \end{aligned}\right\}$$
where $\alpha \in C^1([0,\pi])$, $\beta \in C((0,\pi])$, $\alpha, \beta > 0$ on $[0,\pi]$, and a_0, b_0, a_1, b_1 are such that $(a_0^2 + b_0^2)(a_1^2 + b_1^2) \ne 0$.
In fact, it is known **[D1]** that for all $h \in X := C([0,\pi])$ there exists a unique $u \in C^2([0,\pi])$ satisfying (0.6) and hence the map $K : h \to K(h)$ (is linear and) as an operator from X into itself is compact. It is also known that such a K has a sequence of positive, *simple* eigenvalues $\mu_1 > \mu_2 > \ldots > \mu_k \ldots$, such that $\mu_k \to 0$ as $k \to \infty$. Correspondingly,

the linear Sturm–Liouville eigenvalue problem

$$\left.\begin{aligned} -\frac{\mathrm{d}}{\mathrm{d}x}\left(\alpha\frac{\mathrm{d}}{\mathrm{d}x}u\right)+\beta u=\lambda u(x)\quad (0<x<\pi),\\ a_0u(0)+b_0u'(0)=a_1u(\pi)+b_1u'(\pi)=0,\end{aligned}\right\}$$

has a sequence of *simple* eigenvalues $\lambda_k = 1/\mu_k \to \infty$.

Another classical result we will need is the *Maximum Principle*.

Theorem 0.8 *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with smooth boundary and let $\lambda < \lambda_1$. If $u \in C^2(\Omega) \cup C(\overline{\Omega})$ is such that*

$$-\mathcal{L}u \geq \lambda u \text{ in } \Omega,$$
$$u \geq 0 \text{ on } \partial\Omega,$$

then $u \geq 0$ in Ω.

1

Differential calculus

This introductory chapter is mainly devoted to the differential calculus in Banach spaces. In addition to being a fundamental tool later on, the treatment of the calculus at this level permits better understanding at certain aspects, which might otherwise be neglected.

We discuss in Section 1 the Fréchet and Gâteaux derivatives as well as their elementary properties. The differentiability of the Nemitski operator is investigated in Section 2 and higher and partial derivatives are introduced in Sections 3 and 4, respectively.

1 Fréchet and Gâteaux derivatives

The Fréchet-differential is nothing else than the natural extension to Banach spaces of the usual definition of differential of a map in Euclidean spaces.

Let U be an open subset of X and consider a map $F : U \to Y$.

Definition 1.1 Let $u \in U$. We say that F is *(Fréchet-) differentiable* at u if there exists $A \in L(X, Y)$ such that, if we set

$$R(h) = F(u+h) - F(u) - A(h),$$

there results

$$R(h) = o(\|h\|), \tag{1.1}$$

that is

$$\frac{\|R(h)\|}{\|h\|} \to 0 \text{ as } \|h\| \to 0.$$

Such an A is uniquely determined and will be called the *(Fréchet) differential* of F at u and denoted by

$$A = dF(u).$$

If F is differentiable at all $u \in U$ we say that F is differentiable in U.

Hereafter, when there is no possible misunderstanding, *Fréchet differentiability* will be referred to simply as *differentiability*. A few comments on the preceding definition are in order.

(i) Let us verify that A is unique. Supposing the contrary, let $B \in L(X,Y)$ satisfy Definition 1.1 and $A \neq B$. It follows that

$$\frac{\|Ah - Bh\|}{\|h\|} \to 0 \quad \text{as } \|h\| \to 0. \tag{1.2}$$

If $A \neq B$ there exists $h^* \in X$ such that $a := \|Ah^* - Bh^*\| \neq 0$. Taking $h = t\,h^*$, $t \in \mathbb{R} - \{0\}$, one has

$$\frac{\|A(th^*) - B(th^*)\|}{\|th^*\|} = \frac{\|Ah^* - Bh^*\|}{\|h^*\|} = \frac{a}{\|h^*\|},$$

a constant, in contradiction with (1.2).

(ii) If F is differentiable at u then

$$F(u+h) = F(u) + \mathrm{d}F(u)h + o(\|h\|)$$

and F is continuous at the same point. Conversely if $F \in C(U,Y)$ then it is not necessary to require in Definition 1.1 the continuity of A. In fact (1.1) yields

$$A(h) = F(u+h) - F(u) - o(\|h\|)$$

and the continuity of F implies the continuity of A.

(iii) The definition of differentiability depends not on the norms but on the topology of X and Y only. That is if, for example, $\|.\|$ and $|||.|||$ are two equivalent norms on X then F is differentiable at u in $(X, \|.\|)$ if and only if F is in $(X, |||.|||)$ and the differential is the same.

Remark 1.2 The preceding comment (iii) could suggest the idea of extending the notion of Fréchet differentiability to locally convex topological spaces. The most natural way would be the following: let the topology of X (respectively Y) be induced by an infinite family of seminorms $|.|_{X,i}$ (resp. $\|.\|_{Y,j}$); define the differential of F as the linear continuous map A with the property that for all $|.|_{Y,j}$ there exists a seminorm $|.|_{X,i}$ such that $|F(u+h) - F(u) - Ah|_{Y,j} = o(|h|_{X,i})$. With such a definition all the main properties of the differential (below) hold true. Unfortunately, in dealing with the higher derivatives, there are strong difficulties and people introduced new classes of spaces, such as the "pseudotopological spaces", where a differential calculus suitable for the purposes of analysis can be carried out. These kind of topics, however, are beyond the purposes of our book.

Examples 1.3

(a) The constant map $F(u) = c$ is differentiable at any u and $\mathrm{d}F(u) = 0$ for all $u \in X$.

(b) Let $A \in L(X,Y)$. Since $A(u+h) - A(u) = A(h)$, it follows that A is differentiable in X and $\mathrm{d}A(u) = A$.

(c) Let $B : X \times Y \to Z$ be a bilinear continuous map. There results

$$B(u+h, v+k) - B(u,v) = B(h,v) + B(u,k) + B(h,k).$$

From the continuity at the origin it follows that

$$\|B(h,k)\| \leq c\|h\|\,\|k\|.$$

Then B is differentiable at any $(u,v) \in X \times Y$ and $\mathrm{d}B(u,v)$ is the map $(h,k) \to B(h,v) + B(u,k)$.

(d) Let H be a Hilbert space with scalar product $(.|.)$ and consider the map $F : u \to \|u\|^2 = (u|u)$. From

$$\|u+h\|^2 - \|u\|^2 = 2(u|h) + \|h\|^2$$

it follows that F is differentiable at any u and $\mathrm{d}F(u)h = 2(u|h)$. Note that $\|.\|$ is not differentiable at $u = 0$. For, otherwise, $\|h\| = Ah + o(\|h\|)$ for some $A \in L(H, \mathbb{R})$. Replacing h with $-h$ we would deduce that $\|h\| = -Ah + o(\|h\|)$ and hence $\|h\| = o(\|h\|)$, a contradiction.

(e) If $X = \mathbb{R}$, $U = (a,b)$ and $F : U \to Y$ is differentiable at $t \in U$, the differential $\mathrm{d}F(t)$ can be identified with $\mathrm{d}F(t)[1] \in Y$ though the canonical isomorphism $i : L(\mathbb{R}, Y) \to Y$, $i(A) = A(1)$. For example, if $Y = \mathbb{R}^n$ and $F(t) = (f_i(t))_{i=1,\dots,n}$, $\mathrm{d}F(t)$ "is" the vector with components $\mathrm{d}f_i/\mathrm{d}t$.

The main differentiation rules are collected in the following proposition.

Proposition 1.4

(i) *Let $F, G : U \to Y$. If F and G are differentiable at $u \in U$ then $aF + bG$ is differentiable at u for any $a, b \in \mathbb{R}$ and*

$$\mathrm{d}(aF + bG)(u)h = a\,\mathrm{d}F(u)h + b\,\mathrm{d}G(u)h.$$

(ii) (Composite-map formula) *Let $F : U \to Y$ and $G : V \to Z$ with*

$V \supset F(U)$, U and V open subsets of X and Y, respectively, and consider the composite map

$$G \circ F : U \to Z,\ G \circ F(u) := G(F(u)).$$

If F is differentiable at $u \in U$ and G is differentiable at $v := F(u) \in V$, then $G \circ F$ is differentiable at u and

$$\mathrm{d}(G \circ F)(u)h = \mathrm{d}G(v)[\mathrm{d}F(u)h].$$

In other words the differential of $G \circ F$ at u is the composition of the linear maps $\mathrm{d}F(u)$ and $\mathrm{d}G(v)$, with $v = F(u)$.

The proofs of (i) and (ii) do not differ from those of the differentiation rules in $\mathbb{R}^n$.

Definition 1.5 Let $F : U \to Y$ be a differentiable in U. The map

$$F' : U \to L(X,Y),\ F' : u \to \mathrm{d}F(u),$$

is called the *(Fréchet) derivative* of F.

If F' is continuous as a map from U to $L(X,Y)$ we will say that F is C^1 and write $F' \in C^1(U,Y)$.

Let us introduce the concept of *variational* (or *potential*) operator. If $Y = \mathbb{R}$, maps $J : U \to \mathbb{R}$ are usually called *functionals* and J' turns out to be a map from U to $L(X,\mathbb{R}) = X^*$ (the dual of X). In particular, if $X = H$ is a Hilbert space, $J'(u) \in H^*$ for all u and the Riesz Representation Theorem allows us to identify $J'(u)$ with an element of H. To be precise, we give the following definition.

Definition 1.6 Given a differentiable functional $J : U \to \mathbb{R}$ the *gradient* of J at u, denoted by $\nabla J(u)$, is the element of H defined by

$$(\nabla J(u)|h) = \mathrm{d}J(u)h, \text{ for all } h \in H. \tag{1.3}$$

A map $F : U \to H$ with the property that there exists a differentiable functional $J : U \to \mathbb{R}$ such that $F = \nabla J$ is called a *variational* (or *potential*) *operator*.

As for maps in $\mathbb{R}^n$, we can also define here a directional derivative, usually called the Gâteaux differential (for short, G-differential).

Definition 1.7 Let $F : U \to Y$ be given and let $x \in U$. We say that F is *G-differentiable* at u if there exists $A \in L(X,Y)$ such that for all $h \in X$ there results

$$\frac{F(u+\varepsilon h) - F(u)}{\varepsilon} \to Ah \text{ as } \varepsilon \to 0. \tag{1.4}$$

The map A is uniquely determined, called the *G-differential* of F at u and denoted by $\mathrm{d}_G F(u)$.

Clearly, if F is Fréchet-differentiable at u then F is G-differentiable there and the two differentials coincide. Conversely, the G-differentiability does not imply the continuity of F, even: recall the elementary example $F : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$F(s,t) = \left[\frac{s^2 t}{s^4 + t^2}\right]^2 \text{ if } t \neq 0,$$
$$F(s,0) = 0.$$

The following result replaces the elementary "Mean-Value Theorem" and plays a fundamental role in what follows.

For $u, v \in U$ let $[u,v]$ denote the segment $\{tu + (1-t)v : t \in [0,1]\}$.

Theorem 1.8 *Let $F : U \to Y$ be G-differentiable at any point of U. Given $u, v \in U$ such that $[u,v] \subset U$, there results*

$$\|F(u) - F(v)\| \le \sup\{\|\mathrm{d}_G F(w)\| : w \in [u,v]\}\ \|u - v\|.$$

Proof. Without loss of generality we can assume that $F(u) \neq F(v)$. By a well-known corollary of the Hahn–Banach Theorem there exists $\psi \in Y^*$, $\|\psi\| = 1$, such that

$$\langle \psi, F(u) - F(v)\rangle = \|F(u) - F(v)\|. \tag{1.5}$$

Let $\gamma(t) = tu + (1-t)v$, $t \in [0,1]$, and consider the map $h : [0,1] \to \mathbb{R}$ defined by setting

$$h(t) = \langle \psi, F[\gamma(t)]\rangle = \langle \psi, F(tu + (1-t)v)\rangle.$$

From $\gamma(t+\tau) = \gamma(t) + \tau(u - v)$ it follows that

$$\frac{h(t+\tau) - h(t)}{\tau} = \left\langle \psi, \frac{F[\gamma(t+\tau)] - F[\gamma(t)]}{\tau}\right\rangle$$
$$= \left\langle \psi, \frac{F[\gamma(t) + \tau(u-v)] - F[\gamma(t)]}{\tau}\right\rangle. \tag{1.6}$$

Since F is G-differentiable in U, passing to the limit in (1.6) as $\tau \to 0$, we find

$$h'(t) = \langle \psi, \mathrm{d}_G F(tu + (1-t)v)(u - v)\rangle. \tag{1.7}$$

Applying the Mean–Value Theorem to h one has

$$h(1) - h(0) = h'(\theta) \text{ for some } \theta \in (0,1). \tag{1.8}$$

Substituting (1.5) and (1.7) into (1.8) we get

$$\begin{aligned}\|F(u)-F(v)\| &= h(1)-h(0)=h'(\theta)\\ &= \langle \psi, \mathrm{d_G}F(\theta u+(1-\theta)v)(u-v)\rangle\\ &\le \|\psi\|\ \|\mathrm{d_G}F(\theta u+(1-\theta)v)\|\ \|u-v\|.\end{aligned}$$

Since $\|\psi\|=1$ and $\theta u+(1-\theta)v\in[u,v]$ the theorem follows.

As a consequence we can find a classic criterion of Fréchet differentiability.

Theorem 1.9 *Suppose $F:U\to Y$ is G-differentiable in U and let*

$$F'_{\mathrm G}:U\to L(X,Y),\ \ F'_{\mathrm G}(u)=\mathrm{d_G}F(u),$$

be continuous at u^*. Then F is Fréchet-differentiable at u^* and $\mathrm{d}F(u^*)=\mathrm{d_G}F(u^*)$.

Proof. We set

$$R(h):=F(u^*+h)-F(u^*)-\mathrm{d_G}F(u^*)h.$$

Plainly, R is G-differentiable in B_ε, for $\varepsilon>0$ small enough, and

$$\mathrm{d_G}R(h):k\to \mathrm{d_G}F(u^*+h)k-\mathrm{d_G}F(u^*)k. \tag{1.9}$$

Applying Theorem 1.8 with $[u,v]=[0,h]$, we find (note that $R(0)=0$)

$$\|R(h)\|\le \sup_{0\le t\le 1}\|\mathrm{d_G}R(th)\|\ \|h\|. \tag{1.10}$$

From (1.9) with th instead of h, we deduce

$$\|\mathrm{d_G}R(th)\|=\|\mathrm{d_G}F(u^*+th)-\mathrm{d_G}F(u^*)\|.$$

Substituting into (1.10) we find

$$\|R(h)\|\le \sup_{0\le t\le 1}\|\mathrm{d_G}F(u^*+th)-\mathrm{d_G}F(u^*)\|\ \|h\|.$$

Since $F'_{\mathrm G}$ is continuous,

$$\sup_{0\le t\le 1}\|\mathrm{d_G}F(u^*+th)-\mathrm{d_G}F(u^*)\|\to 0 \text{ as } \|h\|\to 0$$

and therefore $R(h)=o(\|h\|)$.

In view of Theorem 1.8, to find the Fréchet differential of F one can determine $\mathrm{d_G}F$ and show that $F'_{\mathrm G}$ is continuous.

Let $F\in C([a,b],X)$ and set (see subsection 0.3)

$$\Phi(t)=\int_a^t F(\xi)\mathrm{d}\xi.$$

It is immediately verifiable that Φ is differentiable and $\Phi'(t)=F(t)$

(we are using the canonical identification between $L(\mathbb{R},X)$ and X; see Example 1.3 (e)). Φ is called a primitive of F.

From Theorem 1.8 it follows that

$$\|\Phi(t)-\Phi(s)\|\le \sup\{\|F(\xi)\|\,(t-s): s\le\xi\le t\}.$$

Hence, if $F(\xi)=0$ for all $\xi\in[a,b]$, one has Φ =constant. In particular, Φ is, up to a constant, the unique primitive of F.

As a consequence we can obtain the following useful formula. Suppose that $[u,v]\in U$ and let $F\in C^1(U,Y)$. The map $F\circ\gamma:[0,1]\to Y$, $F\circ\gamma(t)=F(tu+(1-t)v)$ is C^1 and

$$(F\circ\gamma)'(t)=F'(tu+(1-t)v)\ [u-v].$$

Integrating from 0 to 1 we get

$$\begin{aligned}F(v)-F(u) &= \int_0^1 F'(tu+(1-t)v)[u-v]dt\\ &= \left[\int_0^1 F'(tu+(1-t)v)dt\right](u-v).\end{aligned} \tag{1.11}$$

Note that in the last integral F' is meant to take values in $L(X,Y)$.

2 Continuity and differentiability of Nemitski operators

In this section we want to study the differentiability of an important class of operators arising in nonlinear analysis: the so called "Nemitski operators" we are going to define.

Nemitski operators

Let Ω be an open bounded subset of $\mathbb{R}^n$ and let $M(\Omega)$ denote the class of real-valued functions $u:\Omega\to\mathbb{R}$ that are measurable on Ω. Here, and always hereafter, the measure is the Lebesgue one and will be denoted by μ; all the functions we will deal with in this section are taken in $M(\Omega)$.

Let $f:\Omega\times\mathbb{R}\to\mathbb{R}$ be given.

Definition 2.1 The *Nemitski operator* associated to f is the map defined on $M(\Omega)$ by setting

$$u(x)\to f(x,u(x)).$$

The same symbol f will be used to denote both f and its Nemitski operator.

We shall assume that f is a Carathéodory function. More precisely, we will say that f satisfies (C) if

(i) $s \to f(x,s)$ is continuous for almost every $x \in \Omega$,
(ii) $x \to f(x,s)$ is measurable for all $s \in \mathbb{R}$.

For the purpose of analysis it is particularly interesting when the Nemitski operators act on Lebesgue spaces $L^p = L^p(\Omega)$ (hereafter we will write L^p for $L^p(\Omega)$) and we shall discuss this case in some detail. Let us start by noticing that

$$f(u) \in M(\Omega) \text{ for all } u \in M(\Omega). \tag{2.1}$$

Indeed, if $u \in M(\Omega)$ there is a sequence χ_n of simple functions such that $\chi_n \to u$ a.e. in Ω. From (C) it follows that

$$f(\chi_n) \text{ is measurable and } f(\chi_n) \to f(u) \text{ a.e. in } \Omega,$$

and from this we deduce that $f(u) \in M(\Omega)$.

Continuity of Nemitski operators

Let $p, q \geq 1$ and suppose

$$|f(x,s)| \leq a + b|s|^\alpha, \quad \alpha = \frac{p}{q}, \tag{2.2}$$

for some constants $a, b > 0$.

Theorem 2.2 *Let $\Omega \subset \mathbb{R}^n$ be bounded and suppose f satisfies* (C) *and* (2.2). *Then the Nemitski operator f is a continuous map from L^p to L^q.*

For the proof we need the following measure-theoretic result (see, for example, [**Br**], Theorem IV.9).

Theorem 2.3 *Let $\mu(\Omega) < \infty$ and let $u_n \to u$ in L^p. Then there exist a sub-sequence u_{n_k} and $h \in L^p$ such that*

$$u_{n_k} \to u \quad \text{a.e. in } \Omega, \tag{2.3}$$

$$|u_{n_k}| \leq h \quad \text{a.e. in } \Omega. \tag{2.4}$$

Proof of Theorem 2.2 From (2.1)–(2.2) it follows immediately that $f(u) \in L^q$ whenever $u \in L^p$.
To show that f is continuous from L^p to L^q, let $u_n, u \in L^p$ be such that

$$\| u_n - u \|_{L^p} \to 0.$$

Using Theorem 2.3 we can find a sub-sequence $\{u_{n_k}\}$ of $\{u_n\}$ and $h \in L^p$ satisfying (2.3)–(2.4). Since u_{n_k} converges almost everywhere to u, it readily follows from (C) that

$$f(u_{n_k}) \to f(u) \text{ a.e. in } \Omega. \tag{2.5}$$

Moreover, from assumption (2.2) and (2.4) we infer

$$|f(u_{n_k})| \leq a + b\,|u_{n_k}|^\alpha \leq a + b\,|h|^\alpha \in L^q. \tag{2.6}$$

As an immediate consequence of the Lebesgue Dominated-Convergence Theorem, (2.5)–(2.6) yield

$$\|f(u_{n_k}) - f(u)\|^q_{L^q} = \int_\Omega |f(u_{n_k}) - f(u)|^q \to 0.$$

Since any sequence u_n converging to u in L^p has a sub-sequence u_{n_k} such that $f(u_{n_k}) \to f(u)$ in L^q, we can conclude that f is continuous at u, as a map from L^p to L^q.

Remark 2.4 Theorem 2.2 can be proved assuming that f satisfies (2.2) with a replaced by $a(x) \in L^q$.

Remark 2.5 It is possible to show that, if (C) holds and $f(u) \in L^q$ for all $u \in L^p$, then $f \in C(L^p, L^q)$. For this and other kinds of arguments we refer to [**Va**], p.154 and following.

Differentability of Nemitski operators

Our next result deals with the differentiability of Nemitski operators. First some remarks are in order.

Let $p > 2$ and suppose f has partial derivative $f_s = \partial f/\partial s$ satisfying (C) and such that

$$|f_s(x,s)| \leq a + b|s|^{p-2} \tag{2.7}$$

for some constants $a, b > 0$. Since f_s satisfies (2.7), Theorem 2.2 applies and the Nemitski operator f_s is continuous from L^p to L^r, with $r = p/(p-2)$. As a consequence, for the function $f_s(u)v$ defined by

$$f_s(u)v : x \to f_s(x, u(x))v(x)$$

one has that $f_s(u)v \in L^{p'}$ for $u, v \in L^p$, where $p' = p/(p-1)$ is the conjugate exponent of p.

Theorem 2.6 *Let $\Omega \subset \mathbb{R}^n$ be bounded and suppose that $p > 2$ and f satisfies* (C). *Moreover, we suppose that $f(x,0)$ is bounded and that f has partial derivative f_s satisfying* (C) *and* (2.7).

Then $f : L^p \to L^{p'}$ is Fréchet-differentiable on L^p with differential

$$df(u) : v \to f_s(u)v.$$

Proof. Integrating (2.7) we find constants $c, d > 0$ such that

$$|f(x,s)| \le c + d|s|^{p-1},$$

and another application of Theorem 2.2 yields the continuity of f as a map from L^p to $L^{p'}$, with $p' = p/(p-1)$.

For $u, v \in L^p$ we evaluate

$$\omega(u,v) = \|f(u+v) - f(u) - f_s(u)v\|_{L^{p'}}$$

$$= \left[\int_\Omega |f(x,u(x)+v(x)) - f(x,u(x)) - f_s(x,u(x))v(x)|^{p'} \right]^{1/p'}.$$

By the Mean-Value Theorem one has (for almost every $x \in \Omega$)

$$|f(x,u(x)+v(x)) - f(x,u(x)) - f_s(x,u(x))v(x)|$$

$$= |v(x) \int_0^1 [f_s(x,u(x)+\zeta v(x)) - f_s(x,u(x))] \mathrm{d}\zeta| = |v(x)\, w(x)|,$$

where

$$w(x) = \int_0^1 [f_s(x,u(x)+\zeta v(x)) - f_s(x,u(x))] \mathrm{d}\zeta.$$

With this notation and using the Hölder inequality we get that

$$\omega(u,v) = \left[\int_\Omega |v(x)w(x)|^{p'} \mathrm{d}x \right]^{1/p'}$$

$$\le \|v\|_{L^p} \|w\|_{L^r} \quad \left(r = \frac{p}{p-2} \right). \tag{2.8}$$

Now, the norm $\|w\|_{L^r}$ can be estimated as follows:

$$\|w\|^r_{L^r} \le \int_\Omega \mathrm{d}x \int_0^1 |f_s(x,u(x)+\zeta v(x)) - f_s(x,u(x))|^r \mathrm{d}\zeta$$

$$= \int_0^1 \mathrm{d}\zeta \int_\Omega |f_s(x,u(x)+\zeta v(x)) - f_s(x,u(x))|^r \mathrm{d}x$$

$$= \int_0^1 \|f_s(u+\zeta v) - f_s(u)\|^r_r \mathrm{d}\zeta. \tag{2.9}$$

As remarked before, f_s is continuous from L^p to L^r. Hence

$$\|f_s(u+\zeta v) - f_s(u)\|^r_{L^r} \to 0 \quad \text{as } \|v\|_{L^p} \to 0,\ \zeta \in [0,1]. \tag{2.10}$$

From (2.8),(2.9) and (2.10) it follows that $\omega(u,v) = o(\|v\|_{L^p})$.

For $p = 2$ the above result does not hold, in general. Indeed, under

the preceding assumptions, the Nemitski operator f is G-differentiable but, possibly, not Fréchet-differentiable. To be precise, let us assume that (C) holds for f, f_s and

$$|f_s(x,s)| \le \text{const.} \tag{2.11}$$

As before, it follows plainly that f is continuous from L^2 to L^2 and the map $v \to f_s(u)v$ from L^2 to L^2 is linear and bounded. In addition one has the following

Theorem 2.7 *Let $\Omega \subset \mathbb{R}^n$ be bounded and let f and f_s satisfy* (C) *and* (2.11). *Then $f : L^2 \to L^2$ is G-differentiable and $\mathrm{d}_G f(u)[v] = f_s(u)v$.*

Proof. According to Definition 1.7 we have to show that for all $u, v \in L^2$ there results

$$\left\| \frac{f(u+tv) - f(u)}{t} - f_s(u)v \right\|_{L^2} \to 0 \text{ as } t \to 0. \tag{2.12}$$

As in the proof of theorem 2.6 one finds

$$\frac{f(u+tv) - f(u)}{t} - f_s(u)v = v \int_0^1 [f_s(u+\zeta tv) - f_s(u)] \mathrm{d}\zeta.$$

Letting

$$w_t = w_t(u,v) = \int_0^1 [f_s(u+\zeta tv) - f_s(u)] \mathrm{d}\zeta,$$

one has

$$\left\| \frac{f(u+tv) - f(u)}{t} - f_s(u)v \right\|^2_{L^2} = \int_\Omega v^2 w_t^2$$

$$\le \int_\Omega v^2 \mathrm{d}x \int_0^1 |f_s(u+\zeta tv) - f_s(u)|^2 \mathrm{d}\zeta.$$

When $t \to 0$ then $t\zeta v \to 0$ a.e. in Ω and hence

$$f_s(u+t\zeta v) - f_s(u) \to 0 \text{ a.e. in } \Omega.$$

Since

$$|f_s(x,u(x)+t\zeta v(x)) - f_s(x,u(x))|^2 \le \text{const.}$$

the Lebesgue Dominated-Convergence Theorem implies

$$\int_0^1 |f_s(u+\zeta tv) - f_s(u)|^2 \mathrm{d}\zeta \to 0, \text{ as } t \to 0, \tag{2.13}$$

and (2.12) follows.

The preceding theorem is completed by the following proposition.

Proposition 2.8 *Let $\Omega \subset \mathbb{R}^n$ be bounded, suppose f and f_s satisfy* (C) *and* (2.11) *and let f be Fréchet-differentiable at some $u^* \in L^2$. Then there exists $a(x), b(x) \in M(\Omega)$ such that*

$$f(x,u) = a(x)u + b(x).$$

Proof. Suppose first that $u^* = 0$ and $f(x,0) = 0$.

Let $D(y,\delta)$ denote the ball centred at $y \in \Omega$ with measure δ. Given $x^* \in \Omega$ and $\lambda \in \mathbb{R}$, consider the functions $v_\delta(x) \in L^2(\Omega)$ given by

$$v_\delta(x) = \lambda, \text{ for } x \in D(x^*,\delta),$$
$$v_\delta(x) = 0, \text{ for } x \in \Omega \backslash D(x^*,\delta).$$

Obviously $v_\delta \to 0$ in L^2 as $\delta \to 0$.

Recall that if f is Fréchet-differentiable at $u^* = 0$ then $f'(0) = \mathrm{d}_G f(0)$. Hence $f'(0)v = f_s(0)v$, and therefore

$$\frac{\|f(v_\delta) - f_s(0)v_\delta\|_{L^2}}{\|v_\delta\|_{L^2}} \to 0, \text{ as } \delta \to 0.$$

By a direct calculation one finds

$$\frac{\|f(v_\delta) - f_s(0)v_\delta\|_{L^2}}{\|v_\delta\|_{L^2}} = \frac{1}{|\lambda|\sqrt{\delta}}\left[\int_{D(x^*,\delta)} |f(x,\lambda) - f_s(x,0)\lambda|^2\right]^{1/2}$$

and hence

$$\frac{1}{|\lambda|\sqrt{\delta}}\left[\int_{D(x^*,\delta)} |f(x,\lambda) - f_s(x,0)\lambda|^2\right]^{1/2} \to 0, \text{ as } \delta \to 0. \tag{2.14}$$

If we set

$$\psi_\lambda(x) = \left|\frac{f(x,\lambda) - f_s(x,0)\lambda}{\lambda}\right|^2,$$

(2.14) becomes

$$\frac{1}{\delta}\int_{D(x^*,\delta)} \psi_\lambda(x)\mathrm{d}x \to 0, \text{ as } \delta \to 0. \tag{2.15}$$

Since for all $\psi \in L^1$ and almost every $y \in \Omega$ there results

$$\frac{1}{\delta}\int_{D(y,\delta)} \psi(x)\mathrm{d}x \to \psi(y), \text{ as } \delta \to 0,$$

we deduce from (2.15) that for all $\lambda \in \mathbb{R}$ there exists a null set N_λ such that for all $y \notin N_\lambda$ one has $\psi_\lambda(y) = 0$.

Taking λ in a countable dense subset Λ of $\mathbb{R}$ and letting $N = \cup_{\lambda\in\Lambda} N_\lambda$ one infers that for all $y \notin N$ and all $\lambda \in \Lambda$ there results $\psi_\lambda(y) = 0$ that is,

$$f(y,\lambda) = f_s(y,0)\lambda. \tag{2.16}$$

Using the continuity of $f(x,.)$, one deduces that (2.16) holds true for all $\lambda \in \mathbb{R}$ and almost every y.

Lastly, let us set

$$g(x,u) = f(x,u+u^*) - f(u^*).$$

One has that g is Fréchet-differentiable at 0 and $g(x,0) = 0$. Applying the preceding arguments to g, we find

$$f(x,u+u^*) - f(u^*) = f_s(x,u^*)u$$

and the proposition follows.

Potential operators

We end this section by dealing with potential operators (Definition 1.6). The results will not be used in the remainder of this book but are important in connection with variational problems.

Let $H_0^1 = H_0^1(\Omega)$ (where Ω is a bounded domain of $\mathbb{R}^n$) denote the usual Sobolev space (see Subsection 0.5) with scalar product $(.|.)_{H^{1,2}}$ and norm $\|.\|_{H^{1,2}}$. Let $n > 2$ (if $n = 1, 2$, the arguments we are going to expand apply as well, with some modifications; see Remark 2.10 below).

Suppose f satisfies (C) and

$$|f(x,s)| \le a + b|s|^\sigma \quad \text{with } \sigma \le \frac{n+2}{n-2} = 2^* - 1, \tag{2.17}$$

Here $2^* = 2n/(n-2)$ and, by the Sobolev Embedding Theorem (see Theorem 0.3 (i)), $H_0^1 \hookrightarrow L^{2^*}$ and

$$\|v\|_{L^{2^*}} \le \text{const.}\|v\|_{H^{1,2}}$$

By Theorem 2.2 it follows that

$$f \in C(L^{2^*}, L^q) \text{ with } q = \frac{2^*}{\sigma} \ge \frac{2n}{n+2}. \tag{2.18}$$

In particular one has $f(u) \in L^{2n/(n+2)}$ for all $u \in H_0^1$. As a consequence, $f(u)v \in L^1$ for all $u, v \in H_0^1$ and the equality

$$(N(u)|v)_{H^{1,2}} = \int_\Omega f(x,u(x))v(x)\mathrm{d}x, \quad u, v \in H_0^1, \tag{2.19}$$

defines an operator $N : H_0^1 \to H_0^1$. Note that N is continuous. To see this we evaluate

$$\|N(u) - N(v)\|_{H^{1,2}} = \sup\left\{ \left|\int_\Omega [f(x,u) - f(x,v)]w\mathrm{d}x\right| : \|w\|_{H^{1,2}} \le 1\right\}$$

$$\le \sup\{\|f(u)-f(v)\|_{L^{2n/(n+2)}}\|w\|_{L^{2^*}} : \|w\|_{H^{1,2}} \le 1\}$$
$$\le c\|f(u)-f(v)\|_{L^{2n/(n+2)}}.$$

If $u_m \to v$ in H_0^1 one has (by the Sobolev Embedding Theorem) $u_m \to v$ in L^{2^*} and (by (2.18)) $f(u_m) \to f(v)$ in $L^{2n/(n+2)}$.

Set

$$F(x,s) = \int_0^s f(x,t)\mathrm{d}t.$$

From (2.17) it follows that there exist $c, d > 0$ such that

$$|F(x,s)| \le c + d|s|^{2^*}. \tag{2.20}$$

Then $F(.,u(.)) \in L^1$ for all $u \in H_0^1(\subset L^{2^*})$ and it makes sense to define a functional $\phi : H_0^1 \to \mathbb{R}$ by setting

$$\phi(u) = \int_\Omega F(x,u(x))dx.$$

The functional ϕ can be obtained by composition according to the following diagram

$$\begin{array}{ccccccccc} H_0^1 & \xrightarrow{\alpha} & L^{2^*} & \xrightarrow{F} & L^{2n/(n+2)} & \xrightarrow{\beta} & L^1 & \to & \mathbb{R} \\ u & \to & u & \to & F(.,u(.)) & \to & F(.,u(.)) & \to & \int_\Omega F(x,u(x))\mathrm{d}x \end{array}$$

where α and β stand for the embedding of H_0^1 into L^{2^*}, of $L^{2n/(n+2)}$ into L^1 respectively. Since F satisfies (2.20), Theorem 2.6 applies to F as a map from L^{2^*} to $L^{2n/(n+2)}$. From Proposition 1.4(b) (derivative of the composite map) it follows that ϕ is differentiable with

$$\phi'(u) : v \longrightarrow \int_\Omega f(x,u(x))v(x)\mathrm{d}x.$$

Then, recalling the definition of "gradient" (Definition 1.6), we have the following

Theorem 2.9 *Let $\Omega \subset \mathbb{R}^n$ be bounded and suppose f satisfies* (C) *and* (2.17). *Then ϕ is a C^1 functional on H_0^1 with gradient*

$$\nabla\phi(u) = N(u),$$

where N is defined in (2.19).

Remark 2.10 If $n = 2$ then the same result holds assuming f satisfies (2.17) with any $\sigma < \infty$. It suffices to repeat the above arguments using the stronger form of the Sobolev Embeddings when $\Omega \subset \mathbb{R}^2$.

3 Higher derivatives

Let $F \in C(U,Y)$ be differentiable in the open set $U \subset X$ and consider $F' : U \to L(X,Y)$

Definition 3.1 Let $u^* \in U$: F is *twice (Fréchet-) differentiable* at u^* if F' is differentiable at u^*. The *second (Fréchet) differential* of F at u^* is defined as

$$\mathrm{d}^2F(u^*) = \mathrm{d}F'(u^*).$$

If F is twice differentiable at all points of U we say that F is twice differentiable in U.

According to the above definition $\mathrm{d}^2F(u^*)$ is a linear continuous map from X to $L(X,Y)$:

$$\mathrm{d}^2F(u^*) \in L(X, L(X,Y)).$$

It is convenient to see $\mathrm{d}^2F(u^*)$ as a bilinear map on X. For this, let $L_2(X,Y)$ denote the space of continuous bilinear maps from $X \times X \to Y$. To any $A \in L(X, L(X,Y))$ we can associate $\Phi_A \in L_2(X,Y)$ given by $\Phi_A(u_1,u_2) = [A(u_1)](u_2)$. Conversely, given $\Phi \in L_2(X,Y)$ and $h \in X$, $\Phi(h,.) : k \to \Phi(h,k)$ is a continuous linear map from X to Y; hence to any $\Phi \in L_2(X,Y)$ is associated the linear application $X \to L(X,Y)$,

$$\tilde{\Phi} : h \to \Phi(h,.) \in L(X,Y).$$

It is easy to see that in this way we define an isomorphism between $L(X, L(X,Y))$ and $L_2(X,Y)$. Actually, such an isomorphism is an isometry because there results

$$\|\tilde{\Phi}\|_{L(X,L(X,Y))} = \sup_{\|h\|\le 1} \|\tilde{\Phi}(h)\|_{L(X,Y)}$$
$$= \sup_{\|h\|\le 1} \sup_{\|k\|\le 1} \|\Phi(h,k)\| = \|\Phi\|_{L_2(X,Y)}.$$

In the following we will use the same symbol $\mathrm{d}^2F(u^*)$ to denote the continuous bilinear map obtained by the preceding isometry. The value of $\mathrm{d}^2F(u^*)$ at a pair (h,k) will be denoted by

$$\mathrm{d}^2F(u^*)[h,k].$$

If F is twice differentiable in U, the *second (Fréchet) derivative* of F is the map $F'' : U \to L_2(X,Y)$,

$$F'' : u \to \mathrm{d}^2F(u).$$

If F'' is continuous from U to $L_2(X,Y)$ we say that $F \in C^2(U,Y)$.

Examples 3.2

(i) If $A \in L(X,Y)$ then $A \in C^2(X,Y)$ and $\mathrm{d}^2A[h,k] = 0$ for all $(h,k) \in X \times X$.
(ii) Let $X = C([0,1])$ and $F : X \to X$, $F : u(t) \to u^2(t)$. $F \in C^2(X,X)$ and

$$\mathrm{d}^2F(u) : (h(t),k(t)) \to 2h(t)k(t).$$

The following proposition can be useful for evaluating $\mathrm{d}^2F(u)$.

Proposition 3.3 *Let $F : U \to Y$ be twice differentiable at $u^* \in U$. Then for all fixed $h \in X$ the map $F_h : X \to Y$ defined by setting*

$$F_h(u) = \mathrm{d}F(u)h$$

is differentiable at u^* and $\mathrm{d}F_h(u^*)k = F''(u^*)[h,k]$.

Proof. F_h is obtained by composition

$$U \xrightarrow{\mathrm{d}F} L(X,Y) \xrightarrow{\mathcal{E}_h} Y$$
$$u \longrightarrow \mathrm{d}F(u) \longrightarrow \mathrm{d}F(u)h$$

between the derivative $u \to \mathrm{d}F(u)$ and the "evaluation map" $\mathcal{E}_h$ which associates to each $A \in L(X,Y)$ the value $A(h) \in Y$. Since $\mathcal{E}_h$ is linear, the result follows by the composite mapping formula 1.4-(ii).

We have seen that $F''(u)$ can be regarded as a bilinear map. More precisely one has the following

Theorem 3.4 *If $F : U \to Y$ is twice differentiable at $u \in U$, then $F''(u) \in L_2(X,Y)$ is symmetric.*

Proof. For $h,k \in X$ with $h,k \in B(\varepsilon)$ (ε small enough), we set

$$\psi(h,k) = F(u+h+k) - F(u+h) - F(u+k) + F(u),$$
$$\gamma_h(\xi) = F(u+h+\xi) - F(u+\xi),$$

and consider, for h fixed, the map $g_h : B(\varepsilon) \to Y$,

$$g_h : k \to \psi(h,k) - F''(u)[h,k] = \gamma_h(k) - \gamma_h(0) - F''(u)[h,k].$$

Since F is differentiable in U and $F''(u)(h) : k \to F''(u)[h,k]$ is linear (as a map from X to $L(X,Y)$), Theorem 1.8 yields

$$\begin{aligned}&\|\psi(h,k) - F''(u)[h,k]\|\\ &\le \sup\{\|\mathrm{d}\gamma_h(tk) - F''(u)(h)\| : 0 \le t \le 1\}\|\,k\|\\ &= \sup\{\|\mathrm{d}F(u+h+tk) - \mathrm{d}F(u+tk)\\ &\quad - F''(u)(h)\| : 0 \le t \le 1\}\|k\|. \qquad (3.1)\end{aligned}$$

Since F is twice differentiable at $u \in U$, one has

$$F'(u+h+tk) = F'(u) + F''(u)(h+tk) + \omega(h+tk),$$
$$F'(u+tk) = F'(u) + F''(u)(tk) + \omega(tk),$$

with $\omega(v) = o(\|v\|)$. Hence

$$F'(u+h+tk) - F'(u+tk) = F''(u)(h) + \omega(h+tk) - \omega(tk). \qquad (3.2)$$

Using (3.1) and (3.2) and taking into account that $\omega(v) = o(\|v\|)$ we get, that

$$\begin{aligned}\|\psi(h,k) - F''(u)[h,k]\| &\le \sup\{\|\omega(h+tk) - \omega(tk)\| : 0 \le t \le 1\}\|k\|\\ &\le \varepsilon(\|h\| + 2\|k\|)\,\|k\|, \qquad (3.3)\end{aligned}$$

provided $\|h\|$ and $\|k\|$ are sufficiently small.

Exchanging the roles of h,k we get (for $\|h\|,\|k\|$ small)

$$\begin{aligned}\|\psi(k,h) - F''(u)[k,h]\| &\le \sup\{\|\omega(k+th) - \omega(th)\| : 0 \le t \le 1\}\ \|h\|\\ &\le \varepsilon(\|k\| + 2\|h\|)\,\|h\|. \qquad (3.4)\end{aligned}$$

Since $\psi(h,k) = \psi(k,h)$ we deduce from (3.3) and (3.4)

$$\begin{aligned}\|F''(u)[h,k] - F''(u)[k,h]\| &\le \varepsilon(2\|k\|^2 + 2\|h\|^2 + 2\|h\|\,\|k\|)\\ &\le 3\,\varepsilon(\|k\|^2 + \|h\|^2). \qquad (3.5)\end{aligned}$$

Inequality (3.5) has been proved for $\|h\|,\|k\|$ small enough, but holds true for all $\|h\|,\|k\|$, because $F''(u)[h,k]$ is homogeneous of degree 2. Since ε is arbitrary, (3.5) implies that $F''(u)[h,k] = F''(u)[k,h]$ for all h,k.

To define $(n+1)$-th derivatives ($n \ge 2$) we can proceed by induction.

Given $F : U \to Y$, let F be n times differentiable in U. The nth differential at a point $x \in U$ will be identified with a continuous n-linear map from $X \times X \times \ldots \times X$ (n times) to Y (recall that, as before, there is an isometry between $L(X,\ldots,L(X,Y))\ldots$ and $L_n(X,Y)$).

Let $F^{(n)} : U \to L_n(X,Y)$ denote the map

$$F^{(n)} : u \to \mathrm{d}^nF(u).$$

The $(n+1)$-th differential at u^* will be defined as the differential of $F^{(n)}$, namely

$$\mathrm{d}^{(n+1)}F(u^*) = \mathrm{d}F^{(n)}(u^*) \in L(X, L_n(X,Y)) \approx L_{n+1}(X,Y).$$

We will say that $F \in C^n(U,Y)$ if F is n times (Fréchet) differentiable in U and the nth derivative $F^{(n)}$ is continuous from U to $L_n(X,Y)$. The value of $\mathrm{d}^nF(u^*)$ at $(h_1,\ldots,h_n)$ will be denoted by

$$\mathrm{d}^nF(u^*)[h_1,\ldots,h_n].$$

If $h = h_1 = \ldots = h_n$ we will write for short $\mathrm{d}^nF(u^*)[h]^n$.

In order to extend Theorem 3.4 to higher derivatives some preliminaries are in order. Given a map $G : U \to L_m(X,Y)$ and the point $\mathbf{h} = (h_1,\dots,h_m) \in X \times \dots \times X$, we can associate G with the map $G[\mathbf{h}] : U \to Y$ defined by setting

$$G[\mathbf{h}](u) = G(u)[h_1,\dots,h_n].$$

We can immediately see (see Proposition 3.3) that if G is differentiable at u then $G[\mathbf{h}]$ is differentiable at u and there results

$$d(G[\mathbf{h}])(u) : v \to dG(u)[v,h_1,\dots,h_n]. \tag{3.6}$$

Let F be n times differentiable on U and set $\mathbf{h} = (h_2,\dots,h_n)$. Applying (3.6) to $G = d^{n-1}F$, we find that

$$d(d^{n-1}F[\mathbf{h}])(u)[h_1] = d^nF(u^*)[h_1,\dots,h_n]. \tag{3.7}$$

Theorem 3.5 *If $F : U \to Y$ is n times differentiable in U, then the map*

$$(h_1,\dots,h_n) \to d^nF(u^*)[h_1,\dots,h_n]$$

is symmetric.

Proof. The result is true for $n = 2$ (Theorem 3.4). By induction on n, let the claim hold for $n-1 \geq 2$. Then

$$d^{n-1}F(u)[h_2,\dots,h_i,\dots,h_j,\dots,h_n]$$
$$= d^{n-1}F(u)[h_2,\dots,h_j,\dots,h_i,\dots,h_n].$$

Applying (3.7) to $h(u) = d^{n-1}F(u)[h_2,\dots,h_i,\dots,h_j,\dots,h_n]$ we get that

$$d^nF(u^*)[h_1,h_2,\dots,h_i,\dots,h_j,\dots,h_n]$$
$$= d^nF(u^*)[h_1,h_2,\dots,h_j,\dots,h_i,\dots,h_n]. \tag{3.8}$$

Similarly, letting $G(u) = d^{n-2}F(u)[h_3,\dots,h_n]$, one has

$$d^2G(u^*)[h_1,h_2] = d^nF(u^*)[h_1,h_2,h_3,\dots,h_n],$$

and from Theorem 3.4 it follows that

$$d^nF(u^*)[h_1,h_2,h_3,\dots,h_n] = d^2G(u^*)[h_1,h_2]$$
$$= d^2G(u^*)[h_2,h_1] = d^nF(u^*)[h_2,h_1,h_3,\dots,h_n]. \tag{3.9}$$

The symmetry of $d^nF(u^*)$ is an immediate consequence of (3.8) and (3.9).

4 Partial derivatives, Taylor's formula

Let us consider two Banach spaces X, Y and let $(u^*,v^*) \in X \times Y$. Define

mappings $\sigma_{v^*} : X \to X \times Y$ and $\tau_{u^*} : Y \to X \times Y$ as follows.

$$\sigma_{v^*}(u) = (u,v^*);$$
$$\tau_{u^*}(y) = (u^*,v).$$

Notice that the derivatives of σ_{v^*} and τ_{u^*} are respectively, the linear maps

$$\sigma := d\sigma_{v^*} : h \to (h,0),$$
$$\tau := d\tau_{u^*} : k \to (0,k).$$

Let Q be an open subset of $X \times Y$, $(u^*,v^*) \in Q$ and $F : Q \to Z$.

Definition 4.1 If the map $F \circ \sigma_{v^*}$ is differentiable at u^* we say that F is *differentiable with respect to u* at (u^*,v^*). The linear map $d[F \circ \sigma_{v^*}](u^*) \in L(X,Z)$ is called the *partial derivative* of F at (u^*,v^*) with respect to u and denoted by $d_uF(u^*,v^*)$.

Similarly, if $F \circ \tau_{u^*}$ is differentiable at v^* we say that F is *differentiable with respect to v* at (u^*,v^*) and the linear map $d[F \circ \tau_{u^*}](v^*) \in L(Y,Z)$ is called the *v-partial derivative* of F at (u^*,v^*) and denoted by $d_vF(u^*,v^*)$.

The preceding definition is equivalent to requiring that there exist a linear map $A_u \in L(X,Z)$ (resp. $A_v \in L(Y,Z)$), such that

$$F(u^*+h,v^*) - F(u^*,v^*) = A_u(h) + o(\|h\|),$$
$$F(u^*,v^*+k) - F(u^*,v^*) = A_v(k) + o(\|k\|).$$

The following result is an immediate consequence of Definition 4.1 and the differentation rule 1.4 (ii).

Proposition 4.2 *If F is differentiable at (u^*,v^*) then F has partial derivatives with respect to u and v at (u^*,v^*) and we have*

$$d_uF(u^*,v^*)(h) = dF(u^*,v^*)\sigma(h) = dF(u^*,v^*)(h,0),$$
$$d_vF(u^*,v^*)(k) = dF(u^*,v^*)\tau(k) = dF(u^*,v^*)(0,k).$$

In quite similar way one can define higher partial derivatives. For example, if F has u-partial derivative at all $(u,v) \in Q$, we can define the map $F_u : Q \to L(X,Z)$ by setting

$$F_u(u,v) = d_uF(u,v).$$

Then the partial derivative $d_{u,v}F(u^*,v^*)$ is the v-derivative at (u^*,v^*) of F_u, namely

$$d_{u,v}F(u^*,v^*) = d_v[F_u](u^*,v^*).$$

The map $F_{u,v} : Q \to L(Y, L(X,Z))$ will be defined by setting

$$F_{u,v}(u,v) = d_{u,v}F(u,v).$$

Moreover, if F is twice differentiable at (u^*, v^*), then $\mathrm{d}_{u,v}F(u^*,v^*)$ is the bilinear map from $X \times Y$ to Z given by

$$(h,k) \to F''(u^*,v^*)[\sigma h, \tau k]. \tag{4.1}$$

The notation $\mathrm{d}^m_{u^\ell, v^{m-\ell}}$ will be employed to indicate

$$\mathrm{d}^m_{u^\ell, v^{m-\ell}} = \mathrm{d}^\ell_{u^\ell}\left(\mathrm{d}^{m-\ell}_{v^{m-\ell}}\right).$$

The definition of partial derivative given above permits us to obtain in a rather straightforward way all the classical results of calculus.

For example one can prove the following.

Theorem 4.3 *Suppose that*

(i) *F has u- and v-derivatives in a neighbourhood N of $(u^*, v^*) \in Q$,*
(ii) *F_u and F_v are continuous in N.*
Then F is differentiable at (u^*, v^*).

As another example, we can use (4.1) and Theorem 3.4 to show

$$\begin{aligned}\mathrm{d}_{u,v}F(u^*,v^*)[h,k] &= F''(u^*,v^*)[\sigma h, \tau k] \\ &= F''(u^*,v^*)[\tau k, \sigma h] = \mathrm{d}_{v,u}F(u^*,v^*)[k,h],\end{aligned}$$

which is nothing else than the classical Schwarz Theorem.

Taylor's formula

Let $F \in C^n(Q,Y)$ and let $u, u+v \in Q$ be such that the interval $[u, u+v] \subset Q$.

Set $\gamma(t) = u + tv$, $t \in [0,1]$ and let $\phi : [0,1] \to Y$ be defined by

$$\phi(t) = F(\gamma(t)).$$

Using Proposition 1.4 (ii) and (3.7) it follows readily that the function ϕ is C^n and there result

$$\begin{aligned}\phi'(t) &= \mathrm{d}F(u+tv)[v], \\ \phi''(t) &= \mathrm{d}^2F(u+tv)[v]^2, \\ &\dots\dots\dots\dots \\ \phi^{(n)}(t) &= \mathrm{d}^nF(u+tv)[v]^n.\end{aligned}$$

By elementary calculations one has

$$\phi(1) = \phi(0) + \phi'(0) + \frac{1}{2!}\phi''(0) + \cdots + \frac{1}{(n-1)!}\phi^{(n)}(0)$$

$$+ \frac{1}{(n-1)!}\int_0^1 (1-t)^{n-1}\phi^{(n)}(t)\mathrm{d}t,$$

and hence

$$F(u+v) = F(u) + \mathrm{d}F(u)[v] + \cdots$$

$$+ \frac{1}{(n-1)!}\int_0^1 (1-t)^{n-1}\mathrm{d}^{(n)}F(u+tv)[v]^n\mathrm{d}t.$$

The last integral can be written as

$$\frac{1}{(n-1)!}\int_0^1 (1-t)^{n-1}\mathrm{d}^{(n)}F(u+tv)\mathrm{d}t[v]^n$$

$$= \frac{1}{n!}\mathrm{d}^nF(u)[v]^n + \varepsilon(u,v)[v]^n, \tag{4.2}$$

where

$$\varepsilon(u,v) = \frac{1}{(n-1)!}\int_0^1 (1-t)^{n-1}[\mathrm{d}^{(n)}F(u+tv) - \mathrm{d}^{(n)}F(u)]\mathrm{d}t \to 0 \text{ as } v \to 0.$$

Lastly, let us write explicitly the form of (4.2) when $F = F(u,v)$ is defined on $Q \subset X \times Y$ with values in Z and is C^n, that is, has continuous partial derivatives up to order n. We write (u,v) instead of u and set $w = (h,k) = \sigma h + \tau k$. If we use Proposition 4.2 the mth term in (4.2) becomes

$$\begin{aligned}\frac{1}{m!}\mathrm{d}^{(m)}F(u,v)[w]^m &= \frac{1}{m!}\mathrm{d}^{(m)}F(u,v)[\sigma h + \tau k]^m \\ &= \frac{1}{m!}\mathrm{d}^{(m)}F(u,v)\sum\binom{m}{\ell}[\sigma h]^\ell[\tau k]^{m-\ell} \\ &= \frac{1}{m!}\sum\binom{m}{\ell}\mathrm{d}^{(m)}F(u,v)[\sigma h]^\ell[\tau k]^{m-\ell} \\ &= \frac{1}{m!}\sum\binom{m}{\ell}\mathrm{d}^m_{u^\ell,v^{m-\ell}}F(u,v)[h]^\ell[k]^{m-\ell}.\end{aligned}$$

Remark 4.4 (on notation) Hereafter we will often deal with maps $F : \mathbb{R} \times X \to Y$ depending on a real parameter λ. In such a case the mixed derivative $F_{u,\lambda}(\lambda_o, u_o)$ is a linear map from $\mathbb{R}$ to $L(X,Y)$: $F_{u,\lambda}(\lambda_o, u_o) \in L(\mathbb{R}, L(X,Y))$. Then, in accordance with what we remarked in Example 1.3 (c), we can and will identify $F_{u,\lambda}(\lambda_o, u_o)$ with the linear map $h \to F_{u,\lambda}(\lambda_o, u_o)[h, 1]$.