TEXTS AND READINGS IN MATHEMATICS 28

Nonlinear Functional Analysis
A First Course

Already Published Volumes

R. B. Bapat: Linear Algebra and Linear Models (Second Edition)
Rajendra Bhatia: Fourier Series (Second Edition)
C. Musili: Representations of Finite Groups
H. Helson: Linear Algebra (Second Edition)
D. Sarason: Notes on Complex Function Theory
M. G. Nadkarni: Basic Ergodic Theory (Second Edition)
H. Helson: Harmonic Analysis (Second Edition)
K. Chandrasekharan: A Course on Integration Theory
K. Chandrasekharan: A Course on Topological Groups
R. Bhatia (ed.): Analysis, Geometry and Probability
K. R. Davidson: C* - Algebras by Example
M. Bhattacharjee *et al.*: Notes on Infinite Permutation Groups
V. S. Sunder: Functional Analysis - Spectral Theory
V. S. Varadarajan: Algebra in Ancient and Modern Times
M. G. Nadkarni: Spectral Theory of Dynamical Systems
A. Borel: Semisimple Groups and Riemannian Symmetric Spaces
M. Marcolli: Seiberg-Witten Gauge Theory
A. Bottcher and S. M. Grudsky: Toeplitz Matrices, Asymptotic Linear Algebra and Functional Analysis
A. R. Rao and P. Bhimasankaram: Linear Algebra (Second Edition)
C. Musili: Algebraic Geometry for Beginners
A. R. Rajwade: Convex Polyhedra with Regularity Conditions and Hilbert's Third Problem
S. Kumaresan: A Course in Differential Geometry and Lie Groups
Stef Tijs: Introduction to Game Theory
B. Sury: The Congruence Subgroup Problem
R. Bhatia (ed.): Connected at Infinity
K. Mukherjea: Differential Calculus in Normed Linear Spaces
Satya Deo: Algebraic Topology: A Primer

Nonlinear Functional Analysis
A First Course

S. Kesavan
Institute of Mathematical Sciences
Chennai

HINDUSTAN BOOK AGENCY

Published by

Hindustan Book Agency (India)
P 19 Green Park Extension
New Delhi 110 016
India

email: hba@vsnl.com
http://www.hindbook.com

Produced from camera ready copy supplied by the Author.

ISBN 81-85931-46-1

Preface

Nonlinear Functional Analysis studies the properties of (continuous) mappings between normed linear spaces and evolves methods to solve nonlinear equations involving such mappings. Two major approaches to the solution of nonlinear equations could be described as *topological* and *variational*. Topological methods are derived from fixed point theorems and are usually based on the notion of the topological degree. Variational methods describe the solutions as critical points of a suitable functional and study ways of locating them.

The aim of this book is to present the basic theory of these methods. It is meant to be a primer of nonlinear analysis and is designed to be used as a text or reference book by students at the masters or doctoral level in Indian universities. The prerequisite for following this book is knowledge of functional analysis and topology, usually part of the curriculum at the masters level in most universities in India.

The first chapter covers the preliminaries needed from the differential calculus in normed linear spaces. It introduces the notion of the Fréchet derivative, which generalizes the notion of the derivative of a real valued function of a single real variable. Some classical theorems which are repeatedly used in the sequel, like the implicit function theorem and Sard's theorem, are proved here.

The second chapter develops the theory of the topological degree in finite dimensions. The Brouwer fixed point theorem and Borsuk's theorem are proved and some of their applications are presented.

The next chapter extends the notion of the topological degree to infinite dimensional spaces for a special class of mappings known as compact perturbations of the identity. Again, fixed point theorems (in particular, Schauder's theorem) are proved and applications are given.

The fourth chapter deals with bifurcation theory. This studies the nature of the set of solutions to equations dependent on a parameter, in the neighbourhood of a 'trivial solution'. Science and engineering are full of instances of such problems. A variety of methods for the identification of bifurcation points - topological and variational - are presented.

The concluding chapter deals with the existence and multiplicity of critical points of functionals defined on Banach spaces. While minimization is one method, other critical points, like saddle points are found by using results like the mountain pass theorem, or, more generally, what are known as min - max theorems.

Nonlinear Analysis, today, has a bewildering array of tools. In selecting the above topics, a conscious choice has been made with the following objectives in mind:

- to provide a text book which can be used for an introductory one - semester course covering classical material;
- to be of interest to a *general* student of higher mathematics.

The examples and exercises that are found throughout the text have been chosen to be in tune with these objectives (though, from time to time, my own bias towards differential equations does show up.)

It is for this reason that some of the tools developed more recently have been (regrettably) omitted. Two examples spring to ones mind. The first is the *method of concentration compactness* (which won the Fields Medal for P. L. Lions). It deals with the convergence of sequences in Sobolev spaces. Its main application is in the study of minimizing sequences for functionals associated to some semilinear elliptic partial differential equations into which

a certain 'lack of compactness' has been built. Another instance is the theory of Γ - *convergence*. This theory studies the convergence of the minima and minimizers of a family of functionals. Again, while the theory can be developed in the very general context of a topological space, alot of technical results in Sobolev spaces are needed in order to present reasonably interesting results. The applications of this theory are myriad, ranging from nonlinear elasticity to homogenization theory. Such topics, in my opinion, would be ideal for a sequel to this volume, meant for an advanced course on nonlinear analysis, specifically aimed at students working in applications of mathematics.

The material presented here is classical and no claim is made towards originality of presentation (except for some of my own work included in Chapter 4). My treatment of the subject has been greatly influenced by the works of Cartan [4], Deimling [7], Kavian [11], Nirenberg [19] and Rabinowitz [20].

This book grew out of the notes prepared for courses that I gave on various occasions to doctoral students at the TIFR Centre, Bangalore, India (where I worked earlier), the Dipartimento di Matematica G. Castelnuovo, Università degli Studi di Roma "La Sapienza", Rome, Italy and the Laboratoire MMAS, Université de Metz, Metz, France. I would like to take this oppurtunity to thank these institutions for their facilities and hospitality.

I would like to thank the Institute of Mathematical Sciences, Chennai, India, for its excellent facilities and research environment which permitted me to bring out this book. I also thank the publishers, Hindustan Book Agency and the Managing Editor of their TRIM Series, Prof. Rajendra Bhatia, for their cooperation and support. I also thank the two anonymous referees who read through the entire manuscript and made several helpful suggestions which led to the improvement of this text. At a personal level, I would like to thank my friend and erstwhile colleague, Prof.

V. S. Borkar, who egged me on to give the lectures at Bangalore in the first place and kept insisting that I publish the notes. I also wish to thank one of my summer students, Mr. Shivanand Dwivedi, who, while learning the material from the manuscript, also did valuable proof reading. Finally, for numerous personal reasons, I thank the members of my family and fondly dedicate this book to them.

Chennai.
October, 2003.

S. Kesavan

Contents

1 Differential Calculus on Normed Linear Spaces **1**
- 1.1 The Fréchet Derivative 1
- 1.2 Higher Order Derivatives 16
- 1.3 Some Important Theorems 22
- 1.4 Extrema of Real Valued Functions 29

2 The Brouwer Degree **36**
- 2.1 Definition of the Degree 36
- 2.2 Properties of the Degree 45
- 2.3 Brouwer's Theorem and Applications 50
- 2.4 Borsuk's Theorem 57
- 2.5 The Genus 63

3 The Leray - Schauder Degree **69**
- 3.1 Preliminaries 69
- 3.2 Definition of the Degree 72
- 3.3 Properties of the Degree 74
- 3.4 Fixed Point Theorems 78
- 3.5 The Index 82
- 3.6 An Application to Differential Equations 87

4 Bifurcation Theory **92**
- 4.1 Introduction 92
- 4.2 The Lyapunov - Schmidt Method 97
- 4.3 Morse's Lemma 99
- 4.4 A Perturbation Method 107

4.5 Krasnoselsk'ii's Theorem 111
4.6 Rabinowitz' Theorem 113
4.7 A Variational Method 117

5 Critical Points of Functionals **125**
5.1 Minimization of Functionals 125
5.2 Saddle Points 132
5.3 The Palais - Smale Condition 137
5.4 The Deformation Lemma 144
5.5 The Mountain Pass Theorem 150
5.6 Multiplicity of Critical Points 155
5.7 Critical Points with Constraints 159

Bibliography **171**

Index **175**

Chapter 1

Differential Calculus on Normed Linear Spaces

1.1 The Fréchet Derivative

In this chapter we will review some of the important results of the differential calculus on normed linear spaces.

Given a function $f : \mathbb{R} \to \mathbb{R}$ we know what is meant by its derivative (if it exists) at a point $a \in \mathbb{R}$. It is a number denoted by $f'(a)$ (or $Df(a)$ or $\frac{df}{dx}(a)$) such that

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = f'(a) \tag{1.1.1}$$

or, equivalently,

$$|f(a+h) - f(a) - f'(a)h| = o(h) \tag{1.1.2}$$

where, by the symbol $o(h)$ we understand that the right-hand side is equal to a function $\varepsilon(h)$ such that

$$\frac{|\varepsilon(h)|}{|h|} \to 0 \text{ as } |h| \to 0. \tag{1.1.3}$$

If we wish to generalize this notion of the derivative to a function defined in an open set of $\mathbb{R}^n$ or, more generally, to a function defined in an open set of a normed linear space E and taking values in another normed linear space F, it will be convenient to regard $f'(a)h$ as the result of a linear operation on h. Thus, $f'(a)$ is now considered as a bounded linear operator on $\mathbb{R}$ which satisfies (1.1.2). We now define the notion of differentiability for functions defined on a normed linear space.

Let E and F be normed linear spaces (over $\mathbb{R}$).We denote by $\mathcal{L}(E,F)$ the space of bounded linear transformations of E into F.

Definition 1.1.1 *Let $\mathcal{U} \subset E$ be an open set and let $f : \mathcal{U} \to F$ be a given function. The function f is said to be* **differentiable** *at $a \in \mathcal{U}$ if there exists a bounded linear transformation $f'(a) \in \mathcal{L}(E,F)$ such that*

$$\|f(a+h) - f(a) - f'(a)h\| = o(\|h\|). \tag{1.1.4}$$

Equivalently, we can write

$$f(a+h) - f(a) - f'(a)h = \varepsilon(h) \tag{1.1.5}$$

where $\frac{\varepsilon(h)}{\|h\|} \to 0$ as $\|h\| \to 0$.■

Remark 1.1.1 The following facts are simple consequences of the above definition: (i)if f is differentiable at $a \in \mathcal{U}$, then f is continuous at that point; (ii)if f is differentiable at $a \in \mathcal{U}$, then the derivative $f'(a) \in \mathcal{L}(E,F)$ is uniquely defined. It is for the uniqueness of the derivative that it is convenient to assume that the domain of definition is an open set. ■

The derivative defined above is called the Fréchet derivative of f at the point a. We can also define the Gâteau derivative of f at a *along a given vector* $h \in E$ by means of the limit

$$\lim_{t\to 0} \frac{f(a+th) - f(a)}{t}. \tag{1.1.6}$$

Remark 1.1.2 If f is Fréchet differentiable at a point a, then, for every $h \in E$, it is Gâteau differentiable at that point along h and the Gâteau derivative is given by $f'(a)h$. The converse is not true. A function may possess a Gâteau derivative at a point along every direction but can fail to be Fréchet differentiable at

that point. ■

Example 1.1.1 Let $E = \mathbb{R}^2$ and $F = \mathbb{R}$. Define

$$f(x,y) = \begin{cases} \frac{x^5}{(y-x^2)^2+x^4} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

Then if $(x,y) \to (0,0)$ along any direction h (*i.e.* along the line joining h to the origin), we get that the limit in (1.1.6) exists and is equal to zero. Thus, if f were differentiable, $f'(0,0) = 0$. However, if we pass to the same limit along the parabola $y = x^2$, the limit turns out to be unity, which contradicts (1.1.5). Thus, f is not differentiable at the origin even though it possesses a Gâteau derivative at that point along every direction. ■

Definition 1.1.2 *Let $f : \mathcal{U} \subset E \to F$ be a given function. If $f'(a)$ exists for each $a \in \mathcal{U}$, we say that f is differentiable in $\mathcal{U}$. If the mapping $a \longmapsto f'(a)$ is continuous from $\mathcal{U}$ into $\mathcal{L}(E,F)$, we say that f is of class C^1.*■

We now give examples to illustrate the Fréchet derivative.

Example 1.1.2 Let E, F be normed linear spaces and let $C \in \mathcal{L}(E,F)$. For $b \in F$, define

$$f(x) = Cx + b.$$

Then f is differentiable in E and

$$f'(x)h = Ch, \text{ for every } x, h \in E.$$

Thus $f' \equiv C$.■

Example 1.1.3 Let E be a Hilbert space and $a : E \times E \to \mathbb{R}$ a symmetric and continuous bilinear form on E. Let $b \in E$. Define

$$f(x) = \frac{1}{2}a(x,x) - (b,x), \text{ for } x \in E.$$

where (.,.) stands for the inner-product in E. Then

$$f(x+h)-f(x)=a(x,h)+\frac{1}{2}a(h,h)-(b,h).$$

Since $a(.,.)$ is continuous,

$$|a(h,h)|\leq M||h||^2.$$

Hence it follows that f is differentiable in E and that

$$f'(x)h=a(x,h)-(b,h), \text{ for every } x,h\in E. \blacksquare$$

Example 1.1.4 (The Nemytskii Operator) Let $\Omega\subset\mathbb{R}^N$ be a domain and let $f:\Omega\times\mathbb{R}\to\mathbb{R}$ be a given function such that the mapping $x\mapsto f(x,t)$ is measurable for all fixed $t\in\mathbb{R}$ and the mapping $t\mapsto f(x,t)$ is continuous for almost all $x\in\Omega$. Such a function is called a *Carathéodory function*. Let W be a vector space of real - valued functions on Ω. The *Nemytskii operator* associated to f is a nonlinear mapping defined on W by

$$N(u)(x) \;=\; f(x,u(x)).$$

A remarkable theorem, due to Krasnoselsk'ii [14] (see also Joshi and Bose [10] for a proof), is that if $(1/p)+(1/q)=1$, where $1\leq p,q\leq\infty$, and if N maps $L^p(\Omega)$ into $L^q(\Omega)$, then this mapping is continuous and bounded, *i.e.* it maps bounded sets into bounded sets. A typical condition on f would be a growth condition of the type

$$|f(x,t)| \;\leq\; a(x)+b|t|^{p/q}$$

where a is a non-negative function in $L^q(\Omega)$ and b is a positive constant.

Let Ω be a bounded domain and let $p=q=2$. Assume that, in addition, f is in $\Omega\times\mathbb{R}$ and that $\frac{\partial f}{\partial t}(x,u(x))$ is in $L^\infty(\Omega)$ if $u\in L^2(\Omega)$. Thus, the mapping $u\mapsto\frac{\partial u}{\partial t}(.,u(.))$ is, by Krasnoselsk'ii's result, continuous from $L^2(\Omega)$ into itself (since, Ω being bounded, $L^\infty(\Omega)$ is contained in $L^2(\Omega)$). Then, N is also differentiable and if $h\in L^2(\Omega)$, then the function $N'(u)h\in L^2(\Omega)$ is given by

$$N'(u)(h)(x) \;=\; \frac{\partial f}{\partial t}(x,u(x))h(x).$$

To see this, notice that by the classical mean value theorem for functions of several variables, there exists $\theta(x)$ such that $0<\theta(x)<1$ and

$$f(x,u(x)+h(x))-f(x,u(x)) \;=\; \frac{\partial f}{\partial t}(x,u(x)+\theta(x)h(x))h(x).$$

Hence, denoting the norm in $L^2(\Omega)$ by $||.||$, we get

$$\frac{||N(u+h)-N(u)-N'(u)h||}{||h||} \;\leq\; \left|\left|\frac{\partial f}{\partial t}(.,u+\theta h)-\frac{\partial f}{\partial t}(.,u)\right|\right|.$$

By the continuity of the Nemytskii operator associated to $\frac{\partial f}{\partial t}$, it follows that the term on the right tends to zero as $h\to 0$ in $L^2(\Omega)$ and this proves our claim. $\blacksquare$

Example 1.1.5 Let $\Omega\subset\mathbb{R}^N$ be a bounded domain and let $f:\Omega\times\mathbb{R}\to\mathbb{R}$ be function as in the preceding example. Let $H_0^1(\Omega)$ be the usual Sobolev space (cf. Kesavan [13]) of functions in $L^2(\Omega)$ all of whose first derivatives are also in that space and which vanish, in the sense of trace on the boundary $\partial\Omega$. Then, the following problem has a unique (weak) solution (cf. Kesavan [13]):

$$\begin{aligned} -\Delta w(x) &= f(x,u(x)) & x\in\Omega \\ w(x) &= 0 & x\in\partial\Omega. \end{aligned}$$

Since $H_0^1(\Omega)\subset L^2(\Omega)$, we can thus define the mapping $T:L^2(\Omega)\to L^2(\Omega)$ via the relation $T(u)=w$. Let $h\in L^2(\Omega)$. let $z\in H_0^1(\Omega)$ be the unique solution of the problem:

$$\begin{aligned} -\Delta z(x) &= \tfrac{\partial f}{\partial t}(x,u(x))h(x) & x\in\Omega \\ z(x) &= 0 & x\in\partial\Omega. \end{aligned}$$

We claim that T is differentiable and that $T'(u)h = z$. Indeed, if $v = T(u+h)$, then, $\zeta = v - w - z$ vanishes on $\partial\Omega$ and satisfies

$$-\Delta\zeta = f(.,u+h) - f(.,u) - \frac{\partial f}{\partial t}(.,u)h$$

in Ω. By standard estimates, we know that $\|\zeta\|_{L^2(\Omega)}$ is bounded by the norm in $L^2(\Omega)$ of the expression in the right-hand side, which, in turn, is of the order $o(\|h\|_{L^2(\Omega)})$ as seen in the preceding example. This establishes our claim. ■

Exercise 1.1.1 Let $\Omega \subset \mathbb{R}^n$ be a bounded domain, where $n \leq 3$. Then it is known that (cf. Kesavan [13]) $H_0^1(\Omega) \subset L^p(\Omega)$, with continuous inclusion, if $1 \leq p \leq 6$. Let $f \in L^2(\Omega)$ be given. Show that the functional J defined for $v \in H_0^1(\Omega)$ by

$$J(v) = \frac{1}{2}\int_\Omega |\nabla v|^2 dx - \frac{1}{6}\int_\Omega v^6 dx - \int_\Omega fv dx$$

is differentiable and that

$$< J'(u), v > = \int_\Omega \nabla u.\nabla v dx - \int_\Omega u^5 v dx - \int_\Omega fv dx$$

where $< .,. >$ denotes the duality bracket between $H_0^1(\Omega)$ and its dual (denoted by $H^{-1}(\Omega)$). ■

Exercise 1.1.2 Let $M(n,\mathbb{R})$ denote the space of all $n\times n$ matrices with real entries. Let $GL(n,\mathbb{R})$ be the set of all invertible matrices in $M(n,\mathbb{R})$.
(i) Show that $GL(n,\mathbb{R})$ is an open set in $M(n,\mathbb{R})$ (provided with the usual topology of $\mathbb{R}^{n^2}$).
(ii) Show that the mapping $f : GL(n,\mathbb{R}) \to M(n,\mathbb{R})$ defined by $f(A) = A^{-1}$ is differentiable and that

$$f'(A)H = -A^{-1}HA^{-1}.■$$

Exercise 1.1.3 Let E be a normed linear space. Show that the map $f : E \to \mathbb{R}$ defined by $f(x) = \|x\|$ for all $x \in E$ is never

differentiable at the origin. ■

The Fréchet derivative follows the usual rules of the calculus. For instance, if f and g are two functions which are differentiable at a point a and if we define $f+g$ and λf (for $\lambda \in \mathbb{R}$) by

$$(f+g)(x) = f(x) + g(x)\ ,\ (\lambda f)(x) = \lambda f(x),$$

then

$$(f+g)'(a) = f'(a) + g'(a)\ ,\ (\lambda f)'(a) = \lambda f'(a)$$

as can be easily seen. Another important rule relates to the derivative of the composition of two differentiable functions.

Proposition 1.1.1 *Let E, F and G be normed linear spaces, $\mathcal{U}$ an open set in E and $\mathcal{V}$ an open set in F. Let $f : \mathcal{U} \to F, g : \mathcal{V} \to G$ such that for a given point $a \in \mathcal{U}$, we have $f(a) = b \in \mathcal{V}$. On the open set $\mathcal{U}' = f^{-1}(\mathcal{V})$, which contains a, define*

$$h = g \circ f : \mathcal{U}' \to G.$$

If f is differentiable at a and g at b, then h is differentiable at a and

$$h'(a) = g'(f(a)) \circ f'(a). \tag{1.1.7}$$

Proof: We have

$$\begin{aligned} f(x) &= f(a) + f'(a)(x-a) + \varepsilon(x-a), \\ g(y) &= g(b) + g'(b)(y-b) + \eta(y-b), \end{aligned}$$

where $\varepsilon(x-a) = o(\|x-a\|)$ and $\eta(y-b) = o(\|y-b\|)$. Now,

$$\begin{aligned} h(x) - h(a) &= g(f(x)) - g(f(a)) \\ &= g'(f(a))(f(x) - f(a)) + \eta(f(x) - f(a)). \end{aligned}$$

Thus

$$h(x)-h(a) = g'(f(a))f'(a)(x-a)+g'(f(a))\varepsilon(x-a)+\eta(f(x)-f(a)). \tag{1.1.8}$$

But

$$||g'(f(a))\varepsilon(x-a)|| \leq ||g'(f(a))||.||\varepsilon(x-a)|| = o(||x-a||). \quad (1.1.9)$$

Further, if $M > ||f'(a)||$, then, for $||x-a||$ small enough,

$$||f(x) - f(a)|| \leq M||x-a||,$$

and so $||f(x) - f(a)|| \to 0$ as $||x-a|| \to 0$. Thus

$$\frac{||\eta(f(x) - f(a))||}{||x-a||} \leq M\frac{||\eta(f(x) - f(a))||}{||f(x) - f(a)||} \to 0 \text{ as } ||x-a|| \to 0$$

which proves that

$$||\eta(f(x) - f(a))|| = o(||x-a||). \quad (1.1.10)$$

The relations (1.1.8)-(1.1.10) prove (1.1.7). ■

We look at some special situations where E and F are product spaces. Let us assume that $F = F_1 \times ... \times F_m$, the product of normed linear spaces. For $1 \leq i \leq m$, define the projection

$$p_i : F \to F_i,$$

and let $u_i : F_i \to F$ be the injection defined by

$$u_i(x_i) = (0,..,0,x_i,0,..,0),$$

(with 0 everywhere except in the i-th place). Then

$$\begin{cases} p_i \circ u_i = I_{F_i} \\ \sum_{i=1}^m u_i \circ p_i = I_F \end{cases} \quad (1.1.11)$$

(where I_E denotes the identity map in a normed linear space E).

Proposition 1.1.2 *Let $\mathcal{U} \subset E$ be an open set and $f : \mathcal{U} \to F$ be a given map. Then f is differentiable at $a \in \mathcal{U}$ if, and only if, $f_i = p_i \circ f : \mathcal{U} \to F_i$ is differentiable at a for each $i, 1 \leq i \leq m$. In this case,*

$$f'(a) = \sum_{i=1}^m u_i \circ f_i'(a). \quad (1.1.12)$$

Proof: If f is differentiable, so is f_i, since it is the composition of f and a continuous linear map (which is always differentiable; cf. Example 1.1.2).Thus, by Proposition 1.1.1,

$$f_i'(a) = p_i \circ f'(a). \quad (1.1.13)$$

If, conversely, f_i is differentiable for each i, we get, from (1.1.11), that

$$f = \sum_{i=1}^m u_i \circ f_i$$

and again, as u_i is a linear map, it is differentiable and (1.1.12) follows.This completes the proof. ■

Let us now consider the case where E is the product of normed linear spaces. Let $E = E_1 \times ... \times E_n$ and $\mathcal{U} \subset E$ an open set and $f : \mathcal{U} \to F$ a given map. Given $a = (a_1, ..., a_n) \in E$, we define $\lambda_i : E_i \to E$ by

$$\lambda_i(x_i) = (a_1, ..a_{i-1}, x_i, a_{i+1}, ..a_n).$$

Proposition 1.1.3 *If f is differentiable at $a \in \mathcal{U}$, then for each $i, f \circ \lambda_i$ is differentiable at a_i. Further,*

$$f'(a)(h_1, ...h_n) = \sum_{i=1}^n (f \circ \lambda_i)'(a_i)h_i \quad (1.1.14)$$

for any $h = (h_1, ..., h_n) \in E_1 \times ... \times E_n = E$.

Proof: If u_i is the injection of E_i into E as defined previously, we have

$$\lambda_i(x_i) = a + u_i(x_i - a_i).$$

Then (cf. Example 1.1.2),

$$\lambda_i'(x_i) = u_i, \text{ for all } x_i \in E_i.$$

Hence if f is differentiable, so is $f \circ \lambda_i$ and

$$(f \circ \lambda_i)'(a_i) = f'(a) \circ u_i.$$

Once again we have

$$\sum_{i=1}^{n} u_i \circ p_i = I_E$$

which gives

$$\sum_{i=1}^{n} (f'(a) \circ u_i) \circ p_i = f'(a)$$

which is just a reformulation of (1.1.14). ■

Definition 1.1.3 *The derivative of $f \circ \lambda_i$ at a_i is called the i-th* **partial derivative** *of f at a and is denoted by $\frac{\partial f}{\partial x_i}(a)$ or by $\partial_i f(a)$.*■

Example 1.1.6 Let $\mathcal{U} \subset \mathbb{R}^n$ be an open set and let $f : \mathcal{U} \to \mathbb{R}$ be a given function differentiable at a point $a \in \mathcal{U}$. Then the partial derivatives of f at a are the usual ones we know from the calculus of functions of several variables. Further the relation (1.1.14) implies that $f'(a)$ can be represented as follows (since $\mathcal{L}(\mathbb{R}^n, \mathbb{R}) \cong \mathbb{R}^n$):

$$f'(a) = \left(\frac{\partial f}{\partial x_1}(a), ..., \frac{\partial f}{\partial x_n}(a)\right)$$

(which is also denoted by $\nabla f(a)$). It can also be seen that $\frac{\partial f}{\partial x_i}(a)$ is the Gâteau derivative of f along e_i, the i-th standard basis vector of $\mathbb{R}^n$.■

Example 1.1.7 Let $\mathcal{U}$ be as in the previous example and let $f : \mathcal{U} \to \mathbb{R}^m$ be differentiable at $a \in \mathcal{U}$. Then $f'(a) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ which can be represented by an $m \times n$ matrix. Indeed, if $f(x) = (f_1(x), ..., f_m(x))$,then, by (1.1.12) and (1.1.14), we deduce that $f'(a)$ is given by the usual Jacobian matrix,

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ .. & .. & .. \\ .. & .. & .. \\ \frac{\partial f_m}{\partial x_1}(a) & \cdots & \frac{\partial f_m}{\partial x_n}(a) \end{bmatrix}.■$$

Remark 1.1.3 As shown by Example 1.1.1, the converse of Proposition 1.1.3 is false; all partial derivatives of f may exist at a point but f could fail to be differentiable there. However, we will prove (cf. Proposition 1.1.4) the differentiability of f given the existence of partial derivatives under some additional hypotheses. ■

We now discuss an important result of differential calculus, *viz.* the mean value theorem. One of the earliest forms of this theorem states that if $f : \mathbb{R} \to \mathbb{R}$ is differentiable in an interval containing $[a, a+h]$, then there exists a $\theta \in (0,1)$ such that

$$f(a+h) - f(a) = f'(a+\theta h)h. \tag{1.1.15}$$

It is clear that (1.1.15) cannot be true in more general situations for arbitrary normed linear spaces E and F. Indeed, even if we take $E = \mathbb{R}$ and $F = \mathbb{R}^2$ and set

$$f(t) = (\cos t, \sin t),$$

then it follows from Proposition 1.1.2 that

$$f'(t) = (-\sin t, \cos t).$$

Thus, while $f(0) = f(2\pi)$, we can never have $f'(t) = 0$ for any $t \in (0, 2\pi)$. However, there are other versions of this result which are true.

Indeed, from (1.1.15) we deduce that if f is a real valued function of a real variable which is differentiable in an interval containing $[a, a+h]$, then

$$|f(a+h) - f(a)| \leq \sup_{0 \leq \theta \leq 1} |f'(a+\theta h)||h|. \tag{1.1.16}$$

This form of the mean value theorem can be readily generalized.

Definition 1.1.4 *Let E be a normed linear space and let $a, b \in E$. Then by the interval $[a, b]$ we mean the set*

$$\{x \in E | x = (1-t)a + tb, 0 \leq t \leq 1\}.$$

The open interval (a, b) is similarly defined using $0 < t < 1$.■

Theorem 1.1.1 *(Mean Value Theorem)Let E and F be normed linear spaces and $\mathcal{U}$ an open set in E. Let $f : \mathcal{U} \to F$ be differentiable in $\mathcal{U}$ and let $[a, b] \subset \mathcal{U}$. Then*

$$||f(b) - f(a)|| \leq ||b - a|| \sup_{x \in [a,b]} ||f'(x)||. \tag{1.1.17}$$

Proof: Step 1. We first prove the result when $f : \mathcal{U} \subset \mathbb{R} \to F$,where F is a normed linear space. Clearly, it suffices to show that, given $\varepsilon > 0$, we have

$$||f(x) - f(a)|| \leq (k + \varepsilon)(x - a) + \varepsilon \tag{1.1.18}$$

where $x \in [a, b] \subset \mathcal{U}$ and $k = \sup_{x \in [a,b]} ||f'(x)||$; the relation (1.1.17) follows on letting $\varepsilon \to 0$ and then setting $x = b$.

Assuming the contrary, let $\mathcal{V}$ be the set of all $x \in [a, b]$ such that

$$||f(x) - f(a)|| > (k + \varepsilon)(x - a) + \varepsilon.$$

On one hand, by the continuity of the members on both sides of the above inequality, we have that $\mathcal{V}$ is an open subset of $[a, b]$. Hence, if c is its greatest lower bound, then $c \notin \mathcal{V}$.

By continuity, (1.1.18) is true for all x near a and so $c \neq a$. Again, $c \neq b$, for, otherwise, we would have $\mathcal{V} = \{b\}$ which is not possible. Hence, $a < c < b$ and $||f'(c)|| \leq k$. Thus, for some $\eta > 0$, we have

$$k \geq \frac{||f(x) - f(c)||}{x - c} - \varepsilon$$

for all $c \leq x \leq c + \eta$, or,

$$||f(x) - f(c)|| \leq (k + \varepsilon)(x - c).$$

But since $c \notin \mathcal{V}$,

$$||f(c) - f(a)|| \leq (k + \varepsilon)(c - a) + \varepsilon.$$

Combining the two inequalities above,we deduce that for $c \leq x \leq c + \eta$, $x \notin \mathcal{V}$ which contradicts the definition of c. This shows that $\mathcal{V} = \emptyset$ and hence proves (1.1.17).

Step 2. Define, for $0 \leq t \leq 1$,

$$h(t) = f((1-t)a + tb).$$

Thus, $h : [0, 1] \to F$ and, by Proposition 1.1.2, it is differentiable. Further,

$$h'(t) = f'((1-t)a + tb)(b - a)$$

whence

$$||h'(t)|| \leq ||f'((1-t)a + tb)||.||b - a||.$$

The result now follows on applying Step 1 to h.■

Corollary 1.1.1 *If $\mathcal{U} \subset E$ is an open convex set and $f : \mathcal{U} \to F$ is differentiable at all points of $\mathcal{U}$ and verifies*

$$||f'(x)|| \leq k, \text{ for every } x \in \mathcal{U},$$

then, for any x_1 and x_2 in $\mathcal{U}$, we have

$$||f(x_1) - f(x_2)|| \leq k||x_1 - x_2||. \blacksquare \tag{1.1.19}$$

Remark 1.1.4 A function which satisfies (1.1.19) is called a Lipschitz continuous function with Lipschitz constant k.■

Corollary 1.1.2 *Let $f : \mathcal{U} \subset E \to F$ be differentiable on the set $\mathcal{U}$. Assume that $\mathcal{U}$ is connected and that $f'(x) = 0$ for every $x \in \mathcal{U}$. Then f is a constant function.*

Proof: Let $x_o \in \mathcal{U}$ and let $B(x_o; \varepsilon) \subset \mathcal{U}$ be the open ball with centre at x_o and radius $\varepsilon > 0$. Then $B(x_o; \varepsilon)$ is an open convex set and, by (1.1.19), we get that $f(x) = f(x_o)$ on this set since

$k = 0$. Thus f is locally constant. Then for any $b \in F, f^{-1}(b)$ is an open set and since F is Hausdorff, it is also closed. Thus, by the connectedness of $\mathcal{U}, f^{-1}(b) = \emptyset$ or $\mathcal{U}$. If we choose $b = f(x_o)$ for some $x_o \in \mathcal{U}$, we get that $f(x) = b$ for all $x \in \mathcal{U}$.■

The mean value theorem has numerous applications. We present below one such result, the converse of Proposition 1.1.3, as promised earlier (cf. Remark 1.1.3). Some other important applications will be discussed in Section 1.3.

Proposition 1.1.4 *Let $E = E_1 \times ... \times E_n$, the product of normed linear spaces and let F be a normed linear space. Let $\mathcal{U} \subset E$ be an open set and let $f : \mathcal{U} \to F$ be given. If the partial derivatives $\frac{\partial f}{\partial x_i}(x)$ exist at every point of $\mathcal{U}$ and if the maps $x \longmapsto \frac{\partial f}{\partial x_i}(x)$ are continuous at $a \in \mathcal{U}$ for each $1 \le i \le n$, then f is differentiable at a.*

Proof: If f were differentiable, then we know what $f'(a)$ should be, *i.e.* we need to show that

$$f'(a)h = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)h_i. \tag{1.1.20}$$

Consider

$$\begin{aligned} & f(x_1, .., x_n) - f(a_1, .., a_n) - \textstyle\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i) \\ = \ & f(x_1, .., x_n) - f(a_1, x_2, .., x_n) - \tfrac{\partial f}{\partial x_1}(a)(x_1 - a_1) \\ & + f(a_1, x_2, .., x_n) - f(a_1, a_2, x_3, .., x_n) - \tfrac{\partial f}{\partial x_2}(a)(x_2 - a_2) \\ & + ... + f(a_1, .., a_{n-1}, x_n) - f(a_1, .., a_n) - \tfrac{\partial f}{\partial x_n}(a)(x_n - a_n). \end{aligned} \tag{1.1.21}$$

Let $\varepsilon > 0$ be any arbitrarily small positive quantity. Let

$$g(\xi_1) = f(\xi_1, x_2, .., x_n) - \frac{\partial f}{\partial x_1}(a)(\xi_1 - a_1).$$

Then

$$f(x_1, x_2, .., x_n) - f(a_1, x_2, .., x_n) - \frac{\partial f}{\partial x_1}(a)(x_1 - a_1) = g(x_1) - g(a_1).$$

But g is differentiable and

$$g'(\xi_1) = \frac{\partial f}{\partial x_1}(\xi_1, x_2, .., x_n) - \frac{\partial f}{\partial x_1}(a_1, a_2, .., a_n).$$

Since $\frac{\partial f}{\partial x_1}$ is continuous, there exists $\eta > 0$ such that $||g'(\xi_1)|| < \varepsilon$ whenever $||x_i - a_i|| < \eta$ for each $1 \le i \le n$.Hence,by the mean value theorem,

$$||g(x_1) - g(a_1)|| \le \varepsilon ||x - a||.$$

We have similar estimates for the terms in each row of (1.1.21) which in turn proves that

$$||f(x) - f(a) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i)|| = o(||x - a||).$$

This shows that f is differentiable at a and that its derivative at a is given by (1.1.20). ■

Exercise 1.1.4 If $f : \mathcal{U} \subset E \to \mathbb{R}$, then, show that, under the hypotheses of Thoerem 1.1.1, there exists a point $c \in (a, b)$ such that

$$f(b) - f(a) = f'(c)(b - a).■$$

Exercise 1.1.5 We say that $f : \mathcal{U} \subset E \to F$ is Gâteau differentiable at a point $a \in \mathcal{U}$ if there exists a continuous linear operator $df(a) : E \to F$ such that

$$\lim_{t \to 0} \frac{f(a + th) - f(a)}{t} = df(a)h, \quad \text{for all } \ h \in E.$$

If $F = \mathbb{R}$, and f is Gâteau differentiable at all points of $\mathcal{U}$, show that we can still have the same conclusion as in the preceding exerecise. If the mapping $x \mapsto df(x)$ is continuous at $a \in \mathcal{U}$, show that, in that case, f is Fréchet differentiable at a and that $f'(a) = df(a)$.■

There are several applications of the mean value theorem which interested readers can find in Cartan [4]. We will use it again in Section 1.3 to prove some important results.

1.2 Higher Order Derivatives

Let E and F be normed linear spaces and let $\mathcal{U} \subset E$ be an open set. Let $f : \mathcal{U} \to F$ be a given mapping which is differentiable in $\mathcal{U}$. Consider the map

$$x \longmapsto f'(x) \in \mathcal{L}(E, F).$$

This is a mapping from $\mathcal{U}$ into $\mathcal{L}(E,F)$ and one could again investigate its differentiability. Its derivative at a point a, if it exists, is a linear map of E into $\mathcal{L}(E,F)$ and thus belongs to $\mathcal{L}(E,\mathcal{L}(E,F))$. This is called the second derivative of f at a and is denoted by $f''(a)$ (or $D^2f(a)$).

Remark 1.2.1 To define the second derivative at a point $a \in \mathcal{U}$, we need not assume that it is differentiable at every point of $\mathcal{U}$. It suffices to assume that it is differentiable in a neighbourhood $\mathcal{V}$ of a, and in that neighbourhood, the mapping

$$x \longmapsto f'(x)$$

is differentiable at a.■

Definition 1.2.1 *The function f is said to be twice differentiable in $\mathcal{U}$ if $f''(a)$ exists at every point a of $\mathcal{U}$. If the map $x \longmapsto f''(x)$ is continuous from $\mathcal{U}$ into $\mathcal{L}(E,\mathcal{L}(E,F))$, we say that f is of class $\mathcal{C}^2$.*■

Recall that the space $\mathcal{L}(E,\mathcal{L}(E,F))$ is isomorphic to $\mathcal{L}_2(E,F)$, the space of continuous bilinear forms from $E \times E$ into F. Thus $f''(a)$ can be thought of as a bilinear form and if $(h,k) \in E \times E$,

$$f''(a)(h,k) = (f''(a)h)k. \tag{1.2.1}$$

Note that $f''(a)h \in \mathcal{L}(E,F)$ and so (1.2.1) makes sense.

As an application of the mean value theorem, one can prove that (cf.Cartan [4])

$$f''(a)(h,k) = f''(a)(k,h), \tag{1.2.2}$$

i.e. the bilinear form is symmetric. We omit the proof.

Example 1.2.1 Let us consider the function defined in Example 1.1.3. Let E be a Hilbert space and let $a(.,.)$ be a symmetric bilinear form on E and $b \in E$. We have

$$f(x) = \frac{1}{2}a(x,x) - (b,x).$$

Thus

$$f'(x_o)h = a(x_o,h) - (b,h).$$

It is now easy to see that

$$f''(x_o)(h,k) = a(h,k).■$$

Let us now assume that $E = E_1 \times ... \times E_n$, the product of normed linear spaces and that $\mathcal{U} \subset E$. Assume that f is twice differentiable at $a \in \mathcal{U}$. Then f' exists at each point in a neighbourhood of a. Now

$$f'(x)(h_1,...,h_n) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x)h_i.$$

In the same way,

$$f''(a)(k_1,...,k_n) = \sum_{i=1}^{n} \frac{\partial f'}{\partial x_i}(a)k_i.$$

Hence,

$$(f''(a)(k_1,...,k_n))(h_1,...,h_n) = (\sum_{i=1}^{n} \frac{\partial f'}{\partial x_i}(a)k_i)(h_1,...,h_n). \tag{1.2.3}$$

Note that $\frac{\partial f'}{\partial x_i}(a) \in \mathcal{L}(E_i,\mathcal{L}(E,F))$ and so (1.2.3) defines an element of F. Now,

$$(\frac{\partial f'}{\partial x_i}(a)k_i)(h_1,...,h_n) = \sum_{j=1}^{n}(\frac{\partial}{\partial x_i}(\frac{\partial f}{\partial x_j}(a))k_i)h_j.$$

We denote by $\frac{\partial^2 f}{\partial x_i \partial x_j}(a)$ (or by $\partial_{ij} f(a)$) the term $\frac{\partial}{\partial x_i}(\frac{\partial f}{\partial x_j}(a))$ which is an element of $\mathcal{L}(E_i, \mathcal{L}(E_j, F))$, *i.e.* a bilinear form on $E_i \times E_j$ with values in F. Thus we can rewrite (1.2.3) as

$$(f''(a)k)h = \sum_{i,j=1}^{n} (\frac{\partial^2 f}{\partial x_i \partial x_j}(a)k_i)h_j. \tag{1.2.4}$$

But $(f''(a)k)h = (f''(a)h)k$, which gives

$$\begin{aligned} \textstyle\sum_{i,j=1}^{n}(\frac{\partial^2 f}{\partial x_i \partial x_j}(a)k_i)h_j &= \textstyle\sum_{i,j=1}^{n}(\frac{\partial^2 f}{\partial x_i \partial x_j}(a)h_i)k_j \\ &= \textstyle\sum_{i,j=1}^{n}(\frac{\partial^2 f}{\partial x_j \partial x_i}(a)h_j)k_i. \end{aligned} \tag{1.2.5}$$

In particular, if $E = \mathbb{R}^n$,*i.e.* $E_i = \mathbb{R}$ for $1 \leq i \leq n$, then we can identify $\mathcal{L}(\mathbb{R}, F)$ with F and so $\mathcal{L}(\mathbb{R}, \mathcal{L}(\mathbb{R}, F))$ with F again. Then the bilinear map $\frac{\partial^2 f}{\partial x_i \partial x_j}(a) : \mathbb{R} \times \mathbb{R} \to F$ is just given by

$$(\lambda, \mu) \longmapsto c_{ij}\lambda_i \mu_j$$

where $c_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \in F$. Then, we deduce from (1.2.5) that

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a). \tag{1.2.6}$$

As in the case of Proposition 1.1.4, the existence of the second order partial derivatives implies that f is twice differentiable at a point provided these partial derivatives exist at all points of $\mathcal{U}$ and are continuous at the given point.

We can now successively define higher order derivatives. By considering the differentiability of the map

$$x \in \mathcal{U} \longmapsto f''(x) \in \mathcal{L}_2(E, F),$$

we can define the third derivative of f and so on. In general, we can define the n-th derivative of f at a point a, denoted by $f^{(n)}(a)$ or $D^n f(a)$, by induction: if the concept of the first $(n-1)$ derivatives has already been defined, then f is said to be n-times differentiable at $a \in \mathcal{U}$ if it is $(n-1)$-times differentiable in a neighbourhood $\mathcal{V} \subset \mathcal{U}$ of a and the map

$$x \longmapsto f^{(n-1)}(x) \in \mathcal{L}_{n-1}(E, F)$$

is differentiable at a. The derivative $f^{(n)}(a)$ belongs to $\mathcal{L}_n(E, F)$,the space of n-linear forms on E with values in F.If the map $x \longmapsto f^{(n)}(x)$ is continuous, we say that f is of class C^n.

Definition 1.2.2 *We say that f is of class C^0 if it is continuous. We say that f is of class C^∞ if it is of class C^n for all positive integers n.*■

Remark 1.2.2 As in the case of the second derivative, the n-th derivative is a symmetric n-linear form. Thus, if σ is a permutation of the symbols $\{1, 2, ..., n\}$, then,

$$f^{(n)}(a)(h_1, ..., h_n) = f^{(n)}(a)(h_{\sigma(1)}, ..., h_{\sigma(n)}).\blacksquare$$

For completeness, we state below the various generalizations of the mean value theorem and the Taylor expansion formulae known for real valued functions of a real variable. For detailed proofs, see Cartan [4].

Theorem 1.2.1 *(Taylor's Formula) Let E and F be normed linear spaces and $\mathcal{U} \subset E$ an open set. Let $f : \mathcal{U} \to F$ be $(n-1)$ -times differentiable in $\mathcal{U}$ and let f be n-times differentiable at $a \in \mathcal{U}$. Then,*

$$||f(a+h) - f(a) - f'(a)h - ... - \frac{1}{n!}f^{(n)}(a)(h)^n|| = o(||h||^n), \tag{1.2.7}$$

where $f^{(n)}(a)(h)^n = f^{(n)}(a)(\underbrace{h, ..., h}_{n \text{ times}})$.■

Theorem 1.2.2 *(Mean Value Theorem) Let $f : \mathcal{U} \subset E \to F$ be $(n+1)$-times differentiable in $\mathcal{U}$ and assume that*

$$||f^{(n+1)}(x)|| \leq M \textit{ for every } x \in \mathcal{U}.$$

Then

$$||f(a+h) - f(a) - f'(a)h - ... - \frac{1}{n!}f^{(n)}(a)(h)^n|| \leq \frac{M||h||^{n+1}}{(n+1)!}. \blacksquare \quad (1.2.8)$$

Remark 1.2.3 The relation (1.2.8) is stronger than (1.2.7). While (1.2.7) is asymptotic in the sense that it just tells us what happens as $||h|| \to 0$,(1.2.8) gives us an estimate of the error in the n-th order expansion. It is sometimes called the Taylor's formula with Lagrange form of the remainder. Note however that the hypotheses of Theorem 1.2.2 are also stronger than those of Theorem 1.2.1. ■

Example 1.2.2 Let $E = F = \mathbb{R}$ and consider the function

$$f(x) = \begin{cases} x^3 \sin(1/x) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

It is easy to check that f is *not* twice differentiable at the origin. However,

$$f(x) = o(||x||^2)$$

and so it possesses a "Taylor expansion" of order 2 at the origin, *i.e.*

$$f(x) = a_0 + a_1 x + a_2 x^2 + o(|x|^2)$$

with $a_0 = a_1 = a_2 = 0$.■

Remark 1.2.4 As the above example shows, a function $f : \mathcal{U} \subset \mathbb{R} \to \mathbb{R}$ may possess an n-th order Taylor expansion at a point in the form

$$f(a+h) = a_0 + a_1 h + \frac{a_2}{2!}h^2 + ... + \frac{a_n}{n!}h^n + o(|h|^n) \quad (1.2.9)$$

but fail to be n-times differentiable at a. However, if it is n-times differentiable at a and admits a Taylor expansion of the form given in (1.2.9), then necessarily,

$$a_0 = f(a) \ , \ a_i = f^{(i)}(a) \ , \ 1 \leq i \leq n.\blacksquare$$

Finally, we describe the Taylor formula with the "integral form of the remainder".

If F is a Banach space and $f : [a,b] \subset \mathbb{R} \to F$ is a continuous map, then we can define the integral

$$y = \int_a^b f(t)dt$$

as a vector in F. This can be done by defining it as the limit ,as $n \to \infty$, of appropriate Riemann sums. Using the continuity of f, we can show that the Riemann sums form a Cauchy sequence and hence converge, since F is *complete*. The integral also turns out to be the *unique* vector $y \in F$ (unique, by virtue of the Hahn-Banach Theorem) such that for any $\varphi \in F^*$ (the dual of F),

$$\varphi(y) = \int_a^b \varphi(f(t))dt$$

where the integral on the right-hand side is now that of a real valued continuous function on $[a,b]$.

Theorem 1.2.3 *Let F be a Banach space and E a normed linear space and $\mathcal{U} \subset E$ an open set. Let $f : \mathcal{U} \to F$ be of class $\mathcal{C}^{(n+1)}$. Let $[a, a+h] \subset \mathcal{U}$. Then,*

$$\begin{aligned} f(a+h) &= f(a) + f'(a)h + ... + \tfrac{1}{n!}f^{(n)}(a)(h)^n \\ &+ \int_0^1 \tfrac{(1-t)^n}{n!} f^{(n+1)}(a+th)(h)^{(n+1)}dt.\blacksquare \end{aligned} \quad (1.2.10)$$

Remark 1.2.5 Observe now that we need even stronger hypotheses on f.We need that the derivative of f of order $(n+1)$ be continuous and that F be complete. ■

If f is of class $\mathcal{C}^1$, a particular case of (1.2.10) gives

$$f(a+h) = f(a) + \int_0^1 f'(a+th)hdt. \quad (1.2.11)$$

1.3 Some Important Theorems

In this section we will present some very important results in Analysis which will also be frequently used in the sequel. It will also be seen that the mean value theorem will play an important role in the proof of these results.

The first result we will examine concerns the solution set of the equation

$$f(x,y)=0 \tag{1.3.1}$$

where $f : E\times F \to G$ is a continuous function, E, F and G being normed linear spaces. Of course, the very general nature of the problem prevents us from having a precise general result on the global structure of the set of solutions to (1.3.1). However, given a solution, say, (a, b), of the equation, under some reasonable conditions, we can describe *locally* the set of solutions of (1.3.1). In fact, we will show that the solutions "close" to (a, b) lie on a "curve". More precisely, we have the following result.

Theorem 1.3.1 *(Implicit Function Theorem) Let E_1, E_2 and F be normed linear spaces and assume that E_2 is complete. Let $\Omega\subset E_1\times E_2$ be an open set and let $f:\Omega\to F$ be a function such that*
(i) f is continuous;
(ii)for every $(x_1,x_2)\in\Omega$, $\frac{\partial f}{\partial x_2}(x_1,x_2)$ exists and is continuous on Ω;
(iii)$f(a,b)=0$ and $A=\frac{\partial f}{\partial x_2}(a,b)$ is invertible with continuous inverse.
Then, there exist neighbourhoods $\mathcal{U}$ of a and $\mathcal{V}$ of b and a continuous function $\varphi:\mathcal{U}\to\mathcal{V}$ such that $\varphi(a)=b$ and

$$f(x,\varphi(x))=0 \tag{1.3.2}$$

and these are the only solutions of (1.3.1) in $\mathcal{U}\times\mathcal{V}$. Further, if f is differentiable at (a,b), then φ is differentiable at a and

$$\varphi'(a) = -\left[\frac{\partial f}{\partial x_2}(a,b)\right]^{-1}\left[\frac{\partial f}{\partial x_1}(a,b)\right]. \tag{1.3.3}$$

Proof: Step 1. Define $g:\Omega\to E_2$ by

$$g(x_1,x_2)=x_2-A^{-1}f(x_1,x_2). \tag{1.3.4}$$

Looking for solutions of (1.3.1) is the same as finding fixed points x_2 of g for given x_1. Now,

$$\frac{\partial g}{\partial x_2}(x_1,x_2)=I-A^{-1}\frac{\partial f}{\partial x_2}(x_1,x_2)$$

and, by (ii) above, $\frac{\partial g}{\partial x_2}$ is continuous on Ω. Further,

$$\frac{\partial g}{\partial x_2}(a,b)=0\ ,\ g(a,b)=b. \tag{1.3.5}$$

Hence, by the continuity of the partial derivative, there exist neighbourhoods $\mathcal{U}_1$ of a and $\mathcal{V}$ of b such that

$$\left\|\frac{\partial g}{\partial x_2}(x_1,x_2)\right\|\le\frac{1}{2} \tag{1.3.6}$$

for every $(x_1,x_2)\in\mathcal{U}_1\times\mathcal{V}$. Without loss of generality, we may assume that $\mathcal{V}=B(b;r)$, the closed ball with centre at b and radius $r>0$. By the continuity of g, we deduce the existence of a neighbourhood $\mathcal{U}\subset\mathcal{U}_1$ of a such that

$$||g(x_1,b)-g(a,b)||\le r/2,\ \text{for every } x_1\in\mathcal{U}. \tag{1.3.7}$$

Step 2. Let $x_1\in\mathcal{U}$ be fixed. Define $g_{x_1}:\mathcal{V}\to E_2$ by

$$g_{x_1}(x_2)=g(x_1,x_2).$$

Then

$$\begin{aligned}
||g_{x_1}(x_2)-b|| &= ||g(x_1,x_2)-b|| \\
&= ||g(x_1,x_2)-g(a,b)|| \\
&\le ||g(x_1,x_2)-g(x_1,b)||+||g(x_1,b)-g(a,b)|| \\
&\le \tfrac{1}{2}||x_2-b||+\tfrac{r}{2} \\
&\le r
\end{aligned}$$

by virtue of (1.3.6) and the mean value theorem and also by (1.3.7). Thus, g_{x_1} maps $\mathcal{V}$ into itself and, again by (1.3.6) and the mean value theorem, g_{x_1} is a contraction (with Lipschitz constant equal to 1/2) and so there exists a unique fixed point $x_2 = \varphi(x_1)$ in $\mathcal{V}$. Thus the only solutions to (1.3.1) in $\mathcal{U} \times \mathcal{V}$ are given by $(x_1, \varphi(x_1))$. Clearly, $\varphi(a) = b$.

Step 3. (Continuity of φ) Let $x_1^o \in \mathcal{U}$. Then

$$\begin{aligned} ||\varphi(x_1) - \varphi(x_1^o)|| &= ||g(x_1, \varphi(x_1)) - g(x_1^o, \varphi(x_1^o))|| \\ &\leq ||g(x_1, \varphi(x_1)) - g(x_1, \varphi(x_1^o))|| \\ &\quad + ||g(x_1, \varphi(x_1^o)) - g(x_1^o, \varphi(x_1^o))|| \\ &\leq \tfrac{1}{2}||\varphi(x_1) - \varphi(x_1^o)|| \\ &\quad + ||g(x_1, \varphi(x_1^o)) - g(x_1^o, \varphi(x_1^o))||. \end{aligned}$$

Thus,

$$||\varphi(x_1) - \varphi(x_1^o)|| \leq 2||g(x_1, \varphi(x_1^o)) - g(x_1^o, \varphi(x_1^o))||$$

and the result follows from the continuity of g.

Step 4. (Differentiability of φ) Assume now that f is differentiable at (a, b). Let $h \in E$, small enough such that $a + h \in \mathcal{U}$. Set

$$k = \varphi(a + h) - \varphi(a).$$

Now,

$$\begin{aligned} 0 &= f(a + h, \varphi(a + h)) - f(a, \varphi(a)) \\ &= \tfrac{\partial f}{\partial x_1}(a, b)h + \tfrac{\partial f}{\partial x_2}(a, b)k + \varepsilon(h, k)(||h|| + ||k||) \end{aligned}$$

where $\varepsilon(h, k) \to 0$ when $||h|| \to 0$ and $||k|| \to 0$. Then,

$$k = -A^{-1}\left[\frac{\partial f}{\partial x_1}(a, b)\right]h - (||h|| + ||k||)A^{-1}\varepsilon(h, k).$$

To prove (1.3.3), it suffices to show that

$$(||h|| + ||k||)||A^{-1}\varepsilon(h, k)|| = \eta(h)||h|| \qquad (1.3.8)$$

where $\eta(h) \to 0$ as $||h|| \to 0$. But

$$||k|| \leq \alpha||h|| + \beta(||h|| + ||k||)||\varepsilon(h, k)||$$

where

$$\alpha = \left\|A^{-1}\frac{\partial f}{\partial x_1}(a, b)\right\| \text{ and } \beta = ||A^{-1}||.$$

Since φ is continuous, $||k|| \to 0$ as $||h|| \to 0$ and so $\varepsilon(h, k) \to 0$ as $||h|| \to 0$. Thus there exists $r_o > 0$ such that if $||h|| \leq r_o$, then $\beta||\varepsilon(h, k)|| \leq 1/2$ and so

$$||k|| \leq \alpha||h|| + \frac{1}{2}(||h|| + ||k||)$$

or

$$||k|| \leq (2\alpha + 1)||h||$$

which proves (1.3.8) and completes the proof of the theorem. ■

Remark 1.3.1 If f is a $\mathcal{C}^2$ function, then the right-hand side of (1.3.3) is $\mathcal{C}^1$ and so φ will be $\mathcal{C}^2$. By induction, if f is of class $\mathcal{C}^p$, then φ will also be of class $\mathcal{C}^p$.■

Remark 1.3.2 We can derive (1.3.3) heuristically by implicit differentiation. We have

$$f(x_1, \varphi(x_1)) = 0.$$

Differentiating this w.r.t x_1, we get

$$\frac{\partial f}{\partial x_1}(a, b) + \frac{\partial f}{\partial x_2}(a, b)\varphi'(a) = 0$$

which yields (1.3.3). ■

The following consequence of the implicit function theorem tells us when a mapping is a local homeomorphism.

Theorem 1.3.2 *(Inverse Function Theorem) Let E and F be Banach spaces and $f : \Omega \subset E \to F$ be a $\mathcal{C}^p$- map, for some $p \geq 1$.Let*

$a \in \Omega$ with $f(a) = b$ and let $f'(a) : E \to F$ be an isomorphism. Then there exists a neighbourhood $\mathcal{V}$ of b in F and a unique C^p function $g : \mathcal{V} \to E$ such that

$$\left\{ \begin{array}{rcl} a & = & g(b) \\ f(g(y)) & = & y \end{array} \right. \tag{1.3.9}$$

for every $y \in \mathcal{V}$.

Proof: Define $\psi : \Omega \times F \to F$ by

$$\psi(x, y) = f(x) - y.$$

The result now follows immediately on applying the implicit function theorem to ψ.■

A stronger version of Theorem 1.3.2 exists, wherein sufficient conditions are given for f to be a global homeomorphism. We state it without proof (cf. Schwartz [23]).

Theorem 1.3.3 *Let $f : E \to F$ be a C^1 map such that for every $x \in E$, we have that $f'(x) : E \to F$ is an isomorphism of E onto F. Assume further that there exists a constant $K > 0$ such that $||(f'(x))^{-1}|| \leq K$ for every $x \in E$. Then f is a homeomorphism of E onto F.*■

Example 1.3.1 The above result is not true if $(f'(x))^{-1}$ is not uniformly bounded in $\mathcal{L}(F, E)$.Consider $E = F = \mathbb{R}^2$ and

$$f(x) = (e^{x_1} \sin x_2, e^{x_1} \cos x_2) \text{ , for } x = (x_1, x_2).$$

Then

$$f'(x) = \left[\begin{array}{cc} e^{x_1} \sin x_2 & e^{x_1} \cos x_2 \\ e^{x_1} \cos x_2 & -e^{x_1} \sin x_2 \end{array} \right]$$

and $\det(f'(x)) = -e^{2x_1} \neq 0$ and so $f'(x)$ is invertible for each $x \in \mathbb{R}^2$ but the norm of its inverse is unbounded. The mapping f is neither one-one nor onto for

$$f(x_1, x_2 + 2n\pi) = f(x_1, x_2) \text{ for all } n$$

and

$$f(x) \neq 0 \text{ for any } x \in \mathbb{R}^2.■$$

We conclude with one more classical result.

Theorem 1.3.4 *(Sard's Theorem) Let $\Omega \subset \mathbb{R}^n$ be an open set and let $f : \Omega \to \mathbb{R}^n$ be a C^1 function. Let*

$$S = \{x \in \Omega \mid J_f(x) = 0\},$$

where $J_f(x) = \det(f'(x))$. Then $f(S)$ is of measure zero in $\mathbb{R}^n$.

Proof: Step 1. Let C be a cube of side a contained in Ω. We divide it into k^n sub-cubes each of side a/k. Since f is of class C^1, the map $x \longmapsto f'(x)$ is uniformly continuous on C. Thus, for $\varepsilon > 0$, there exists $\delta > 0$ such that

$$||x - y|| < \delta \Longrightarrow ||f'(x) - f'(y)|| < \varepsilon. \tag{1.3.10}$$

We choose k large enough so that $\sqrt{n}a/k < \delta$, *i.e.* the diameter of each sub-cube is less than δ.

Since f' is bounded on C, the function f is Lipschitz continuous on C by the mean value theorem. Thus,

$$||f(x_1) - f(x_2)|| \leq L||x_1 - x_2||, \text{ for every } x_1, x_2 \in C, \tag{1.3.11}$$

where

$$L = \sup_{x \in C} ||f'(x)||.$$

Step 2. Let $x \in C \cap S$. Then there exists a sub-cube $\tilde{C}$ such that $x \in \tilde{C}$. Then for every $y \in \tilde{C}$,

$$||f(x) - f(y)|| \leq L\sqrt{n}a/k, \tag{1.3.12}$$

by (1.3.11). On the other hand,

$$f(y) - f(x) - f'(x)(y - x) = \int_0^1 (f'(x + t(y - x)) - f'(x))(y - x)dt$$

which, by (1.3.10), yields

$$||f(y) - f(x) - f'(x)(y-x)|| \leq \varepsilon||y-x|| \leq \varepsilon\sqrt{n}a/k. \quad (1.3.13)$$

Step 3. Now $f'(x)$ is singular and so $f'(x)(\mathbb{R}^n) = H$ is a subspace of dimension $\leq n-1$. Hence, by (1.3.13), we have

$$\rho(f(y), f(x)+H) \leq \varepsilon\sqrt{n}a/k, \text{ for every } y \in \tilde{C} \quad (1.3.14)$$

where $\rho(a, X)$ denotes the distance of a point a from a set X. Combining (1.3.12) and (1.3.14), we deduce that $f(\tilde{C})$ is contained in a cylindrical block of radius $L\sqrt{n}a/k$ and height $2\varepsilon\sqrt{n}a/k$. Thus

$$m(f(\tilde{C})) \leq 2A\varepsilon\sqrt{n}\frac{a}{k}(2L\sqrt{n}\frac{a}{k})^{n-1}$$

where A is a constant depending only on n and $m(X)$ is the (Lebesgue) measure of a set X. Thus

$$\begin{aligned} m(f(C\cap S)) &\leq \textstyle\sum_{\tilde{C}\cap S\neq\emptyset} m(f(\tilde{C})) \\ &\leq 2^n A L^{n-1} a^n n^{n/2} \varepsilon \\ &= K(n, C)\varepsilon. \end{aligned}$$

Since ε is arbitrary, $m(f(C\cap S)) = 0$ for any cube C contained in Ω. But Ω can be covered by a countable number of such cubes and so $f(S)$ is of measure zero. ■

Example 1.3.2 If $f(x) = Tx$, where T is a linear operator on $\mathbb{R}^n$ such that $\det T = 0$,then $f'(x) = T$ and so $S = \mathbb{R}^n$. But $f(S)$ is now a subspace of dimension $\leq n-1$ in $\mathbb{R}^n$ and hence $f(S)$ is of measure zero. ■

Remark 1.3.3 More generally,we can show that

$$m(f(X)) \leq \int_X |J_f(x)|dx$$

for any set $X \subset \Omega$ (cf. Schwartz [23]). ■

Remark 1.3.4 A more general version of Sard's theorem states that if $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ is a mapping of class C^{n-m+1} and if $n \geq m$, then $f(S)$ is of measure zero, where now S is the set of all points $x \in \Omega$ such that the rank of $f'(x)$ is $< m$ (cf. Sard [22]). The result is not true in general if f is only of class C^{n-m} (cf. Whitney [26]). ■

Definition 1.3.1 *If $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ and $x \in \Omega$ is such that $f'(x)$ is of rank $< m$, then x is said to be a* **critical point** *of f. If not, x is a regular point. A vector $y \in \mathbb{R}^m$ is called a* **critical value** *if there exists a critical point $x \in \Omega$ such that $f(x) = y$. Otherwise, it is called a* **regular value.** ■

Thus, Sard's theorem states that if $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ is of class C^1, then almost every value f takes is regular, *i.e.* the set of critical values is of measure zero.

1.4 Extrema of Real Valued Functions

Let $\mathcal{U}$ be an open set in a normed linear space E and let $f : \mathcal{U} \to \mathbb{R}$ be a given function.

Definition 1.4.1 *We say that f attains a* **relative maximum** *(resp.* **minimum***) at $u \in \mathcal{U}$ if there exists a neighbourhood $\mathcal{V} \subset \mathcal{U}$ of u such that for all $v \in \mathcal{V}$,*

$$f(v) \leq f(u) \text{ (resp. } f(v) \geq f(u)).$$

If the inequality is strict for all $v \in \mathcal{V}, v \neq u$, then we say that f attains a **strict relative maximum** *(resp. minimum) at $u \in \mathcal{U}$.* ■

We will now present results which generalize the well known necessary conditions for the existence of a relative extremum (*i.e.* maximum or minimum) at a point in terms of its first and second order derivatives at that point.

Theorem 1.4.1 *Let $f : \mathcal{U} \subset E \to \mathbb{R}$ admit a relative extremum at $u \in \mathcal{U}$. If f is differentiable at u, then*

$$f'(u) = 0. \tag{1.4.1}$$

Proof: Let $v \in E$ be an arbitrary vector. Since $\mathcal{U}$ is open, we can find an open interval $J \subset \mathbb{R}$ containing the origin such that for all $t \in J$, $u + tv \in \mathcal{U}$. Define

$$\varphi(t) = f(u+tv).$$

Then φ is differentiable at the origin and

$$\varphi'(0) = f'(u)v.$$

Assume that f attains a relative minimum at u. Then

$$0 \geq \lim_{t\to 0^-} \frac{\varphi(t)-\varphi(0)}{t} = \varphi'(0) = \lim_{t\to 0^+} \frac{\varphi(t)-\varphi(0)}{t} \geq 0.$$

Thus, $f'(u)v = 0$ and since v was arbitrary, (1.4.1) follows. The case of a relative maximum is similar. ■

Remark 1.4.1 The relation (1.4.1) is called *Euler's equation* corresponding to the extremal problem. ■

Remark 1.4.2 If $E = \mathbb{R}^n$, then $f'(u) = 0$ is equivalent to the system of equations

$$\partial_i f(u_1, u_2, ..., u_n) = 0 \text{ for all } 1 \leq i \leq n$$

where $u = (u_1, u_2, ..., u_n)$ and this is the well known necessary condition for the existence of a relative extremum. ■

We now consider the case of extrema under constraints. Let E and F be normed linear spaces and let $\mathcal{U} \subset E$ be an open subset. Let

$$K = \{v \in \mathcal{U} \mid \varphi(v) = 0\} \tag{1.4.2}$$

where $\varphi : \mathcal{U} \subset E \to F$ is a given mapping. We then look for a relative extremum of $f : \mathcal{U} \subset E \to \mathbb{R}$ in K. Notice that K is not an open set. In fact, if φ is continuous, then K is closed. Thus the previous theorem cannot be applied.

Exercise 1.4.1 Let E be a vector space and let g and $g_i, 1 \leq i \leq m$ be linear functionals on E. Assume that

$$\cap_{i=1}^m \mathrm{Ker}(g_i) \subset \mathrm{Ker}(g).$$

Show that there exist scalars $\lambda_i, 1 \leq i \leq m$ such that

$$g = \sum_{i=1}^m \lambda_i g_i. \blacksquare$$

Theorem 1.4.2 *Assume that E is a Banach space and that $\varphi \in C^1(E;\mathbb{R})$. Let*

$$K = \{v \in E \mid \varphi(v) = 0\}.$$

Assume, further, that for all $v \in K$, we have $\varphi'(v) \neq 0$. If $f \in C^1(E;\mathbb{R})$ and if $u \in K$ is such that

$$f(u) = \inf_{v\in K} f(v),$$

then there exists $\lambda \in \mathbb{R}$ such that $f'(u) = \lambda\varphi'(u)$.

Proof: Since $u \in K$, by hypothesis, $\varphi'(u) \neq 0$. Hence, there exists $w_0 \in E$ such that $\|w_0\| = 1$ and $< \varphi'(u), w_0 >= 1$ where $< .,. >$ denotes the duality bracket between E and its dual. Set $E_0 = \mathrm{Ker}(\varphi'(u))$. Then, clearly, $E = E_0 \oplus \mathbb{R}\{w_0\}$. Define $\Phi : E_0 \times \mathbb{R} \to \mathbb{R}$ by

$$\Phi(v,t) = \varphi(u+v+tw_0).$$

Then $\Phi(0,0) = 0$. Further, we also have that

$$\begin{array}{rclcl} \partial_v\Phi(0,0) & = & \varphi'(u)|_{E_0} & = & 0 \\ \partial_t\Phi(0,0) & = & < \varphi'(u), w_0 > & = & 1. \end{array}$$

Thus, by the implicit function theorem, there exists a neighbourhood $\mathcal{V}$ of 0 in E_0 and a $\mathcal{C}^1$ function $\zeta : \mathcal{V} \to \mathbb{R}$ such that $\zeta(0) = \zeta'(0) = 0$ and if $\Omega = u + \mathcal{V}$, then the only points in $\Omega \cap K$ are of the form $w = u + v + \zeta(v)w_0$ with $v \in \mathcal{V}$. Now, setting $\widetilde{f}(v) = f(u + v + \zeta(v)w_0)$, for $v \in \mathcal{V}$, we see that $\widetilde{f}$ attains its minimum over the open set $\mathcal{V}$ at 0. Thus, $< \widetilde{f}'(0), v >= 0$ for every $v \in E_0$. Since $\zeta'(0) = 0$, we deduce that for all $v \in E_0$, we have $0 =< \widetilde{f}'(0), v >=< f'(u), v >$. Thus, it follows that $\mathrm{Ker}(\varphi'(u)) \subset \mathrm{Ker}(f'(u))$ and the conclusion follows from the preceding exercise. ■

Remark 1.4.3 In the same way, it can be shown that if $\varphi_i, 1 \leq i \leq m$ are in $\mathcal{C}^1(E;\mathbb{R})$ and if we define K by

$$K = \{v \in E \mid \varphi_i(v) = 0, 1 \leq i \leq m\},$$

then, if f attains a relative extremum at $u \in K$ and if $\varphi_i'(u), 1 \leq i \leq m$ are linearly independent, there exist scalars $\lambda_i, 1 \leq i \leq m$ such that

$$f'(u) = \sum_{i=1}^{m} \lambda_i \varphi_i'(u). ■$$

Example 1.4.1 Let $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$. Let $\varphi_i : \mathcal{U} \to \mathbb{R}$, $1 \leq i \leq m$, $1 \leq m < n$ be given functions. Let

$$K = \{v \in \mathcal{U} \mid \varphi_i(v) = 0,\ 1 \leq i \leq m\}.$$

Thus, if at a point $u \in \mathcal{U}$, the m vectors $\varphi_i'(u)$, $1 \leq i \leq m$ are linearly independent, a necessary condition that f attains a relative extremum at u, by the preceding theorem, is the existence of $\lambda_i \in \mathbb{R}, 1 \leq i \leq m$ such that $f'(u) - \sum_{i=1}^m \lambda_i \varphi_i'(u) = 0$. Let $u = (u_1, ..., u_n)$. To find the $n + m$ unknowns u_i, $1 \leq i \leq n$ and

λ_j, $1 \leq j \leq m$, we solve the system of $n + m$ equations

$$\left.\begin{array}{rcl}
\partial_1 f(u) - \lambda_1 \partial_1 \varphi_1(u) - ... - \lambda_m \partial_1 \varphi_m(u) & = & 0 \\
 & \cdot & \\
 & \cdot & \\
\partial_n f(u) - \lambda_1 \partial_n \varphi_1(u) - ... - \lambda_m \partial_n \varphi_m(u) & = & 0 \\
\varphi_1(u) & = & 0 \\
 & \cdot & \\
 & \cdot & \\
\varphi_m(u) & = & 0
\end{array}\right\}$$

which is exactly the well known method of *Lagrange multipliers* in the calculus of several variables. ■

Exercise 1.4.2 Let A be a symmetric $n \times n$ matrix with real entries and let B be a symmetric and positive definite matrix. Characterize the relative extrema of the functional

$$f(v) = \frac{1}{2}(Av, v)$$

on the set K given by

$$K = \{v \in \mathbb{R}^n \mid (Bv, v) = 1\}$$

where $(.,.)$ denotes the usual inner product in $\mathbb{R}^n$. ■

We will now take into account the second order derivatives of f to characterize extremal points.

Theorem 1.4.3 *Let $f : \mathcal{U} \subset E \to \mathbb{R}$ be differentiable in $\mathcal{U}$ and twice differentiable at $u \in \mathcal{U}$. If f admits a relative minimum at u, then, for all $v \in E$,*

$$f''(u)(v, v) \geq 0. \tag{1.4.3}$$

Proof: Let $v \neq 0$ be an arbitrary vector in E. Then, there exists an interval $J \subset \mathbb{R}$ containing the origin such that for all $t \in J$, we have that $u + tv \in \mathcal{U}$. Thus for all $t \in J$, it follows that

$f(u+tv) \geq f(u)$. By Taylor's formula (cf. Theorem 1.2.1) and the fact that $f'(u) = 0$, we get

$$0 \leq f(u+tv) - f(u) = \frac{t^2}{2}(f''(u)(v,v) + \varepsilon(t))$$

where $\varepsilon(t) \to 0$ as $t \to 0$, from which (1.4.3) follows. ■

Remark 1.4.4 If f admits a relative maximum at $u \in \mathcal{U}$, then, under the above conditions, the inequality in (1.4.3) will be reversed. ■

Theorem 1.4.4 *(Sufficient conditions) Let $f : \mathcal{U} \subset E \to \mathbb{R}$ be differentiable in $\mathcal{U}$ and let $u \in \mathcal{U}$ be such that $f'(u) = 0$.*
(i) If f is twice differentiable at u and if there exists $\alpha > 0$ such that

$$f''(u)(v,v) \geq \alpha\|v\|^2 \tag{1.4.4}$$

for all $v \in E$, then f admits a strict relative minimum at u.
(ii) If f is twice differentiable in $\mathcal{U}$ and there exists a ball $B(u;r) \subset \mathcal{U}$ such that

$$f''(v)(w,w) \geq 0 \tag{1.4.5}$$

for all $v \in B(u;r)$ and all $w \in E$, then f admits a relative minimum at u.

Proof: (i) For all w with sufficiently small norm, we have, by Taylor's formula,

$$f(u+w)-f(u) = \frac{1}{2}(f''(u)(w,w)+\|w\|^2\varepsilon(w)) \geq \frac{1}{2}(\alpha-\varepsilon(w))\|w\|^2$$

where $\varepsilon(w) \to 0$ as $\|w\| \to 0$. Thus there exists $r > 0$ such that as soon as $\|w\| < r$ we have $\varepsilon(w) < \alpha$. Then $f(u+w) > f(u)$ for all $u+w \in B(u;r)$ and so f admits a strict relative minimum at u.

(ii) Since f is real valued, there exists a v in the open interval $(u, u+w) \subset B(u;r)$ such that

$$f(u+w) = f(u) + \frac{1}{2}f''(v)(w,w).$$

Hence, by (1.4.5) we have that $f(u+w) \geq f(u)$ for all $u+w \in B(u;r)$. ■

Exercise 1.4.3 Let A be a symmetric $n \times n$ matrix with real entries. Define

$$f(v) = \frac{1}{2}(Av, v) - (b, v)$$

where $b \in \mathbb{R}^n$ is a given vector.
(i) Show that f admits a strict minimum in $\mathbb{R}^n$ if, and only if, A is positive definite.
(ii) Show that f attains its minimum if, and only if, A is positive semi-definite and the set $S = \{w \in \mathbb{R}^n \mid Aw = b\}$ is non-empty.
(iii) If the matrix A is positive semi-definite and the set S is empty, show that

$$\inf_{v \in \mathbb{R}^n} f(v) = -\infty.$$

(iv) If the infimum of f over $\mathbb{R}^n$ is a real number, show that the matrix A is positive semi-definite and that the set S is non-empty.
■

Chapter 2

The Brouwer Degree

2.1 Definition of the Degree

The topological degree is a useful tool in the study of existence of solutions to nonlinear equations. In this chapter, we will study the finite dimensional version of the degree, known as the Brouwer degree.

Let $\Omega \subset \mathbb{R}^n$ be a bounded open set. By $C^k(\overline{\Omega};\mathbb{R}^n)$, we denote the space of functions $f : \Omega \to \mathbb{R}^n$ which are k times differentiable in Ω such that these functions and all their derivatives upto order k can be extended continuously to $\overline{\Omega}$. We denote the boundary of Ω by $\partial\Omega$.

Let $f \in C^1(\overline{\Omega};\mathbb{R}^n)$. Recall that $f'(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ and hence $f'(x)$ can be represented by an $n \times n$ matrix. Let S be the set of critical points of f (cf. Definition 1.3.1).

Definition 2.1.1 *Let $f : \Omega \to \mathbb{R}^n$ be a function in $C^1(\overline{\Omega};\mathbb{R}^n)$ and let $b \notin f(S) \cup f(\partial\Omega)$. Then we define the* **degree** *of f in Ω with respect to b as*

$$d(f,\Omega,b) = \begin{cases} 0, & \text{if } f^{-1}(b) = \emptyset, \\ \sum_{x \in f^{-1}(b)} \operatorname{sgn}(J_f(x)), & \text{otherwise.} \end{cases} \quad ■ \qquad (2.1.1)$$

The function *sgn* denotes the sign (= +1 if positive and = −1 if negative) and $J_f(x)$ denotes the determinant of $f'(x)$.

Remark 2.1.1 We will now verify that the above definition makes sense. Since $b \notin f(S) \cup f(\partial\Omega)$, we know that $f'(x)$ is well defined for $x \in f^{-1}(b)$ and that $J_f(x) \neq 0$. Thus, $J_f(x)$ has a definite sign and, by the inverse function theorem, f is invertible in a neighbourhood of x. Consequently, since $\overline{\Omega}$ is compact, the set $f^{-1}(b)$ is finite and so (2.1.1) makes sense. ■

Example 2.1.1 Let $I : \mathbb{R}^n \to \mathbb{R}^n$ be the identity map and set $f = I|_\Omega$ for $\Omega \subset \mathbb{R}^n$. Then

$$d(f,\Omega,b) = \begin{cases} 1 & \text{if } b \in \Omega \\ 0 & \text{if } b \notin \overline{\Omega}. \end{cases}$$

More generally, if $T : \mathbb{R}^n \to \mathbb{R}^n$ is a nonsingular linear operator, and if $f = T|_\Omega$, we have

$$d(f,\Omega,b) = \begin{cases} \operatorname{sgn}(\det T) & \text{if } b \in T(\Omega) \\ 0 & \text{if } b \notin T(\overline{\Omega}). \end{cases}$$

Notice that $\operatorname{sgn}(\det T) = (-1)^\beta$, where β is the sum of the (algebraic) multiplicities of the negative eigenvalues of T.■

Example 2.1.2 Let $\Omega = (-1, 1)$ and define

$$f(x) = x^2 - \varepsilon^2, \text{ for } \varepsilon < 1.$$

Then, $f'(x) = 2x$ and $f^{-1}(0) = \{+\varepsilon, -\varepsilon\}$. Thus,

$$d(f,\Omega,0) = 0.■$$

Remark 2.1.2 We defined the degree to be zero if the value b were not attained by f. The converse, as seen by the above example, is false. However, if $d(f,\Omega,b) \neq 0$, the the solution set to the equation $f(x) = b$ is indeed non-empty. ■

We wish to extend the definition of the degree to functions which are merely continuous on $\overline{\Omega}$. To do this we need some preliminary results. We start with another formula for the degree.

Proposition 2.1.1 *Let $f \in C^1(\overline{\Omega}; \mathbb{R}^n)$ and let $b \notin f(S) \cup f(\partial\Omega)$. Then there exists ε_o such that, for all $0 < \varepsilon < \varepsilon_o$,*

$$d(f,\Omega,b) \;=\; \int_\Omega \varphi_\varepsilon(f(x)-b)J_f(x)dx \qquad (2.1.2)$$

where $\varphi_\varepsilon : \mathbb{R}^n \to \mathbb{R}$ is a C^∞ function whose support is contained in the ball $B(0;\varepsilon)$ with centre at 0 and radius ε and such that

$$\int_{\mathbb{R}^n} \varphi_\varepsilon(x)dx \;=\; 1. \qquad (2.1.3)$$

Proof: If $f^{-1}(b) = \emptyset$, then we choose $\varepsilon_o < \rho(b, f(\overline{\Omega}))$ (where $\rho(x, A)$ denotes the distance of the point x from the set A). If φ_ε is as above, we then have $\varphi_\varepsilon(f(x) - b) = 0$ and so (2.1.2) is trivially true.

Let us now assume that $f^{-1}(b) = \{x_1, x_2, ..., x_m\}$. For each $1 \leq i \leq m$, we have $J_f(x_i) \neq 0$, and so, by the inverse function theorem, there exists a neighbourhood $\mathcal{U}_i$ of x_i and a neighbourhood $\mathcal{V}_i$ of b such that the $\mathcal{U}_i$ are all pairwise disjoint and

$$f|_{\mathcal{U}_i} : \mathcal{U}_i \to \mathcal{V}_i$$

is a homeomorphism. Further, by shrinking the neighbourhoods if necessary, we can also ensure that $J_f|_{\mathcal{U}_i}$ has a constant sign. Now choose $\varepsilon_o > 0$ such that

$$B(b;\varepsilon_o) \subset \cap_{i=1}^m \mathcal{V}_i.$$

Set $W_i = f^{-1}(B(b;\varepsilon_o)) \cap \mathcal{U}_i$. Then the W_i are all pairwise disjoint and J_f is of constant sign in each of them. Hence, if $0 < \varepsilon < \varepsilon_o$, as $\varphi_\varepsilon(f(x) - b) = 0$ outside the sets W_i, we have

$$\begin{aligned}\int_\Omega \varphi_\varepsilon(f(x)-b)J_f(x)dx \;&=\; \sum_{i=1}^m \int_{W_i} \varphi_\varepsilon(f(x)-b)J_f(x)dx \\ &=\; \sum_{i=1}^m \mathrm{sgn}(J_f(x_i)) \int_{W_i} \varphi_\varepsilon(f(x)-b)|J_f(x)|dx \\ &=\; \sum_{i=1}^m \mathrm{sgn}(J_f(x_i)) \int_{B(0;\varepsilon)} \varphi_\varepsilon(y)dy\end{aligned}$$

by an obvious change of variable in each of the sets W_i and the right-hand side is exactly $d(f,\Omega,b)$ thanks to (2.1.3).■

We use the formula (2.1.2) to prove the robustness of the degree in the sense that it remains stable when b or f is slightly perturbed. In order to do this, we need a technical result.

Lemma 2.1.1 *Let $g \in C^2(\overline{\Omega}; \mathbb{R}^{n-1})$. Set*

$$B_i \;=\det(\partial_1 g, ..., \partial_{i-1}g, \partial_{i+1}g, ..., \partial_n g).$$

Then

$$\sum_{i=1}^n (-1)^i \partial_i B_i \;=\; 0. \qquad (2.1.4)$$

Proof: Let $1 \leq i \leq n$. Set $C_{ii} = 0$. If $j < i$, define

$$C_{ij} \;=\; \det(\partial_1 g, ..., \partial_{j-1}g, \partial_{ij}g, \partial_{j+1}g, ..., \partial_{i-1}g, \partial_{i+1}g, ..., \partial_n g)$$

and, if $j > i$, set

$$C_{ij} \;=\; \det(\partial_1 g, ..., \partial_{i-1}g, \partial_{i+1}g, ..., \partial_{j-1}g, \partial_{ij}g, \partial_{j+1}g, ..., \partial_n g).$$

Then, clearly, $\partial_i B_i = \sum_{j=1}^n C_{ij}$, by the rule for differentiating determinants. Thus the left hand side of (2.1.4) equals

$$\sum_{i,j=1}^n (-1)^i C_{ij}.$$

Since g is C^2, $\partial_{ij}g = \partial_{ji}g$ and so, by the property of determinants relating to transposition of columns, it is easy to see that

$$C_{ij} \;=\; (-1)^{j+i-1}C_{ji}$$

and the lemma follows easily.■

Lemma 2.1.2 *Let $f \in C^2(\overline{\Omega}; \mathbb{R}^n)$. Let $A_{ij}(x)$ denote the cofactor of the entry $\partial_i f_j(x)$ in $J_f(x)$. Then for all $1 \leq j \leq n$,*

$$\sum_{i=1}^n \partial_i A_{ij} \;=\; 0. \qquad (2.1.5)$$

Proof: Recall that A_{ij} is given by

$$A_{ij} = (-1)^{i+j} \det(\partial_l f_k)_{k\neq j, l\neq i}.$$

For fixed j, we apply the preceding lemma to

$$g = (f_1, ..., f_{j-1}, f_{j+1}, ..., f_n)$$

to get the desired result.■

Remark 2.1.3 The above lemma is essentially a consequence of the fact that the order of derivation is immaterial for C^2 functions (and this was used in the proof of Lemma 2.1.1). For instance, if $n = 2$, then

$$A_{11} = \partial_2 f_2, \quad A_{21} = -\partial_1 f_2$$
$$A_{12} = -\partial_2 f_1, \quad A_{22} = \partial_1 f_1,$$

and we readily verify that, if f is C^2,

$$\partial_1 A_{11} + \partial_2 A_{21} = \partial_1 A_{12} + \partial_2 A_{22} = 0.■$$

Proposition 2.1.2 *Let $f \in C^2(\overline{\Omega}; \mathbb{R}^n)$ and let $b \notin f(\partial\Omega)$. Let $\rho_o = \rho(b, f(\partial\Omega)) > 0$. Let $b_i \in B(b; \rho_o)$ for $i = 1, 2$. If $b_i \notin f(S)$, we have*

$$d(f, \Omega, b_1) = d(f, \Omega, b_2).$$

Proof: Clearly, by choice, $b_i \notin f(\partial\Omega)$. Thus, by hypothesis, the degree $d(f, \Omega, b_i)$ is well-defined for $i = 1, 2$. Let

$$\delta < \rho_o - |b - b_i|, \ i = 1, 2.$$

Then there exists $\varepsilon < \delta$ such that

$$d(f, \Omega, b_i) = \int_\Omega \varphi_\varepsilon(f(x) - b_i) J_f(x) dx, \ i = 1, 2,$$

where φ_ε is as in Proposition 2.1.1. Then

$$\begin{aligned} \varphi_\varepsilon(y - b_2) - \varphi_\varepsilon(y - b_1) &= \textstyle\int_0^1 \frac{d}{dt} \varphi_\varepsilon(y - b_1 + t(b_1 - b_2)) dt \\ &= (b_1 - b_2). \textstyle\int_0^1 \nabla \varphi_\varepsilon(y - b_1 + t(b_1 - b_2)) dt \\ &= \operatorname{div}(w(y)) \end{aligned}$$

where

$$w(y) = \left(\int_0^1 \varphi_\varepsilon(y - b_1 + t(b_1 - b_2)) dt \right) .(b_1 - b_2).$$

Now, if $y \in f(\partial\Omega)$,

$$\begin{aligned} |y - (1-t)b_1 - tb_2| &= |(y - b) + (1-t)(b - b_1) + t(b - b_2)| \\ &> \rho_o - (1-t)(\rho_o - \delta) - t(\rho_o - \delta) \\ &= \delta > \varepsilon. \end{aligned}$$

Since the support of φ_ε is contained in $B(0; \varepsilon)$, it follows that $w(y) = 0$ for $y \in f(\partial\Omega)$. Now, for $1 \le i \le n$, define

$$v_i = \begin{cases} \sum_{j=1}^n w_j(f(x)) A_{ij}(x), & x \in \overline{\Omega} \\ 0, & x \in \mathbb{R}^n \backslash \overline{\Omega}. \end{cases}$$

By the preceding considerations, $v_i = 0$ on $\partial\Omega$. Now

$$\frac{\partial v_i}{\partial x_i} = \begin{cases} \sum_{j,k=1}^n \frac{\partial w_j}{\partial x_k}(f(x)) \frac{\partial f_k}{\partial x_i}(x) A_{ij}(x) \\ \\ + \sum_{j=1}^n w_j(f(x)) \frac{\partial}{\partial x_i} A_{ij}(x). \end{cases}$$

Thus,

$$\operatorname{div}(v(x)) = \begin{cases} \sum_{j,k=1}^n \frac{\partial w_j}{\partial x_k}(f(x)) \left(\sum_{i=1}^n \frac{\partial f_k}{\partial x_i}(x) A_{ij}(x) \right) \\ + \sum_{j=1}^n w_j(f(x)) \left(\sum_{i=1}^n \frac{\partial}{\partial x_i} A_{ij}(x) \right). \end{cases}$$

By Lemma 2.1.2, the second term on the right-hand side vanishes. Notice that, by the definition of the A_{ij},

$$\sum_{i=1}^n \frac{\partial f_k}{\partial x_i}(x) A_{ij}(x) = \delta_{jk} J_f(x).$$

Thus,

$$\operatorname{div}(v(x)) = \sum_{j=1}^n \frac{\partial w_j}{\partial x_j}(f(x)) J_f(x) = \operatorname{div}(w(f(x)) J_f(x).$$

Hence it follows that

$$\begin{aligned} d(f,\Omega,b_2) - d(f,\Omega,b_1) &= \int_\Omega \operatorname{div}(w(f(x))J_f(x)dx \\ &= \int_\Omega \operatorname{div}(v(x))dx \;=\; 0 \end{aligned}$$

since v vanishes on $\partial\Omega$.■

Let $f \in C^2(\overline{\Omega};\mathbb{R}^n)$ and $b \notin f(\partial\Omega)$. Let ρ_o be as in the previous proposition. Since, by Sard's theorem (cf. Theorem 1.3.4), the singular values of f are of measure zero, there exist regular values in the ball $B(b;\rho_o)$ and the degrees of all such points are the same by the previous proposition. We are thus led to the following definition.

Definition 2.1.2 *Let $f \in C^2(\overline{\Omega};\mathbb{R}^n)$ and $b \notin f(\partial\Omega)$. Set $\rho_o = \rho(b, f(\partial\Omega))$. The* **degree** *of f in Ω with respect to b is defined as*

$$d(f,\Omega,b) \;=\; d(f,\Omega,b') \tag{2.1.6}$$

where b' is any regular value in $B(b,\rho_o)$.■

Exercise 2.1.1 Let f, b and ρ_o be as above. If $|b_1 - b| < \rho_o/2$, show that

$$d(f,\Omega,b_1) \;=\; d(f,\Omega,b).■$$

Exercise 2.1.2 Let $\mathbf{C}$ denote the complex plane and let $\Omega \subset \mathbf{C}$ be a bounded open set. Let $f : \overline{\Omega} \to \mathbf{C}$ be a function which is holomorphic in Ω. We identify $\mathbf{C}$ with $\mathbb{R}^2$ in the usual way using the canonical correspondence $z = x + iy \in \mathbf{C} \leftrightarrow (x,y) \in \mathbb{R}^2$. Thus Ω can be considered as a bounded open set in $\mathbb{R}^2$ and we can consider f as a map from Ω into $\mathbb{R}^2$ via the correspondence $f = u + iv \leftrightarrow f = (u,v)$.
(i) Show that $J_f(x,y) = |f'(z)|^2$, for $(x,y) \in \Omega$.
(ii) Compute $d(f,D,0)$ where D is the unit disc in the complex plane and $f(z) = z^n$.■

Proposition 2.1.3 *Let $f, g \in C^2(\overline{\Omega};\mathbb{R}^n)$ and let $b \notin f(\partial\Omega)$. Then there exists $\varepsilon = \varepsilon(f,g,\Omega)$, such that for $0 < |t| < \varepsilon$,*

$$d(f+tg,\Omega,b) \;=\; d(f,\Omega,b). \tag{2.1.7}$$

Proof: Case 1. Let $b \notin f(\overline{\Omega})$. Then $\widetilde{\rho} = \rho(b, f(\overline{\Omega})) > 0$. Set $\varepsilon = \widetilde{\rho}/2\|g\|_\infty$ (where $\|.\|_\infty$ denotes the norm in $C(\overline{\Omega};\mathbb{R}^n)$). If $|t| < \varepsilon$, then $\rho(b,(f+tg)(\overline{\Omega})) \geq \widetilde{\rho}/2 > 0$ and thus $b \notin (f+tg)(\overline{\Omega})$. Hence (2.1.7) is trivially true as both sides vanish.

Case 2. Let $b \notin f(S)$ and let $f^{-1}(b) = \{x_1,...,x_m\}$ so that $J_f(x_i) \neq 0$ for $1 \leq i \leq m$. Define

$$h(t,x) \;=\; f(x) + tg(x) - b.$$

Then, for $1 \leq i \leq m$,

$$\begin{aligned} h(0,x_i) &= 0 \\ \partial_x h(0,x_i) &= f'(x_i) \end{aligned}$$

and $f'(x_i)$ is invertible, by assumption. Hence, by the implicit function theorem, there exist neighbourhoods $(-\varepsilon_i,\varepsilon_i)$ of 0 in $\mathbb{R}$ and pairwise disjoint neighbourhoods $\mathcal{U}_i$ of x_i in Ω and functions $\varphi_i : (-\varepsilon_i,\varepsilon_i) \to \mathcal{U}_i$ such that the only solutions of $h(t,x) = 0$ in $(-\varepsilon_i,\varepsilon_i) \times \mathcal{U}_i$ are of the form $(t,\varphi_i(t))$. Further, by shrinking the neighbourhoods if necessary, we can ensure that $sgn(J_{f+tg}(x)) = sgn(J_f(x_i))$ in each $\mathcal{U}_i$. Now set

$$\varepsilon \;=\; \min_{1\leq i\leq m} \varepsilon_i.$$

The relation (2.1.7) now follows from the definition of the degree in the regular case.

Case 3. Assume now that $b \in f(S)$. Let ρ_o be the distance of b from $f(\partial\Omega)$. Choose $b_1 \in B(b,\rho_o/3)$ such that b_1 is regular and so there exists $\varepsilon_o > 0$ such that for all $0 < |t| < \varepsilon_o$,

$$d(f+tg,\Omega,b_1) \;=\; d(f,\Omega,b_1) \;=\; d(f,\Omega,b).$$

Now choose $\varepsilon < \min\{\varepsilon_o, \rho_o/3\|g\|_\infty\}$. Clearly, $b \notin (f+tg)(\partial\Omega)$ for $|t| < \varepsilon$. In fact, $\rho(b,(f+tg)(\partial\Omega)) \geq 2\rho_o/3$ while

$$|b - b_1| \;<\; \rho_o/3 \;\leq\; \frac{1}{2}\rho(b,(f+tg)(\partial\Omega)).$$

Consequently (cf. Exercise 2.1.1),

$$d(f + tg, \Omega, b_1) = d(f + tg, \Omega, b)$$

and the proof is complete. ■

We are now in a position to define the degree for all continuous functions. Let $f \in C(\overline{\Omega}; \mathbb{R}^n)$ and let $b \notin f(\partial\Omega)$. Let $\rho_o = \rho(b, f(\partial\Omega))$. We can always find $g \in C^2(\overline{\Omega}; \mathbb{R}^n)$ such that $\|g - f\|_\infty < \rho_o/2$. Then clearly, $b \notin g(\partial\Omega)$ and the degree $d(g, \Omega, b)$ is well defined. If g_1 and g_2 are two such functions, set $\widetilde{g} = g_1 - g_2$. Then, for $0 < t < 1$, we have $\|f - (g_2 + t\widetilde{g})\|_\infty < \rho_o$ and, by Proposition 2.1.3, the function

$$d(t) \;=\; d(g_2 + t\widetilde{g}, \Omega, b)$$

is locally constant, and hence, by the connectedness of $[0, 1]$, is constant on this interval. Thus

$$d(g_1, \Omega, b) \;=\; d(g_2, \Omega, b).$$

This paves the way for the following definition.

Definition 2.1.3 *Let f, b and ρ_o be as above. Then the* **degree** *of f in Ω with respect to b is given by*

$$d(f, \Omega, b) \;=\; d(g, \Omega, b) \tag{2.1.8}$$

for any $g \in C^2(\overline{\Omega}; \mathbb{R}^n)$ such that $\|f - g\|_\infty < \rho_o/2$. ■

Remark 2.1.4 Compare this with the result of Exercise 2.1.1. ■

Proposition 2.1.4 *Let $f \in C(\overline{\Omega}; \mathbb{R}^n)$ and let $b \notin f(\partial\Omega)$. Then*

$$d(f, \Omega, b) \;=\; d(f - b, \Omega, 0). \tag{2.1.9}$$

Proof: If $\rho_o = \rho(b, f(\partial\Omega)) = \rho(0, (f - b)(\partial\Omega))$ and if g is C^2 such that $\|g - f\|_\infty < \rho_o/2$, then

$$\|(g - b) - (f - b)\|_\infty \;=\; \|g - f\|_\infty \;<\; \rho_o/2$$

and so, by definition,

$$d(f, \Omega, b) \;=\; d(g, \Omega, b) \;\text{ and }\; d(f - b, \Omega, 0) \;=\; d(g - b, \Omega, 0).$$

If b is a singular value of g, then we can find a regular value b_1 of g such that

$$|b - b_1| \;<\; \rho(b, g(\partial\Omega))/2$$

so that

$$d(g - b_1, \Omega, 0) \;=\; d(g - b, \Omega, 0) \;\text{ and }\; d(g, \Omega, b_1) \;=\; d(g, \Omega, b).$$

Since b_1 is a regular value of g, it is trivial to see that

$$d(g, \Omega, b_1) \;=\; d(g - b_1, \Omega, 0)$$

and the proof is complete. ■

2.2 Properties of the Degree

In this section, we prove the basic properties of the Brouwer degree and look at some of their simple consequences.

Theorem 2.2.1 *(i) (Continuity with respect to the function) Let $f \in C(\overline{\Omega}; \mathbb{R}^n)$ and let $b \notin f(\partial\Omega)$. There exists a neighbourhood $\mathcal{U}$ of f in $C(\overline{\Omega}; \mathbb{R}^n)$ such that for every $g \in \mathcal{U}$,*

$$d(g, \Omega, b) \;=\; d(f, \Omega, b). \tag{2.2.1}$$

(ii) (Invariance under homotopy) Let $H \in C(\overline{\Omega} \times [0, 1]; \mathbb{R}^n)$ such that $b \notin H(\partial\Omega \times [0, 1])$. Then $d(H(., t), \Omega, b)$ is independent of t.
(iii) The degree is constant, with respect to b, in each connected component of $\mathbb{R}^n \backslash f(\partial\Omega)$.
(iv) (Additivity) Let $\Omega_1 \cap \Omega_2 = \emptyset$ and $b \notin f(\partial\Omega_1) \cup f(\partial\Omega_2)$, where $f \in C(\overline{\Omega}; \mathbb{R}^n)$, $\Omega = \Omega_1 \cup \Omega_2$. Then

$$d(f, \Omega, b) \;=\; d(f, \Omega_1, b) + d(f, \Omega_2, b). \tag{2.2.2}$$

Proof: (i) Define

$$\mathcal{U} = \{g \in C(\overline{\Omega}; \mathbb{R}^n) \mid \|f - g\|_\infty < \rho_o/4\}$$

where $\rho_o = \rho(b, f(\partial\Omega))$. If $g \in \mathcal{U}$, then $\rho(b, g(\partial\Omega)) \geq 3\rho_o/4$. Thus $b \notin g(\partial\Omega)$ and the degree is well-defined. Let $h \in C^2(\overline{\Omega}; \mathbb{R}^n)$ such that $\|f - h\|_\infty < \rho_o/8$. Then

$$\|g - h\| < \frac{3}{8}\rho_o \leq \frac{1}{2}\rho(b, g(\partial\Omega)).$$

Hence, by definition,

$$d(g, \Omega, b) = d(h, \Omega, b) = d(f, \Omega, b).$$

(ii) By the preceding step, $d(H(.,t), \Omega, b)$ is locally constant and hence continuous and therefore constant on $[0, 1]$ by connectedness.

(iii) By virtue of (2.1.9), $d(f, \Omega, b) = d(f - b, \Omega, 0)$ and so if $|b - b_1|$ is small, $d(f - b, \Omega, 0) = d(f - b_1, \Omega, 0)$. Thus, the degree is locally constant and thus continuous and so constant on connected components.

(iv) Let ρ_o be as in Step (i) and let g be a C^2 function such that $\|f - g\|_\infty < \rho_o/2$.Then, it is clear that

$$\begin{aligned} d(g, \Omega, b) &= d(f, \Omega, b) \\ d(g, \Omega_i, b) &= d(f, \Omega_i, b), \ i = 1, 2. \end{aligned}$$

Now $B = B(b; \rho_o/2)$ is connected and is contained in $\mathbb{R}^n \backslash g(\partial\Omega)$ as well as in $\mathbb{R}^n \backslash g(\partial\Omega_i)$ for $i = 1, 2$ and hence in one connected component of each of these sets. By Sard's theorem, there exists $c \in B$ such that it is a regular value of g and so

$$d(g, \Omega, c) = d(g, \Omega, b) \quad \text{and} \quad d(g, \Omega_i, c) = d(g, \Omega_i, b)$$

for $i = 1, 2$. Since g is C^2 and c is regular, it readily follows from the definition of the degree that

$$d(g, \Omega, c) = d(g, \Omega_1, c) + d(g, \Omega_2, c)$$

and the result follows. ■

Exercise 2.2.1 Let $f : [a, b] \to \mathbb{R}$ be continuous and such that $f(a)f(b) \neq 0$. Show that

$$d(f, (a, b), 0) = \frac{1}{2}[sgn(f(b)) - sgn(f(a))].■$$

Proposition 2.2.1 *If $f \in C(\overline{\Omega}; \mathbb{R}^n)$ and $b \notin f(\overline{\Omega})$, then $d(f, \Omega, b) = 0$. Equivalently, if $d(f, \Omega, b) \neq 0$, then there exists $x \in \Omega$ such that $f(x) = b$.*

Proof: Let $\rho_o = \rho(b, f(\overline{\Omega}))$. If g is C^2 such that $\|f - g\|_\infty < \rho_o/2$, then $b \notin g(\overline{\Omega})$. Thus, as b is now a regular 'value' of g, we have that $d(g, \Omega, b) = 0$ and the result follows. ■

Corollary 2.2.1 *If $d(f, \Omega, b) \neq 0$, then $f(\Omega)$ is a neighbourhood of b.*

Proof: Let C_b be the connected component of $\mathbb{R}^n \backslash f(\partial\Omega)$ containing b. Then, for all $c \in C_b$, we have $d(f, \Omega, c) \neq 0$ and so, by the preceding proposition, $C_b \subset f(\Omega)$ and the conclusion follows as C_b is open. ■

Exercise 2.2.2 If $f(\Omega)$ is contained in a proper subspace of $\mathbb{R}^n$, show that, for all $b \notin f(\partial\Omega)$, $d(f, \Omega, b) = 0$.■

Exercise 2.2.3 Let $f(z) = z^n + a_1 z^{n-1} + a_2 z^{n-2} + ... + a_n$ be a polynomial with complex coefficients in the complex variable z.
(i) Assume that $|a_1| + |a_2| + ... + |a_n| < 1$. Then using the properties of the degree (cf. Exercise 2.1.2), show that f has a root in the unit disc $D \subset \mathbb{C}$.
(ii) Using the change of variable $z = cw$ where $c > 0$, show that we can reduce the search for a root of a general polynomial f to the preceding case and thus prove the fundamental theorem of algebra. ■

Proposition 2.2.2 *(Excision) Let $K \subset \overline{\Omega}$ be a closed set and let $b \notin f(\partial\Omega) \cup f(K)$. Then*

$$d(f, \Omega, b) \;=\; d(f, \Omega \backslash K, b).$$

Proof: Choose g, a C^2 function such that $d(g, \Omega, b) = d(f, \Omega, b)$ and such that $b \notin g(K)$. Now choose c, a regular value of g, close enough to b such that $c \notin g(K)$ and belonging to the same connected component as b in $\mathbb{R}^n \backslash g(\partial\Omega)$ and $\mathbb{R}^n \backslash g(\partial(\Omega \backslash K))$. The result now follows from the definition of the degree in the regular case. ■

The following two exercises can also be solved by reducing the problem to the regular case.

Exercise 2.2.4 Let $\{\Omega_j\}_{j \in J}$ be a family of pairwise disjoint open sets in $\mathbb{R}^n$ whose union is contained in a bounded open set Ω. Let $f \in C(\overline{\Omega}; \mathbb{R}^n)$ and b such that $f^{-1}(b) \subset \cup_{j \in J} \Omega_j$. Show that $d(f, \Omega_j, b) = 0$ for all but a finite number of $j \in J$ and that

$$d(f, \Omega, b) \;=\; \sum_{j \in J} d(f, \Omega_j, b). \blacksquare$$

Exercise 2.2.5 (Product Formula) For $i = 1, 2$, let $\varphi_i \in C(\overline{\Omega}_i; \mathbb{R}^{n_i})$ where $\Omega_i \subset \mathbb{R}^{n_i}$ is a bounded open set. Let $b_i \notin \varphi_i(\partial\Omega_i)$. Show that

$$d((\varphi_1, \varphi_2), \Omega_1 \times \Omega_2, (b_1, b_2)) \;=\; d(\varphi_1, \Omega_1, b_1).d(\varphi_2, \Omega_2, b_2). \blacksquare$$

Proposition 2.2.3 *Let $f, g \in C(\overline{\Omega}; \mathbb{R}^n)$ such that $f = g$ on $\partial\Omega$. Let $b \notin f(\partial\Omega)$. Then*

$$d(f, \Omega, b) \;=\; d(g, \Omega, b).$$

Proof: Define $H \in C(\overline{\Omega} \times [0,1]; \mathbb{R}^n)$ by

$$H(x,t) \;=\; tf(x) + (1-t)g(x).$$

Then $H(.,t) = f = g$ on the boundary and so $d(H(.,t), \Omega, b)$ is defined and independent of t and the result follows by successively setting $t = 0$ and $t = 1$.■

Corollary 2.2.2 *Let $f, g \in C(\overline{\Omega}; \mathbb{R}^n)$. Assume that there exists $H \in C(\partial\Omega \times [0,1]; \mathbb{R}^n)$ such that H never assumes the value b and such that $H(.,0) = f|_{\partial\Omega}$ and $H(.,1) = g|_{\partial\Omega}$. Then*

$$d(f, \Omega, b) \;=\; d(g, \Omega, b).$$

Proof: By Tietze's theorem, we can extend H to $\widetilde{H} \in C(\overline{\Omega} \times [0,1]; \mathbb{R}^n)$. Set $\widetilde{H}(.,0) = \widetilde{f}$ and $\widetilde{H}(.,1) = \widetilde{g}$. Then, by homotopy invariance of the degree,

$$d(\widetilde{f}, \Omega, b) \;=\; d(\widetilde{g}, \Omega, b).$$

The result now follows from the previous proposition since $f = \widetilde{f}$ and $g = \widetilde{g}$ on the boundary. ■

Remark 2.2.1 The above proposition and its corollary imply that, as long as the value b is not attained on the boundary along a homotopy, the degree is essentially determined by homotopy classes of continuous functions defined on the boundary. If S^n is the unit sphere in R^{n+1}, and if 0 is not attained on it for a continuous function $f : \overline{B}^{n+1} \to \mathbb{R}^{n+1}$, where B^{n+1} is the open unit ball in $\mathbb{R}^{n+1}$, we can consider $\widetilde{f}(x) = f(x)/|f(x)|$ which then maps S^n into itself. We can define

$$d(\widetilde{f}) \;=\; d(f, B^{n+1}, 0).$$

Then $d(.)$ will be constant on homotopy classes of continuous maps of S^n into itself. This gives rise to a theory of a topological degree for such maps. We can also define a degree of continuous maps $\widetilde{f} : S^n \to S^n$ in another way. We know that the singular homology groups of S^n are given by

$$H_m(S^n) \;=\; \begin{cases} \mathbf{Z} & \text{if } m = 0, \text{ or } n \\ \mathbf{0} & \text{otherwise.} \end{cases}$$

Thus $\widetilde{f}$ generates a homomorphism

$$f_{\#} : H_n(S^n) \to H_n(S^n)$$

and, as $H_n(S^n) \cong \mathbf{Z}$, $f_\#$ is completely determined by $f_\#(1) \in \mathbf{Z}$. It turns out that $d(\widetilde{f}) = f_\#(1)$. In the case $n = 2$, this is the familiar winding number for functions on S^1.■

Exercise 2.2.6 If n is odd, show that there does not exist a homotopy $H : S^{n-1} \times [0,1] \to S^{n-1}$ such that $H(x,0) = x$ and $H(x,1) = -x$ for all $x \in S^{n-1}$.■

Proposition 2.2.4 *(Hairy Ball Theorem) If n is odd, there is no non-vanishing vector field on S^{n-1}, i.e., there is no continuous map $\varphi : S^{n-1} \to \mathbb{R}^n$ such that $\varphi(x) \neq 0$ and $(\varphi(x), x) = 0$ (where $(.,.)$ denotes the usual inner-product in $\mathbb{R}^n$) for all $x \in S^{n-1}$.*

Proof: If such a map were to exist, set $\psi(x) = \varphi(x)/|x|$. Then

$$H(x,t) \;=\; \cos(\pi t)x + \sin(\pi t)\psi(x)$$

defines a homotopy as in the preceding exercise, which is impossible. ■

Remark 2.2.2 The map $(x,y) \mapsto (y,-x)$ is a non-vanishing vector field on S^1.■

2.3 Brouwer's Theorem and Applications

Proposition 2.3.1 *There is no retraction from the closed unit ball $\overline{B}^n$ in $\mathbb{R}^n$ onto S^{n-1}, i.e., there does not exist a continuous map $\varphi : \overline{B}^n \to S^{n-1}$, such that $\varphi(x) = x$ for all $x \in S^{n-1}$.*

Proof If such a map existed, then $0 \notin \varphi(\overline{B}^n)$ and since $\varphi = I$, the identity map, on S^{n-1}, we have

$$0 \;=\; d(\varphi, B^n, 0) \;=\; d(I, B^n, 0) \;=\; 1$$

which is impossible. ■

Theorem 2.3.1 *(Brouwer's Fixed Point Theorem) Let $f : \overline{B}^n \to \overline{B}^n$ be continuous. Then f has a fixed point.*

Proof: Assume that f has no fixed point. Then $f(x) \neq x$ for every $x \in \overline{B}^n$. The line segment starting at $f(x)$ and going to x is then well-defined and can be produced in the same direction to meet S^{n-1} at a point that we denote by $\varphi(x)$. Then $\varphi : \overline{B}^n \to S^{n-1}$ is clearly a retraction and we thus get a contradiction to the previous result. ■

Remark 2.3.1 We can describe the mapping φ above analytically as follows. We look for $\lambda \geq 1$ such that

$$|\lambda x + (1-\lambda)f(x)|^2 \;=\; 1$$

which yields

$$|x - f(x)|^2\lambda^2 + 2(x - f(x), f(x))\lambda + (|f(x)|^2 - 1) \;=\; 0.$$

Since $f(x) - x \neq 0$ for all x, this quadratic equation in λ has exactly two roots. The product of the roots is non-positive since $|f(x)| \leq 1$. Hence there are two real roots, one positive and the other negative. Since, at $\lambda = 1$, the value of the quadratic expression is $|x|^2 - 1 \leq 0$, the positive root is greater than or equal to unity and this root is continuously dependent on x and equal to unity on S^{n-1}.■

Remark 2.3.2 Obviously, Brouwer's theorem holds for any closed ball in $\mathbb{R}^n$.■

Exercise 2.3.1 Show that the following statements are equivalent:

(i) There is no retraction from a closed ball in $\mathbb{R}^n$ onto its boundary.

(ii) Every continuous map of a closed ball in $\mathbb{R}^n$ into itself has a fixed point.

(iii) Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be continuous. Let $R > 0$ such that for all $|x| = R$, we have $(f(x), x) \geq 0$. Then there exists x_o such that $|x_o| \leq R$ and $f(x_o) = 0$.■

Exercise 2.3.2 Prove Brouwer's theorem directly from the properties of the degree. ■

Corollary 2.3.1 *Let $K \subset \mathbb{R}^n$ be a compact and convex subset. Let $f : K \to K$ be continuous. Then f has a fixed point.*

Proof: If K is compact, there exists a ball $\overline{B}(0;R)$ containing K. Since K is closed and convex, let $P_K : \mathbb{R}^n \to K$ be the projection map, *i.e.* given $x \in \mathbb{R}^n$, $P_K(x) \in K$ is the unique point such that

$$|x - P_K(x)| = \min_{y\in K}|x-y|.$$

Define $\widetilde{f} : \overline{B}(0;R) \to K \subset \overline{B}(0;R)$ by $\widetilde{f}(x) = f(P_K(x))$. Then $\widetilde{f}$ has a fixed point and as the image of this map is contained in K, it follows that this fixed point x_o is in K. But then $P_K(x_o) = x_o$ and so

$$x_o = f(P_K(x_o)) = f(x_o)$$

which proves the result. ■

We now illustrate the use of Brouwer's theorem via some examples.

Example 2.3.1 Let A be an $n \times n$ matrix such that all its coefficients are non-negative. Then A has a non-negative eigenvalue with an associated eigenvector whose components are all non-negative as well. To see this, set

$$K = \{x \in \mathbb{R}^n \mid x_i \geq 0, 1 \leq i \leq n, \text{ and } \sum_{i=1}^n x_i = 1.\}$$

This is a compact convex set in $\mathbb{R}^n$. If there exists $x_o \in K$ such that $Ax_o = 0$, then 0 is an eigenvalue and we are through. If not, $Ax \neq 0$ for all $x \in K$and so $\sum_{i=1}^n (Ax)_i$ attains a strict positive minimum in K. Define $f : K \to K$ by

$$f_i(x) = \frac{(Ax)_i}{\sum_{j=1}^n (Ax)_j}.$$

Thus, f has a fixed point $x_o \in K$ and we have $Ax_o = \lambda x_o$ where $\lambda = \sum_{j=1}^n (Ax_o)_j$.■

Remark 2.3.3 The Perron - Fröbenius theorem states that if, in addition A satisfies a condition called irreducibility, then , in fact, the spectral radius is itself a (simple) eigenvalue and we have an eigenvector whose components are all strictly positive. ■

Example 2.3.2 (Periodic solutions, cf. Deimling [7]) Let $f : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be ω - periodic in t, *i.e.* $f(t+\omega, x) = f(t,x)$ for every $(t,x) \in \mathbb{R}\times\mathbb{R}^n$. Consider the system of ordinary differential equations,

$$u'(t) = f(t,u). \tag{2.3.1}$$

Let us assume that f is continuous and that there exists a ball $B(0;r) \subset \mathbb{R}^n$ such that for every $x \in \overline{B}(0;r)$, the initial value problem

$$\left.\begin{array}{rcl} u'(t) &=& f(t,u) \\ u(0) &=& x \end{array}\right\} \tag{2.3.2}$$

has a unique solution $u(t;x)$ on $[0,\infty)$ which continuously depends on the initial value x. Thus the map $P_t : \overline{B}(0;r) \to \mathbb{R}^n$ defined by $P_t(x) = u(t;x)$ is continuous. Now assume further that the following condition holds:
(H) For every $t \in [0,\omega]$, and for every x such that $|x| = r$, we have

$$(f(t,x),x) < 0.$$

Then $P_t : \overline{B}(0;r) \to \overline{B}(0;r)$. For, if $|u(t;x)| = r$, then

$$\frac{d}{dt}(|u(t;x)|^2) = 2(u'(t;x),u(t;x)) = (f(t,u(t;x)),u(t;x)) < 0.$$

Hence, by the Brouwer fixed point theorem, P_ω (in particular) has a fixed point, *i.e.* there exists $x_o \in \overline{B}(0;r)$ such that $u(\omega;x_o) = x_o$. Now define

$$v(t) = u(t-k\omega;x_o) \quad \text{for} \quad t \in [k\omega,(k+1)\omega].$$

By the ω-periodicity of f, it follows that v is a ω-periodic solution of (2.3.1). P_ω is called the Poincaré operator associated to (2.3.1). ■

We conclude with an example of the Galerkin method. This method is very useful in constructing solutions of nonlinear equations. The theorem which follows is an abstract result with applications to nonlinear partial differential equations (cf. Lions [16]).

Theorem 2.3.2 *Let H be a separable Hilbert space with scalar product $(.,.)$ and let $A : H \to H$ be a map such that*
(i) A is monotone, i.e. for every $u, v \in H$,

$$(Au - Av, u - v) \ \geq\ 0; \qquad (2.3.3)$$

(ii) A is hemi-continuous, i.e. for every $u, v \in H$ fixed, the map $\lambda \mapsto A(u + \lambda v)$ is continuous;
(iii) A maps bounded sets into bounded sets.
Then, given any $f \in H$, there exists a unique solution $u \in H$ of the equation,

$$u + Au \ =\ f \qquad (2.3.4)$$

and, further,

$$\|u\| \ \leq\ \|A0 - f\|. \qquad (2.3.5)$$

Proof: Step 1. (Uniqueness) If u_1 and u_2 were two solutions of (2.3.4), we have

$$u_1 - u_2 + Au_1 - Au_2 \ =\ 0.$$

Thus,

$$\|u_1 - u_2\|^2 + (Au_1 - Au_2, u_1 - u_2) \ =\ 0$$

and hence, by virtue of (2.3.3), we have $u_1 - u_2 = 0$.

Step 2. (*A priori* Estimate) If $u \in H$ is a solution of (2.3.4), then

$$\|u\|^2 + (Au - A0, u) \ =\ (f - A0, u)$$

and the estimate (2.3.5) follows, again thanks to (2.3.3).

Step 3. Let $W \subset H$ be a finite dimensional subspace with an orthonormal basis $\{w_1, w_2, ..., w_k\}$. Given $\mathbf{v} = (v_1, v_2, ..., v_k) \in \mathbb{R}^k$, we associate with it $v \in W$ by setting $v = \sum_{i=1}^k v_i w_i$. Define $T : \mathbb{R}^k \to \mathbb{R}^k$ by

$$(T\mathbf{v})_i = (v, w_i) + (Av, w_i) - (f, w_i).$$

Then, T is continuous and (denoting the usual inner product in $\mathbb{R}^k$ by $(.,.)_k$) we have

$$\begin{aligned}(T\mathbf{v}, \mathbf{v})_k &= \|v\|^2 + (Av, v) - (f, v)\\ &= \|v\|^2 + (Av - A0, v) - (f - A0, v)\\ &\geq \|v\|^2 - \|A0 - f\|.\|v\|.\end{aligned}$$

Setting $R = \|A0 - f\|$, we have that $(T\mathbf{v}, \mathbf{v})_k \geq 0$ for all $|\mathbf{v}| = \|v\| = R$. Hence, by Brouwer's theorem (cf. Exercise 2.3.1), there esists a $\tilde{\mathbf{u}} \in \mathbb{R}^k$ such that $|\tilde{\mathbf{u}}| \leq R$ and $T\tilde{\mathbf{u}} = 0$. Thus, $\tilde{u} \in W$ satisfies

$$(\tilde{u}, w_i) + (A\tilde{u}, w_i) \ =\ (f, w_i),\ 1 \leq i \leq k$$

i.e.

$$(\tilde{u}, v) + (A\tilde{u}, v) \ =\ (f, v) \ \text{ for all } \ v \in W$$

and $\|\tilde{u}\| \leq \|A0 - f\|$.

Step 4. Let $\{w_n\}$ be a complete orthonormal basis for H (which is separable) and set $W_k = span\{w_1, w_2, ..., w_k\}$. Let $u_n \in W_n$ verify

$$\|u_n\| \ \leq\ \|A0 - f\| \qquad (2.3.6)$$

and

$$(u_n, v) + (Au_n, v) \ =\ (f, v) \ \text{ for all } \ v \in W_n \qquad (2.3.7)$$

as guaranteed by the result of Step 3. Thus, upto the extraction of a subsequence, $u_n \rightharpoonup u$ weakly in H.

Step 5. Given $v \in H$, the sequence $\{v_n\}$ defined by

$$v_n = \sum_{i=1}^{n}(v, w_i)w_i$$

is such that $v_n \in W_n$ for each n and $v_n \to v$ strongly in H. From (2.3.7), we get

$$(u_n, v_n) + (Au_n, v_n) = (f, v_n). \qquad (2.3.8)$$

As $\{Au_n\}$ is bounded, we can also assume (after taking a further subsequence if necessary) that $Au_n \rightharpoonup \chi$ weakly in H. Passing to the limit as $n \to \infty$ in (2.3.8), we get

$$(u, v) + (\chi, v) = (f, v) \text{ for all } v \in H. \qquad (2.3.9)$$

Step 6. By (2.3.3) we have for any $v \in H$,

$$\begin{aligned} 0 \le X_n &= (Au_n - Av, u_n - v) \\ &= (Au_n, u_n) - (Au_n, v) - (Av, u_n - v) \\ &= (f, u_n) - \|u_n\|^2 - (Au_n, v) - (Av, u_n - v), \end{aligned}$$

using (2.3.8). Thus,

$$\begin{aligned} 0 \le \overline{\lim}_{n\to\infty} X_n &= (f, u) - \underline{\lim}_{n\to\infty}\|u_n\|^2 - (\chi, v) - (Av, u - v) \\ &\le (f, u) - \|u\|^2 - (f, v) + (u, v) - (Av, u - v) \end{aligned}$$

using (2.3.9). Thus,

$$(f - u - Au, u - v) + (Au - Av, u - v) \ge 0. \qquad (2.3.10)$$

Let $\lambda > 0$ and $w \in H$. Set $v = u - \lambda w$ in (2.3.10) to get

$$(f - u - Au, w) + (Au - A(u - \lambda w), w) \ge 0.$$

As $\lambda \to 0$, by the hemi-continuity of A, the second term on the left-hand side tends to zero. Thus $(f - u - Au, w) \ge 0$ for all $w \in H$ and, by considering $-w$ in place of w, we conclude that u satisfies (2.3.4). ■

2.4 Borsuk's Theorem

Let $\Omega \subset \mathbb{R}^n$ be a bounded open set which is symmetric with respect to the origin, *i.e.* if $x \in \Omega$, then $-x \in \Omega$. Let $f \in C^1(\overline{\Omega}; \mathbb{R}^n)$ be an odd function, *i.e.* $f(-x) = -f(x)$ for all $x \in \Omega$. Assume that $0 \notin f(S) \cup f(\partial\Omega)$. Assume further that $0 \notin \overline{\Omega}$. If $f^{-1}(0)$ is empty, then $d(f, \Omega, 0) = 0$. If not, the solution set has to be of the form

$$\cup_{i=1}^{m}\{x_i, -x_i\}.$$

Since f' is now an even function, we have $J_f(-x) = J_f(x)$. Thus,

$$\begin{aligned} d(f, \Omega, 0) &= \textstyle\sum_{i=1}^{m}(sgn(J_f(x_i)) + sgn(J_f(-x_i))) \\ &= 2\textstyle\sum_{i=1}^{m} sgn(J_f(x_i)). \end{aligned}$$

Thus, if $0 \notin \overline{\Omega}$, the degree is an even integer.

If $0 \in \Omega$, we do have $f(0) = 0$. Thus the solution set is now of the form

$$\{0\} \text{ or } \{0\} \cup_{i=1}^{m} \{x_i, -x_i\}.$$

In the former case, the degree is ± 1 and in the latter it is $\pm 1 + 2\sum_{i=1}^{m} sgn(J_f(x_i))$. Thus, in either case, the degree $d(f, \Omega, 0)$ is an odd integer.

Borsuk's theorem generalizes this result to the case when $f \in C(\overline{\Omega}; \mathbb{R}^n)$ and $0 \notin f(\partial\Omega)$. In order to prove it, we need a few technical lemmas which essentially deal with the extension of functions to larger sets retaining special properties.

Lemma 2.4.1 *Let $K \subset \mathbb{R}^n$ be compact. Let $\varphi \in C(K; \mathbb{R}^m)$ where $m > n$. Let $0 \notin \varphi(K)$. Then, if Q is any cube containing K, there exists $\varphi_Q \in C(Q; \mathbb{R}^n)$ extending φ and such that $0 \notin \varphi_Q(Q)$.*

Proof: Step 1. Since K is compact, and $0 \notin \varphi(K)$, we have

$$\alpha = \inf_{x\in K} |\varphi(x)| > 0.$$

Let $0 < \varepsilon < \alpha/2$. Let $\psi \in C^1(Q; \mathbb{R}^m)$ such that

$$\sup_{x\in K} |\varphi(x) - \psi(x)| < \varepsilon/2.$$

If $0 \notin \psi(Q)$, set $\psi_1 = \psi$. If, on the other hand, $0 \in \psi(Q)$, define $\Psi \in C^1(Q \times \mathbb{R}^{m-n}; \mathbb{R}^m)$ by

$$\Psi(x,y) = \psi(x) \text{ for } x \in Q,\ y \in \mathbb{R}^{m-n}.$$

Then for all $z = (x,y) \in Q \times \mathbb{R}^{m-n}$, $J_\Psi(z) = 0$. Hence, by Sard's theorem, the range of ψ is of measure zero. Thus, there exists $a \notin \psi(Q)$ with $|a| < \varepsilon/2$. Set $\psi_1 = \psi - a$.

Thus, in either case, there exists $\psi_1 \in C^1(Q;\mathbb{R}^m)$ such that $\|\psi_1 - \varphi\|_{\infty,K} < \varepsilon$ and $0 \notin \psi_1(Q)$.

Step 2. Let $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ be defined by

$$\eta(t) = \begin{cases} 1, & t > \alpha/2 \\ \frac{2t}{\alpha}, & t \le \alpha/2. \end{cases}$$

Define $\varphi_1(x) = \psi_1(x)/\eta(|\psi_1(x)|)$ for $x \in Q$. By definition, $|\varphi_1(x)| \ge \alpha/2$ and so $0 \notin \varphi_1(Q)$. If $x \in K$,

$$|\psi_1(x)| \ \ge\ |\varphi(x)| - |\varphi(x) - \psi_1(x)| \ \ge\ \alpha - \varepsilon \ \ge\ \alpha/2.$$

Hence $\eta(|\psi_1(x)|) = 1$ and so $\varphi_1 = \psi_1$ on K.

Step 3. Let $\theta = \varphi_1 - \varphi$. By the Tietze extension theorem, we can extend θ to $\widetilde{\theta}$ on Q such that $|\widetilde{\theta}| < \varepsilon$ on Q (since $|\theta| = |\psi_1 - \varphi| < \varepsilon$ on K). Set $\varphi_Q = \varphi_1 - \widetilde{\theta}$. Then φ_Q extends φ and

$$|\varphi_Q| \ \ge\ |\varphi_1| - |\widetilde{\theta}| \ \ge\ \alpha/2 - \varepsilon \ >\ 0.$$

Hence $0 \notin \varphi_Q(Q)$, which completes the proof. ■

Lemma 2.4.2 *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set symmetric with respect to the origin and such that $0 \notin \Omega$. Let $\varphi \in C(\partial\Omega;\mathbb{R}^m)$ where $m > n$. Assume that φ is odd and that $0 \notin \varphi(\partial\Omega)$. Then there exists $\Phi \in C(\overline{\Omega};\mathbb{R}^m)$ extending φ which is odd and non-vanishing.*

Proof: We will proceed by induction on n. Let $n = 1$. We set

$$\overline{\Omega} = [-\delta, -\varepsilon] \cup [\varepsilon, \delta],\ 0 < \varepsilon < \delta.$$

By the preceding lemma, we can extend φ to φ_1 on $[\varepsilon,\delta]$ such that it is non-vanishing. Now define

$$\Phi(x) = \begin{cases} \varphi_1(x), & x \in [\varepsilon,\delta] \\ -\varphi_1(-x), & x \in [-\delta,-\varepsilon]. \end{cases}$$

Thus Φ has the required properties.

We now assume the result to hold for all dimensions $\le (n-1)$. Let $\overline{\Omega} \subset \mathbb{R}^n$. Set

$$\Omega_+ = \{x \in \Omega \mid x_N > 0\},\ \Omega_- = \{x \in \Omega \mid x_N < 0\}.$$

Now, $\partial(\Omega \cap \mathbb{R}^{n-1}) = \partial\Omega \cap \mathbb{R}^{n-1}$ and, by induction, we can extend φ to $\widetilde{\varphi}$ on $\overline{\Omega} \cap \mathbb{R}^{n-1}$, with $\widetilde{\varphi}$ odd and non-vanishing. Now, let Q be a cube in $\mathbb{R}^n_+$ containing $\overline{\Omega}_+$ and $\overline{\Omega} \cap \mathbb{R}^{n-1}$. Consider

$$\widetilde{\varphi}_1 = \begin{cases} \varphi & \text{on } \partial\Omega \cap \overline{\Omega}_+ \\ \widetilde{\varphi} & \text{on } \overline{\Omega} \cap \mathbb{R}^{n-1}. \end{cases}$$

Then $\widetilde{\varphi}_1$ can be extended to a non-vanishing function φ_Q on Q. Now define

$$\Phi(x) = \begin{cases} \varphi_Q(x), & x \in \Omega_+ \\ -\varphi_Q(-x), & x \in \Omega_-. \end{cases}$$

It is now immediate to see that Φ has the required properties. ■

Lemma 2.4.3 *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set which is symmetric with respect to the origin and such that $0 \notin \overline{\Omega}$. Let $\varphi \in C(\partial\Omega;\mathbb{R}^n)$ be an odd and non-vanishing mapping. Then, there exists $\Phi \in C(\overline{\Omega};\mathbb{R}^n)$ which extends φ and is odd and such that $0 \notin \Phi(\overline{\Omega} \cap \mathbb{R}^{n-1})$.*

Proof: Consider φ restricted to $\partial\Omega \cap \mathbb{R}^{n-1}$. It is odd and nonvanishing and belongs to $C(\partial(\Omega \cap \mathbb{R}^{n-1});\mathbb{R}^n)$. Thus, by the preceding lemma, we can extend it to a continuous, odd and nonvanishing mapping on $\overline{\Omega} \cap \mathbb{R}^{n-1}$ taking values in $\mathbb{R}^n$. Let $\widetilde{\varphi}$ be equal to this function on $\overline{\Omega} \cap \mathbb{R}^{n-1}$ and to φ on $\partial\Omega_+ \cap \partial\Omega$. By Tietze's theorem, we can extend it to a function on $\overline{\Omega}_+$ and then, as usual, extend it as an odd map to all of $\overline{\Omega}$. ■

Lemma 2.4.4 *Let Ω be as in the preceding lemma and let $\varphi \in C(\overline{\Omega}; \mathbb{R}^n)$ be odd and non-vanishing on $\partial\Omega$. Then $d(\varphi, \Omega, 0)$ is even.*

Proof: By the preceding lemma, we can find $\Phi \in C(\overline{\Omega}; \mathbb{R}^n)$ which is equal to φ on $\partial\Omega$, is odd and which does not vanish on $\overline{\Omega} \cap \mathbb{R}^{n-1}$. Then, (cf. Proposition 2.2.3)

$$d(\varphi, \Omega, 0) = d(\Phi, \Omega, 0).$$

There exists $\varepsilon > 0$ such that if $\|\Phi - \psi\|_\infty < \varepsilon$, then the degrees of Φ and ψ in Ω with respect to 0 are the same. Now choose ψ a C^2 function such that $\|\Phi - \psi\|_\infty < \varepsilon$ and set $\widehat{\varphi}(x) = \frac{1}{2}(\psi(x) - \psi(-x))$. Then $\widehat{\varphi}$ is C^2, is odd, and $\|\Phi - \widehat{\varphi}\|_\infty < \varepsilon$ so that $d(\Phi, \Omega, 0) = d(\widehat{\varphi}, \Omega, 0)$. If we further choose $\varepsilon < \frac{1}{2}\rho(0, \Phi(\overline{\Omega} \cap \mathbb{R}^{n-1}))$, we also have that $\widehat{\varphi}$ is non-vanishing on $\overline{\Omega} \cap \mathbb{R}^{n-1}$. Hence, by excision and additivity,

$$\begin{aligned} d(\widehat{\varphi}, \Omega, 0) &= d(\widehat{\varphi}, \Omega \backslash (\overline{\Omega} \cap \mathbb{R}^{n-1}), 0) \\ &= d(\widehat{\varphi}, \Omega_+, 0) + d(\widehat{\varphi}, \Omega_-, 0). \end{aligned}$$

We can now find a regular value b of $\widehat{\varphi}$ such that

$$\begin{aligned} d(\widehat{\varphi}, \Omega_+, 0) &= d(\widehat{\varphi}, \Omega_+, b) = \textstyle\sum_{\widehat{\varphi}(x)=b} sgn(J_{\widehat{\varphi}}(x)) \\ d(\widehat{\varphi}, \Omega_-, 0) &= d(\widehat{\varphi}, \Omega_-, -b) = \textstyle\sum_{\widehat{\varphi}(x)=b} sgn(J_{\widehat{\varphi}}(-x)) \\ &= \textstyle\sum_{\widehat{\varphi}(x)=b} sgn(J_{\widehat{\varphi}}(x)) \\ &= d(\widehat{\varphi}, \Omega_+, 0) \end{aligned}$$

since $J_{\widehat{\varphi}}$ is now even. Thus $d(\varphi, \Omega, 0) = d(\widehat{\varphi}, \Omega, 0)$ is even. ■

Theorem 2.4.1 *(Borsuk's Theorem) Let $\Omega \subset \mathbb{R}^n$ be a bounded open set, symmetric with respect to the origin and such that $0 \in \Omega$. Let $\varphi \in C(\overline{\Omega}; \mathbb{R}^n)$ be odd and non-vanishing on the boundary. Then $d(\varphi, \Omega, 0)$ is an odd integer.*

Proof: Since $0 \in \Omega$, let $B(0; r)$ be a ball of (sufficiently small) radius contained in Ω. By Tietze's theorem, let $\psi \in C(\overline{\Omega}; \mathbb{R}^n)$ such that

$$\psi(x) = \begin{cases} \varphi(x), & x \in \partial\Omega \\ x, & x \in B(0; r) \end{cases}$$

Then ψ does not vanish for $|x| = r$ and so by Proposition 2.2.3, excision and additivity we have

$$\begin{aligned} d(\varphi, \Omega, 0) &= d(\psi, \Omega, 0) \\ &= d(\psi, \Omega \backslash \partial B(0; r), 0) \\ &= d(\psi, \Omega \backslash \overline{B}(0; r), 0) + d(\psi, B(0; r), 0). \end{aligned}$$

By Lemma 2.4.4, the first term on the right is an even integer and since $\psi = I$ on the boundary of $B(0; r)$, the second term is unity. This proves the theorem. ■

Corollary 2.4.1 *Let Ω be as in the preceding theorem and let $\varphi \in C(\overline{\Omega}; \mathbb{R}^n)$ be odd on the boundary. Then there exist $x, y \in \overline{\Omega}$ such that $\varphi(x) = 0$ and $\varphi(y) = y$.*

Proof: If φ vanishes on the boundary, we have $x \in \partial\Omega$ such that $\varphi(x) = 0$. If not, the degree $d(\varphi, \Omega, 0)$ is well-defined and is an odd integer, and therefore, non-zero. Thus, there exists $x \in \Omega$ where φ vanishes. Now, consider $\psi(x) = x - \varphi(x)$ which is also odd and continuous, and therefore must vanish at a point $y \in \overline{\Omega}$, which completes the proof. ■

Corollary 2.4.2 *There is no retraction of the closed unit ball in $\mathbb{R}^n$ onto its boundary.*

Proof: The identity map is odd on S^{n-1} and, so any retraction must vanish, which is impossible. ■

Corollary 2.4.3 *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set containing the origin and symmetric with respect to it. Let $\varphi \in C(\overline{\Omega}; \mathbb{R}^n)$ be non-vanishing on the boundary. Assume further that for each $x \in \partial\Omega$, $\varphi(x)$ and $\varphi(-x)$ do not point in the same direction. The $d(\varphi, \Omega, 0)$ is odd and thus the image of φ is a neighbourhood of the origin.*

Proof: Define $H(x,t) = \varphi(x) - t\varphi(-x)$ for $(x,t) \in \partial\Omega \times [0,1]$. By hypothesis, H does not vanish on the boundary and the degree is thus well-defined and independent of t. We have $H(.,0) = \varphi$ while $H(.,1)$ is odd. The result now follows from Borsuk's theorem. ■

Corollary 2.4.4 *Let Ω be as in the preceding corollary and let $\varphi \in C(\partial\Omega; \mathbb{R}^n)$ be odd and non-vanishing. Then, there does not exist a homotopy $H \in C(\partial\Omega \times [0,1]; \mathbb{R}^n)$ which is non-vanishing and such that $H(.,0) = \varphi$ and $H(.,1) \equiv x_o \in \mathbb{R}^n \backslash \{0\}$.*

Proof: If such a H existed, we can extend it, by Tietze's theorem, to $H \in C(\overline{\Omega} \times [0,1]; \mathbb{R}^n)$ and while $d(H(.,0), \Omega, 0)$ is odd, we will have $d(H(.,1), \Omega, 0) = 0$, which is impossible. ■

Exercise 2.4.1 Show that no sphere in $\mathbb{R}^n$ can be deformed within itself to a single point. ■

Corollary 2.4.5 *Let Ω be as in the preceding corollary and let $\varphi \in C(\partial\Omega; \mathbb{R}^n)$ be odd and such that its image is contained in a proper subspace of $\mathbb{R}^n$. Then there exists $x_o \in \partial\Omega$ such that $\varphi(x_o) = 0$.*

Proof: If not, φ would be odd and non-vanishing on the boundary and hence its image would be a neighbourhood of the origin which is not possible. ■

Corollary 2.4.6 *(Borsuk - Ulam) Let Ω be as above and let $\varphi \in C(\partial\Omega; \mathbb{R}^n)$ be such that its image is contained in a proper subspace of $\mathbb{R}^n$. Then there exists $\xi \in \partial\Omega$ such that $\varphi(\xi) = \varphi(-\xi)$.*

Proof: Apply the preceding corollary to $\psi(x) = \varphi(x) - \varphi(-x)$.■

Example 2.4.1 Assume that the surface of the earth is spherical and that the temperature and atmospheric pressure vary continuously on it. Then there exist a pair of antipodal points with the same temperature and the same pressure. ■

Example 2.4.2 (Sandwich theorem) Given three regions in $\mathbb{R}^3$, there exists a single plane which divides each region into two parts of equal volume. (A single knife stroke can halve a piece of bread, a piece of cheese and a piece of ham, placed arbitrarily in space!!) The result is true for any n regions in $\mathbb{R}^n$. Consider $x = (x', x_{n+1}) \in S^n$, where $x' \in \mathbb{R}^n$, and the hyperplane defined by

$$H_x = \{y \in \mathbb{R}^n \mid y.x' = x_{n+1}\}.$$

Let

$$H_x^+ = \{y \in \mathbb{R}^n \mid y.x' > x_{n+1}\}.$$

If μ is the n-dimensional Lebesgue measure, define

$$\varphi_i(x) = \mu(A_i \cap H_x^+),\ 1 \leq i \leq n$$

where $\{A_i\}_{i=1}^n$ are the given regions. By the Borsuk - Ulam theorem, there exists $x_o \in S^n$ such that $\varphi_i(x_o) = \varphi_i(-x_o)$ for all $1 \leq i \leq n$ which gives the result. ■

Exercise 2.4.2 Show that there does not exist an odd continuous map $f : S^n \to S^m$ for $m < n$.■

Remark 2.4.1 We can tell two finite sets apart by counting their elements. Two finite dimensional spaces can be compared by looking at their dimensions. We have dim $E >$ dim F if, and only if, there is no injective linear map from E into F. The above exercise is, in some sense, a result in this spirit, to compare two spheres. More generally, we can compare two sets that are symmetric with respect to the origin and not containing it by examining the existence of continuous odd maps from one to the other. This leads us to the notion of the *genus* of such sets which will be discussed in the next section. ■

2.5 The Genus

Let E be a (real) Banach space and let $\Sigma(E)$ denote the collection of all closed subsets in E that are symmetric with respect to the origin and not containing it.

Definition 2.5.1 *Let $A \in \Sigma(E)$. We denote by $\gamma(A)$ the* **genus** *of A which is the smallest positive integer n such that there exists a continuous odd map of A into $\mathbb{R}^n\backslash\{0\}$. We set $\gamma(\emptyset) = 0$ and if no such n exists for A, we set $\gamma(A) = \infty$.*■

Example 2.5.1 Let $E = \mathbb{R}^n$ and let Ω be a bounded open set, symmetric with respect to the origin and containing it. Then $\gamma(\partial\Omega) \leq n$ since we have the identity map $I : \partial\Omega \to \mathbb{R}^n\backslash\{0\}$ which is odd. But by Corollary 2.4.5, there is no odd non-vanishing map into a proper subspace of $\mathbb{R}^n$. Thus, $\gamma(\partial\Omega) = n$. In particular,

$$\gamma(S^{n-1}) \;=\; n. \blacksquare$$

Example 2.5.2 Let E be a Banach space and let $x \in E$, $x \neq 0$. Set $A = \overline{B}(x;r) \cup \overline{B}(-x;r)$ where $0 < r < \|x\|$. Then $A \in \Sigma(E)$ and $\gamma(A) = 1$ since we have the odd map $f : A \to \mathbb{R}\backslash\{0\}$ given by $f \equiv 1$ on $\overline{B}(x;r)$ and $f \equiv -1$ on $\overline{B}(-x;r)$.■

In general, any disconnected set in $\Sigma(E)$ will have genns equal to unity. If $\gamma(A) \geq 2$, then clearly, A has to be an infinite set.

We now list the important properties of the genus.

Theorem 2.5.1 *Let E be Banach and let $A, B \in \Sigma(E)$.*
(i) If there exists an odd continuous map $f : A \to B$, then

$$\gamma(A) \;\leq\; \gamma(B). \tag{2.5.1}$$

(ii) If $A \subset B$, then (2.5.1) is again true.
(iii) If $h : A \to B$ is an odd homeomorphism, then $\gamma(A) = \gamma(B)$.
(iv) Subadditiviy:

$$\gamma(A \cup B) \;\leq\; \gamma(A) + \gamma(B). \tag{2.5.2}$$

(v) Let $\gamma(B) < \infty$. Then

$$\gamma(\overline{A\backslash B}) \;\geq\; \gamma(A) - \gamma(B). \tag{2.5.3}$$

(vi) If A is compact, then $\gamma(A) < \infty$.
(vii) Let A be compact. Define

$$N_\delta(A) \;=\; \{x \in E \mid \rho(x, A) \leq \delta\}. \tag{2.5.4}$$

Then, for sufficiently small δ,

$$\gamma(N_\delta(A)) \;=\; \gamma(A). \tag{2.5.5}$$

Proof: (i) Let $\gamma(B) = n < \infty$ (otherwise the result is trivially true). If $\varphi : B \to \mathbb{R}^n\backslash\{0\}$ is odd, then $\varphi \circ f : A \to \mathbb{R}^n\backslash\{0\}$ is odd and the result follows.
(ii) Set $f = I$, the identity map, in (i).
(iii) Follows from (2.5.1); we have $\gamma(A) \leq \gamma(B) \leq \gamma(A)$.
(iv) If either $\gamma(A)$ or $\gamma(B)$ is infinite, then the result is trivially true. Let $\gamma(A) = m$ and $\gamma(B) = n$ and $\varphi : A \to \mathbb{R}^m\backslash\{0\}$ and $\psi : B \to \mathbb{R}^n\backslash\{0\}$ the corresponding odd maps. Extend these maps to all of E as $\widetilde{\varphi}$ and $\widetilde{\psi}$ respectively and we can assume that they are odd (if, for instance, $\widetilde{\varphi}$ is not odd, we can replace it by $\frac{1}{2}(\widetilde{\varphi}(x) - \widetilde{\varphi}(-x))$). Now define $h = (\widetilde{\varphi}|_{A\cup B}, \widetilde{\psi}|_{A\cup B}) \in \mathbb{R}^{m+n}$. Then h is odd and non-vanishing on $A \cup B$. This proves (2.5.2).
(v)Note that $\overline{A\backslash B} \in \Sigma(E)$ and $A \subset \overline{A\backslash B} \cup B$. Now (2.5.3) follows from (ii) and (2.5.2).
(vi) Let $x \in A$ and let $r < \|x\|$. Then $\gamma(\overline{B}(x;r) \cup \overline{B}(-x;r)) = 1$ and we can cover A by a finite number of such sets and the result follows from (ii) and (iv).
(vii) Let $\gamma(A) = n$ and let $\varphi : A \to \mathbb{R}^n\backslash\{0\}$ be the corresponding odd map. As before, we can extend it to an odd map $\widetilde{\varphi}$ on E. Since A is compact, $|\varphi(x)| \geq \alpha > 0$ for all $x \in A$ and so, for sufficiently small δ, $\widetilde{\varphi}(x) \neq 0$ for $x \in N_\delta(A)$. Thus $\gamma(N_\delta(A) \leq n$ while the reverse inequality is true by virtue of (ii) and (2.5.5) is proved. ■

We now prove a stronger version of Corollary 2.4.5 regarding the zero set of an odd continuous map on a symmetric domain containing the origin.

Proposition 2.5.1 *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set which contains the origin and is symmetric with respect to it. Let $\varphi : \partial\Omega \to \mathbb{R}^m$ be a continuous and odd map, and let $m < n$. Let*

$$A = \{x \in \partial\Omega \mid \varphi(x) = 0\}.$$

Then $\gamma(A) \geq n - m$.

Proof: Let $N_\delta(A)$ be such that (2.5.5) holds. We claim that for some $\varepsilon > 0$, We have $Z_\varepsilon \subset N_\delta(A)$, where

$$Z_\varepsilon = \{x \in \partial\Omega \mid |\varphi(x)| \leq \varepsilon\}.$$

If not, we have a sequence ε_n decreasing to zero and $x_n \in Z_{\varepsilon_n}$ with $|\varphi(x_n)| \leq \varepsilon_n$ and $x_n \notin N_\delta(A)$. Thus, $\rho(x_n, A) \geq \delta$. Since $\partial\Omega$ is compact, for a subsequence, $x_n \to x$ and so $\rho(x, A) \geq \delta$. On the other hand, $\varphi(x) = 0$ and so $x \in A$, a contradiction and so the claim holds.

Thus, $A \subset Z_\varepsilon \subset N_\delta(A)$ and so $\gamma(Z_\varepsilon) = \gamma(A)$. Now, for $\eta > 0$, let

$$C_\eta = \{x \in \partial\Omega, \mid |\varphi(x)| \geq \eta\}.$$

If $P(y) = y/\|y\|$ is the radial projection in $\mathbb{R}^m$, then

$$P \circ \varphi : C_\eta \to \mathbb{R}^m \backslash \{0\}$$

is odd and continuous and so $\gamma(C_\eta) \leq m$. Thus

$$\gamma(\overline{\partial\Omega \backslash C_\eta}) \geq \gamma(\partial\Omega) - \gamma(C_\eta) \geq n - m.$$

But $\overline{\partial\Omega \backslash C_\eta} = Z_\eta$ and the result follows on setting $\eta = \varepsilon$.■

Corollary 2.5.1 *Let Ω be as above and let $m < n$. Let $\psi \in C(\partial\Omega; \mathbb{R}^m)$. If*

$$A = \{x \in \partial\Omega \mid \psi(x) = \psi(-x)\}$$

then $\gamma(A) \geq n - m$.

Proof: Apply the preceding proposition to the function $\varphi(x) = \psi(x) - \psi(-x)$.■

Lemma 2.5.1 *There exists a covering of S^{n-1} by n closed antipodal sets, i.e., $S^{n-1} = \cup_{i=1}^n B_i$ where $B_i = C_i \cup (-C_i)$, $C_i \cap (-C_i) = \emptyset$, $1 \leq i \leq n$.*

Proof: If $n = 1$, we have $S^0 = \{-1, +1\}$ and so $B_1 = \{-1\} \cup \{+1\}$. If $n = 2$, $S^1 = B_1 \cup B_2$ where $B_1 = \{(x, y) \in S^1 \mid |x| \geq 1/2\}$ and $B_2 = \{(x, y) \in S^1 \mid |y| \geq 1/2\}$. Assume the result upto n. Set $S^{n-1} = \cup_{i=1}^n B_i'$. Let $x = (x', x_{n+1}) \in \mathbb{R}^{n+1}$, with $x' \in \mathbb{R}^n$. Identify the hyperplane $\{x_{n+1} = 0\}$ with $\mathbb{R}^n$. Define

$$C_{n+1} = \{(x', x_{n+1}) \in S^n \mid x_{n+1} \geq 1/4\}.$$

For $1 \leq i \leq n$, define

$$C_i = \{(x', x_{n+1}) \in S^n \mid |x_{n+1}| \leq 1/2,\ x'/\sqrt{1 - x_{n+1}^2} \in C_i'\}$$

where $B_i' = C_i' \cup (-C_i')$, $C_i' \cap (-C_i') = \emptyset$. Then the C_i for $1 \leq i \leq n+1$ are closed, $C_i \cap (-C_i) = \emptyset$ and the $B_i = C_i \cup (-C_i)$ cover S^n.■

The above lemma is used to prove a result which will allow us to calculate the genus of a set made up of sets of genus unity.

Theorem 2.5.2 *Let $A \in \Sigma(E)$. Then $\gamma(A) = n$ if, and only if, n is the least integer such that there exist sets $A_i \in \Sigma(E)$ for $1 \leq i \leq n$ such that $\gamma(A_i) = 1$ for all such i and $A \subset \cup_{i=1}^n A_i$.*

Proof: If $\gamma(A) = n$, then there exist $D_1, ..., D_n$ in $\Sigma(E)$ covering A and each of them having genus unity. For, if φ is the odd non-vanishing map into $\mathbb{R}^n$ from A, and if P is the radial projection in $\mathbb{R}^n$, then $P \circ \varphi$ maps A into S^{n-1}. If B_i and C_i are as in the statement of the preceding lemma, then $\{(P \circ \varphi)^{-1}(B_i)\}$ covers A. Further,

$$D_i = (P \circ \varphi)^{-1}(B_i) = (P \circ \varphi)^{-1}(C_i) \cup (P \circ \varphi)^{-1}(-C_i)$$

and so, as the two sets on the right are disjoint, $\gamma(D_i) = 1$.

Sufficiency: If the $\{A_i\}$ exist as in the statement of the theorem, then clearly $\gamma(A) \leq n$. If $\gamma(A) = m < n$, then by the preceding argument, there exist m sets D_j with the same properties as the A_i, contradicting the minimality of n.

Necessity: By our initial argument, we know that A can be covered by n sets of genus unity. If n were not minimal, then A would be covered by m sets of genus unity, where $m < n$ and then $\gamma(A) \leq m$, a contradiction. ■

Inspired by the above theorem, we can define a notion analogous to the genus in topological spaces.

Definition 2.5.2 *Let X be a topological space and $A \subset X$ a closed subset. A is said to be of* **category 1** *in X ($cat_X(A) = 1$) if it can be deformed continuously to a single point, i.e., there exist $H \in C(A \times [0,1]; X)$ such that $H(x,0) = x$ for all $x \in A$ and $H(x,1) = x_o \in A$ for all $x \in A$.*■

Definition 2.5.3 *Let X be a topological space and let $A \subset X$ a closed subset. We say that $cat_X(A) = n$ if, and only if, n is the least integer such that there exist closed sets A_i for $1 \leq i \leq n$ covering A and such that $cat_X(A_i) = 1$ for each such i. If no such n exists, we say that $cat_X(A) = \infty$.*■

The category defined above, called the *Lyusternik - Schnirelman Category*, has properties analogous to the genus. It is more flexible and more general than the genus. But its properties are more difficult to prove. The genus and the category give information on the size of solution sets of nonlinear equations.

Chapter 3

The Leray - Schauder Degree

3.1 Preliminaries

Let X be a (real) Banach space. Henceforth, unless otherwise stated, all mappings of X into itself, or any other space, will be assumed to be continuous and mapping bounded sets into bounded sets.

Definition 3.1.1 *Let X and Y be Banach spaces. Let Ω be an open set in X. Let $T : \overline{\Omega} \to Y$ be continuous. Then T is said to be* **compact** *if it maps bounded sets (in X) into relatively compact sets (in Y).* ■

Example 3.1.1 By Ascoli's theorem, the injection $C^1([0,1];\mathbb{R}) \to C([0,1];\mathbb{R})$ is compact. ■

Example 3.1.2 Let $K \in C([0,1] \times [0,1];\mathbb{R})$. Let $f \in C([0,1];\mathbb{R})$. Define

$$T(f)(x) = \int_0^1 K(x,y) f(y) dy.$$

Then T is a compact linear operator on $C([0,1];\mathbb{R})$. To see this, notice that K is uniformly continuous. Hence, given $\varepsilon > 0$, there exists $\delta > 0$ such that $|x_1 - x_2| < \delta$ imples that, for all $y \in [0,1]$,

$$|K(x_1,y) - K(x_2,y)| < \varepsilon / C$$

where $C > 0$ is fixed. Hence for all $\|f\|_\infty \leq C$, we have

$$|T(f)(x_1) - T(f)(x_2)| \ < \ \varepsilon$$

and so the the image under T of the ball of radius C is equicontinuous. Clearly it is also bounded. Thus, the result follows, once again, from Ascoli's theorem. ■

Example 3.1.3 Let $\Omega \subset \mathbb{R}^n$ be a bounded open set. Then, by the Rellich - Kondrasov theorem (cf. for instance, Kesavan [13]) we have that the injection

$$H_0^1(\Omega) \to L^2(\Omega)$$

is compact. ■

All the above examples deal with compact linear operators.

Example 3.1.4 Let X and Y be Banach spaces and let $T : X \to Y$ be such that $T(X)$ is contained in a finite dimensional subspace of Y. Clearly such a map is compact. Such maps are called maps of finite rank. ■

Exercise 3.1.1 Let $T_n, T : X \to Y$ be bounded linear maps such that all the T_n are of finite rank and $\|T_n - T\| \to 0$ as $n \to \infty$. Show that T is compact. ■

Henceforth, throughout this chapter, Ω will denote a bounded open subset of a Banach space X. The identity operator on X will, as usual, be denoted by I.

Definition 3.1.2 *Let $T : \overline{\Omega} \to X$ be a compact map. The mapping $\varphi = I - T$ is called a* **compact perturbation of the identity**. ■

Proposition 3.1.1 *A compact perturbation of the identity in X is closed (i.e. maps closed sets into closed sets) and proper (i.e. inverse images of compact sets are compact).*

Proof: Let $\varphi = I - T$ be a compact perturbation of the identity. Let $A \subset X$ be closed. Let $y_n = \varphi(x_n) \in \varphi(A)$ and let $y_n \to y$ in X. Thus, $y_n = x_n - Tx_n$. Since $\{x_n\}$ is bounded, we have, for a subsequence, $Tx_n \to z$ and so $x_n \to y + z$ and $y + z \in A$, since A is closed. It then follows that $y = \varphi(y + z)$ and thus φ is closed.

Let $A \subset X$ be compact. Let $\{x_n\}$ be a sequence in $\varphi^{-1}(A)$. Thus, $y_n = x_n - Tx_n \in A$ and since A is compact, we have , for a subsequence, $y_n \to y \in A$. Since Ω is bounded, again, for a further subsequence, $Tx_n \to z$. Again, it follows that, for that subsequence in question, $x_n \to y + z$ and thus $\varphi^{-1}(A)$ is compact and so φ is proper. ■

We will try to generalize the notion of the finite dimensional degree to proper maps in infinite dimensional Banach spaces. We will do this by approximating, in a suitable sense, a compact map by a map of finite rank. To do this we will later need the following technical result.

Lemma 3.1.1 *Let $K \subset X$ be compact. Given $\varepsilon > 0$, there exists a finite dimensional subspace $V_\varepsilon \subset X$ and a map $g_\varepsilon : K \to V_\varepsilon$ such that, for every $x \in K$,*

$$\|g_\varepsilon(x) - x\| \ < \ \varepsilon. \tag{3.1.1}$$

Proof: Given $\varepsilon > 0$, there exist $x_1, ..., x_n \in K$, where $n = n(\varepsilon)$, such that

$$K \ \subset \ \cup_{i=1}^n B(x_i; \varepsilon).$$

Set

$$V_\varepsilon \ = \ span\{x_1, ..., x_n\}.$$

Then $\dim V_\varepsilon \leq n < \infty$. Define, for $x \in K$,

$$b_i(x) \ = \ \begin{cases} \varepsilon - \|x - x_i\|, & \text{if } x \in B(x_i; \varepsilon) \\ 0, & \text{otherwise.} \end{cases} \tag{3.1.2}$$

Since the $B(x_i;\varepsilon)$ cover K, we have $\sum_{i=1}^n b_i(x) \neq 0$ for each $x \in K$. Hence we can define, for $x \in K$,

$$g_\varepsilon(x) = \frac{\sum_{i=1}^n b_i(x)x_i}{\sum_{i=1}^n b_i(x)} \in V_\varepsilon. \tag{3.1.3}$$

If $b_i(x) \neq 0$, then $\|x - x_i\| < \varepsilon$. Thus, if $x \in K$,

$$\|g_\varepsilon(x) - x\| = \left\| \frac{\sum_{i=1}^n b_i(x)(x - x_i)}{\sum_{i=1}^n b_i(x)} \right\| < \varepsilon. \blacksquare$$

3.2 Definition of the Degree

Let V be a finite dimensional space of dimension n. Given a basis for V, we can identify V with $\mathbb{R}^n$. Given two different bases, a vector $x \in V$ may be expressed as $x^{(1)} \in \mathbb{R}^n$ or as $x^{(2)} \in \mathbb{R}^n$ depending on the base chosen. There exists an invertible matrix M such that $Mx^{(2)} = x^{(1)}$. If Ω is a bounded open set in V, and if $b \in V$, then given $\varphi \in C(\overline{\Omega}; V)$, we can consider it as $\varphi_i \in C(\overline{\Omega_i}; \mathbb{R}^n)$, $i = 1$ or 2, as the case may be. We have

$$\varphi_2(x^{(2)}) = M^{-1}\varphi_1(Mx^{(2)}). \tag{3.2.1}$$

Now, if $\varphi \in C^1(\overline{\Omega}; V)$, then (3.2.1) shows that $J_\varphi(x)$ is independent of the basis chosen. Let $b \notin \varphi(\partial\Omega)$. Let us write b as $b^{(1)}$ or $b^{(2)}$ depending on the base chosen. We see that if b is regular, then the degree is independent of the base:

$$d(\varphi_1, \Omega_1, b^{(1)}) = d(\varphi_2, \Omega_2, b^{(2)}).$$

By reducing to the regular case, we see easily that this is true even for values b that are not regular. Thus we have proved the following result.

Lemma 3.2.1 *Let V be finite dimensional of dimension n and let $\varphi \in C(\overline{\Omega}; V)$ where Ω is a bounded open subset of V. Let $b \notin \varphi(\partial\Omega)$. Then $d(\varphi, \Omega, b)$ is independent of the base chosen to identify V with $\mathbb{R}^n$.* $\blacksquare$

Let Ω be a bounded open set in $\mathbb{R}^n$ and let $m < n$. Let $T \in C(\overline{\Omega}; \mathbb{R}^m)$ and let $b \in \mathbb{R}^m$. We imbed $\mathbb{R}^m$ in $\mathbb{R}^n$ by setting the last $n - m$ coordinates as zero. Thus, $b = (b_1, ..., b_m, 0, .., 0) \in \mathbb{R}^n$ and $Tx = (T_1x, ..., T_mx, 0, ..., 0) \in \mathbb{R}^n$. If T is C^1 and $b \notin \varphi(\partial\Omega)$ is a regular value of $\varphi = I - T$, then, if $x \in \varphi^{-1}(b)$, it follows that $x \in \mathbb{R}^m \subset \mathbb{R}^n$ and

$$\varphi'(x) = \begin{bmatrix} (\varphi|_{\Omega\cap\mathbb{R}^m})'(x) & 0 \\ 0 & I_{n-m} \end{bmatrix}$$

and so

$$J_\varphi(x) = J_{\varphi|_{\Omega\cap\mathbb{R}^m}}(x).$$

Thus

$$d(\varphi, \Omega, b) = d(\varphi|_{\overline{\Omega}\cap\mathbb{R}^m}, \Omega\cap\mathbb{R}^m, b). \tag{3.2.2}$$

It now follows that (3.2.2) also holds for $\varphi \in C(\overline{\Omega}; \mathbb{R}^m)$ and any $b \in \mathbb{R}^m \backslash \varphi(\partial\Omega)$. Now if F_1 and F_2 are two finite dimensional subspaces containing $T(\overline{\Omega})$ and b, then $F_1 \cap F_2$ has the same properties. Further, it follows from the preceding considerations that

$$d(\varphi|_{\overline{\Omega}\cap F_i}, \Omega\cap F_i, b) = d(\varphi|_{\overline{\Omega}\cap F_1\cap F_2}, \Omega\cap F_1\cap F_2, b) \tag{3.2.3}$$

for $i = 1, 2$. Thus, we are naturally led to the following definition.

Definition 3.2.1 *Let $\Omega \subset X$ be a bounded open set in a Banach space X and let $T : \overline{\Omega} \to X$ be a map of finite rank. Then, if $b \notin \varphi(\partial\Omega)$ where $\varphi = I - T$, we define*

$$d(\varphi, \Omega, b) = d(\varphi|_{\overline{\Omega}\cap F}, \Omega\cap F, b) \tag{3.2.4}$$

where $F \subset X$ is a finite dimensional subspace containing $T(\overline{\Omega})$ and b. $\blacksquare$

The preceding considerations (cf. (3.2.3)) show that the above definition is independent of the choice of the subspace F.

Definition 3.2.2 *Let Ω be a bounded open set in a Banach space X and let $\varphi = I - T : \overline{\Omega} \to X$ be a compact perturbation of the* *identity. Let $b \notin \varphi(\partial\Omega)$. Let $\rho_o = \rho(b, \varphi(\partial\Omega))$. Let $\widehat{T} : \overline{\Omega} \to X$ be a map of finite rank such that $\|Tx - \widehat{T}x\| < \rho_o/2$ for all $x \in \overline{\Omega}$. We define the* **Leray - Schauder degree** *of φ in Ω with respect to b by*

$$d(\varphi, \Omega, b) = d(\widehat{\varphi}, \Omega, b) \tag{3.2.5}$$

where $\widehat{\varphi} = I - \widehat{T}$.■

We now ensure that the above definition makes sense. First of all, by Proposition 3.1.1, $\varphi(\partial\Omega)$ is closed and so $\rho_o > 0$. Next, there do exist maps of finite rank as in the above definition. For instance, set

$$\widehat{T} = g_{\rho_o/2} \circ T$$

where $g_{\rho_o/2}$ is the map described in Lemma 3.1.1 for $\varepsilon = \rho_o/2$.

If $\widehat{T}$ is any mapping as in the definition, then, for $x \in \partial\Omega$, we have $\|\varphi(x) - \widehat{\varphi}(x)\| < \rho_o/2$ so that $\rho(b, \widehat{\varphi}(\partial\Omega)) \geq \rho_o/2 > 0$. Thus $b \notin \widehat{\varphi}(\partial\Omega)$ and so $d(\widehat{\varphi}, \Omega, b)$ is well-defined. Finally, we need to check that this definition is independent of the choice of $\widehat{T}$.

Let T_1 and T_2 be two mappings of finite rank such that, for $i = 1, 2$ and for all $x \in \overline{\Omega}$, we have $\|Tx - T_ix\| < \rho_o/2$. Set $\varphi_i = I - T_i$. Let $F_i \subset X$ be a finite dimensional subspace containing $T_i(\overline{\Omega})$ and b, for $i = 1, 2$. Let F be a finite dimensional subspace of X containing $F_1 + F_2$ and b. Then for $i = 1, 2$,

$$d(\varphi_i, \Omega, b) = d(\varphi_i|_{\overline{\Omega}\cap F}, \Omega \cap F, b).$$

If $H(x, \theta) = \theta\varphi_1(x) + (1 - \theta)\varphi_2(x)$, for $x \in \overline{\Omega}$ and $\theta \in [0, 1]$, then, for $x \in \partial\Omega$, we have $\|b - H(x, \theta)\| \geq \rho_o/2$ since $\|\varphi(x) - H(x, \theta)\| < \rho_o/2$. Hence by the homotopy invariance of the (Brouwer) degree,

$$d(\varphi_1|_{\overline{\Omega}\cap F}, \Omega \cap F, b) = d(\varphi_2|_{\overline{\Omega}\cap F}, \Omega \cap F, b)$$

and so the degree given by (3.2.5) is indeed well-defined.

3.3 Properties of the Degree

Henceforth, we will denote by $Q(\overline{\Omega}; X)$ the space of all compact mappings from $\overline{\Omega}$ into X, where Ω is a bounded open set in a Banach space X.

Theorem 3.3.1 *(i) Let $T \in Q(\overline{\Omega}; X)$ and let $b \notin (I - T)(\partial\Omega)$. There exists a neighbourhood $\mathcal{U}$ of T in $Q(\overline{\Omega}; X)$ such that for all $S \in \mathcal{U}$, we have $b \notin (I - S)(\partial\Omega)$ and*

$$d(I - S, \Omega, b) = d(I - T, \Omega, b). \tag{3.3.1}$$

(ii) Let $H \in C(\overline{\Omega} \times [0, 1]; X)$ be defined by $H(x, t) = x - S(x, t)$ where $S \in Q(\overline{\Omega} \times [0, 1]; X)$. If $b \notin H(\partial\Omega \times [0, 1])$, then $d(H(., t), \Omega, b)$ is independent of t.
(iii) The degree is constant on connected components of $X \backslash (I - T)(\partial\Omega)$.
(iv) If $\Omega_1 \cap \Omega_2 = \emptyset$, $\Omega = \Omega_1 \cup \Omega_2$ and $b \notin (I - T)(\partial\Omega_i)$, $i = 1, 2$, then

$$d(I - T, \Omega, b) = d(I - T, \Omega_1, b) + d(I - T, \Omega_2, b). \tag{3.3.2}$$

Proof: Let $\rho_o = \rho(b, (I - T)(\partial\Omega)) > 0$. We set

$$\|S - T\|_\infty = \sup_{x \in \overline{\Omega}} \|Sx - Tx\|.$$

Let $\mathcal{U}$ be given by

$$\mathcal{U} = \{S \in Q(\overline{\Omega}; X) \mid \|S - T\|_\infty < \rho_o/2\}.$$

If $S \in \mathcal{U}$, then clearly b is not a boundary value of $I - S$. Choose finite dimensional maps T_1 and S_1 such that

$$\|T - T_1\|_\infty < \rho_o/4, \ \|S - S_1\|_\infty < \rho_o/4.$$

Let F be a finite dimensional subspace of X containing $T_1(\overline{\Omega})$, $S_1(\overline{\Omega})$ and b. Then

$$\left.\begin{array}{rcl} d(I - T, \Omega, b) &=& d((I - T_1)|_{\overline{\Omega}\cap F}, \Omega \cap F, b) \\ d(I - S, \Omega, b) &=& d((I - S_1)|_{\overline{\Omega}\cap F}, \Omega \cap F, b). \end{array}\right\} \tag{3.3.3}$$

Set

$$H(x, \theta) = \theta(I - T_1)x + (1 - \theta)(I - S_1)x$$

for $x \in \Omega \cap F$. Then for $x \in \partial(\Omega \cap F)$ and for $0 < \theta < 1$,

$$\begin{aligned} \|H(x,\theta) - b\| &\geq \|b - (I-T)x\| - \theta\|(I-T_1)x - (I-T)x\| \\ &\quad -(1-\theta)\|(I-S_1)x - (I-S)x\| \\ &\quad -(1-\theta)\|(I-T)x - (I-S)x\| \\ &\geq \rho_o - \theta\rho_o/4 - (1-\theta)\rho_o/4 - (1-\theta)\rho_o/2 \\ &= 3\rho_o/4 - (1-\theta)\rho_o/2 > \rho_o/4 > 0. \end{aligned}$$

Thus $b \notin H(.,\theta)(\partial(\Omega \cap F))$. Then (3.3.1) follows from the homotopy invariance of the Brouwer degree and (3.3.3).
(ii) Since, by (i), the degree is locally constant, the result follows from the connectedness of $[0,1]$.
(iii) Clearly, by definition, it follows that

$$d(\varphi,\Omega,b) = d(\varphi - b,\Omega,0) \qquad (3.3.4)$$

where $\varphi = I - T$. Now the result follows from (i).
(iv) The relation (3.3.2) follows directly from the definition and from the additivity of the Brouwer degree. ■

Remark 3.3.1 Notice that if $t \mapsto S(.,t)$ is continuous from $[0,1]$ into $Q(\overline{\Omega};X)$, then $S \in Q(\overline{\Omega} \times [0,1];X)$.■

Proposition 3.3.1 *Let $\Omega \subset X$ be a bounded open set. Then*

$$d(I,\Omega,b) = \begin{cases} 1, & \text{if } b \in \Omega \\ 0, & \text{if } b \notin \overline{\Omega}. \end{cases} \qquad (3.3.5)$$

Proof: Set $T = 0$ and note that $T(\overline{\Omega})$ and b are contained in the one-dimensional space $\mathbb{R}b$. Thus $d(I,\Omega,b) = d(I,\Omega \cap \mathbb{R}b, b)$ and the result follows. ■

Proposition 3.3.2 *(Excision) Let $K \subset \Omega$ be closed and let $\varphi = I - T$ be a compact perturbation of the identity. Let $b \notin \varphi(K) \cup \varphi(\partial\Omega)$. Then*

$$d(\varphi,\Omega,b) = d(\varphi,\Omega\backslash K,b). \qquad (3.3.6)$$

Proof: By Proposition 3.1.1, $\varphi(K)$ is closed and so

$$\rho_1 = \min\{\rho(b,\varphi(K)),\rho(b,\varphi(\partial\Omega))\} > 0.$$

Let T_1 be a mapping of finite rank such that $\|T - T_1\|_\infty < \rho_1/2$. Then, by definition, $d(\varphi,\Omega,b) = d(\psi,\Omega,b)$, where $\psi = I - T_1$. If F is a finite dimensional subspace containing $T_1(\overline{\Omega})$ and b,

$$\begin{aligned} d(\psi,\Omega,b) &= d(\psi_F,\Omega \cap F,b) \\ &= d(\psi_F,(\Omega\backslash K) \cap F,b) \\ &= d(\psi,\Omega\backslash K,b) \\ &= d(\varphi,\Omega\backslash K,b) \end{aligned}$$

where $\psi_F = \psi|_{\overline{\Omega}\cap F}$.■

Proposition 3.3.3 *Let $S,T \in Q(\overline{\Omega};X)$ such that $S = T$ on $\partial\Omega$. Then if $b \notin \varphi(\partial\Omega) = \psi(\partial\Omega)$, where $\varphi = I - T$ and $\psi = I - S$, we have*

$$d(\varphi,\Omega,b) = d(\psi,\Omega,b).$$

Proof: Set $H(x,t) = t\varphi(x) + (1-t)\psi(x)$ for $t \in [0,1]$. The result now follows from Theorem 3.3.1. ■

Proposition 3.3.4 *Let $F \subset X$ be a closed subspace containing $T(\overline{\Omega})$ and b, where $T \in Q(\overline{\Omega};X)$ and $b \notin \varphi(\partial\Omega)$, $\varphi = I - T$. Then*

$$d(\varphi,\Omega,b) = d(\varphi_F,\Omega \cap F,b) \qquad (3.3.7)$$

where $\varphi_F = \varphi|_{\overline{\Omega}\cap F}$.

Proof: Let $K = \overline{T(\overline{\Omega})}$, which is compact. Let $\rho_o = \rho(b,\varphi(\partial\Omega)) > 0$. Let $V_{\rho_o/2}$ be the finite dimensional space and $g_{\rho_o/2}$ the mapping as described in Lemma 3.1.1. Then, $V_{\rho_o/2} \subset F$. Let

$$\rho_1 = \rho(b,\varphi_F(\partial_F(\Omega \cap F))) > 0.$$

Then $\rho_1 \geq \rho_o$. Further, if $\varphi_{\rho_o} = I - g_{\rho_o/2} \circ T$, then

$$\|\varphi_F - \varphi_{\rho_o}\|_\infty < \rho_o/2 \leq \rho_1/2.$$

Consequently,

$$d(\varphi_F, \Omega \cap F, b) = d(\varphi_{\rho_o}, \Omega \cap F, b) = d(\varphi_{\rho_o}|_{\overline{\Omega} \cap V_{\rho_o/2}}, \Omega \cap V_{\rho_o/2}, b)$$

which is exactly $d(\varphi, \Omega, b)$ by definition. ■

Exercise 3.3.1 Let $T \in Q(\overline{\Omega}; X)$ and $b \notin \varphi(\partial\Omega)$, where $\varphi = I - T$.
(i) If $b \notin \varphi(\overline{\Omega})$, show that $d(\varphi, \Omega, b) = 0$.
(ii) If $d(\varphi, \Omega, b) \neq 0$, show that $\varphi(\Omega)$ is a neighbourhood of b.
(iii) If $\varphi(\Omega)$ is included in a proper subspace of X, then show that $d(\varphi, \Omega, b) = 0$.■

Exercise 3.3.2 Let φ be as in the preceding exercise. If $\{\Omega_j\}_{j \in J}$ is a family of pairwise disjoint open sets in Ω with $\varphi^{-1}(b) \subset \cup_{j \in J} \Omega_j$, show that $d(\varphi, \Omega_j, b)$ is zero except for a finite number of j and that

$$d(\varphi, \Omega, b) = \sum_{j \in J} d(\varphi, \Omega, b).$$

Exercise 3.3.3 Let X_1 and X_2 be Banach spaces and let $\Omega_i \subset X_i$ be bounded open subsets, for $i = 1, 2$. Assume that $T_i \in Q(\overline{\Omega}_i; X_i)$, and that $b_i \notin \varphi_i(\partial\Omega_i)$ where $\varphi_i = I - T_i$, for $i = 1, 2$. Show that

$$d((\varphi_1, \varphi_2), \Omega_1 \times \Omega_2, (b_1, b_2)) = d(\varphi_1, \Omega_1, b_1).d(\varphi_2, \Omega_2, b_2).■$$

3.4 Fixed Point Theorems

The Brouwer fixed point theorem (Theorem 2.3.1) states that any continuous function of a closed ball in $\mathbb{R}^n$ into itself has atleast one fixed point. It was generalized (cf. Corollary 2.3.1) to the case of a compact conevex set. We cannot relax these conditions further to have the fixed point property in general. For instance, $f : S^n \to S^n$ given by $f(x) = -x$ does not have a fixed point and S^n, while being compact, is not convex. Similarly, $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x + 1$ has no fixed point and here $\mathbb{R}$ is convex but not compact. One other generalization is that if K is homeomorphic

to a closed ball in $\mathbb{R}^n$, then any continuous function $f : K \to K$ has a fixed point, as it is trivial to see.

In the infinite dimensional case, the classical version of Brouwer's theorem is false as the following example shows.

Example 3.4.1 Let $H = l^2$, the space of square summable sequences. Thus, if $x = \{x_i\} \in l^2$, we have

$$\|x\| = \sum_{i=1}^{\infty} |x_i|^2 < \infty.$$

Let B be the closed unit ball in H. Define $T : B \to B$ by

$$Tx = \{\sqrt{1 - \|x\|^2}, x_1, x_2, ...\}.$$

Then T is continuous and its range is the unit sphere. Thus, if x were a fixed point, it follows that $\|x\| = \|Tx\| = 1$ and then it would follow from the definition of T that

$$0 = x_1 = x_2 = x_3 = ...$$

Hence $x = 0$, which is absurd. Thus, T cannot have a fixed point. ■

However, the version of the theorem as in Corollary 2.3.1 holds.

Theorem 3.4.1 *(Schauder Fixed Point Theorem) Let X be a Banach space and let $K \subset X$ be a compact and convex subset. If $f : K \to K$ is continuous, then f has a fixed point.*

Proof: Let $\varepsilon > 0$ and consider the pair $(g_\varepsilon, V_\varepsilon)$ as in Lemma 3.1.1. Then, for $x \in K$, $g_\varepsilon(x)$ is a convex combination of the basis vectors $\{x_1, x_2, ..., x_n\}$ where $n = n(\varepsilon)$. Hence, $g_\varepsilon(x) \in K_\varepsilon \subset K$ where K_ε is the closed convex hull of $\{x_1, x_2, ..., x_n\}$. Consider the continuous map $\varphi_\varepsilon : K_\varepsilon \to K_\varepsilon$ defined by

$$\varphi_\varepsilon(x) = g_\varepsilon(f(x)).$$

Since $K_\varepsilon \subset K$, it follows that K_ε is also compact and is thus a compact convex set in the finite dimensional space V_ε as well. Hence, by Corollary 2.3.1, there exists a fixed point $x_\varepsilon \in K_\varepsilon$ of φ_ε. Again, since K is compact, $\{x_\varepsilon\}$ has a convergent subsequence converging to some $x \in K$. Now,

$$\|x - f(x)\| \leq \|x - x_\varepsilon\| + \|x_\varepsilon - f(x_\varepsilon)\| + \|f(x_\varepsilon) - f(x)\|. \quad (3.4.1)$$

The first and last terms on the right-hand side tend to zero as $\varepsilon \to 0$ by the definition of x and the continuity of f. Further,

$$\|x_\varepsilon - f(x_\varepsilon)\| = \|g_\varepsilon(f(x_\varepsilon)) - f(x_\varepsilon)\| < \varepsilon.$$

Thus, the right-hand side of (3.4.1) can be made arbitrarily small and so $f(x) = x$ and the proof is complete.■

A minor variation of the above result is as follows.

Corollary 3.4.1 *Let K be a closed, bounded and convex subset of X and let $f : K \to K$ be compact. Then f has a fixed point.*

Proof: Since $\overline{f(K)}$ is compact, so is its closed convex hull $\widehat{K}$. Since K is closed and convex, and as f maps K into itself, it follows that $\widehat{K} \subset K$. Now $f|_{\widehat{K}}$ maps $\widehat{K}$ into itself and thus has a fixed point which is also a fixed point for f in K.■

Theorem 3.4.2 *(Schaeffer) Let $f : X \to X$ be compact. Assume that there exists $R > 0$ such that if $u = \sigma f(u)$ for some $\sigma \in [0, 1]$, then $\|u\| < R$. Then f has a fixed point in the ball $B(0; R)$.*

Proof: Consider $I - \sigma f : \overline{B}(0; R) \to \overline{B}(0; R)$. By hypothesis, $0 \notin (I - \sigma f)(\partial B(0; R))$ for $\sigma \in [0, 1]$. Consequently, the degree $d(I - \sigma f, B(0; R), 0)$ is well-defined and is independent of σ. Thus,

$$d(I - f, B(0; R), 0) = d(I, B(0; R), 0) = 1.$$

Thus, the degree is non-zero and so the equation $(I - f)(x) = 0$ has a solution in the ball. ■

Example 3.4.2 Let $f : \mathbb{R} \to \mathbb{R}$ be a bounded and continuous function. Let $\Omega \subset \mathbb{R}^n$ be a bounded domain. Consider the semi-linear elliptic boundary value problem:

$$\left.\begin{aligned} -\Delta u &= f(u) \quad \text{in } \Omega \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned}\right\} \qquad (3.4.2)$$

We will show that this problem has a solution in $H_0^1(\Omega)$.

Given $\varphi \in L^2(\Omega)$, define $G(\varphi) \in H_0^1(\Omega)$ to be the unique solution w of the problem:

$$\left.\begin{aligned} -\Delta w &= \varphi \quad \text{in } \Omega \\ w &= 0 \quad \text{on } \partial\Omega. \end{aligned}\right\}$$

Now, let $|f(x)| \leq M$ for all $x \in \mathbb{R}$. Clearly, $f(u) \in L^2(\Omega)$ and so, by the dominated convergence theorem, it is easy to see that the mapping $u \mapsto f(u)$ is continuous from $L^2(\Omega)$ into itself. Thus, using Rellich's theorem which states that the injection of $H_0^1(\Omega)$ into $L^2(\Omega)$ is compact for Ω bounded, it can be deduced that the map $u \mapsto T(u) = G(f(u))$ is a compact map of $H_0^1(\Omega)$ into itself. Clearly u is a solution of (3.4.2) if, and only if, u is a fixed point of T.

If $v = \sigma T v$ for some $\sigma \in [0, 1]$, then

$$\left.\begin{aligned} -\Delta v &= \sigma f(v) \quad \text{in } \Omega \\ v &= 0 \quad \text{on } \partial\Omega. \end{aligned}\right\}$$

Hence,

$$\begin{aligned} \int_\Omega |\nabla v|^2 dx = \sigma \int_\Omega f(v) v dx &\leq M|\Omega|^{1/2}\|v\|_{L^2(\Omega)} \\ &\leq C\left(\int_\Omega |\nabla v|^2 dx\right)^{1/2} \end{aligned}$$

by Poincaré's inequality. Thus,

$$\|v\|_{H_0^1(\Omega)} \leq C < C + \eta$$

for any $\eta > 0$ and so, by Schaeffer's theorem, there exists a fixed point of T satisfying

$$\|u\|_{H_0^1(\Omega)} \leq C. \blacksquare$$

Exercise 3.4.1 Let $\Omega \subset \mathbb{R}^n$ be a bounded domain. Let $f : \mathbb{R} \to \mathbb{R}$ be a Lipschitz continuous map with Lipschitz constant $K > 0$. Let

$$A = -\sum_{i,j=1}^{n} \frac{\partial}{\partial x_i}\left(a_{ij}(x)\frac{\partial}{\partial x_j}\right)$$

where the coefficients $a_{ij} \in L^\infty(\Omega)$ and satisfy

$$\alpha|\xi|^2 \leq \sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \leq \beta|\xi|^2$$

for almost every x in Ω and for all $\xi \in \mathbb{R}^n$, where α and β are positive constants. If K is sufficiently small, show that there exists a solution to the problem:

$$\left.\begin{array}{rcll} Au & = & f(u) & \text{in } \Omega \\ u & = & 0 & \text{on } \partial\Omega. \end{array}\right\} \blacksquare$$

3.5 The Index

Let X be a Banach space and let Ω be a bounded open subset of X. Let $T \in Q(\overline{\Omega}; X)$. Let $\varphi = I - T$.

Definition 3.5.1 *We say that $x_o \in X$ is an* **isolated solution** *of the equation $\varphi(x) = 0$ if there exists $\varepsilon_o > 0$ such that x_o is the only solution of this equation in the ball $B(x_o; \varepsilon_o)$.*$\blacksquare$

If x_o is an isolated solution of the equation $\varphi(x) = 0$ and if ε_o is as in the above definition, then , for every $0 < \varepsilon < \varepsilon_o$, the degree $d(\varphi, B(x_o; \varepsilon), 0)$ is well-defined and, by the excision property, this degree is independent of ε. Thus, if $\varepsilon_n \to 0$, the sequence $\{d(\varphi, B(x_o; \varepsilon_n), 0)\}$ is stationary.

Definition 3.5.2 *The* **index** *of an isolated solution x_o of the equation $\varphi(x) = 0$, denoted by $i(\varphi, x_o, 0)$ is given by the relation*

$$i(\varphi, x_o, 0) = \lim_{\varepsilon \to 0} d(\varphi, B(x_o; \varepsilon), 0).\blacksquare \qquad (3.5.1)$$

Remark 3.5.1 If $b \in X$ is any point, and if x_o is an isolated solution of the equation $\varphi(x) = b$, then we can define the index with respect to b via the relation

$$i(\varphi, x_o, b) = i(\varphi - b, x_o, 0).\blacksquare$$

Remark 3.5.2 If X were finite dimensional, then the above definitions make sense for any $\varphi \in C(\overline{\Omega}; X)$.$\blacksquare$

Proposition 3.5.1 *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map such that 1 is not an eigenvalue, i.e. $\varphi = I - T$ is invertible. Then*

$$i(\varphi, 0, 0) = (-1)^\beta$$

where β is the sum of the (algebraic) multiplicities of the characteristic values of T lying in the interval $(0, 1)$.

Proof: We have $\varphi'(0) = I - T$ which is invertible and so

$$i(\varphi, 0, 0) = \text{sgn}(\det(I - T)).$$

So, if $\{\lambda_i\}$, $1 \leq i \leq n$, are the eigenvalues of T, we need to compute the sign of $p(1)$ where

$$p(\lambda) = \det(\lambda I - T) = \Pi_{i=1}^{n}(\lambda - \lambda_i).$$

If $\lambda_i = 0$ for some i, then it does not contribute to the sign of $p(1)$. If λ_i is complex, then $\overline{\lambda}_i$ is also an eigenvalue and the product $(1 - \lambda_i)(1 - \overline{\lambda}_i) = |1 - \lambda_i|^2$ also does not contribute to the sign of $p(1)$. For non-zero and real λ_i, we call $\mu_i = 1/\lambda_i$ as the corresponding characteristic value. Again, if $\mu_i < 0$ or > 1, the term $1 - \lambda_i = 1 - 1/\mu_i$ does not contribute to the sign of $p(1)$. Thus

$$\text{sgn}(p(1)) = (-1)^\beta$$

where β is as in the statement of the proposition. $\blacksquare$

To generalize this result to infinite dimensions, we recall the following facts about the spectrum of a compact linear operator on a Banach space (cf. Dieudonné [8], Limaye [15] or Sunder [25]).

- The spectrum of a compact linear operator T on a Banach space X is atmost countable with 0 as its only possible accumulation point. If $\lambda \neq 0$ is in the spectrum, then it has to be an eigenvalue. Its reciprocal $\mu = \lambda^{-1}$ is called a characteristic value.

- The sequence

$$\mathrm{Ker}(I - \mu T) \subset \mathrm{Ker}(I - \mu T)^2 \subset \mathrm{Ker}(I - \mu T)^3 \subset ...$$

is stationary, *i.e.* there exists a positive integer k such that

$$\mathrm{Ker}(I - \mu T)^{k-1} \neq \mathrm{Ker}(I - \mu T)^k = \mathrm{Ker}(I - \mu T)^l \quad (3.5.2)$$

for all $l \geq k$. The space $\mathrm{Ker}(I - \mu T)^k$ is finite dimensional and its dimension is called the algebraic multiplicity of μ. If T were symmetric, then $k = 1$, *i.e.* the algebraic and geometric multiplicities are the same.

- If μ were not a characteristic value, $I - \mu T$ is invertible with continuous inverse.

Proposition 3.5.2 *Let X be a Banach space and let $T \in Q(\overline{\Omega}; X)$, where $\Omega \subset X$ is a neighbourhood of the origin. Assume that T is differentiable at the origin. Then $T'(0) : X \to X$ is a compact linear operator.*

Proof: If not, we can find a sequence $\{x_n\}$ in X and an $\varepsilon > 0$ such that $\|x_n\| \leq 1$ and $\|T'(0)(x_n - x_m)\| \geq \varepsilon$ for all n amd m. By the definition of differentiability,

$$\delta^{-1}\|T(\delta x_n) - T(0) - \delta T'(0)x_n\| \to 0$$

uniformly in n as $\delta \to 0$. Choose $\delta > 0$ small enough such that

$$\|T(\delta x_n) - T(0) - \delta T'(0)x_n\| \leq \delta\varepsilon/4$$

for all n so that

$$\|T(\delta x_n) - T(\delta x_m) - \delta T'(0)(x_n - x_m)\| \leq \delta\varepsilon/2.$$

Thus

$$\begin{aligned} \delta\varepsilon/2 &\geq \delta\|T'(0)(x_n - x_m)\| - \|T(\delta x_n) - T(\delta x_m)\| \\ &\geq \delta\varepsilon - \|T(\delta x_n) - T(\delta x_m)\|. \end{aligned}$$

Whence,

$$\|T(\delta x_n) - T(\delta x_m)\| \geq \delta\varepsilon/2.$$

Hence $\{\delta x_n\}$ will be a bounded sequence while $\{T(\delta x_n)\}$ will have no cluster point, contradicting the compactness of T.■

Proposition 3.5.3 *Let X be a Banach space and let Ω be a bounded open subset of X. Let $T \in Q(\overline{\Omega}; X)$ be differentiable at the origin and assume that $T(0) = 0$. If 1 is not a characteristic value of $T'(0)$ (so that zero is an isolated solution of $(I - T)(x) = 0$), we have*

$$i(\varphi, 0, 0) = (-1)^\beta$$

where $\varphi = I - T$ and β is the sum of the (algebraic) multiplicities of the characteristic values of $T'(0)$ lying in the interval $(0, 1)$.

Proof: Since $H(x,t) = x - T(xt)/t$ is an admissible homotopy connecting $I - T$ and $I - T'(0)$, it sufficies to show that

$$i(I - T'(0), 0, 0) = (-1)^\beta$$

where β is as in the statement of the proposition.

Since the only accumulation point of the characteristic values is at infinity, the interval $(0, 1)$ contains only a finite number of characteristic values, say, $\mu_1, \mu_2, ..., \mu_p$. Let $N_i = \mathrm{Ker}(I - \mu_i T)^{k_i}$, the characteristic subspace as in (3.5.2). Set $N = \oplus_{i=1}^p N_i$ which is finite dimensional. Hence, it admits a complement, *i.e.* a closed subspace F such that $X = N \oplus F$. Both N and F are invariant under $I - T'(0)$. Thus, we can now consider $I - T'(0) : X \to X$ as

$$((I - T'(0))|_N, (I - T'(0))|_F) : N \times F \to N \times F$$

and for $\varepsilon > 0$ sufficiently small, we have

$$\begin{aligned} i(I - T'(0), 0, 0) &= d(I - T'(0), B(0;\varepsilon), 0) \\ &= d((I - T'(0))|_N, B(0;\varepsilon) \cap N, 0) \\ &\quad \times d((I - T'(0))|_F, B(0;\varepsilon) \cap F, 0) \end{aligned}$$

by the product formula (cf. Exercise 3.3.3).

For any $t \in [0,1]$, $I - tT'(0)$ is invertible in F and so 0 is the only solution of the equation $(I - tT'(0))x = 0$ and thus the degree $d((I - tT'(0))|_F, B(0;\varepsilon) \cap F, 0)$ is independent of t and is thus equal to $d(I, B(0;\varepsilon) \cap F, 0) = 1$. Since N is finite dimensional, the degree $d((I - T'(0))|_N, B(0;\varepsilon) \cap N, 0)$ is none other than $i((I - T'(0))|_N, 0, 0)$ and, by Proposition 3.5.1, is equal to $(-1)^\beta$, where β is as defined previously since the characteristic values of $(I - T'(0))|_N$ are precisely $\mu_1, \mu_2, ..., \mu_p$ and this completes the proof. ■

If x_o is an isolated solution of $\varphi(x) = 0$ and if $0 \notin \varphi(\partial\Omega)$, we have

$$d(\varphi, \Omega, 0) \;=\; d(\varphi, \Omega\backslash\overline{B}(x_o;\varepsilon), 0) + i(\varphi, x_o, 0) \qquad (3.5.3)$$

for sufficiently small $\varepsilon > 0$. This is useful in getting information on the solution set as illustrated by the following example.

Example 3.5.1 Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a C^1 function such that $|\varphi(x)| \leq 1$ and $\varphi'(x) > 0$ for all $x \in \mathbb{R}$. Assume that $\varphi(1) = 0$. Then the system

$$\left.\begin{array}{rcl} x^3 - 3xy^2 + \varphi(x) &=& 1 \\ -y^3 + 3x^2y &=& 0 \end{array}\right\} \qquad (3.5.4)$$

has atleast three solutions in the ball $B(0;2)$. To see this, define $H : \overline{B}(0;2) \times [0,1] \to \mathbb{R}^2$ by

$$H((x,y),t) \;=\; (x^3 - 3xy^2 + t\varphi(x), -y^3 + 3x^2y).$$

We first verify that this does not assume the value $(1,0)$ on the circle $x^2+y^2 = 4$. Indeed, if $H((x,y),t) = (1,0)$, then either $y = 0$ or $y \neq 0$ and $y^2 = 3x^2$. On the said circle, these conditions are met only at the points $(\pm 2, 0)$ and $(\pm 1, \pm\sqrt{3})$. At these points we must further have $t\varphi(x) = 1 - x^3 - 3xy^2$. But $|t\varphi(x)| < 1$ while at these points the right-hand side is of absolute value $>$ 1. Thus, there are no solutions on the circle and so the degree

$d(H(.,t), B(0;2), (1,0))$ is well-defined and is independent of t. At $t = 0$, the degree is equal to 3, as can be easily seen. Thus, if $\Psi = H(.,1)$, we have

$$d(\Psi, B(0;2), (1,0)) \;=\; 3. \qquad (3.5.5)$$

Now, $\Psi(1,0) = (1,0)$. Further,

$$J_\Psi(x,y) = \begin{vmatrix} 3x^2 - 3y^2 + \varphi'(x) & -6xy \\ 6xy & 3x^2 - 3y^2 \end{vmatrix}$$

and so $J_\Psi(1,0) = 9 + 3\varphi'(1) > 0$. Thus $(1,0)$ is an isolated solution and

$$i(\Psi, (1,0), (1,0)) \;=\; 1. \qquad (3.5.6)$$

Since, by hypotheses, $x^3 + \varphi(x) = 1$ has only one solution, *viz.* $x = 1$, there are no solutions other than $(1,0)$ on the line $y = 0$ for the original system. Now, using (3.5.5) and (3.5.6) in (3.5.3), we deduce that there has to be atleast one solution to (3.5.4) with $y \neq 0$. But if (x,y) is one such solution, it is easy to see that $(x,-y)$ is also a solution. Thus there are at least three solutions in all to the system (3.5.4). ■

3.6 An Application to Differential Equations

Let $\Omega \subset \mathbb{R}^N$ be a bounded domain and let $J \subset \mathbb{R}$ an interval. Let $f : J \times \Omega \to \mathbb{R}^n$ be continuous. Consider the initial value problem:

$$\left.\begin{array}{rcl} u'(t) &=& f(t,u),\ t \in J \\ u(t_o) &=& u_o \end{array}\right\} \qquad (3.6.1)$$

where $(t_o, u_o) \in J \times \Omega$ is given. If f were Lipschitz continuous, then we have a unique local solution to (3.6.1). If f were merely continuous, even then we have the existence of a local solution but the uniqueness is no longer valid. For instance, the problem

$$\left.\begin{array}{rcl} u'(t) &=& u^{2/3} \\ u(0) &=& 0 \end{array}\right\}$$

has atleast two solutions, *viz.* $u = 0$ and $u = (t/3)^3$.

In fact, using the Leray - Schauder degree, we can get more information on the solution set. We can show that at any instant t, the set of all values $u(t)$ taken by solutions u to (3.6.1) is a connected set (this result is due to Kneser and Hukuhara). In particular, if we have two distinct solutions, then we must have an infinity of solutions!

We will follow the treatment of Rabinowitz [20]. We first prove an abstract result.

Theorem 3.6.1 *(Krasnoselsk'ii - Perov) Let X be a Banach space and let Ω be a bounded open subset of X. Let $T \in Q(\overline{\Omega}; X)$ and let $\varphi = I - T$. Assume that the following conditions hold:*
(i) For each $\varepsilon > 0$, there exists $T_\varepsilon \in Q(\overline{\Omega}; X)$ such that for all $u \in \overline{\Omega}$,

$$\|Tu - T_\varepsilon u\| \ < \ \varepsilon.$$

(ii) Whenever $\|b\| < \varepsilon$, the equation

$$u \ = \ T_\varepsilon u + b$$

admits at most one solution.

Let $0 \notin \varphi(\partial\Omega)$ and assume that $d(\varphi, \Omega, 0) \neq 0$. Then the set of solutions

$$S \ = \ \{u \in \Omega \mid \varphi(u) = 0\}$$

is connected.

Proof: Since the degree $d(\varphi, \Omega, 0) \neq 0$, the solution set S must be non-empty. Since φ is proper, S is compact. If S were not connected, then it can be written as the disjoint union of two non-empty compact sets. Thus, there exist non-empty open sets $\mathcal{U}$ and $\mathcal{V}$ such that $\overline{\mathcal{U}} \cap \overline{\mathcal{V}} = \emptyset$, $S \subset \mathcal{U} \cup \mathcal{V}$, $S \cap \mathcal{U} \neq \emptyset$, $S \cap \mathcal{V} \neq \emptyset$.

Since $S \cap \mathcal{U} \neq \emptyset$, there exists $u \in S \cap \mathcal{U}$. Hence $Tu = u$. Define, for $\varepsilon > 0$ and $v \in \overline{\mathcal{V}}$,

$$\psi_\varepsilon(v) \ = \ (v - T_\varepsilon v) - (u - T_\varepsilon u),$$

where T_ε is as in the hypotheses. Let

$$H(t, v) \ = \ t\psi_\varepsilon(v) + (1 - t)\varphi(v)$$

for $t \in [0, 1]$.

Since $0 \notin \varphi(\partial\mathcal{V})$, and since φ is closed, we have $\inf_{v \in \partial\mathcal{V}} \|\varphi(v)\| \geq \alpha > 0$. So, for $v \in \partial\mathcal{V}$,

$$\begin{aligned} \|H(t, v)\| &\geq \ \|\varphi(v)\| - \|Tv - T_\varepsilon v\| - \|u - T_\varepsilon u\| \\ &\geq \ \alpha - \varepsilon - \varepsilon \end{aligned}$$

(since $Tu = u$) by hypotheses. Thus, choosing $\varepsilon < \alpha/4$, we see that $H(t, .)$ does not vanish on the boundary of $\mathcal{V}$ and so the degree $d(H(t, .), \mathcal{V}, 0)$ is well-defined and is independent of t. Consequently,

$$d(\varphi, \mathcal{V}, 0) \ = \ d(\psi_\varepsilon, \mathcal{V}, 0). \tag{3.6.2}$$

But the solution set for the equation $\psi_\varepsilon(v) = 0$ is empty in $\mathcal{V}$. For, if $v \in \mathcal{V}$ were a solution, then

$$v - T_\varepsilon v \ = \ u - T_\varepsilon u$$

and setting $b = u - T_\varepsilon u = Tu - T_\varepsilon u$, we have $\|b\| < \varepsilon$ and so as $u \in \mathcal{U}$ already solves $u - T_\varepsilon u = b$, we cannot have any solution in $\mathcal{V}$ which is disjoint from $\mathcal{U}$. Thus, from (3.6.2) it follows that

$$d(\varphi, \mathcal{V}, 0) \ = \ 0.$$

Similarly, $d(\varphi, \mathcal{U}, 0) = 0$. But then, by the additivity and excision properties of the degree, we have

$$0 \ \neq d(\varphi, \Omega, 0) \ = \ d(\varphi, \mathcal{U}, 0) + d(\varphi, \mathcal{V}, 0) \ = \ 0$$

and we have a contradiction. Thus the set S is connected. ■

We now apply this result to an initial value problem of the type (3.6.1). Let

$$f : [-a, a] \times \overline{B}(0; c) \to \mathbb{R}^n \tag{3.6.3}$$

be continuous, where $B(0;c)$ is the (open) ball of radius c and centre at the origin in $\mathbb{R}^n$. Let

$$\left.\begin{array}{rcl} M & = & \sup\{|f(t,u)| \ : \ |t| \leq a, \ |u| \leq c\} \\ \alpha & = & \min\{a, c/M\}. \end{array}\right\} \quad (3.6.4)$$

Then by the Cauchy - Peano existence theorem, there exists atleast one solution to the problem

$$\left.\begin{array}{rcl} u'(t) & = & f(t,u) \\ u(0) & = & 0 \end{array}\right\} \quad (3.6.5)$$

for $|t| \leq \alpha$.

We can extend f to a continuous function $\overline{f} : [-a,a]\times\mathbb{R}^n \to \mathbb{R}^n$ such that

$$\sup\{|\overline{f}(t,u)| \ : \ |t| \leq \alpha, \ u \in \mathbb{R}^n\} \ = \ M.$$

If $E = C([-\alpha,\alpha];\mathbb{R}^n)$ which is a Banach space with the sup-norm, a simple application of Ascoli's theorem shows that $T : E \to E$ defined by

$$Tu(t) \ = \ \int_0^t \overline{f}(\tau, u(\tau))d\tau$$

is compact. If $u = Tu$, then

$$\|u\| \ \leq \ M|t| \ \leq \ M\alpha \ \leq \ c$$

and thus $\overline{f}(t,u) = f(t,u)$. It then follows that solutions to $\varphi(u) = 0$ where $\varphi = I - T$, are precisely solutions of (3.6.5).

We now show that we are in the situation of Theorem 3.6.1. Let Ω be the ball of radius $c+1$ and centre at the origin in E. Since the solutions of $\varphi(u) = 0$ verify the estimate $\|u\| \leq c$, as seen above, it follows that $0 \notin \varphi(\partial\Omega)$. This is also true for all solutions of the equation $u - \sigma Tu = 0$ where $\sigma \in [0,1]$. Hence,

$$d(\varphi,\Omega,0) \ = \ d(I,\Omega,0) \ = 1.$$

Given $\varepsilon > 0$, there exists $f_\varepsilon \in C^1([-\alpha,\alpha]\times\mathbb{R}^n;\mathbb{R}^n)$ such that for all $|t| \leq \alpha$, and for all $|z| \leq c+1$,

$$|f_\varepsilon(t,z) - \overline{f}(t,z)| \ \leq \ \varepsilon/\alpha.$$

Set

$$T_\varepsilon u(t) \ = \ \int_0^t f_\varepsilon(\tau, u(\tau))d\tau.$$

Then, again, T_ε is compact and

$$|T_\varepsilon u(t) - Tu(t)| \ \leq \ (\varepsilon/\alpha)|t| \ \leq \ \varepsilon.$$

Finally, if $b \in E$, and v and w are solutions of $u = T_\varepsilon u + b$, then

$$v(t) - w(t) \ = \ \int_0^t (f_\varepsilon(\tau, v(\tau)) - f_\varepsilon(\tau, w(\tau))d\tau$$

so that

$$|v(t) - w(t)| \ \leq \ K\int_0^t |v(\tau) - w(\tau)|d\tau$$

where

$$K \ = \ \sup\left\{\left|\frac{\partial f_\varepsilon}{\partial z}(t,z)\right| \ : \ |t| \leq \alpha, \ |z| \leq c+1\right\} \ < \ \infty.$$

Hence, by Gronwall's lemma, $u(t) = v(t)$ for all t.

Thus all the hypotheses of Theorem 3.6.1 are satisfied and so the solution set of the equation $\varphi(u) = 0$ is connected in E.

Theorem 3.6.2 *(Kneser - Hukuhara) Let the conditions (3.6.3) and (3.6.4) hold for f.Define*

$$K_t \ = \ \{u(t) \mid u \ \textit{ solution of } \ (3.6.5)\}.$$

Then K_t is a connected set in $\mathbb{R}^n$.

Proof: If S is the solution set for the equation $\varphi(u) = 0$ in E as described above, then $K_t = \delta_t(S)$, where $\delta_t : E \to \mathbb{R}^n$ is the evaluation map $\delta_t(u) = u(t)$. Since S is connected in E and since δ_t is continuous, it follows that K_t is also connected. ■

Chapter 4

Bifurcation Theory

4.1 Introduction

Let X and Y be Banach spaces. Let $f \in C(X;Y)$. We are often interested in the set of solutions to the equation

$$f(x) = 0.$$

However, this question is too general to be answered satisfactorily, even when the spaces X and Y are finite dimensional. Very often, we are led to study nonlinear equations dependent on a parameter of the form

$$f(x, \lambda) = 0$$

where $f : X \times Y \to Z$, with X, Y and Z being Banach spaces. Usually, it will turn out that $Y = \mathbb{R}$. It is quite usual for the above equation to possess a 'nice' family of solutions (often called the trivial solutions). However, for certain values of λ, new solutions may appear and hence we use the term 'bifurcation'.

The classical example for this kind of phenomenon is the buckling of a thin rod. Consider a thin rod of unit length lying on the x-axis along the interval $[0,1]$ with its left end point fixed. Consider a compressive force of magnitude P applied at the right end. Upto a certain critical value of P, the rod is merely compressed along the axis. But once P crosses a critical value, the rod buckles out of its original state. Thus, we can consider the zero vertical

displacement as the trivial solution for all values of P while there exist nontrivial solutions to the displacement for some values of P.

There are numerous examples from physics for bifurcation phenomena. The Bénard problem in heat transfer is one such. If an infinite layer of viscous incompressible fluid lies between a pair of parallel and perfectly conducting plates and if a temperature gradient T is maintained between them, the lower plate being warmer, then upto a certain value of T, the heat is transfered purely by conduction and there is no movement of the fluid (trivial solution). When T crosses a critical value, convection currents appear.

In the Taylor problem, a viscous incompressible fluid lies between a pair of coaxial cylinders with vertical axis. The inner cylinder rotates with an angular velocity ω while the outer cylinder is at rest. For small values of ω,the flow - called the Couette flow - consists of circular orbits of particles with velocity proportional to their distance from the axis of rotation. As ω increases beyond a critical value, the fluid breaks up into horizontal bands called Taylor vortices and a new motion in the vertical direction is superimposed on the Couette flow.

Normally, when we wish to approximate a nonlinear equation, we linearize it. However, this is not satisfactory in describing bifurcation phenomena as the example of the buckling rod shows.

Example 4.1.1 (cf. Stakgold [24]) Consider a rod occupying the interval $[0, 1]$ of the x-axis with its left end fixed and right end free to move along the axis. If it is subjected to a compressive load, the rod will buckle. Assume that the buckling takes place in the $x - y$ plane. Let $\varphi(x)$ denote the angle between the tangent at a point $x \in (0,1)$ of the buckled rod and the x- axis, then the function φ satisfies the following equation.

$$\left.\begin{array}{rcl} \varphi'' + \mu \sin\varphi & = & 0 \quad \text{in } (0,1) \\ \varphi'(0) = \varphi'(1) & = & 0. \end{array}\right\}$$

We can compute the vertical displacement $v(x)$ of the rod from

$\varphi(x)$. Notice that $\varphi \equiv 0$ (corresponding to the unbuckled state) is a solution for all μ. If we linearize this equation about this trivial solution, we get

$$\left.\begin{array}{rcll} \varphi'' + \mu\varphi & = & 0 & \text{in } (0,1) \\ \varphi'(0) = \varphi'(1) & = & 0 & \end{array}\right\}$$

and the equation satisfied by v becomes

$$\left.\begin{array}{rcll} v'' + \mu v & = & 0 & \text{in } (0,1) \\ v(0) = v(1) & = & 0. & \end{array}\right\}$$

This linear eigenvalue problem has non-zero solutions only when $\mu = 0$ or $\mu = n^2\pi^2$, $n \in I\!N$. At $\mu = 0$, we get $v =$ constant and hence $v = 0$. Thus there is no deflection. If $0 < \mu < \pi^2$, again we have only the zero solution. Thus the rod buckles at $\mu = \pi^2$ and as we compress further, again returns to the unbuckled state as μ increases further, till it reaches the value $\mu = 4\pi^2$ and so on. This is clearly unacceptable physically. Hence we need a proper nonlinear theory to study bifurcation phenomena. We will return to this example in Remark 4.3.4. ■

Henceforth we assume that X, Y and Z are Banach spaces and that $f : X \times Y \to Z$ is continuous. Further assume that for all $\lambda \in Y$,

$$f(0,\lambda) = 0.$$

Definition 4.1.1 *A point* $(0,\lambda_0) \in X \times Y$ *is said to be a* **bifurcation point** *if every neighbourhood of this point in* $X \times Y$ *contains a solution* (x,λ), $x \neq 0$ *of the equation*

$$f(u,\mu) = 0. \blacksquare \qquad (4.1.1)$$

bf Remark 4.1.1 Note that, essentially, only small neighbourhoods count. The definition ensures the existence of a sequence $\{(x_n,\lambda_n)\}$ of nontrivial solutions such that $x_n \to 0$ and $\lambda_n \to \lambda_0$ as $n \to \infty$. It does not guarantee the existence of a continuous branch of solutions $(x(\lambda).\lambda)$ with $x(\lambda) \to 0$ as $\lambda \to \lambda_0$.■

The following basic questions can thus be asked.
(i) When is a point $(0,\lambda) \in X \times Y$ a bifurcation point?
(ii) Do there exist branches of solutions emanating from this point, and if so, how many?
(iii) Can we describe the dependence of these branches on λ at least in a neighbourhood of the bifurcation point?
(iv) In case of several branches of solutions, which branch does the system follow?

The basic results of bifurcation theory presented in the sequel attempt to answer questions (i) - (iii) above. Question (iv) is related to the study of stability of solutions. We first prove a necessary condition for the existence of a bifurcation point and then consider several simple examples to illustrate the various possibilities that one could expect.

Proposition 4.1.1 *Let* $f : X \times Y \to Z$ *be differentiable. If* $(0,\lambda_0) \in X \times Y$ *is a bifurcation point, then* $\partial_x f(0,\lambda_0) : X \to Z$ *is not an isomorphism.*

Proof: Let $\{(x_n,\lambda_n)\}$ be a sequence of solutions to (4.1.1) converging to $(0,\lambda_0)$. Then

$$0 = f(x_n,\lambda_n) = f(0,\lambda_n) + \partial_x f(0,\lambda_n)x_n + R(x_n,\lambda_n)$$

where $\|R(x_n,\lambda)\|/\|x_n\| \to 0$ as $n \to \infty$. If $\partial_x f(0,\lambda_0)$ is an isomorphism, then so is $\partial_x f(0,\lambda_n)$ for sufficiently large n and its inverse is bounded and independent of n. Since $f(0,\lambda_n) = 0$ as well, we have

$$x_n = -(\partial_x f(0,\lambda_n))^{-1} R(x_n,\lambda_n)$$

and so, for a constant $C > 0$, independent of n, we have

$$\|x_n\| \leq C\|R(x_n,\lambda_n)\|$$

which is impossible. ■

Example 4.1.2 Let $X = Y = Z = \mathbb{R}$. Let $f(x,\lambda) = x - x^2\lambda$. For any $\lambda \in \mathbb{R}$, the point $(0,\lambda)$ is not a bifurcation point. The solution set is given by the x-axis *i.e.* the set of points $(0, \lambda)$ for all $\lambda \in \mathbb{R}$ and the two branches of the rectangular hyperbola $x\lambda = 1$. Notice that $\partial_x f(0, \lambda) = 1$ which gives the identity map on $\mathbb{R}$.■

Example 4.1.3 Let $X = Z = \mathbb{R}^2$ and $Y = \mathbb{R}$. Let

$$f(x, \lambda) = (1 - \lambda) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} x_2^3 \\ -x_1^3 \end{pmatrix}.$$

Then $\partial_x f(0, \lambda) = (1 - \lambda)I$ which fails to be an isomorphism only for $\lambda = 1$. However, $(0, 1)$ is not a bifurcation point. Indeed it is immediate to see that if (x, λ) is a solution to (4.1.1), then $x_1^4 + x_2^4 = 0$ so that the only solutions are the trivial ones. Thus, the condition given in Proposition 4.1.1 is only necessary but not sufficient. ■

Example 4.1.4 Let $X = Y = Z = \mathbb{R}$. Let $f(x, \lambda) = x + x^3 - \lambda x$. Then $(0, \lambda)$ is always a solution of (4.1.1). Further, if $x \neq 0$ and if (x, λ) is a solution, then,

$$x^2 = \lambda - 1.$$

Thus, there are no nontrivial solutions for $\lambda \leq 1$ while there is a branch of nontrivial solutions given by the above parabola bifurcating from the trivial branch at $(0, 1)$. Notice again that $\partial_x f(0, \lambda) = 1 - \lambda$ which fails to be an isomorphism only at $\lambda = 1$.■

Example 4.1.5 Let $X = Z = \mathbb{R}^2$ and $Y = \mathbb{R}$. Let $f(x, \lambda) = Ax - \lambda x$ where $Ax = \beta x + C(x)$, and $C(x)$ being given by

$$C(x) = \begin{pmatrix} \gamma x_1(x_1^2 + x_2^2) \\ \gamma x_2(x_1^2 + x_2^2) \end{pmatrix}, \gamma \neq 0.$$

Thus $\partial_x f(0, \lambda) = (\beta - \lambda)I$ and so bifurcation can only occur at $(0, \beta)$. If (x, λ) with $x \neq 0$. $\lambda \neq \beta$ is a solution of (4.1.1), then, for $i = 1, 2$,

$$\gamma x_i(x_1^2 + x_2^2) = (\lambda - \beta)x_i.$$

Hence $x_1^2 + x_2^2 = (\lambda - \beta)/\gamma$ which is a paraboloid of revolution about the λ- axis. In this case, we have a continuum of branches emanating from the point $(0, \beta)$. Again, there are no nontrivial solutions for $\lambda < \beta$.■

Example 4.1.6 It can happen that a single branch emanates from a bifurcation point involving a multiple eigenvalue of the linearized operator. Let $X = Z = \mathbb{R}^2$ and $Y = \mathbb{R}$. Let

$$A(x) = \begin{pmatrix} x_1 + 2x_1x_2 \\ x_2 + x_1^2 + 2x_2^2 \end{pmatrix}.$$

Let $f(x, \lambda) = A(x) - \lambda x$. Then $\partial_x f(0, \lambda) = (1 - \lambda)I$. Thus, $\lambda = 1$ is a double eigenvalue. If (x, λ) is a solution of (4.1.1), then

$$\left.\begin{array}{rcl} x_1 + 2x_1x_2 - \lambda x_1 & = & 0 \\ x_2 + x_1^2 + 2x_2^2 - \lambda x_2 & = & 0. \end{array}\right\}$$

Clearly $x_2 = 0$ implies that $x_1 = 0$. Thus, for a nontrivial solution, $x_2 \neq 0$. If $x_1 \neq 0$, then by the first equation above, $x_2 = (\lambda - 1)/2$ and substituting in the second equation we get $x_1 = 0$, a contradiction. Thus, $x_1 = 0$ and then $x_2 = (\lambda - 1)/2$. Thus the only branch emanating from $(0, 1)$ is the line $x_1 = 0$, $x_2 = (\lambda - 1)/2$.■

4.2 The Lyapunov - Schmidt Method

Henceforth, we consider equations of the form (4.1.1) where $f : X \times \mathbb{R} \to Y$ is a continuous function, X and Y being Banach spaces.

Definition 4.2.1 *A linear operator $T : X \to Y$ between Banach spaces X and Y is said to be a* **Fredhölm operator** *if* $\mathrm{Ker}(T)$ *has finite dimension and the range $\mathcal{R}(T)$ has finite codimension (i.e.* $\dim(Y/\mathcal{R}(T)) < \infty$*).* ■

Let $f : X \times \mathbb{R} \to Y$ be a C^p map for some $p \geq 1$. Assume that

$$f(0, \lambda) = 0$$

for all $\lambda \in \mathbb{R}$. Assume further that $\partial_x f(0,\lambda_0) : X \to Y$ is a Fredhölm map. Thus

$$X_1 = \mathrm{Ker}(\partial_x f(0,\lambda_0)), \qquad \dim(X_1) < \infty$$
$$Y_1 = \mathcal{R}(\partial_x f(0,\lambda_0)), \qquad \mathrm{codim}(Y_1) < \infty.$$

Hence, there exists a closed subspace X_2 of X and a finite dimensional subspace Y_2 of Y such that

$$X = X_1 \oplus X_2, \quad Y = Y_1 \oplus Y_2.$$

Let P be the projection from Y onto Y_1. Thus, to solve (4.1.1), we need to solve the following equivalent set of equations

$$\left.\begin{aligned} Pf(x,\lambda) &= 0 \\ (I-P)f(x,\lambda) &= 0. \end{aligned}\right\} \qquad (4.2.1)$$

Consider the map $g : (X_1 \times \mathbb{R}) \times X_2 \to Y_1$ defined by

$$g((x_1,\lambda),x_2) = Pf(x_1+x_2,\lambda).$$

Then

$$\partial_{x_2} g((0,\lambda_0),0) = P\partial_x f(0,\lambda_0)|_{X_2} = \partial_x f(0,\lambda_0)|_{X_2}$$

which is clearly an isomorphism. Hence, by the implicit function theorem, there exists a neighbourhood $\mathcal{U}$ of $(0,\lambda_0)$ in $X_1 \times \mathbb{R}$ and a neighbourhood $\mathcal{V}$ of 0 in X_2 and a C^p function $u : \mathcal{U} \to \mathcal{V}$ such that

$$Pf(x_1+u(x_1,\lambda),\lambda) = 0$$

for all $(x_1,\lambda) \in \mathcal{U}$ and these are the only solutions in that neighbourhood. Thus the first equation in (4.2.1) is already satisfied and hence we are reduced to solving the equation

$$(I-P)f(x_1+u(x_1,\lambda),\lambda) = 0. \qquad (4.2.2)$$

Notice that the above equation involves the space $X_1 \times \mathbb{R}$ and Y_2 which are both finite dimensional. Equation (4.2.2) is called the *bifurcation equation.*

If $\dim(Y_2) = 1$, then there exists $y^* \in Y'$, where Y' is the dual space of Y, such that the bifurcation equation (4.2.2) is reduced to

$$< y^*, f(x_1+u(x_1,\lambda),\lambda) > = 0 \qquad (4.2.3)$$

where $< .,. >$ denotes the duality bracket between Y' and Y.

By the implicit function theorem, we can also calculate the derivatives of u. We have

$$\begin{aligned} \frac{\partial u}{\partial x_1}(0,\lambda_0) &= \left(\frac{\partial g}{\partial x_2}((0,\lambda_0),0)\right)^{-1}\left(\frac{\partial g}{\partial x_1}((0,\lambda_0),0)\right), \\ \frac{\partial u}{\partial \lambda}(0,\lambda_0) &= \left(\frac{\partial g}{\partial x_2}((0,\lambda_0),0)\right)^{-1}\left(\frac{\partial g}{\partial \lambda}((0,\lambda_0),0).\right) \end{aligned}$$

But, since $\mathrm{Ker}(\partial_x f(0,\lambda_0)) = X_1$, we have

$$\frac{\partial g}{\partial x_1}((0,\lambda_0),0) = P\partial_x f(0,\lambda_0)|_{X_1} = 0.$$

Thus

$$\frac{\partial u}{\partial x_1}(0,\lambda_0) = 0; \quad \frac{\partial u}{\partial \lambda}(0,\lambda_0) = [\partial_x f(0,\lambda_0)|_{X_2}]^{-1}[P\partial_\lambda f(0,\lambda_0)].$$

Remark 4.2.1 If $\partial_x f(0,\lambda_0)$ is surjective, then $P = I$ and so there is no bifurcation equation. The solution set near $(0,\lambda_0)$ is a finite dimensional submanifold of X described by

$$\{(x_1+u(x_1,\lambda),\lambda) \mid (x_1,\lambda) \in \mathcal{U}\}.\blacksquare$$

4.3 Morse's Lemma

The following result, due to Morse, allows us to describe the structure of the set of zeros of a nonlinear equation near a bifurcation point in some cases. We follow the treatment given by Nirenberg [19].

Theorem 4.3.1 *(Morse's Lemma) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^p map, for some $p \geq 2$. Assume that $f(0) = 0$, $f'(0) = 0$ and that $f''(0)$ is a non-singular matrix. Then, in a neighbourhood of the origin,* *there exists a change of coordinates $x \mapsto y(x)$ which is a C^{p-2} map such that $y(0)=0$, $y'(0)=I$ and*

$$f(x) = \frac{1}{2}(f''(0)y(x), y(x)). \tag{4.3.1}$$

Proof: We will look for a matrix $R(x)$ such that $R(0)=I$ and set $y(x)=R(x)x$. Then, clearly, $y(0)=0$ and $y'(0)=I$. We will then try to express $f(x)$ in the form $\frac{1}{2}(R(x)^T f''(0)R(x)x, x)$ where A^T stands for the transpose of a matrix A.
Step 1. By the fundamental theorem of calculus, we have

$$f(x) = \int_0^1 f'(tx)x\,dt.$$

Integrating by parts, we get

$$f(x) = f'(x)x - \int_0^1 t(f''(tx)x, x)dt = \int_0^1 (1-t)(f''(tx)x, x)dt.$$

Step 2. Set

$$B(x) = 2\int_0^1 (1-t)f''(tx)dt.$$

Notice that $B(x)$ is a symmetric matrix and that $B(0)=f''(0)$. Let $\mathcal{M}$ be the space of all $n\times n$ matrices and let $\mathcal{S}$ denote the subspace of symmetric matrices of order n. Consider the C^{p-2} map $g:\mathbb{R}^n\times\mathcal{M}\to\mathcal{S}$ defined by

$$g(x,R) = R^T f''(0)R - B(x).$$

Then $g(0,I)=0$. Further,

$$\frac{\partial g}{\partial R}(0,I)S = S^T f''(0) + f''(0)S \in \mathcal{S}.$$

If $S\in\mathcal{S}$, then $T=\frac{1}{2}(f''(0))^{-1}S$ is such that $\frac{\partial g}{\partial R}(0,I)T=S$. Thus $\frac{\partial g}{\partial R}(0,I)$ is surjective.

Step 3. Let $\mathcal{M}_1=\mathrm{Ker}(\frac{\partial g}{\partial R}(0,I))$. Set $\mathcal{M}=\mathcal{M}_1\oplus\mathcal{M}_2$. Let $A\in\mathcal{M}$ and consider the operator $P(A)=\frac{1}{2}(A-(f''(0))^{-1}A^T f''(0))$. Then it is immediate to see that $P(A)\in\mathcal{M}_1$ for all $A\in\mathcal{M}$ and also that for $A\in\mathcal{M}_1$, we have $P(A)=A$. Thus, P is the projection of $\mathcal{M}$ onto $\mathcal{M}_1$. Since $P(I)=0$, it follows that $I\in\mathcal{M}_2$.

Step 4. Now consider $\varphi:(\mathcal{M}_1\times\mathbb{R}^n)\times\mathcal{M}_2\to\mathcal{S}$ defined by

$$\varphi((R_1,x),R_2) = g(x,R_1+R_2).$$

Then $\varphi((0,0),I)=g(0,I)=0$ and

$$\frac{\partial\varphi}{\partial R_2}((0,0),I) = \left.\frac{\partial g}{\partial R}(0,I)\right|_{\mathcal{M}_2}$$

which is an isomorphism from $\mathcal{M}_2$ onto $\mathcal{S}$. Thus, by the implicit function theorem,there exists a neighbourhood of the origin in $\mathcal{M}_1\times\mathbb{R}^n$ and a C^{p-2} map u from this neighbourhood into $\mathcal{M}_2$ such that $u(0,0)=I$ and $g(x,R_1+u(R_1,x))=0$ for all (R_1,x) in that neighbourhood. Now set $R(x)=u(0,x)$ so that, in a neighbourhood of the origin in $\mathbb{R}^n$, we have $g(x,R(x))=0$, *i.e.* $R(x)^T f''(0)R(x)=B(x)$, which proves the result. ■

Corollary 4.3.1 *Let $n=2$ and let f be as in the preceding theorem. If $f''(0)$ is an indefinite matrix then the set of solutions to the equation $f(x)=0$ near the origin is a pair of curves which intersect only at the origin. If $p>2$, these curves are C^1 and they cut transversally.* ■

Remark 4.3.1 In general if $n>2$, and if $f''(0)$ is indefinite, then the solution set near the origin is in the form of a deformed cone. ■

We will now consider some applications of Morse's lemma.

Theorem 4.3.2 *Let f be a C^p map from a Banach space X into a Banach space Y, for some $p\geq 2$. Assume that $f(0)=0$ and that $f'(0)$ is a Fredhölm operator. Let $X_1=\mathrm{Ker}(f'(0))$ be of dimension* *n and let $Y_1 = \mathcal{R}(f'(0))$ be of codimension 1, so that there exists $y^* \in Y'$ such that*

$$Y_1 \;=\; \{y \in Y \mid \; < y^*, y >= 0\}.$$

Assume that $y^* \frac{\partial^2 f}{\partial x_1^2}(0)$ is a nonsingular and indefinite matrix. Then, in a neighbourhood of the origin, the set of solutions of $f(x) = 0$ consists of a deformed cone of dimension $n-1$ with vertex at the origin. In particular, if $n = 2$, then it consists of two C^{p-2} curves crossing only at the origin (transversally, if $p > 2$).

If $y^* \frac{\partial^2 f}{\partial x_1^2}(0)$ is positive (or negative) definite, then the origin is the only local solution of the equation.

Proof: Proceeding as in the Lyapunov- Schmidt method, there exists a C^p map u from the neighbourhood of the origin in X_1 into a neighbourhood of the origin of its complement X_2 such that the only solutions in a neighbourhood of the origin in X of $f(x) = 0$ are given by those of (cf. (4.2.3))

$$g(x_1) \;=\; < y^*, f(x_1 + u(x_1)) > \;=\; 0.$$

We apply the Morse lemma to the above equation to deduce the desired result.

First of all, we know that (cf. Section 4.2) $u(0) = 0$ and that $u'(0) = 0$. Thus $g(0) = 0$. Now, if $z \in X_1$,

$$g'(0)z \;=\; < y^*, f'(0) \circ (I + u'(0))z > \;=\; < y^*, f'(0)z > \;=\; 0$$

since X_1 is the kernel of $f'(0)$.

Finally, we compute $g''(0)$. We recall the formula for the second derivative of a composite function (cf. Cartan [4]). If $H = G \circ F$ and if $F(a) = b$, then

$$H''(a)(z_1, z_2) \;=\; G'(b)(F''(a)(z_1, z_2)) + G''(b)(F'(a)z_1, F'(a)z_2).$$

Setting $\varphi(x_1) = f(x_1 + u(x_1))$, we immediately see that

$$g''(0)(x_1, x_1') \;=\; < y^*, \varphi''(0)(x_1, x_1') >$$

for any $x_1, x_1' \in X_1$. Again $\varphi = f \circ (I + u)$ and so

$$\varphi''(0)(x_1, x_1') \;=\; f'(0)(u''(0)(x_1, x_1')) + f''(0)((I + u'(0))x_1, (I + u'(0))x_1').$$

Since y^* vanishes on the image of $f'(0)$, we easily see that

$$g''(0)(x_1, x_1') \;=\; < y^*, f''(0)(x_1, x_1') > .$$

The result now follows from the indefiniteness of $g''(0)$, as guaranteed by the hypotheses, as a direct consequence of the lemma of Morse. ■

Remark 4.3.2 If $n = 2$, and $p > 2$, then the solution set near the origin consists of two curves cutting each other transversally. Their slopes are given by the vectors which make the indefinite form $< y^*, \frac{\partial^2 f}{\partial x_1^2}(0)(v, v) >$ vanish. ■

Let $f : X \times \mathbb{R} \to Y$ be a C^p function for some $p \geq 2$. Assume that $f(0, \lambda) = 0$ for all $\lambda \in \mathbb{R}$. We have seen that in order that $(0, \lambda_0)$ be a bifurcation point it is necessary that $\partial_x f(0, \lambda_0)$ is not an isomorphism and that this condition is not sufficient. We now prove a result which gives further conditions to ensure that such a point is a bifurcation point.

Theorem 4.3.3 *Let $f : X \times \mathbb{R} \to Y$ be a C^p map for some $p \geq 2$. Assume that $f(0, \lambda_0) = 0$. Assume further that*
(i) $\partial_\lambda f(0, \lambda_0) = 0$.
(ii) $\mathrm{Ker}(\partial_x f(0, \lambda_0))$ is one dimensional and spanned by $x_0 \in X$.
(iii) $\mathcal{R}(\partial_x f(0, \lambda_0)) = Y_1$ which has codimension 1.
(iv) With the obvious identifications,

$$\partial_{\lambda\lambda} f(0, \lambda_0) \in Y_1 \quad \text{and} \quad \partial_{\lambda x} f(0, \lambda_0)x_0 \notin Y_1.$$

Then, $(0, \lambda_0)$ is a bifurcation point and the set of solutions to $f(x, \lambda) = 0$ near $(0, \lambda_0)$ consists of two C^{p-2} curves Γ_1 and Γ_2

cutting only at $(0,\lambda_0)$. *Further, if* $p>2$, Γ_1 *is tangent to the* λ-*axis at* $(0,\lambda_0)$ *and can be parametrized by* λ); *i.e.*

$$\Gamma_1 = \{(x(\lambda),\lambda) \mid |\lambda| \le \varepsilon\}.$$

Γ_2 *can be parametrized as*

$$\Gamma_2 = \{(sx_0 + x_2(s), \lambda(s)) \mid |s| \le \varepsilon\}$$

with $x_2(0)=0,\ x_2'(0)=0,\ \lambda(0)=\lambda_0$.

Proof: Set $\widehat{X} = X \times \mathbb{R}$ and define $F : \widehat{X} \to Y$ by $F(\widehat{x}) = f(x,\lambda)$ where $\widehat{x} = (x,\lambda)$. Then with obvious notation, we have $F'((0,\lambda_0)) = \partial_x f(0,\lambda_0) \oplus \partial_\lambda f(0,\lambda_0)$. Thus, $\text{Ker}(F'((0,\lambda_0))$ is two dimensional and is spanned by $(x_0,0)$ and $(0,1)$ by virtue of conditions (i) and (ii) above. The image of $F'((0,\lambda_0))$ is Y_1 and has codimension 1; let $y^* \in Y'$ annihilate Y_1. Then the matrix $y^*\partial_{\widehat{x}_1\widehat{x}_1}F((0,\lambda_0))$, where $\widehat{x}_1$ is the generic variable in $\widehat{X}_1 = \text{Ker}(F'((0,\lambda_0)))$, is given by (with the obvious identifications)

$$\begin{bmatrix} < y^*, \partial_{x_0x_0} f(0,\lambda_0) > & < y^*, \partial_{x_0\lambda} f(0,\lambda_0) > \\ < y^*, \partial_{x_0\lambda} f(0,\lambda_0) > & < y^*, \partial_{\lambda\lambda} f(0,\lambda) > \end{bmatrix}.$$

By condition (iv) above, the off-diagonal terms (which are equal) are non-zero while the last term in the diagonal vanishes. Thus, the determinant of this matrix is strictly negative and hence the matrix is non-singular and indefinite. Hence, from the previous theorem, we deduce the existence of two branches of solutions which will cut each other transversally when $p>2$. The slopes of these branches at the bifurcation point are given by the vectors which make the quadratic form associated to the above matrix vanish. Since the last diagonal term is zero, one such vector is $(0,1)$. Thus one of the curves is tangent to the λ- axis at the bifurcation point. ■.

Remark 4.3.3 If $f(0,\lambda)=0$ for all $\lambda \in \mathbb{R}$, then Γ_1 is the λ- axis itself. ■

Example 4.3.1 Let $g:\mathbb{R}\to\mathbb{R}$ be a C^1 function with bounded derivative such that $g(0)=0$ and $g'(0)\neq 0$. Let $\Omega \subset \mathbb{R}^n$ be a bounded domain. Consider the problem

$$\left.\begin{aligned} \Delta u &= \lambda g(u) && \text{in } \Omega \\ u &= 0 && \text{on } \partial\Omega. \end{aligned}\right\} \tag{4.3.2}$$

Notice that for all $\lambda \in \mathbb{R}$, we have the trivial solution $u \equiv 0$. We look for solutions in $H_0^1(\Omega)$. By hypotheses, $u \mapsto g(u)$ is a mapping of $L^2(\Omega)$ into itself and so the map $u \mapsto T(u,\lambda) = w \in H_0^1(\Omega)$, where w is the unique solution of the problem

$$\left.\begin{aligned} \Delta w &= \lambda g(u) && \text{in } \Omega \\ w &= 0 && \text{on } \partial\Omega, \end{aligned}\right\}$$

is a compact map of $L^2(\Omega)$ into itself. Thus the solutions of (4.3.2) are just the solutions of

$$u - T(u,\lambda) = 0.$$

It is easy to see that $\partial_u T(0,\lambda)v = z$ where $z \in H_0^1(\Omega)$ is the solution of the problem

$$\left.\begin{aligned} \Delta z &= \lambda g'(0)v && \text{in } \Omega \\ z &= 0 && \text{on } \partial\Omega. \end{aligned}\right\}$$

Again, $\partial_u T(0,\lambda)$ is a compact operator and it is also self-adjoint. Assume that $\lambda_0 \neq 0$ is a simple characteristic value (*i.e.* λ_0^{-1} is a simple eigenvalue) of this operator with normalized eigenfunction $\varphi \in H_0^1(\Omega)$ (*i.e.* $(\varphi,\varphi) = \int_\Omega \varphi^2 dx = 1$). Since the operator is self-adjoint, it follows that $Y_1 = \mathcal{R}(I - \partial_u T(0,\lambda_0))$ has codimension one and is the orthogonal complement (in $L^2(\Omega)$) of φ.

Now $\partial_\lambda T(u,\lambda)$ can be identified with an element of $L^2(\Omega)$ and is, in fact, the solution (in $H_0^1(\Omega)$) of the problem

$$\left.\begin{aligned} \Delta w &= g(u) && \text{in } \Omega \\ w &= 0 && \text{on } \partial\Omega. \end{aligned}\right\}$$

Thus, clearly $\partial_\lambda T(0,\lambda) = 0$ and so $\partial_\lambda(I - T(0,\lambda_0)) = 0$. Also $\partial_{\lambda\lambda}(I - T(0,\lambda_0)) = 0 \in Y_1$. Finally, $\partial_{\lambda u}(I - T(0,\lambda_0))\varphi$ can also be identified with an element of $L^2(\Omega)$ as the solution (in $H_0^1(\Omega)$) of the problem

$$\left.\begin{array}{rcll} \Delta w & = & g'(0)\varphi & \text{in } \Omega \\ w & = & 0 & \text{on } \partial\Omega. \end{array}\right\}$$

It is easy to see that $w \notin Y_1$. For,

$$\begin{array}{rcl} (w,\varphi) & = & \frac{1}{\lambda_0 g'(0)}(w,\Delta\varphi) \;=\; \frac{1}{\lambda_0 g'(0)}(\Delta w,\varphi) \\ & = & \frac{1}{\lambda_0}(\varphi,\varphi) \neq 0. \end{array}$$

Thus, by the preceding theorem, if λ_0 is a simple eigenvalue of the linearized problem, then $(0,\lambda_0)$ is a bifurcation point.■

Remark 4.3.4 We go back to Example 4.1.1. We can treat the equation for φ in exactly the same way as the preceding example. In this case the characteristic values are all simple and so all these values yield bifurcation points. When the rod is subjected to a compressive load, it will just contract and remain in the undeflected state. When the force reaches a critical level corresponding to the first characteristic value, the rod will buckle and there will be a non-zero vertical displacement. Beyond this value, there are several solutions given by the various bifurcation branches at each successive characteristic value. The actual displacement realised by the rod will be the physically acceptable solution which depends on other criteria like the minimization of the strain energy. ■

Theorem 4.3.3 was also proved by Crandall and Rabinowitz [6] using the implicit function theorem. In some cases we can also deduce it by a perturbation method (which is in fact the Lyapunov-Schmidt method in disguise, and thus, again a consequence of the implicit function theorem). We will see an illustration of this in the next section.

4.4 A Perturbation Method

Let H be a separable Hilbert space and let $L : H \to H$ be a compact and self-adjoint bounded linear operator. Let λ_0 be a simple characteristic value and let φ_0 be an eigenvector such that

$$\varphi_0 = \lambda_0 L\varphi_0,\ (L\varphi_0,\varphi_0) = 1. \tag{4.4.1}$$

Let $A : H \to H$ be a compact map such that $A(0) = 0$ and for which the following property holds: There exists a constant $C > 0$ such that if $\|u\| \le r$ and $\|v\| \le r$, where $u, v \in H$, then

$$\|A(u) - A(v)\| \le Cr^2\|u - v\|. \tag{4.4.2}$$

Let us consider solutions $(u,\lambda) \in H \times \mathbb{R}$ of the equation

$$u - \lambda Lu + A(u) = 0. \tag{4.4.3}$$

An example of this situation occurs in the study of buckling of clamped plates via the von Karman equations (cf. Kesavan [12]).

By hypotheses, the trivial solution $u = 0$ is valid for all values of λ. Bifurcation can occur only at the characteristic values of L. It is trivial to check that all the hypotheses of Theorem 4.3.3 are verified and thus $(0,\lambda_0)$ is a bifurcation point and there are two branches, one of which is the λ- axis itself.

We now will prove the existence of the other branch via a perturbation method, which also provides a method for actually computing the branch of nontrivial solutions bifurcating from the trivial branch.

The idea is to look for solutions of the form

$$u = \varepsilon\varphi_0 + v \tag{4.4.4}$$

where ε is a small parameter, $(v,\varphi_0) = 0$ and $\|v\| \le \varepsilon$. Substituting this in (4.4.3), we get

$$\varepsilon\varphi_0 + v - \lambda L(\varepsilon\varphi_0 + v) + A(\varepsilon\varphi_0 + v) = 0$$

or, equivalently,

$$v - \lambda_0 Lv = (\lambda - \lambda_0)Lu - A(u). \tag{4.4.5}$$

But this equation has a solution if, and only if (by the Fredhölm alternative), the right-hand side is orthogonal to φ_0. Further, the solution v will be unique if we impose the additional condition that $(v, \varphi_0) = 0$. Hence

$$(\lambda - \lambda_0)(Lu, \varphi_0) = (A(u), \varphi_0).$$

Now, using (4.4.1), we get

$$\begin{aligned} (Lu, \varphi_0) &= \varepsilon(L\varphi_0, \varphi_0) + (Lv, \varphi_0) \\ &= \varepsilon + (v, L\varphi_0) \\ &= \varepsilon + \lambda_0^{-1}(v, \varphi_0) \\ &= \varepsilon. \end{aligned}$$

Consequently,

$$\lambda = \lambda_0 + \varepsilon^{-1}(A(u), \varphi_0). \tag{4.4.6}$$

Let us define, for $w \in H$,

$$\left.\begin{aligned} \Lambda_\varepsilon w &= \lambda_0 + \varepsilon^{-1}(A(w), \varphi_0) \\ S_\varepsilon w &= (\Lambda_\varepsilon w - \lambda_0)Lw - A(w). \end{aligned}\right\} \tag{4.4.7}$$

If P_0 is the orthogonal projection onto the orthogonal complement of φ_0, *i.e.* the subspace $\{\varphi_0\}^\perp$, let Qw denote the solution $v \in \{\varphi_0\}^\perp$ of the equation

$$v - \lambda_0 Lv = P_0 w. \tag{4.4.8}$$

Notice that Q is a bounded linear map. There exists a constant $C > 0$ such that

$$\|Qw\| \le C\|w\|.$$

(If not, we can find a sequence $\{w_n\}$ in H such that $w_n \to 0$ in H while $z_n = Qw_n$ is such that $\|z_n\| = 1$ for all n. Then, for a subsequence, $z_n \rightharpoonup z$ weakly in H and by the compactness of L, it follows that

$$z - \lambda_0 Lz = 0.$$

Hence $z = \alpha\varphi_0$ for some $\alpha \in \mathbb{R}$. But $(z_n, \varphi_0) = 0$ implies that $(z, \varphi_0) = 0$ and so it follows that $z = 0$. Again, by the compactness of L, it follows that $z_n = \lambda_0 Lz_n + P_0 w_n$ converges strongly in H and so $\|z\| = 1$, which gives a contradiction.) Thus, for $v \in \{\varphi_0\}^\perp$, we define $T_\varepsilon v \in \{\varphi_0\}^\perp$ by

$$T_\varepsilon v = Q(S_\varepsilon(\varepsilon\varphi_0 + v)) \tag{4.4.9}$$

and we are looking for fixed points of T_ε such that $\|v\| \le \varepsilon$ in view of (4.4.5) and (4.4.6).

We will now show that T_ε is a contraction of the closed ball of radius ε in $\{\varphi_0\}^\perp$ into itself. This will then prove the existence of a unique fixed point for each ε and will also provide an algorithm to compute it.

Lemma 4.4.1 *Let*

$$\begin{aligned} B_\varepsilon &= \{v \in H \mid (v, \varphi_0) = 0,\ \|v\| \le \varepsilon\}, \\ U_\varepsilon &= \{\varepsilon\varphi_0 + v \mid v \in B_\varepsilon\}. \end{aligned}$$

There exists a constant $C > 0$, independent of ε such that for all $u, u_1, u_2 \in U_\varepsilon$, we have

$$|\Lambda_\varepsilon u - \lambda_0| \le C\varepsilon^2 \tag{4.4.10}$$

$$|\Lambda_\varepsilon u_1 - \Lambda_\varepsilon u_2| \le C\varepsilon\|u_1 - u_2\| \tag{4.4.11}$$

$$\|S_\varepsilon u\| \le C\varepsilon^3 \tag{4.4.12}$$

$$\|S_\varepsilon u_1 - S_\varepsilon u_2\| \le C\varepsilon^2\|u_1 - u_2\|. \tag{4.4.13}$$

Proof: We have

$$|\Lambda_\varepsilon u - \lambda_0| \le \varepsilon^{-1}\|A(u)\|\|\varphi_0\|.$$

But by (4.4.2),

$$\|A(u)\| = \|A(u) - A(0)\| \le C\varepsilon^2\|u\| \le C\varepsilon^3.$$

This proves (4.4.10). Again,

$$|\Lambda_\varepsilon u_1 - \Lambda_\varepsilon u_2| = \varepsilon^{-1}|(A(u_1) - A(u_2), \varphi_0)|$$

and so (4.4.11) follows from (4.4.2). Next,

$$\begin{aligned} \|S_\varepsilon u\| &\leq |\Lambda_\varepsilon u - \lambda_0|\|Lu\| + \|A(u)\| \\ &\leq C\varepsilon^2\|u\| + C\varepsilon^3 \leq C\varepsilon^3. \end{aligned}$$

Finally,

$$\begin{aligned} \|S_\varepsilon u_1 - S_\varepsilon u_2\| \leq &|\Lambda_\varepsilon u_1 - \lambda_0|\|L(u_1 - u_2)\| + \\ &+|\Lambda_\varepsilon u_1 - \Lambda_\varepsilon u_2|\|Lu_2\| + \|A(u_1) - A(u_2)\| \end{aligned}$$

and (4.4.13) follows from (4.4.10), (4.4.11) and (4.4.2). ■

Proposition 4.4.1 *There exists $\varepsilon_0 > 0$ such that, for all $0 < \varepsilon < \varepsilon_0$, the map T_ε is a contraction of B_ε into itself.*

Proof: By the preceding lemma, if $v \in B_\varepsilon$, we have

$$\|T_\varepsilon v\| = \|Q(S_\varepsilon(\varepsilon\varphi_0 + v))\| \leq C_1\varepsilon^3$$

and, similarly,

$$\|T_\varepsilon v_1 - T_\varepsilon v_2\| \leq C_2\varepsilon^2\|v_1 - v_2\|.$$

Thus if we choose ε_0 such that $C_1\varepsilon_0^2 < 1$ and $C_2\varepsilon_0^2 < 1$, we get that the desired result. ■

We are thus led to the following algorithm for the computation of the branch of nontrivial solutions of (4.4.3) near $(0, \lambda_0)$.

Step 1. Choose $v^0 \in B_\varepsilon$. For instance $v^0 = 0$.
Step 2. Assume that v^i has been computed. Set $u^i = \varepsilon\varphi_0 + v^i$.
Step 3. Compute

$$\begin{aligned} \lambda^{i+1} &= \lambda_0 + \varepsilon^{-1}(A(u^i), \varphi_0) \\ S_\varepsilon u^i &= (\lambda^{i+1} - \lambda_0)Lu^i - A(u^i). \end{aligned}$$

Step 4. Solve for $v^{i+1} \in \{\varphi_0\}^\perp$, the linear equation

$$v^{i+1} - \lambda_0 Lv^{i+1} = S_\varepsilon u^i.$$

Then $u^i \to u_\varepsilon$ and $\lambda^i \to \lambda_\varepsilon$, where $(u_\varepsilon, \lambda_\varepsilon)$ will be a nontrivial solution of (4.4.3) with $\|u_\varepsilon\| \leq C\varepsilon$ and $|\lambda_\varepsilon - \lambda_0| \leq C\varepsilon^2$.

4.5 Krasnoselsk'ii's Theorem

Let X be a Banach space and let $L : X \to X$ be a compact bounded linear operator. Let $g : X \times \mathbb{R} \to X$ be a compact mapping. We are interested in solutions of the equation

$$u - \lambda Lu + g(u, \lambda) = 0. \tag{4.5.1}$$

Assume that $g(0, \lambda) = 0$ for all $\lambda \in \mathbb{R}$ so that we always have the trivial branch of solutions. Let λ_0 be a characteristic value of L and let $g(x, \lambda) = o(\|x\|)$ uniformly in a neighbourhood of λ_0 so that $\partial_u f(0, \lambda_0)$, where $f(u, \lambda) = u - \lambda Lu + g(u, \lambda)$, is given by $I - \lambda_0 L$ which is not an isomorphism. Thus $(0, \lambda_0)$ satisfies the necessary condition for being a bifurcation point. The following theorem, due to Krasnoselsk'ii, provides a sufficient condition for it to be a bifurcation point.

Theorem 4.5.1 *(Krasnoselsk'ii) Let λ_0 be a characteristic value of odd (algebraic) multiplicity of L. Then $(0, \lambda_0)$ is a bifurcation point.*

Proof: Assume the contrary. Then there exist $\varepsilon_0 > 0$ and $\varepsilon_1 > 0$ such that, for $|\lambda - \lambda_0| \leq \varepsilon_0$ and for $\|x\| \leq \varepsilon_1$, $x \neq 0$, we have $f(x, \lambda) \neq 0$. Thus for any fixed λ such that $|\lambda - \lambda_0| < \varepsilon_0$, the degree $d(f(.,\lambda), B(0;\varepsilon_1), 0)$ is well defined ($f(.,\lambda)$ is a compact perturbation of the identity) and this degree is independent of λ by homotopy invariance. Let $\lambda_1 < \lambda_0$ and $\lambda_2 > \lambda_0$ with $|\lambda_j - \lambda_0| < \varepsilon_0$ for $j = 1, 2$. Then (cf. Proposition 3.5.3) we know that

$$d(f(.,\lambda_j), B(0;\varepsilon_1), 0) = i(f(.,\lambda_j), 0, 0) = (-1)^{\beta(\lambda_j)}$$

for $j = 1, 2$, where $\beta(\lambda)$ is the sum of the algebraic multiplicities of the characteristic values of L between 0 and λ. But we can always choose ε_0 sufficently small such that λ_0 is the only characteristic value of L in the interval $(\lambda_0 - \varepsilon_0, \lambda_0 + \varepsilon_0)$. Then, it follows that

$$\beta(\lambda_2) = \beta(\lambda_1) \pm \text{ multiplicity of } \lambda_0$$

and thus

$$d(f(.,\lambda_1), B(0;\varepsilon_1), 0) = -d(f(.,\lambda_2), B(0;\varepsilon_1), 0)$$

which is a contradiction. This proves the theorem. ■

If the multiplicity of λ_0 is even, we cannot guarantee that $(0, \lambda_0)$ is a bifurcation point as can be seen from Example 4.1.3. If g is smooth and if λ_0 is a simple charactersitic value, then it is a simple exercise to check that the hypotheses of Theorem 4.3.3 are verified and that we have two branches of solutions (one of which is the λ-axis) cutting each other transversally. If the multiplicity is not unity, then Theorem 4.3.3 does not hold and while we still have bifurcation at $(0, \lambda_0)$, the non-trivial branch may not cut the trivial one transversally, as the following example shows.

Example 4.5.1 (cf. Nirenberg [19]) Let $v : S^2 \to \mathbb{R}^3$ be a vector field vanishing only at $(0,0,1)$, *i.e.* $(v(y), y) = 0$ for all $y \in S^2$ and it vanishes only at the north pole (cf. Proposition 2.2.4). An example of such a vector field is given by

$$v(y) = (1 - y_3 - y_1^2, -y_1y_2, y_1 - y_1y_3)$$

where $y = (y_1, y_2, y_3) \in S^2$. It is easy to see that $(v(y), y) = 0$ for all $y \in S^2$ and that $|v(y)|^2 = (1 - y_3)^2$ so that $v(y) = 0$ if, and only if $y_3 = 1$ and thus $y_1 = y_2 = 0$.

Define, for $x \in \mathbb{R}^3$,

$$g(x) = \begin{cases} e^{-\frac{1}{|x|^2}} v(x/|x|), & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

Let $L = I$. Consider the equation

$$(1 - \lambda)x + g(x) = 0. \qquad (4.5.2)$$

Then $\lambda = 1$ is a characteristic value of multiplicity 3.

Since $(g(x), x) = 0$, it follows that the only nontrivial solutions of (4.5.2) occur when $\lambda = 1$. In this case, we have $g(x) = 0$ and so $x = (0, 0, x_3)$ where $x_3 > 0$. Thus the branch bifurcating from the λ-axis is the positive half-line parallel to the x_3-axis which does not cut the λ-axis transversally. ■

4.6 Rabinowitz' Theorem

In this section, we present a result which could be described as a global bifurcation result. It investigates the global behaviour of a branch of solutions bifurcating from the trivial branch in the context of the result of Krasnoselsk'ii presented in the previous section.

Let X be a Banach space and let $L : X \to X$ be a compact bounded linear operator. Let $g : X \times R \to X$ be a compact mapping such that $g(x, \lambda) = o(\|x\|)$ in the neighbourhood of a characteristic value λ_0 of L. We consider solutions of the equation

$$f(x, \lambda) \equiv x - \lambda Lx + g(x, \lambda) = 0. \qquad (4.6.1)$$

By Krasnoselsk'ii's theorem, if λ_0 is of odd multiplicity, then $(0, \lambda_0)$ is a bifurcation point. Thus it will lie in the closure S of all nontrivial solutions of (4.6.1) in $X \times \mathbb{R}$. The theorem of Rabinowitz investigates the behaviour of the connected component $\mathcal{C}$ in S that contains the point $(0, \lambda_0)$. We present below the proof due to Ize (cf. Nirenberg [19]).

Since L is compact, we can choose $\varepsilon_0 > 0$ such that λ_0 is the only characteristic value of L in the interval $[\lambda_0 - \varepsilon_0, \lambda_0 + \varepsilon_0]$. Thus $I - \lambda L$ is invertible for $\lambda \neq \lambda_0$ in the above interval. Hence, the indices

$$\begin{aligned} i_+ &= i(I - \lambda L, 0, 0) = d(I - \lambda L, B(0;r), 0), \quad \lambda > \lambda_0 \\ i_- &= i(I - \lambda L, 0, 0) = d(I - \lambda L, B(0;r), 0), \quad \lambda < \lambda_0 \end{aligned}$$

are well defined and are independent of λ and $r > 0$, for r sufficiently small. Now define $H_r : X \times \mathbb{R} \to X \times \mathbb{R}$ by

$$H_r(x, \varepsilon) = (f(x, \lambda_0 + \varepsilon), \|x\|^2 - r^2).$$

Now, if $\lambda = \lambda_0 \pm \varepsilon_0$, we know that $\partial_x f(0, \lambda) = I - \lambda L$ is invertible. Hence $x = 0$ is the only solution in $B(0, r)$ for (4.6.1) if r is sufficiently small. Hence, if we set

$$B_{r,\varepsilon_0} = \{(x, \varepsilon) \mid \|x\|^2 + \varepsilon^2 < r^2 + \varepsilon_0^2\},$$

there will be no zero of H_r on the boundary of B_{r,ε_0}. Thus the degree

$$d(H_r, B_{r,\varepsilon_0}, (0,0))$$

is well defined.

Lemma 4.6.1 *(Ize) We have*

$$d(H_r, B_{r,\varepsilon_0}, (0,0)) = i_- - i_+. \tag{4.6.2}$$

Proof: For $0 \le t \le 1$, define

$$H_r^t(x, \varepsilon) = ((I-(\lambda_0+\varepsilon)L)x + tg(x, \lambda_0+\varepsilon), t(\|x\|^2 - r^2) + (1-t)(\varepsilon_0^2 - \varepsilon^2)).$$

For any such t, H_r^t will not vanish on the boundary of $B(0; r) \times (-\varepsilon_0, \varepsilon_0)$. Thus the degree

$$d(H_r^t, B_{r,\varepsilon_0}, (0,0))$$

is well defined and is independent of t. At $t = 0$,

$$H_r^0(x, \varepsilon) = ((I - (\lambda_0 + \varepsilon)L)x, \ \varepsilon_0^2 - \varepsilon^2)$$

and so its derivative at $(0, \varepsilon)$ is given by

$$(H_r^0)'(0, \varepsilon)(x, \eta) = ((I - (\lambda_0 + \varepsilon)L)x, \ -2\varepsilon\eta).$$

Now, H_r^0 vanishes only at the points $(0, -\varepsilon_0)$ and $(0, \varepsilon_0)$. These are isolated solutions and their indices are got by the product formula as i_- for the point $(0, -\varepsilon_0)$ and as $-i_+$ for the point $(0, \varepsilon_0)$ and we deduce (4.6.2). ■

Theorem 4.6.1 *(Rabinowitz) Let λ_0 be a characteristic value of odd multiplicity for L. Let $\mathcal{C}$ be the connected component of $\mathcal{S}$ containing the point $(0, \lambda_0)$. Then*
(i) either $\mathcal{C}$ is not compact,
(ii) or, $\mathcal{C}$ contains a finite number of points of the form $(0, \lambda_j)$ where the λ_j are characteristic values of L, and the number of such λ_j of odd multiplicity is even.
Consequently, a compact component of $\mathcal{S}$ through $(0, \lambda_0)$ must meet the λ-axis again at a point $(0, \lambda')$, where λ' is a characteristic value of odd multiplicity.

Proof: Assume that $\mathcal{C}$ is compact. Since the only accumulation point of the characteristic values of L (which is compact) is at infinity, it follows that there are atmost finitely many points $(0, \lambda_j)$ in $\mathcal{C}$ where λ_j is a characteristic value of L. Let Ω be a bounded open set in $X \times \mathbb{R}$ such that $\mathcal{C} \subset \Omega$, such that $\partial\Omega$ does not contain any nontrivial solution of (4.6.1) and also such that the only points of the form $(0, \lambda)$ in Ω, where λ is a characteristic value of L, are when λ is one of the λ_j mentioned above. Now, define

$$f_r(x, \lambda) = (f(x, \lambda), \ \|x\|^2 - r^2).$$

Notice that $f_r = H_r^1$ defined in the previous lemma when λ lies in the neighbourhood of a characteristic value of L. If $f_r(x, \lambda) = (0, 0)$ on $\partial\Omega$, then it follows that on one hand, $x = 0$, while on the other hand $\|x\| = r$ which is impossible for $r > 0$. Thus the degree $d(f_r, \Omega, (0, 0))$ is well defined and is independent of $r > 0$.

If $r > 0$ is large, then f_r cannot vanish for it will imply that $\|x\| = r$ which cannot hold in the bounded set Ω. Thus

$$d(f_r, \Omega, (0, 0)) = 0.$$

Now, let $r > 0$ be small. If f_r vanishes at (x, λ), then $\|x\| = r$ and λ must lie close to one of the λ_j. (Fix disjoint and small neighbourhoods $(\lambda_j - \varepsilon_j, \lambda_j + \varepsilon_j)$ of these characteristic values. Outside these neighbourhoods, $\|(I - \lambda L)^{-1}\|$ is uniformly bounded and so, for sufficiently small $r > 0$, the only solutions of (4.6.1) are the trivial ones.) Hence, by the preceding lemma and the additivity of the degree,

$$0 = d(f_r, \Omega, (0,0)) = \sum_j (i_-(j) - i_+(j))$$

where $i_-(j)$ and $i_+(j)$ are the indices associated to the characteristic value λ_j. But $i_-(j) = (-1)^{m_j} i_+(j)$, where m_j is the multiplicity of λ_j. Thus the only terms surviving in the above sum are those corresponding to characteristic values of odd multiplicity and, as the sum is zero, there must be an even number of them. This completes the proof. ■

Both possibilities exist as the following examples show.

Example 4.6.1 Consider the equation $u - \lambda L u = 0$. Then the component attached to $(0, \lambda_0)$ is $\{(x, \lambda_0) \mid x \in \text{the eigenspace of } \lambda_0\}$ which is not compact. ■

Example 4.6.2 (Rabinowitz) Let $X = \mathbb{R}^2$ and consider the equation

$$u - \lambda L u - L g(u) = 0$$

where

$$L = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}, \; g(u) = \begin{bmatrix} -u_2^3 \\ u_1^3 \end{bmatrix}.$$

The characteristic values of L are 1 and 2 and both are of odd multiplicity. We have $L^{-1}u = \lambda u + g(u)$ and so

$$\begin{array}{rcl} (1-\lambda)u_1 & = & -u_2^3 \\ (2-\lambda)u_2 & = & u_1^3 \end{array}$$

whence we get $1 \leq \lambda \leq 2$ and

$$\left.\begin{array}{rcl} u_1 & = & \pm(\lambda-1)^{\frac{1}{8}}(2-\lambda)^{\frac{3}{8}} \\ u_2 & = & \pm(\lambda-1)^{\frac{3}{8}}(2-\lambda)^{\frac{1}{8}} \end{array}\right\}.$$

Thus, there is only one connected component of solutions *i.e.* $\mathcal{S} = \mathcal{C}$ and it is compact and meets the λ-axis at both characteristic values. ■

4.7 A Variational Method

We will now return to the problem studied in Section 4.4 with some additional hypotheses. We assume that H is a separable Hilbert space and look for $(u, \lambda) \in H \times \mathbb{R}$ such that

$$u - \lambda L u + A(u) = 0. \qquad (4.7.1)$$

We make the following hypotheses on L and A.
(H1) The bounded linear operator $L : H \to H$ is compact and self-adjoint. Further, for all $v \in H$,

$$(Lv, v) \geq 0$$

with strict inequality if $v \neq 0$.■

(H2) The nonlinear mapping $A : H \to H$ is compact and for all $v \in H$ and for all $t \in \mathbb{R}$,

$$A(tu) = t^3 A(u). \qquad (4.7.2)$$

(In particular, $A(0) = 0$ and there exists a constant $C > 0$ such that $\|A(v)\| \leq C\|v\|^3$ for all $v \in H$). Further

$$(A(v), v) \geq 0 \qquad (4.7.3)$$

for all $v \in H$ with strict inequality if $v \neq 0$.
Finally, we assume that the functional $j(v) = \frac{1}{4}(A(v), v)$ is differentiable in H and that

$$(j'(v), h) = (A(v), h) \qquad (4.7.4)$$

for all $v, h \in H$.■

The above hypotheses, as well as the condition (4.4.2), are all verified in the case of the weak formulation of the von Karman equations modelling the buckling of a thin elastic plate (cf. Kesavan [12]).

In view of the hypothesis (H1), the operator L has a least characteristic value $\lambda_1 > 0$ and it is characterized by

$$\frac{1}{\lambda_1} = \sup_{v\neq 0} \frac{(Lv, v)}{\|v\|^2}. \tag{4.7.5}$$

In view of hypothesis (H2), it follows that $(0, \lambda)$ is always a solution of (4.7.1) for all $\lambda \in \mathbb{R}$.

Proposition 4.7.1 *If $\lambda \leq \lambda_1$, the equation (4.7.1) has only the trivial solution. If $\lambda > \lambda_1$, for each such λ, there exist at least two nontrivial solutions.*

Proof: Step 1. Let $\lambda \leq \lambda_1$ and let u be a solution of (4.7.1). Then,

$$\|u\|^2 - \lambda(Lu, u) + (A(u), u) = 0.$$

Hence, it follows from (4.7.5) that

$$\left(1 - \frac{\lambda}{\lambda_1}\right)\|u\|^2 + (A(u), u) \leq 0.$$

Thus, $(A(u), u) \leq 0$ from which it follows that $u = 0$ in view of (H2).

Step 2. If $\lambda > \lambda_1$, we can find $w \in H$ such that

$$(Lw, w) > \frac{\|w\|^2}{\lambda}. \tag{4.7.6}$$

Now consider the functional

$$J(v) = \frac{1}{2}\|v\|^2 - \frac{\lambda}{2}(Lv, v) + \frac{1}{4}(A(v), v). \tag{4.7.7}$$

Setting $\varphi(t) = J(tw)$, we get

$$\varphi(t) = \frac{t^2}{2}(\|w\|^2 - \lambda(Lw, w)) + \frac{t^4}{4}(A(w), w)$$

so that $\varphi(t) \to \infty$ when $t \to \pm\infty$. But, in view of (4.7.6), it follows that $\varphi(t) < 0$ for t small. Thus there do exist $v \in H$ such that $J(v) < 0$ and so

$$\inf_{v\in H} J(v) < 0. \tag{4.7.8}$$

Step 3. We claim that J is coercive, *i.e.* $J(v) \to +\infty$ when $\|v\| \to +\infty$. If not, there exists $\alpha > 0$ and a sequence $\{v_n\}$ in H such that $J(v_n) \leq \alpha$ and $\|v_n\| \to +\infty$ as $n \to \infty$. we set $w_n = v_n/\|v_n\|$ so that $\|w_n\| = 1$ for all n. Then

$$\alpha \geq J(v_n) = \frac{1}{2}\|v_n\|^2 - \frac{\lambda}{2}(Lv_n, v_n) + \frac{1}{4}(A(v_n), v_n) \tag{4.7.9}$$

i.e.

$$\frac{\alpha}{\|v_n\|^4} \geq \frac{1 - \lambda(Lw_n, w_n)}{2\|v_n\|^2} + \frac{1}{4}(A(w_n), w_n). \tag{4.7.10}$$

Since $\{w_n\}$ is bounded, working with an appropriate subsequence, we may assume that $w_n \rightharpoonup w$ weakly in H. Passing to the limit, using the compactness of L and A, we deduce from (4.7.10) that $(A(w), w) \leq 0$, so that $w = 0$. Again, since $(A(v_n), v_n) \geq 0$, we deduce from (4.7.9) that

$$\frac{\alpha}{\|v_n\|^2} \geq \frac{1}{2} - \frac{\lambda}{2}(Lw_n, w_n)$$

which yields a contradiction on passing to the limit as $n \to \infty$. Thus J is coercive.

Step 4. Let $\{u_n\}$ be a minimizing sequence in H, *i.e.* $J(u_n) \to \inf_{v\in H} J(v) < 0$. By the coercivity of J, it follows that $\{u_n\}$ is bounded and so working with a weakly convergent subsequence, we may assume that $u_n \rightharpoonup u$ weakly in H. By the compactness of L and A, it is easy to see that J attains its infimum at u and since this infimum is negative, it follows that $u \neq 0$. Thus (cf. Theorem 1.4.1) $J'(u) = 0$ and in view of (4.7.4), this is precisely the equation (4.7.1). Again, thanks to (4.7.2), it follows that $-u$ is also a solution. This completes the proof. ■

Upto now, we have the following information about the solutions of (4.7.1).

- If $\lambda \leq \lambda_1$, then we only have the trivial solution.
- If $\lambda > \lambda_1$, then we have the trivial solution and at least two nontrivial solutions.
- If λ is a characteristic value of L with odd multiplicity, then, by Krasnoselsk'ii's theorem, $(0, \lambda)$ is a bifurcation point.
- If λ is a simple characteristic value, then we can check that all the hypotheses of Theorem 4.3.3 are verified and thus $(0, \lambda)$ is again a bifurcation point and we have a curve of solutions branching from the trivial branch.

We will now show, by a different method, that $(0, \lambda_1)$ is always a bifurcation point. Notice that this is not necessarily covered by the cases listed above.

As observed in the proof of Proposition 4.7.1, we can obtain solutions of (4.7.1) by finding the extrema of the functional J defined by (4.7.7). Another method is to find the extrema of the functional $v \mapsto (Lv, v)$ over the set

$$\partial A_r \;=\; \{v \in H \mid \frac{1}{2}\|v\|^2 + \frac{1}{4}(A(v), v) = r\}. \tag{4.7.11}$$

Then λ will appear in the form of a Lagrange multiplier (cf. Theorem 1.4.2 and Remark 1.4.3). For convenience, let us set

$$F(v) \;=\; \frac{1}{2}\|v\|^2 + \frac{1}{4}(A(v), v).$$

Proposition 4.7.2 *The functional $v \mapsto (Lv, v)$ does not attain a minimum on ∂A_r for $r > 0$.*

Proof: We will show that the infimum of this functional is zero and so, as $0 \notin \partial A_r$, the minimum will not be attained.

Let $\{z_i\}$ be an orthonormal basis for H. Set

$$w_m = \sqrt{2(r+\varepsilon)}z_m$$

where $\varepsilon > 0$ is fixed. Thus, $\|w_m\|^2 = 2(r+\varepsilon)$. Hence

$$F(w_m) \;=\; r + \varepsilon + \frac{1}{4}(A(w_m), w_m) \;>\; r.$$

Now consider the polynomial

$$p_m(t) \;=\; F(tw_m) \;=\; t^2(r+\varepsilon) + \frac{t^4}{4}(A(w_m), w_m).$$

Then p_m is increasing for $t > 0$, $p_m(0) = 0$ and $p_m(1) > r$. Hence, there exists a unique $t_m \in (0,1)$ such that $p_m(t_m) = r$, *i.e.* $t_m w_m \in \partial A_r$. Since $z_m \rightharpoonup 0$ weakly in H, it follows that $t_m w_m \rightharpoonup 0$ weakly in H as well. Since L is compact, $L(t_m w_m) \to 0$ strongly in H and so

$$(L(t_m w_m), t_m w_m) \to 0$$

as $m \to \infty$. Thus the infimum of the given functional (which is non-negative, by hypothesis) is zero and the proof is complete. ■

Proposition 4.7.3 *The functional $v \mapsto (Lv, v)$ attains its maximum on ∂A_r for every $r > 0$. If u_r is a maximizer, so is $-u_r$. Further, $u_r \to 0$ as $r \to 0$.*

Proof: Clearly the supremum of the functional is strictly positive. Thus any maximizer will be non-zero. If u_r is a maximizer, so will be $-u_r$. Since $u_r \in \partial A_r$, we have

$$\|u_r\|^2 \;\leq\; 2r$$

and so $u_r \to 0$ as $r \to 0$. Thus we only need to show the existence of a maximizer.

Let $\{w_n\}$ be a maximizing sequence in ∂A_r. Again, each element of this sequence is bounded in norm by $2r$ and so, we can work with a weakly convergent subsequence. We will assume, therefore, that $w_m \rightharpoonup w$ weakly in H. Then, since L is compact, we immediately see that

$$(Lw, w) = \sup_{v \in \partial A_r} (Lv, v).$$

By the weak lower semicontinuity of the norm (cf. Definition 5.1.2 and Example 5.1.1) and the compactness of A, we have

$$F(w) \leq r.$$

Assume, if possible, that $F(w) < r$. Then (by considering the polynomial $p(t) = F(tw)$) we can easily deduce the existence of a real number $t > 1$ such that $tw \in \partial A_r$. But then

$$(L(tw), tw) = t^2(Lw, w) > (Lw, w) = \sup_{v \in \partial A_r} (Lv, v)$$

which is a contradiction. thus $w \in \partial A_r$ and is a maximizer. ■

Thus, if $u_r \in \partial A_r$ is a maximizer of the above functional, then, we have a Lagrange multiplier $-\nu_r$ such that $Lu_r - \nu_r(u_r + A(u_r)) = 0$, thanks to the relation (4.7.4) of the hypothesis (H2) (cf. Theorem 1.4.2). It is immediate to see that $\nu_r \neq 0$. Thus, setting $\mu_r = 1/\nu_r$, we see that (u_r, μ_r) and $(-u_r, \mu_r)$ are solutions to (4.7.1). Since we know that $u_r \to 0$ as $r \to 0$, we only need to show that $\mu_r \to \lambda_1$ as $r \to 0$ in order to prove that $(0, \lambda_1)$ is a bifurcation point. This we now proceed to do.

Proposition 4.7.4 *With the preceding notations, we have*

$$\lim_{r \to 0} \mu_r = \lambda_1.$$

Proof: Step 1. Since $u_r \neq 0$, we know by Proposition 4.7.1 that $\mu_r \geq \lambda_1$. Thus $0 \leq \nu_r \leq 1/\lambda_1$ and so, for a subsequence, $\nu_r \to \nu$

as $r \to 0$.

Step 2. Let $t_r > 0$ be defined by

$$t_r^2 = \frac{2r}{\|u_r\|^2}.$$

Since $u_r \in \partial A_r$, it follows that

$$\frac{1}{4}(A(u_r), u_r) = r\left(1 - \frac{1}{t_r^2}\right).$$

Further, since $\|u_r\|^2 \leq 2r$, it follows that

$$0 \leq \frac{(A(u_r), u_r)}{4r} \leq Cr$$

where $C > 0$ is a constant independent of r. Thus $t_r \to 1$ as $r \to 0$.

Step 3. Now, since (u_r, μ_r) satisfies (4.7.1), we have

$$\nu_r t_r u_r - L(t_r u_r) + \nu_r t_r A(u_r) = 0.$$

Setting $w_r = t_r u_r$, we get $\|w_r\|^2 = 2r$ and

$$\nu_r \|w_r\|^2 - (Lw_r, w_r) + \nu_r t_r (A(u_r), w_r) = 0.$$

Hence,

$$\nu_r - \frac{(Lw_r, w_r)}{2r} + \frac{\nu_r t_r}{2r}(A(u_r), w_r) = 0. \tag{4.7.12}$$

But

$$\frac{1}{2r}|(A(u_r), w_r)| \leq \frac{C}{r}\|u_r\|^3 \|w_r\| \leq Cr$$

which tends to zero as $r \to 0$. Hence, passing to the limit in (4.7.12) using the result of Step 2, we get

$$\lim_{r \to 0} \frac{(Lw_r, w_r)}{2r} = \nu.$$

Step 4. Let u be a normalized eigenfunction of L corresponding to λ_1, *i.e.*

$$u = \lambda_1 Lu, \; \|u\|^2 = 1.$$

Let $\widetilde{t}_r > 0$ be such that $\widetilde{t}_r u \in \partial A_r$ for $r < 1/2$. Indeed, if $p(t) = F(tu)$, then $p(0) = 0$ and $p(1) \geq 1/2 > r$ and so there does exist such a $\widetilde{t}_r \in (0,1)$. Further,

$$\frac{1}{2}\widetilde{t}_r^2 + \frac{\widetilde{t}_r^4}{4}(A(u),u) = r$$

which implies that $\widetilde{t}_r \to 0$ as $r \to 0$. Hence we deduce that

$$\frac{2r}{\widetilde{t}_r^2} = 1 + \frac{\widetilde{t}_r^2(A(u),u)}{2} \to 1$$

as $r \to 0$.

Step 5. Let $\widetilde{u}_r = \widetilde{t}_r u$. Since u_r maximizes $v \mapsto (Lv,v)$ over ∂A_r, we have

$$\frac{1}{\lambda_1} = (Lu,u) = \frac{(L\widetilde{u}_r, \widetilde{u}_r)}{\widetilde{t}_r^2} \leq \frac{(Lu_r, u_r)}{\widetilde{t}_r^2}$$

whence we get

$$\frac{1}{\lambda_1} \leq \frac{(Lw_r, w_r)}{t_r^2 \widetilde{t}_r^2} = \frac{(Lw_r, w_r)}{2r} \frac{1}{t_r^2} \frac{2r}{\widetilde{t}_r^2}$$

and, by the preceding steps, the right-hand side tends to ν as $r \to 0$. Thus $1/\lambda_1 \leq \nu$, while Step 1 gives the reverse inequality. Thus $\nu = 1/\lambda_1$, *i.e.* $\mu_r \to \lambda_1$ and the proof is complete. ■

Remark 4.7.1 Using the Lyusternik- Schnirelman category (cf. Definition 2.5.3) suitably, it can be proved that every characteristic value λ of L is such that $(0, \lambda)$ is a bifurcation point for the equation (4.7.1) (cf. Berger [2] and Berger and Fife [3]). ■

Chapter 5

Critical Points of Functionals

5.1 Minimization of Functionals

In the last section of the preceding chapter, we have already seen examples of how solutions to certain nonlinear equations could be obtained as critical points of appropriate functionals.

Let H be a Hilbert space and let $F : H \to \mathbb{R}$ be a differentiable functional. Then, for $v \in H$, we have $F'(v) \in \mathcal{L}(H, \mathbb{R}) = H'$, the dual space, and by the Riesz representation theorem, H' can be identified with H. Thus, F' can be thought of as a mapping of H into itself and $F'(v)h = (F'(v), h)$ for all $v, h \in H$, where $(.,.)$ denotes the inner product of H.

Thus, if $f : H \to H$ is a given mapping and if there exists a functional $F : H \to \mathbb{R}$ such that $F' = f$, then looking for the zeros of f is the same as looking for the critical points of F.

One of the principal critical points of a functional is that point where the functional attains a minimum (or maximum) in the absence of constraints. We therefore examine conditions when a functional attains a minimum (analogous results can be formulated for a maximum).

Definition 5.1.1 *Let X be a topological space and let $f : X \to \mathbb{R}$ be a given function. Then f is said to be* **lower semi-continuous** *(l.s.c.) if, for every $c \in \mathbb{R}$, the set $f^{-1}((-\infty, c])$ is closed. It is said to be* **upper semi-continuous** *(u.s.c.) if $-f$ is l.s.c.* ■

Exercise 5.1.1 Let X be a metric space and let $f : X \to \mathbb{R}$ be a given function. Show that f is l.s.c. if, and only if, for every convergent sequence $x_n \to x$, we have

$$f(x) \leq \liminf_{n\to\infty} f(x_n). \blacksquare$$

Definition 5.1.2 *Let E be a Banach space and let $\Omega \subset E$. We say that $J : \Omega \to \mathbb{R}$ is* **weakly l.s.c.** *if $J^{-1}((-\infty, c])$ is weakly closed for all $c \in \mathbb{R}$. We say that J is* **weakly sequentially l.s.c.** *if whenever a sequence $\{x_n\}$ in Ω coverges weakly to $x \in \Omega$, we have*

$$J(x) \leq \liminf_{n\to\infty} J(x_n). \blacksquare \tag{5.1.1}$$

Remark 5.1.1 Obviously, a weakly l.s.c. functional in E is weakly sequentially l.s.c. $\blacksquare$

Definition 5.1.3 *Let E be a normed linear space and let $K \subset E$ be a convex subset. A functional J defined over K is said to be* **convex** *if for every $u, v \in K$ and every $\lambda \in [0, 1]$,*

$$J(\lambda u + (1 - \lambda)v) \leq \lambda J(u) + (1 - \lambda)J(v). \tag{5.1.2}$$

We say that J is strictly convex if strict inequality holds in (5.1.2) whenever $u \neq v$ and $\lambda \in (0, 1)$. We say that J is concave (resp. strictly concave) if $-J$ is convex (resp. strictly convex). $\blacksquare$

Example 5.1.1 Let E be a Banach space and let $K \subset E$ be a closed convex set. Then any convex functional $J : K \to \mathbb{R}$ which is l.s.c. is also weakly l.s.c.; in particular, the norm is a weakly l.s.c. functional. To see this, notice that, by the convexity of J, the set $J^{-1}((-\infty, c])$ is convex and it is also closed since J is l.s.c.. But, by the Hahn-Banach theorem, a closed and convex set is also weakly closed and the result follows. $\blacksquare$

Example 5.1.2 Consider the functional J defined in Section 4.7 by (4.7.7). Since L and A are compact, it follows that if $v_n \rightharpoonup v$ weakly in H,

$$\liminf_{n\to\infty} J(v_n) = \liminf_{n\to\infty} \frac{1}{2}\|v_n\|^2 - \frac{\lambda}{2}(Lv, v) + \frac{1}{4}(A(v), v) \geq J(v).$$

Thus, J is weakly sequentially l.s.c.. $\blacksquare$

Definition 5.1.4 *Let $J : \Omega \subset E \to \mathbb{R}$ be a functional defined on a subset Ω of a Banach space E. We say that J is* **coercive** *if $J(v_n) \to +\infty$ whenever we have $x_n \in \Omega$ such that $\|x_n\| \to +\infty$.* $\blacksquare$

Proposition 5.1.1 *Let E be a reflexive Banach space and let $K \subset E$ be a closed convex subset. Let $J : K \to \mathbb{R}$ be a coercive and weakly sequentially l.s.c. functional. Then J attains its minimum over K, i.e. there exists $u \in K$ such that*

$$J(u) = \min_{v\in K} J(v).$$

Proof: Let $\{u_n\}$ be a minimizing sequence in K. Since J is coercive, it follows that this sequence is bounded. Since E is reflexive, there exists a weakly convergent subsequence. Thus, let us assume that $u_n \rightharpoonup u$ weakly in E. Since, K is closed and convex, it is weakly closed and so $u \in K$. Now,

$$\inf_{v\in K} J(v) \leq J(u) \leq \liminf_{n\to\infty} J(u_n) = \inf_{v\in K} J(v)$$

which completes the proof. $\blacksquare$

In view of Example 5.1.1, we have the following result.

Corollary 5.1.1 *Let E and K be as in the preceding proposition. Let $J : K \to \mathbb{R}$ be a convex and l.s.c. functional which is also coercive. Then J attains a minimum over K and, if, in addition, J is strictly convex, the minimum is attained at a unique point.*

Proof: The weak sequential lower semi-continuity of J follows from Example 5.1.1 and Remark 5.1.1. Hence, the existence of a minimum follows from the preceding proposition. If α is the minimum value of J and if it is attained at two distinct points u_1 and u_2 in K, then $(u_1 + u_2)/2 \in K$ and so

$$\alpha \leq J((u_1 + u_2)/2) < \frac{1}{2}(J(u_1) + J(u_2)) = \alpha$$

which is impossible. ∎

Let E be a normed linear space and let Ω be an open subset on which a functional J is defined. If J attains a minimum at $u \in \Omega$ and if J is differentiable at u, then the Euler equation, which is a necessary condition, reads as $J'(u) = 0$. But if J attains a minimum with respect to a convex subset K at a point $u \in K$, this condition is no longer valid. Instead, we have the following necessary condition (sometimes called a *variational inequality*).

Proposition 5.1.2 *Let E be a normed linear space and let $\Omega \subset E$ be open. Let $J : \Omega \to \mathbb{R}$ be a given functional. Let $K \subset \Omega$ be convex and let J attain a minimum over K at $u \in K$. Assume that J is differentiable at u. Then, for every $v \in K$,*

$$J'(u)(v-u) \;\geq\; 0. \tag{5.1.3}$$

Proof: Let $v \in K$ be an arbitrary point and set $w = v - u$. Then, for $\theta \in [0,1]$, we have $u + \theta w \in K$ since K is convex. Since J is differentiable at u, we have

$$0 \;\leq\; J(u+\theta w) - J(u) \;=\; \theta(J'(u)w + \varepsilon(\theta))$$

where $\varepsilon(\theta) \to 0$ as $\theta \to 0$. Thus, we deduce that $J'(u)w \geq 0$ which is exactly (5.1.3). ∎

Exercise 5.1.2 Under the conditions of the preceding proposition, show that, if, in addition, K is a cone with vertex at the origin, then (5.1.3) is equivalent to the conditions:

$$\begin{aligned} J'(u)u &= 0 \\ J'(u)v &\geq 0. \end{aligned}$$

Deduce that if K is a subspace of E, then (5.1.3) becomes

$$J'(u)v \;=\; 0 \quad \text{for all} \quad v \in K. ∎$$

This necessary condition is also sufficient for convex functionals. We need a preliminary result to prove this.

Proposition 5.1.3 *Let E be a normed linear space and let $\Omega \subset E$ be open. Let $J : \Omega \to \mathbb{R}$ be differentiable in Ω. Let $K \subset \Omega$ be a convex subset.*
(i) J is convex over K if, and only if,

$$J(v) \;\geq\; J(u) + J'(u)(v-u) \tag{5.1.4}$$

for every $u, v \in K$.
(ii) J is strictly convex over K if, and only if, strict inequality holds in (5.1.4) whenever $u, v \in K$ and $u \neq v$.

Proof: Step 1. Let $u, v \in K$ and let $\theta \in (0,1)$. Then, by the definition of convexity,

$$\frac{J(u+\theta(v-u)) - J(u)}{\theta} \;\leq\; J(v) - J(u)$$

and (5.1.4) follows on letting $\theta \to 0$.

Step 2. Unfortunately, we cannot prove the strict version of (5.1.4) in case of strict convexity by the same argument, since on passing to the limit as $\theta \to 0$, we only get the same inequality. Let $\lambda \in (0,1)$. Let $v \neq u$. Then

$$u + \theta(v-u) \;=\; \frac{(\lambda-\theta)}{\lambda}u + \frac{\theta}{\lambda}(u + \lambda(v-u)).$$

Hence, for $0 \leq \theta \leq \lambda$, we have

$$J(u+\theta(v-u)) \;\leq\; \frac{(\lambda-\theta)}{\lambda}J(u) + \frac{\theta}{\lambda}J(u+\lambda(v-u)).$$

Thus, by strict convexity, for $0 < \theta \leq \lambda < 1$, we have

$$\frac{J(u+\theta(v-u)) - J(u)}{\theta} \;\leq\; \frac{J(u+\lambda(v-u)) - J(u)}{\lambda} \;<\; J(v) - J(u)$$

and now we get the strict inequality in (5.1.4) on letting $\theta \to 0$.

Step 3. If (5.1.4) were valid for all $u, v \in K$, then for $\theta \in (0,1)$, we have

$$\begin{aligned} J(v) &\geq J(v+\theta(u-v)) - \theta J'(v+\theta(u-v))(u-v), \\ J(u) &\geq J(v+\theta(u-v)) + (1-\theta)J'(v+\theta(u-v))(u-v). \end{aligned}$$

Multiplying the first inequality by $(1-\theta)$ and the second by θ and adding them, we get that J is convex. The strict convexity will follow from the corresponding strict inequalities. ■

When studying functions of a single real variable, we have seen that a function which is twice differentiable is convex if and only if its second derivative is non-negative, *i.e.* the first derivative is monotonic increasing. This result generalizes as follows.

Proposition 5.1.4 *Let E be a normed linear space and let $\Omega \subset E$ be open. Let $J : \Omega \to \mathbb{R}$ be differentiable in Ω. Let $K \subset \Omega$ be a convex subset. Then J is convex if, and only if, J' is monotone, i.e. for every $u, v \in K$,*

$$(J'(v) - J'(u))(v-u) \geq 0. \tag{5.1.5}$$

Proof: If J is convex, then (5.1.4) holds. Interchanging the roles of u and v in that inequality and adding the two inequalities, we get (5.1.5).

Conversely, assume that J' is monotone. Let $0 < \lambda < 1$. Let $u, v \in K$. Then,

$$\lambda J(u) + (1-\lambda)J(v) - J(\lambda u + (1-\lambda)v) =$$

$$= \lambda(J(u) - J(\lambda u + (1-\lambda)v)) + (1-\lambda)(J(v) - J(\lambda u + (1-\lambda)v)).$$

Now, by the mean value theorem for real valued functions (cf. Exercise 1.1.1) we have

$$J(u) + -J(\lambda u + (1-\lambda)v) = (1-\lambda)J'(z_1)(u-v)$$

where

$$z_1 = \lambda u + (1-\lambda)v + \theta_1(1-\lambda)(u-v)$$

for some $0 < \theta_1 < 1$. Similarly,

$$J(v) - J(\lambda u + (1-\lambda)v) = \lambda J'(z_2)(v-u)$$

where

$$z_2 = \lambda u + (1-\lambda)v + \theta_2\lambda(v-u)$$

for some $0 < \theta_2 < 1$. Thus,

$$\lambda J(u)+(1-\lambda)J(v)-J(\lambda u+(1-\lambda)v) = \lambda(1-\lambda)(J'(z_1)-J'(z_2))(u-v).$$

But

$$z_1 - z_2 = (\theta_1(1-\lambda) + \theta_2\lambda)(u-v)$$

and so, by the monotonicity of J', it follows that

$$\lambda J(u) + (1-\lambda)J(v) - J(\lambda u + (1-\lambda)v) \geq 0,$$

i.e. J is convex. ■

Proposition 5.1.5 *Let Ω be an open subset of a normed linear space E and let $J : \Omega \to \mathbb{R}$ be differentiable in Ω. Let $K \subset \Omega$ be convex and let J be convex over K. Then J admits a minimum over K at $u \in K$ if, and only if, (5.1.3) holds.*

Proof: The necessity has already been established. Conversely, if (5.1.3) holds, then by (5.1.4), we have for any $v \in K$,

$$J(v) - J(u) \geq J'(u)(v-u) \geq 0. \blacksquare$$

Proposition 5.1.6 *Under the conditions of the preceding proposition, assume, further, that J' is hemi-continuous, i.e. the map $\tau \mapsto J'(v + \tau(w-v))$ is continuous on $[0,1]$ for every $v, w \in K$. Then, J attains a minimum at $u \in K$ if, and only if,*

$$J'(v)(v-u) \geq 0 \tag{5.1.6}$$

for every $v \in K$.

Proof: If J attains a minimum at $u \in K$, then (5.1.3) holds and then so does (5.1.6) by the monotonicity of J' (cf. (5.1.5)). Conversely, if (5.1.6) holds, then for $0 < \tau < 1$, we have

$$J'(u + \tau(v-u))(u + \tau(v-u) - u) \ \geq \ 0.$$

Dividing by τ and then letting $\tau \to 0$, we deduce (5.1.3) using the hemi-continuity of J' from which it follows that J attains a minimum at u.■

Exercise 5.1.3 Let E be a normed linear space and let $\Omega \subset E$ be open. Let $J : \Omega \to \mathbb{R}$ be twice differentiable in Ω. Let $K \subset \Omega$ be convex. Show that J is convex if, and only if, for all $u, v \in K$,

$$J''(u)(v-u, v-u) \ \geq \ 0.$$

Show also that strict inequality in the above, whenever $u \neq v$, implies that J is strictly convex. Show, by means of an example, that the converse of the last assertion does not hold. ■

5.2 Saddle Points

In the previous section, we encountered results on functionals which attained their minimum (resp. maximum) on certain sets. But such functionals are therefore necessarily bounded below (resp. above) on those sets. In case of functionals that are bounded neither above nor below, we look for other critical points.

Definition 5.2.1 *Let X and Y be two sets and let $L : X \times Y \to \mathbb{R}$ be a given mapping. A point $(x^*, y^*) \in X \times Y$ is said to be a* **saddle point** *of L over $X \times Y$ if, for every $(x, y) \in X \times Y$, we have*

$$L(x^*, y) \ \leq \ L(x^*, y^*) \ \leq \ L(x, y^*).\blacksquare \qquad (5.2.1)$$

Thus, at (x^*, y^*), the function L attains a maximum in one direction and a minimum in the other.

Example 5.2.1 Let $X = Y = \mathbb{R}$ and let $L(x, y) = x^2 - y^2$. Then it is easy to see that the origin is a saddle point for L.■

Proposition 5.2.1 *Let X and Y be two sets and let $L : X \times Y \to \mathbb{R}$. Then L admits a saddle point over $X \times Y$ if, and only if,*

$$\max_{y \in Y} \inf_{x \in X} L(x, y) \ = \ \min_{x \in X} \sup_{y \in Y} L(x, y). \qquad (5.2.2)$$

Proof: It is obvious that

$$\sup_{y \in Y} \inf_{x \in X} L(x, y) \ \leq \ \inf_{x \in X} \sup_{y \in Y} L(x, y). \qquad (5.2.3)$$

Assume that there exists a saddle point $(x^*, y^*) \in X \times Y$. Then, by definition,

$$\sup_{y \in Y} L(x^*, y) \ = \ L(x^*, y^*) \ = \inf_{x \in X} L(x, y^*).$$

But then

$$\inf_{x \in X} \sup_{y \in Y} L(x, y) \ \leq \ L(x^*, y^*) \ \leq \ \sup_{y \in Y} \inf_{x \in X} L(x, y) \qquad (5.2.4)$$

which shows that we have equality in (5.2.3). Further, we also see from (5.2.4) that

$$\begin{array}{rcccl} L(x^*, y^*) & = & \sup_{y \in Y} L(x^*, y) & = & \min_{x \in X} \sup_{y \in Y} L(x, y) \\ L(x^*, y^*) & = & \inf_{x \in X} L(x, y^*) & = & \max_{y \in Y} \inf_{x \in X} L(x, y) \end{array}$$

which proves (5.2.2).

Conversely, assume that (5.2.2) holds. Let $(x^*, y^*) \in X \times Y$ be defined by

$$\begin{array}{rcccl} \sup_{y \in Y} L(x^*, y) & = & \min_{x \in X} \sup_{y \in Y} L(x, y) & = & m \\ \inf_{x \in X} L(x, y^*) & = & \max_{y \in Y} \inf_{x \in X} L(x, y) & = & m. \end{array}$$

Then, clearly, $m = \inf_{x \in X} L(x, y^*) \leq L(x^*, y^*) \leq \sup_{y \in Y} L(x^*, y) = m$. Thus, $m = L(x^*, y^*)$. Thus for all $(x, y) \in X \times Y$, we have (5.2.1), *i.e.* (x^*, y^*) is a saddle point. ■

Example 5.2.2 Let $X = Y = [0, 2\pi]$ and let $L(x,y) = \sin(x+y)$. Then

$$\sup_{y\in Y}\inf_{x\in X} L(x,y) = -1; \quad \inf_{x\in X}\sup_{y\in Y} L(x,y) = 1.$$

Thus, L cannot have a saddle point. ■

The following theorem gives sufficient conditions for the existence of a saddle point.

Theorem 5.2.1 *(Ky Fan - von Neumann) Let H_1 and H_2 be Hilbert spaces and let $K_i \subset H_i, i = 1,2$ be bounded and closed convex subsets. Let $L : K_1 \times K_2 \to \mathbb{R}$ be such that the following conditions hold:*
(i) For every $x \in K_1$, the map $y \mapsto L(x,y)$ is concave and u.s.c. on K_2.
(ii) For every $y \in K_2$, the map $x \mapsto L(x,y)$ is convex and l.s.c. on K_1.
Then, L admits at least one saddle point over $K_1 \times K_2$.

Proof: Step 1. Let us assume, to start with, that for each $y \in K_2$, the map $L(.,y)$ is strictly convex. Let $y \in K_2$. Set $F(y) = \inf_{x\in K_1} L(x,y)$. By Corollary 5.1.1 (the coerciveness condition is not necessary since K_1 is bounded), there exists a unique point $\varphi(y) \in K_1$ such that

$$L(\varphi(y), y) = F(y) = \min_{x\in K_1} L(x,y).$$

Step 2. It is trivial to check that F is a concave function. It is u.s.c. as well. For, let $c \in \mathbb{R}$. Consider the set $\mathcal{U} = \{y \in K_2 \mid F(y) < c\}$. If $y \in \mathcal{U}$, there exists $x \in K_1$ such that $L(x,y) < c$ and since the function $L(x,.)$ is u.s.c., there is a neighbourhood $\mathcal{V}$ in K_2 of y such that for all $z \in \mathcal{V}$, we have $L(x,z) < c$. Hence $F(z) < c$ for all $z \in \mathcal{V}$ which shows that $\mathcal{U}$ is open in K_2 and so F is u.s.c. (*i.e.* the lower envelope of a concave and u.s.c. function is concave and u.s.c.). Again, by a result analogous to Corollary 5.1.1, F will attain a maximum at a point $y^* \in K_2$. We will show that $(\varphi(y^*), y^*)$ is a saddle point for L.

Step 3. Set $x^* = \varphi(y^*)$. Let $y \in K_2$ and let $t \in [0,1]$. Set $x_t = \varphi((1-t)y^* + ty)$. For any $x \in K_1$, by the concavity of $L(x,.)$, we have

$$L(x,(1-t)y^* + ty) \geq (1-t)L(x,y^*) + tL(x,y).$$

Thus,

$$\begin{aligned} F(y^*) &\geq F((1-t)y^* + ty) = L(x_t,(1-t)y^* + ty) \\ &\geq (1-t)L(x_t,y^*) + tL(x_t,y) \\ &\geq (1-t)F(y^*) + tL(x_t,y) \end{aligned}$$

whence we deduce that, for all $y \in K_2$,

$$F(y^*) \geq L(x_t, y).$$

Let $t_n \downarrow 0$. Since K_1 is bounded, closed and convex, we have that (for a subsequence) $x_{t_n} \rightharpoonup \widetilde{x} \in K_1$ weakly in H_1. Since the map $L(.,y)$ is convex and l.s.c., it is also weakly l.s.c. and thus

$$L(x^*,y^*) = F(y^*) \geq \liminf_{t_n\downarrow 0} L(x_{t_n}, y) \geq L(\widetilde{x}, y). \qquad (5.2.5)$$

On the other hand, for any $x \in K_1$,

$$\begin{aligned} (1-t_n)L(x_{t_n},y^*) + t_nL(x_{t_n},y) &\leq L(x_{t_n},(1-t_n)y^* + t_ny) \\ &\leq L(x,(1-t_n)y^* + t_ny). \end{aligned}$$

Taking the limsup on the right (and using the upper semi-continuity of $L(x,.)$) and liminf on the left (and using (5.2.5)) as $t_n \downarrow 0$, we get, for all $x \in K_1$,

$$L(\widetilde{x}, y^*) \leq L(x,y^*)$$

and so, by the strict convexity of the map $L(.,y^*)$, it follows that $\widetilde{x} = x^*$. Thus, when $t \to 0$, $x_t \rightharpoonup x^*$ weakly in H_1 and again, by (5.2.5),

$$L(x^*,y) \leq L(x^*,y^*) = \inf_{x\in K_1} L(x,y^*) \leq L(x,y^*)$$

for any $(x,y) \in K_1 \times K_2$ and so (x^*,y^*) is a saddle point for L.

Step 4. If the map $L(.,y)$ were not strictly convex, let us consider, for $\varepsilon > 0$,

$$L_\varepsilon(x,y) = L(x,y) + \varepsilon\|x\|^2.$$

Since the norm in a Hilbert space is strictly convex, so is the map $L_\varepsilon(.,y)$ and, by the preceding steps, we have a saddlepoint $(x_\varepsilon^*, y_\varepsilon^*)$ for L_ε. Hence, for every $(x,y) \in K_1 \times K_2$, we have

$$L_\varepsilon(x_\varepsilon^*, y) \le L_\varepsilon(x_\varepsilon^*, y_\varepsilon^*) \le L_\varepsilon(x, y_\varepsilon^*).$$

Since $L(x_\varepsilon^*, y) \le L_\varepsilon(x_\varepsilon^*, y)$, we get

$$L(x_\varepsilon^*, y) \le L(x, y_\varepsilon^*) + \varepsilon\|x\|^2. \tag{5.2.6}$$

Now (for a subsequence), $x_\varepsilon^* \rightharpoonup x^* \in K_1$ weakly in H_1 and $y_\varepsilon^* \rightharpoonup y^* \in K_2$ weakly in H_2. Taking the liminf on the left side and the limsup on the right side of (5.2.6), we deduce that, for all $(x,y) \in K_1 \times K_2$,

$$L(x^*, y) \le L(x, y^*)$$

which is exactly (5.2.1), *i.e.* (x^*,y^*) is a saddle point for L.■

Remark 5.2.1 The above theorem holds in reflexive Banach spaces as well. We need to use the result that, in a reflexive Banach space, we can replace the norm by an equivalent one which is also strictly convex. This will then carry out Step 4 of the proof above. The result is also valid when the sets K_1 and K_2 are unbounded under the following additional (coercivity type) conditions:

- If K_1 is unbounded, then there exists $y_0 \in K_2$ such that

$$\lim_{\|x\|\to\infty} L(x, y_0) = +\infty.$$

- If K_2 is unbounded, then there exists $x_0 \in K_1$ such that

$$\lim_{\|y\|\to\infty} L(x_0, y) = -\infty.$$

The idea of the proof is to first consider the sets $(K_1 \cap B_1(0;n)) \times (K_2 \cap B_2(0;n))$ (where $B_i(0;n)$ is the ball centred at the origin and of radius n in H_i for $i = 1,2$) and consider the corresponding saddle points (x_n^*, y_n^*). The above conditions are to be used to show that $\{x_n^*\}$ and $\{y_n^*\}$ are bounded sequences and then pass to the limit. ■

Exercise 5.2.1 Let $H_i, i = 1,2$ be Hilbert spaces and let $\Omega \subset H_1 \times H_2$ be an open set. Let $L : \Omega \to \mathbb{R}$ be differentiable in Ω. Let $K_i \subset H_i, i = 1,2$ be closed convex subsets such that $K_1 \times K_2 \subset \Omega$. Show that if L satisfies the conditions of Theorem 5.2.1, then (x^*,y^*) is a saddle point for L if, and only if, for every $(x,y) \in K_1 \times K_2$,

$$(\partial_1 L(x^*,y^*), x - x^*) \ge 0; \ (\partial_2 L(x^*,y^*), y - y^*) \le 0.■$$

If K_1 and K_2 are closed subspaces of H_1 and H_2 respectively, it follows from the above exercise that $\partial_1 L(x^*,y^*) = 0$ and $\partial_2 L(x^*,y^*) = 0$, *i.e.* (x^*,y^*) is a critical point of L.

5.3 The Palais - Smale Condition

In proving that a functional attains its minimum, we had to show that a minimizing sequence was compact (*i.e.* it admitted a convergent subsequence) in an appropriate topology. More generally, when looking for critical points of a functional, we will construct sequences which we expect to converge to a critical point. However, we must notice that, even when a functional, J, is bounded below and we have a minimizing sequence $\{u_n\}$, it is not necessary that $J'(u_n) \to 0$. We now give below a fairly strong compactness condition which is relevant, from this point of view, in the search for critical points.

Definition 5.3.1 *Let E be a Banach space and let $J : E \to \mathbb{R}$ be a C^1 functional. Let $c \in \mathbb{R}$. Then J is said to satisfy the* **Palais - Smale condition,** *(PS), at level c if, given a sequence $\{u_n\}$ in E such that $J(u_n) \to c$ and $J'(u_n) \to 0$ in E' (the dual space), there always exists a convergent subsequence $\{u_{n_k}\}$.*■

Remark 5.3.1 If the Palais - Smale condition is satisfied at all levels $c \in \mathbb{R}$, we simply say that J satisfies PS. ■

Example 5.3.1 The function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = e^x$ does not satisfy PS. ■

Example 5.3.2 Let $n \leq 3$ and let $\Omega \subset \mathbb{R}^n$ be a bounded domain. Let $E = H_0^1(\Omega)$, the usual Sobolev space of square integrable functions whose first order (distribution) derivatives are also square integrable functions and which vanish, in the trace sense, on the boundary $\partial\Omega$ (cf. Kesavan [13]). We denote the norm in this space by $|v|_{1,\Omega}$ $(= (\int_\Omega |\nabla v|^2 dx)^{1/2})$ and the norm in $L^2(\Omega)$ by $|v|_{0,\Omega}$ $(= (\int_\Omega |v|^2 dx)^{1/2})$. Let $2 \leq r \leq 4$ be an integer. Let $f \in L^2(\Omega)$ be a given function. Define the functional $J : E \to \mathbb{R}$ by

$$J(v) = \frac{1}{2}\int_\Omega |\nabla v|^2 dx - \frac{1}{r+1}\int_\Omega v^{r+1} dx - \int_\Omega f v dx. \qquad (5.3.1)$$

We claim that J satisfies PS. Indeed let $\{v_m\}$ be a sequence in E such that $J(v_m) \to c$ and $J'(v_m) \to 0$ in $E' = H^{-1}(\Omega)$. If $< .,. >$ denotes the duality bracket between E' and E, we have

$$< J'(u), w > = \int_\Omega \nabla u.\nabla w dx - \int_\Omega u^r w dx - \int_\Omega f w dx. \qquad (5.3.2)$$

Using (5.3.1) and (5.3.2), we get

$$\left.\begin{array}{rcl} < J'(v_m), v_m > & = & \int_\Omega |\nabla v_m|^2 dx - \int_\Omega v_m^{r+1} dx - \int_\Omega f v_m dx \\ & = & (r+1)J(v_m) - \frac{(r-1)}{2}\int_\Omega |\nabla v_m|^2 dx \\ & & + r\int_\Omega f v_m dx \end{array}\right\} \qquad (5.3.3)$$

which gives

$$\frac{(r-1)}{2}|v_m|_{1,\Omega}^2 \leq (r+1)J(v_m) + r|f|_{0,\Omega}|v_m|_{0,\Omega} + \|J'(v_m)\|_{-1,\Omega}|v_m|_{1,\Omega} \qquad (5.3.4)$$

where $\|.\|_{-1,\Omega}$ denotes the norm in $E' = H^{-1}(\Omega)$. We deduce from (5.3.4) that the sequence $\{v_m\}$ is bounded in E (if not, divide throughout by $|v_m|_{1,\Omega}^2$ and pass to the limit as $m \to \infty$ to get a contradiction). Thus, for a subsequence, $v_m \rightharpoonup v$ weakly in $E = H_0^1(\Omega)$ and, by Rellich's theorem, strongly in $L^{r+1}(\Omega)$ and in $L^2(\Omega)$, since $n \leq 3$. Substituting $u = v_m$ in (5.3.2), and passing to the limit, we get

$$0 = \int_\Omega \nabla v.\nabla w dx - \int_\Omega v^r w dx - \int_\Omega f w dx$$

so that

$$\int_\Omega |\nabla v|^2 dx = \int_\Omega v^{r+1} dx + \int_\Omega f v dx.$$

On the other hand, we also have, from the first relation in (5.3.3), that

$$\lim_{m\to\infty}\int_\Omega |\nabla v_m|^2 dx = \int_\Omega v^{r+1} dx + \int_\Omega f v dx = \int_\Omega |\nabla v|^2 dx.$$

Thus the concerned (sub)sequence is, in fact, strongly convergent in $H_0^1(\Omega)$ and so J satisfies PS. ■

Exercise 5.3.1 With the same notations as in the preceding example, show that the following functionals satisfy PS:
(i)

$$J(v) = \frac{1}{2}\int_\Omega |\nabla v|^2 dx + \frac{1}{r+1}\int_\Omega v^{r+1} dx - \int_\Omega f v dx.$$

(ii)

$$J(v) = \frac{1}{2}\int_\Omega |\nabla v|^2 dx + \frac{1}{4}\int_\Omega v^4 dx - \frac{\lambda}{2}\int_\Omega v^2 dx$$

where $\lambda \in \mathbb{R}$.■

We remarked earlier that for a functional bounded below, minimizing sequences need not be such that their gradients tend to zero. However, if the functional satisfies the Palais - Smale condition, it can be shown that it attains a minimum. Before we do this, we need a preliminary result, which is a very general result proved by Ekeland [9].

Lemma 5.3.1 *(Ekeland Variational Principle) Let (X,d) be a complete metric space and let $J : X \to \mathbb{R}$ be a l.s.c. function. Assume, further, that J is bounded below and set $c = \inf_{x\in X} J(x)$. Then, for every $\varepsilon > 0$, there exists $u_\varepsilon \in X$ such that*

$$\left.\begin{array}{rcl} c \ \leq \ J(u_\varepsilon) & \leq & c+\varepsilon \\ J(x) - J(u_\varepsilon) + \varepsilon d(x,u_\varepsilon) & > & 0 \end{array}\right\} \quad (5.3.5)$$

for every $x \in X$, $x \neq u_\varepsilon$.

Proof: Step 1. Fix $\varepsilon > 0$. Consider the epigraph of J, *i.e.* the set

$$A = \{(x,a) \in X \times \mathbb{R} \mid J(x) \leq a\}.$$

Since J is l.s.c., A is closed. Notice that for all $x \in X$, we have $(x, J(x)) \in A$. We define an order relation in $X \times \mathbb{R}$ by

$$(x,a) \preceq (y,b) \Leftrightarrow a - b + \varepsilon d(x,y) \leq 0.$$

Notice that if the above relation holds then, necessarily $b \geq a$. Notice also that, if $a \leq b$, then, $(x,a) \preceq (x,b)$. We will now proceed to construct, inductively, a decreasing family of sets $A_n \subset A$.

Step 2. We can always choose $x_1 \in X$ such that

$$c \leq J(x_1) \leq c+\varepsilon.$$

Set $a_1 = J(x_1)$ and define

$$A_1 = \{(x,a) \in A \mid (x,a) \preceq (x_1,a_1)\}.$$

Assume, now, that, for $1 \leq i \leq n$, we have determined (x_i, a_i) such that $a_i = J(x_i)$ and set

$$A_i = \{(x,a) \in A \mid (x,a) \preceq (x_i,a_i)\}.$$

Define, further,

$$\widetilde{A_n} = \{x \in X \mid \text{there exists } a \in \mathbb{R} \text{ such that } (x,a) \in A_n\}.$$

Let

$$c_n = \inf_{x\in\widetilde{A_n}} J(x).$$

If $(x,a) \in A_i$, by definition $J(x) \leq a$ and further $(x,a) \preceq (x_i,a_i)$ and so $a \leq a_i$. Thus, for all $1 \leq i \leq n$, it follows that $c_i \leq a_i$.

Step 3. Assume, for the moment, that for all $1 \leq i \leq n$, we have $c_i < a_i$. Then, we can choose $x_{n+1} \in \widetilde{A_n}$ such that

$$0 \leq J(x_{n+1}) - c_n \leq \frac{1}{2}(a_n - c_n).$$

Then, there exists $a \in \mathbb{R}$ such that $J(x_{n+1}) \leq a$ and $(x_{n+1}, a) \preceq (x_n, a_n)$ whence it follows that $(x_{n+1}, J(x_{n+1})) \preceq (x_{n+1}, a) \preceq (x_n, a_n)$. Thus we can set $a_{n+1} = J(x_{n+1})$ to get $(x_{n+1}, a_{n+1}) \in A_n$ and then define

$$A_{n+1} = \{(x,a) \in A \mid (x,a) \preceq (x_{n+1},a_{n+1})\}.$$

Step 4. By the transitivity of the order relation, it is evident that $A_{n+1} \subset A_n$. Also, as $\widetilde{A_{n+1}} \subset \widetilde{A_n}$, we also have $c \leq c_n \leq c_{n+1} \leq a_{n+1}$. Thus,

$$0 \leq a_{n+1} - c_{n+1} \leq a_{n+1} - c_n \leq \frac{1}{2}(a_n - c_n)$$

so that, recursively,

$$0 \leq a_{n+1} - c_{n+1} \leq 2^{-n}(a_1 - c_1).$$

Further, $(x,a) \in A_{n+1}$ implies that

$$a - a_{n+1} + \varepsilon d(x, x_{n+1}) \ \leq \ 0.$$

Also, since $c_{n+1} \leq a \leq a_{n+1}$, we deduce that

$$d(x, x_{n+1}) + |a - a_{n+1}| \ \leq \ \left(1 + \frac{1}{\varepsilon}\right) 2^{-n}(a_1 - c_1)$$

which implies that the diameter of A_{n+1} tends to zero. Since A is complete, it follows that there exists a unique point $(u,b) \in A$ such that

$$\{(u,b)\} \ = \ \cap_{n=1}^{\infty} A_n.$$

Step 5. If $(x,a) \in A$ such that $(x,a) \preceq (u,b)$, then, for all $n \geq 1$, we have $(x,a) \preceq (x_n, a_n)$ and so $(x,a) \in A_n$ for all n. Thus $(x,a) = (u,b)$. Hence (u,b) *is a minimal element in* A *in this* sense. Further, since, $J(u) \leq b$, we have $(u, J(u)) \preceq (u,b)$ and so, by the minimality of (u,b), we have that $J(u) = b$. It now follows that if $(x,a) \in A$ and $(x,a) \neq (u,b)$, then

$$a - b + \varepsilon d(x,u) \ > \ 0.$$

In particular, for $x \neq u$, we have $(x, J(x)) \in A$ and so

$$J(x) - J(u) + \varepsilon d(x,u) \ > \ 0.$$

Finally, since $(u, J(u)) \in A_1$, we have

$$J(u) \ \leq \ J(x_1) \ \leq \ c + \varepsilon$$

and we can conclude the proof setting $u_\varepsilon = u$.

Step 6. In case we encounter $n \geq 1$ such that $c_n = a_n = J(x_n)$, let $(x,a) \preceq (x_n, a_n)$. Then, by definition, $a_n = c_n \leq J(x) \leq a$ while $a - a_n + \varepsilon d(x, x_n) \leq 0$. This implies that $(x,a) = (x_n, a_n)$ and so A_n is the singleton $\{(x_n, a_n)\}$ and we can set $(u,b) = (x_n, a_n)$ and verify its minimality and conclude the proof as before. ■

Proposition 5.3.1 *Let E be a Banach space and let $J : E \to \mathbb{R}$ be a C^1 functional. If J satisfies PS and is bounded below, then it attains a minimum over E.*

Proof: By the Ekeland variational principle, we can choose a sequence $\{u_n\}$ such that $c \leq J(u_n) \leq c + 1/n$ and also such that, for all $v \in E$,

$$J(v) + \frac{1}{n}\|v - u_n\| \ \geq \ J(u_n)$$

where $c = \inf_{v \in E} J(v)$. Thus, $J(u_n) \to c$. Further,

$$J(v) \ = \ J(u_n) + J'(u_n)(v - u_n) + o(\|v - u_n\|).$$

Let $w \in E$ such that $\|w\| = 1$ and let $t > 0$. Then, applying the above relations to $v = u_n + tw$, we get

$$-\frac{t}{n} \ \leq \ J(u_n + tw) - J(u_n) \ = \ tJ'(u_n)w + o(t).$$

Dividing by t and letting $t \to 0$, we get, for all w such that $\|w\| = 1$,

$$-\frac{1}{n} \ \leq \ J'(u_n)w$$

and so, taking $-w$ in place of w as well, it follows that,

$$\|J'(u_n)\| \ \leq \ \frac{1}{n},$$

i.e. $J'(u_n) \to 0$. Thus, by PS, we can extract a convergent subsequence $u_{n_k} \to u$ and, since J is of class C^1, we have $J(u) = c$ and that $J'(u) = 0$.■

Example 5.3.3 Let $n \leq 3$ and let $\Omega \subset \mathbb{R}^n$ be a bounded domain. Let $E = H_0^1(\Omega)$. Define $J : E \to \mathbb{R}$ by

$$J(v) \ = \ \frac{1}{2}\int_\Omega |\nabla v|^2 dx + \frac{1}{4}\int_\Omega v^4 dx - \int_\Omega f v dx$$

where $f \in L^2(\Omega)$ is given. We know that (cf. Exercise 5.3.1) J satisfies PS. Further,

$$\begin{aligned} J(v) &\geq \tfrac{1}{2}\int_\Omega |\nabla v|^2 dx - |f|_{0,\Omega}|v|_{0,\Omega} \\ &\geq \tfrac{1}{2}|v|_{1,\Omega}^2 - C|v|_{1,\Omega} \quad \text{(by Poincaré's inequality)} \\ &\geq -C^2/2. \end{aligned}$$

Thus, J is bounded below as well and so J attains a minimum at a point $u \in E$. Thus $u \in H_0^1(\Omega)$ satisfies, for all $v \in H_0^1(\Omega)$,

$$\int_\Omega \nabla u.\nabla v dx + \int_\Omega u^3 v dx = \int_\Omega f v dx$$

which is the weak formulation of the semilinear elliptic boundary value problem:

$$\left. \begin{aligned} -\Delta u + u^3 &= f \quad \text{in } \Omega \\ u &= 0 \quad \text{on } \partial\Omega \end{aligned} \right\}$$

where $\Delta = \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2}$ is the Laplacian. ■

Remark 5.3.2 Of course, in this example, we could have also proved the existence of a minimum by showing that J was coercive and weakly sequentially lower semi-continuous.■

5.4 The Deformation Lemma

Let X be a set and let $J : X \to \mathbb{R}$ be a given mapping. For $c \in \mathbb{R}$, we denote by $\{J \leq c\}$ the set

$$\{x \in X \mid J(x) \leq c\}.$$

In an analogous way we can define the sets $\{J > c\}, \{c_1 < J < c_2\}, \{J = c\}$ and so on.

One of the rich theories in the study of critical points of functionals is that of M. Morse. For an introduction to this fascinating topic, see, for instance, the book of Milnor [18]. One of the principal ideas behind this theory is that the critical values

of a functional on a suitable topological space X with a differentiable structure are precisely those values $c \in \mathbb{R}$ for which, when $\varepsilon > 0$ is sufficiently small, one cannot continuously deform the set $\{J \leq c + \varepsilon\}$ into $\{J \leq c - \varepsilon\}$. In fact these sets can be, topologically, very different.

For example, if $X = \mathbb{R}$ and $J(x) = x^3 - 3x$, the critical points are located at $x = \pm 1$. Let $c_1 = J(1) = -2$ and $c_2 = J(-1) = 2$. If $c < c_1$, then the set $\{J \leq c\}$ is of the type $[-\infty, \alpha]$ for some $\alpha \in \mathbb{R}$. If $c_1 < c < c_2$, then this set is of the form $[-\infty, \alpha] \cup [\beta, \gamma]$ while for $c > c_2$, the set is again of the form $[-\infty, \alpha]$. Thus, in the second case, the set has two connected components while the others have just one connected component.

It can also happen that, while passing through a critical value, though in all cases the sets have the same number of connected components, some could be simply connected while others are multiply connected (cf. Kavian [11]).

The deformation lemma (Lemma 5.4.2, below) gives a precise meaning to what we mean by saying that one set is continuously deformable into another. We need a few preliminaries before we can state and prove this result.

Definition 5.4.1 *Let E be a Banach space and let $J : E \to \mathbb{R}$ be a $\mathcal{C}^1$ functional. Let $u \in E$. A vector $v \in E$ is said to be a* **pseudo-gradient** *for J at u if*

$$\|v\| \leq 2\|J'(u)\| \quad \textit{and} \quad < J'(u), v > \geq \|J'(u)\|^2. ■$$

In the above definition, and throughout the sequel, $< .,. >$ will stand for the duality bracket between E and its dual E'.

Remark 5.4.1 The above conditions clearly imply that if v is a pseudo-gradient for J at u, then

$$\|J'(u)\| \leq \|v\| \leq 2\|J'(u)\|. ■ \tag{5.4.1}$$

Example 5.4.1 If E were a Hilbert space, then $J'(u) \in E$ will itself be a pseudo-gradient for J at u.■

Definition 5.4.2 *Let E be a Banach space and let $J : E \to \mathbb{R}$ be a C^1 functional. Let $E_r = \{u \in E \mid J'(u) \neq 0\}$ denote the set of regular points of E. A mapping $V : E_r \to E$ is called a* **pseudo-gradient vector field** *for J if V is a locally Lipschitz function on E_r and, for every $u \in E_r$, $V(u)$ is a pseudo-gradient for J at u.*■

The condition that a pseudo-gradient vector field be locally Lipschitz makes the search for such vector fields a non-trivial task. Even in the case of a Hilbert space, the map $u \mapsto J'(u)$ need not necessarily be locally Lipschitz. In this context, the following result is relevant.

Lemma 5.4.1 *Let E be a Banach space and let $J : E \to \mathbb{R}$ be a C^1 functional. Then J admits a pseudo-gradient vector field.*

Proof: Let $u \in E_r$. Since $J'(u) \neq 0$, there exists a point $x_u \in E$ such that $\|x_u\| = 1$ and $< J'(u), x_u > > (2/3)\|J'(u)\|$. Let $v_u = (3/2)\|J'(u)\|x_u$. Then, by construction, $\|v_u\| = (3/2)\|J'(u)\| \leq 2\|J'(u)\|$ and $< J'(u), v_u > > \|J'(u)\|^2$. Thus, v_u is a pseudo-gradient for J at u and, since J' is continuous, it is also a pseudo-gradient for J at all points $x \in \mathcal{V}_u$, a neighbourhood of u.

The sets $\mathcal{V}_u$ form a covering of E_r which is a metric space and hence a paracompact space. Thus (cf. Dieudonné [8]) there exists a locally finite refinement $\{\omega_j\}$ and an associated locally Lipschitz partition of unity $\{\theta_j\}$, *i.e.* $\omega_j \subset \mathcal{V}_{u_j}$ for some $u_j \in E_r$ and $\text{supp}(\theta_j) \subset \omega_j$, and, further $\sum_j \theta_j \equiv 1$ (the sum makes sense, irrespective of the cardinality of the indexing set, since the refinement is locally finite). Now we set

$$V(x) = \sum_j \theta_j(x) v_{u_j}$$

and V will be the required pseudo-gradient vector field. ■

Corollary 5.4.1 *If E and J are as in the preceding lemma and if, in addition, J is an even functional, then there exists a pseudo-gradient vector field that is odd.*

Proof: Since J is even, it follows that J' is odd and that E_r is symmetric with respect to the origin. If V_1 is a pseudo-gradient vector field, then so is $V_2(x) = -V_1(-x)$. Then $V(x) = (V_1(x) - V_1(-x))/2$ is also pseudo-gradient vector field, which, in addition, is odd. ■

Lemma 5.4.2 *(Deformation Lemma) Let E be a Banach space and let $J : E \to \mathbb{R}$ be a C^1 functional which satisfies PS. Let $c \in \mathbb{R}$ be a regular value of J. Then, there exists $\varepsilon_0 > 0$ such that, for all $0 < \varepsilon < \varepsilon_0$, there exists a continuous map $\eta : \mathbb{R} \times E \to E$ (called the flow associated to J) satisfying the following conditions:*
(i) For every $u \in E$, $\eta(0, u) = u$.
(ii) For every $t \in \mathbb{R}$, the mapping $\eta(t, .) : E \to E$ is a homeomorphism.
(iii) For every $t \in \mathbb{R}$ and every $u \notin \{c - \varepsilon_0 \leq J \leq c + \varepsilon_0\}$, $\eta(t, u) = u$.
(iv) For every $u \in E$, the function $t \mapsto J(\eta(t, u))$ is decreasing.
(v) If $u \in \{J \leq c + \varepsilon\}$, then $\eta(1, u) \in \{J \leq c - \varepsilon\}$.
(vi) If, in addition, J is even, then, for every $t \in \mathbb{R}$, the map $\eta(t, .)$ is odd.

Proof: Step 1. Since c is not a critical value and since J satisfies PS, it follows that we can find an $\varepsilon_1 > 0$ and a $\delta > 0$ (in fact, we can choose, without loss of generality, $\delta \leq 1$) such that, for all $u \in \{c - \varepsilon_1 \leq J \leq c + \varepsilon_1\}$, we have $\|J'(u)\| \geq \delta$. We now set $\varepsilon_0 = \min\{\varepsilon_1, \delta^2/8\}$ and, for $0 < \varepsilon < \varepsilon_0$, we define

$$A = \{J \leq c - \varepsilon_0\} \cup \{J \geq c + \varepsilon_0\}, \quad B = \{c - \varepsilon \leq J \leq c + \varepsilon\}.$$

Since A and B are disjoint closed sets, the function

$$\varphi(x) = \frac{\rho(x, A)}{\rho(x, A) + \rho(x, B)}$$

(where $\rho(x, S)$ denotes the distance of the point x from the set S) is locally Lipschitz and $\varphi \equiv 0$ on A while $\varphi \equiv 1$ on B. In addition, if J is even, these sets are symmetric with respect to the origin and φ is even as well.

Step 2. Let V be a pseudo-gradient vector field (which can be chosen to be odd, if J is even) for J. For $x \in E$, we define

$$W(x) = \varphi(x) \min\left\{1, \frac{1}{\|V(x)\|}\right\} V(x).$$

Then, W is well defined, locally Lipschitz and such that $\|W(x)\| \leq 1$ for all $x \in E$. If J is even, then W is odd. Now, the initial value problem

$$\left.\begin{array}{rcl} \frac{d\eta}{dt}(t,x) & = & -W(\eta(t,x)) \\ \eta(0,x) & = & x \end{array}\right\} \quad (5.4.2)$$

has a unique solution $\eta(.,x) \in C^1(\mathbb{R}, E)$ for each $x \in E$ and, further, η is locally Lipschitz on $\mathbb{R} \times E$. Since, for $t, s \in \mathbb{R}$,

$$\eta(t, \eta(s,x)) = \eta(t+s, x)$$

(by the uniqueness of the solution to (5.4.2)), it readily follows that, for each $t \in \mathbb{R}$, the map $\eta(t,.) : E \to E$ is a homeomorphism with inverse map given by $\eta(-t,.)$. This is the required flow associated to J and we will now verify the conditions (i) - (vi) in the statement of the lemma.

Step 3. Condition (i) is satisfied,by the definiton of η and we have just proved condition (ii) in the preceding step. Again, by the uniqueness of the solution to (5.4.2), it follows that if J is even, then, as W is odd, that $\eta(t,.)$ is also odd for all $t \in \mathbb{R}$ and thus condition (vi) is verified.

Step 4. If $u \notin \{c - \varepsilon_0 \leq J \leq c + \varepsilon_0\}$, then $\varphi(u) = 0$ and so the unique solution to (5.4.2) when $x = u$ is the constant function $\eta(t,u) = u$ for all $t \in \mathbb{R}$. This proves (iii).

Step 5. Let $u \in E$. Since $V(\eta(t,u))$ is a pseudo-gradient for J at

$\eta(t,u)$, we have

$$\left.\begin{array}{rcl} \frac{d}{dt}J(\eta(t,u)) & = & < J'(\eta(t,u)), \frac{d\eta}{dt}(t,u) > \\ & = & -\varphi(\eta(t,u))\kappa < J'(\eta(t,u)), V(\eta(t,u)) > \\ & \leq & -\varphi(\eta(t,u))\kappa\|J'(\eta(t,u))\|^2 \end{array}\right\} \quad (5.4.3)$$

where

$$\kappa = \min\left\{1, \frac{1}{\|V(\eta(t,u))\|}\right\}$$

which shows that J decreases along the flow starting from any $u \in E$, thus proving (iv).

Step 6. We now verify the crucial condition (v). Let $u \in \{J \leq c + \varepsilon\}$. If, for some $t_0 \in [0,1)$, we have $\eta(t_0, u) \in \{J \leq c - \varepsilon\}$, then, by Step 5, $\eta(t,u)$ will continue in this set for all future time and so $\eta(1,u) \in \{J \leq c - \varepsilon\}$. Assume now that, for all $t \in [0,1)$, we have that $\eta(t,u) \in \{c - \varepsilon < J \leq c + \varepsilon\} \subset B$. Then, by (5.4.1) and (5.4.3),

$$\begin{array}{rcl} \frac{d}{dt}J(\eta(t,u)) & \leq & -\frac{1}{4}\kappa\|V(\eta(t,u))\|^2 \\ & \leq & \left\{\begin{array}{ll} -\frac{1}{4} & \text{if } \|V(\eta(t,u))\| \geq 1 \\ -\frac{\delta^2}{4} & \text{if } \|V(\eta(t,u))\| < 1. \end{array}\right. \end{array}$$

Since $\delta \leq 1$, we get

$$J(\eta(1,u)) \leq -\frac{\delta^2}{4} + J(u) \leq -\frac{\delta^2}{4} + c + \varepsilon$$

and hence, by the definition of ε_0,

$$J(\eta(1,u)) \leq c - \frac{\delta^2}{8} \leq c - \varepsilon$$

which completes the proof. ■

The deformation lemma is the basis of all variational methods which seek critical points via a min-max principle.

Theorem 5.4.1 *(Min-Max Principle) Let E be a Banach space and let $J : E \to \mathbb{R}$ be a $\mathcal{C}^1$ functional satisfying PS. Let $\mathcal{A}$ be a non-empty collection of non-empty subsets of E with the following property: for every $c \in \mathbb{R}$, and sufficiently small $\varepsilon > 0$, the associated flow $\eta(t,u)$ constructed in the deformation lemma is such that whenever $A \in \mathcal{A}$, we have $\eta(1,A) \in \mathcal{A}$.*
Define

$$c^* \;=\; \inf_{A\in\mathcal{A}} \sup_{v\in A} J(v).$$

If $c^* \in \mathbb{R}$, then, c^* is a critical value of J.

Proof: If c^* is not a critical value, choose $A \in \mathcal{A}$ such that

$$\sup_{v\in A} J(v) \;<\; c^* + \varepsilon$$

for sufficiently small $\varepsilon > 0$. Thus, $A \subset \{J \leq c^* + \varepsilon\}$ and so $\eta(1,A) \subset \{J \leq c^* - \varepsilon\}$ and this contradicts the definition of c^* since $\eta(1,A) \in \mathcal{A}$.■

5.5 The Mountain Pass Theorem

The mountain pass theorem, due to Ambrosetti and Rabinowitz [1], is one of the important applications of the min-max principle enunciated in the previous section. It has turned out to be extremely useful in the study of solutions of semilinear elliptic boundary value problems.

Theorem 5.5.1 *(Mountain Pass Theorem) Let E be a Banach space and let $J : E \to \mathbb{R}$ be a $\mathcal{C}^1$ functional satisfying PS. Let $u_0, u_1 \in E$, $c_0 \in \mathbb{R}$ and $R > 0$ such that*
(i) $\|u_0 - u_1\| \;>\; R$; (ii) for all $v \in E$ such that $\|v - u_0\| = R$,

$$\max\{J(u_0), J(u_1)\} \;< c_0 \;\leq J(v).$$

Then J admits a critical value $c \geq c_0$ defined by

$$c \;=\; \inf_{\gamma\in\mathcal{P}} \max_{t\in[0,1]} J(\gamma(t)), \tag{5.5.1}$$

where $\mathcal{P}$ is the collection of all continuous paths $\gamma : [0,1] \to E$ such that $\gamma(0) = u_0$ and $\gamma(1) = u_1$.

Proof: Clearly c is finite and, since any path from u_0 to u_1 must cross the sphere $\{v \in E \mid \|v - u_0\| = R\}$ by virtue of condition (i) in the hypotheses, it follows that $c \geq c_0$. Assume that c is not a critical value. Then we can find $\varepsilon_0 > 0$ and a flow η as in the deformation lemma such that for all $0 < \varepsilon < \varepsilon_0$, we have $\eta(1, \{J \leq c + \varepsilon\}) \subset \{J \leq c - \varepsilon\}$. Choose ε_0 such that

$$\max\{J(u_0), J(u_1)\} \;<\; c - \varepsilon_0 \tag{5.5.2}$$

and for some $0 < \varepsilon < \varepsilon_0$, choose $\gamma \in \mathcal{P}$ such that

$$\max_{t\in[0,1]} J(\gamma(t)) \;<\; c + \varepsilon.$$

Set $\zeta(t) = \eta(1, \gamma(t))$. Then, it follows that

$$\max_{t\in[0,1]} J(\zeta(t)) \;\leq\; c - \varepsilon. \tag{5.5.3}$$

Now, $\zeta(0) = \eta(1, u_0) = u_0$ and $\zeta(1) = \eta(1, u_1) = u_1$ by virtue of condition (iii) of the deformation lemma, in view of (5.5.2). Thus $\zeta \in \mathcal{P}$ and (5.5.3) contradicts the definition of c. This completes the proof.■

The above theorem derives its name from the following geometric analogy. Assume that J represents the height of a place above sea level and let u_0 be a place surrounded by a ring of mountains and u_1 a place on the plains outside this ring. To travel from u_0 to u_1, we will naturally have to cross the mountain range and the ideal path for us to choose would be the one wherein we climb the least which gives us a pass in the mountain range.

Since the critical value c is strictly greater than $J(u_0)$ and $J(u_1)$, if it happens that either of them is already a critical point (for instance, J could attain a local minimum at u_0) we get a new critical point by this theorem. We illustrate this via an example.

Example 5.5.1 Let $n \leq 3$ and let $\Omega \subset \mathbb{R}^n$ be a bounded domain. Let $f \in L^2(\Omega)$ be given. Consider the problem

$$\left.\begin{array}{rcll} -\Delta u &=& u^2 + f & \text{in } \Omega \\ u &=& 0 & \text{on } \partial\Omega. \end{array}\right\} \tag{5.5.4}$$

The weak formulation of this problem is to look for $u \in H_0^1(\Omega)$ such that

$$\int_\Omega \nabla u.\nabla v dx = \int_\Omega u^2 v dx + \int_\Omega f v dx \tag{5.5.5}$$

for all $v \in H_0^1(\Omega)$. Then, a solution u is a critical point of the functional defined by

$$J(v) = \frac{1}{2}\int_\Omega |\nabla v|^2 dx - \frac{1}{3}\int_\Omega v^3 dx - \int_\Omega f v dx \tag{5.5.6}$$

for all $v \in H_0^1(\Omega)$. Thus, J is a C^1 functional on the Hilbert space $H_0^1(\Omega)$ and we have already seen (cf. Example 5.3.2) that it satisfies PS.

If f is a sufficiently smooth function such that $f < 0$, one can use the maximum principle available for second order elliptic operators and show that there exists a solution $u_0 < 0$. (This is called the method of monotone iterations or the method of sub- and super- solutions, or, again, Perron's method, cf. Kesavan [13]). Let $w \in H_0^1(\Omega)$ be such that $|w|_{1,\Omega} = 1$ and consider $v = u_0 + \varepsilon w$ for some fixed $\varepsilon > 0$. Thus $|v - u_0|_{1,\Omega} = \varepsilon > 0$. Now, taking into account (5.5.5) - where u_0 takes the place of u - a simple calculation yields

$$J(v) - J(u_0) = \frac{\varepsilon^2}{2} - \varepsilon^2 \int_\Omega u_0 w^2 dx - \frac{\varepsilon^3}{3}\int_\Omega w^3 dx.$$

Since, $u_0 \leq 0$, we get

$$J(v) - J(u_0) \geq \frac{\varepsilon^2}{2} - \frac{\varepsilon^3}{3}\int_\Omega w^3 dx.$$

Further, since, for $n \leq 3$, we have the continuous inclusion of $H_0^1(\Omega)$ into $L^3(\Omega)$, and we finally get

$$J(v) - J(u_0) \geq \frac{\varepsilon^2}{2} - \frac{\varepsilon^3 C}{3}.$$

Thus, choosing $\varepsilon < 3/4C$, we get that

$$J(v) - J(u_0) \geq \frac{\varepsilon^2}{4} \tag{5.5.7}$$

for all v such that $|v - u_0|_{1,\Omega} = \varepsilon$.

Let us now consider $z = t\varphi_1$ where $\varphi_1 \in H_0^1(\Omega)$ is the positive and normalized eigenfunction corresponding to the first eigenvalue of the Laplacian in Ω with Dirichlet boundary conditions, *i.e.*

$$\left.\begin{array}{rcll} -\Delta\varphi_1 &=& \lambda_1\varphi_1 & \text{in } \Omega \\ \varphi_1 &>& 0 & \text{in } \Omega \\ \varphi_1 &=& 0 & \text{on } \partial\Omega \\ \int_\Omega \varphi_1^2 dx &=& 1 & \end{array}\right\} \tag{5.5.8}$$

and $\lambda_1 > 0$ is the first such real number such that a solution to (5.5.8) holds. (That $\varphi_1 > 0$ is a consequence of the strong maximum principle and this property does not hold for subsequent eigenfunctions, cf. Kesavan [13]). Then,

$$J(z) = \frac{t^2}{2}\lambda_1 - \frac{t^3}{3}\int_\Omega \varphi_1^3 dx - t\int_\Omega f\varphi_1 dx.$$

As $t \to +\infty$, $J(z) \to -\infty$ (in particular, J is not bounded below). Thus, we choose t large enough such that $J(z) < J(u_0) < \inf_{|v-u_0|_{1,\Omega}=\varepsilon} J(v)$. Then, all the hypotheses of the mountain pass theorem are satisfied and we have a critical point u such that $J(u) > J(u_0)$. Thus the problem (5.5.4) has at least two solutions when $f < 0$ and is smooth enough.■

Several generalizations of the mountain pass theorem are available.

Theorem 5.5.2 *(Rabinowitz [21]) Let E be a Banach space and let $E_1 \subset E$ be a finite dimensional subspace. Let E_2 be a closed subspace such that $E = E_1 \oplus E_2$. Let $J : E \to \mathbb{R}$ be a $\mathcal{C}^1$ functional satisfying PS and such that $J(0) = 0$. Assume further that*
(i) there exist $R > 0$ and $a > 0$ such that if $u \in E_2$ and $\|u\| = R$, then, $J(u) \geq a$;
(ii) there exist $u_0 \in E_2$ such that $\|u_0\| = 1$ and real numbers $R_0 > R$, $R_1 > R$ such that $J(u) \leq 0$ for all $u \in \partial\Omega$ where

$$\Omega \;=\; \{u_1 + ru_0 \mid u_1 \in E_1,\; \|u_1\| \leq R_1,\; 0 \leq r \leq R_0\}$$

and $\partial\Omega$ is its boundary in $E_1 \oplus \mathbb{R}\{u_0\}$. Then J admits a critical value $c \geq a$ defined by

$$c \;=\; \inf_{A \in \mathcal{A}} \max_{v \in A} J(v)$$

where

$$\mathcal{A} \;=\; \{\varphi(\Omega) \mid \varphi \in \mathcal{C}(\Omega, E),\; \varphi(u) = u \;\; \text{for all} \;\; u \in \partial\Omega\}.$$

Proof: Observe, first of all, that since $\Omega \in \mathcal{A}$, the collection $\mathcal{A}$ is non-empty.

Step 1. Let $A \in \mathcal{A}$. We claim that there exists $u \in E_2 \cap A$ such that $\|u\| = R$. Let P denote the projection of E onto E_1. Then, looking for such an element u is the same as finding $u \in A$ such that $Pu = 0$ and $\|u - Pu\| = R$. Let $A = \varphi(\Omega)$. Define $F : E_1 \oplus \mathbb{R}\{u_0\} \to E_1 \oplus \mathbb{R}\{u_0\}$ by

$$F(x) \;=\; P\varphi(x) + \|\varphi(x) - P\varphi(x)\|u_0.$$

If $x \in \partial\Omega$, then $\varphi(x) = x$ and so $F(x) = x$. The search for u now amounts to finding $u \in \Omega$ such that $F(x) = Ru_0$ (so that we can set $u = \varphi(x)$).

Now, Ω is a cylinder and $\partial\Omega$ has either elements of the form u_1 or $u_1 + R_0u_0$ with $u_1 \in E_1$, $\|u_1\| \leq R_1$, or, elements of the form $u_1 + ru_2$ with $u_1 \in E_1$, $\|u_1\| = R_1 > R$. Thus, $Ru_0 \notin \partial\Omega$. Hence,

the (Brouwer) degree $d(F, \Omega, Ru_0)$ is well defined and, since $F = I$, the identity, on $\partial\Omega$, we have (cf. Proposiition 2.2.3)

$$d(F, \Omega, Ru_0) \;=\; d(I, \Omega, Ru_0) \;=\; 1$$

since $Ru_0 \in \Omega$. Thus, the degree being non-zero, we do have $x \in \Omega$ such that $F(x) = Ru_0$ and the claim is established.

Step 2. Thus, if $A \in \mathcal{A}$, we have $\max_{v \in A} J(v) \geq a$ by virtue of Step 1 and condition (i) in the hypotheses. Thus, $c \geq a > 0$. If c is not a critical value, we then apply the deformation lemma to find $\varepsilon_0 > 0$ and a flow η such that for all $0 < \varepsilon < \varepsilon_0$, we have $\eta(1, \{J \leq c + \varepsilon\}) \subset \{J \leq c - \varepsilon\}$. Choose $0 < \varepsilon_0 < a/2$ and for $0 < \varepsilon < \varepsilon_0$, choose $A \in \mathcal{A}$ such that $\max_{v \in A} J(v) \leq c + \varepsilon$. If $B = \eta(1, A)$, then $\max_{v \in B} J(v) \leq c - \varepsilon$. The proof of by contradiction will be then complete if we show that $B \in \mathcal{A}$.

Step 3. If $A = \varphi(\Omega)$, define $\psi(v) = \eta(1, \varphi(v))$ so that $B = \psi(\Omega)$. If $u \in \partial\Omega$, then $\varphi(u) = u$. Further, since $J(u) \leq 0$, by hypothesis, we have $u \notin \{c - \varepsilon_0 \leq J \leq c + \varepsilon_0\}$ and so $\eta(1, u) = u$. Thus, it follows that $\psi(u) = u$ on $\partial\Omega$ and so $B \in \mathcal{A}$ which completes the proof. ■

For applications of this result to semilinear elliptic boundary value problems, see Kavian [11] or Rabinowitz [21].

5.6 Multiplicity of Critical Points

In the preceding section, we saw applications of the min-max principle to produce critical points. A combination of the deformation lemma and the notion of the genus of symmetric and closed sets, introduced in Section 2.5, yields results on the existence of several critical points via the min-max principle.

Let E be a Banach space and, as before, let $\Sigma(E)$ denote the collection of subsets of E that are closed and symmetric (with respect to the origin), not containing the origin. Let $\gamma(A)$ denote the genus of a set $A \in \Sigma(E)$ (cf. Definition 2.5.1). For $j \geq 1$, we define

$$\Gamma_j = \{A \in \Sigma(E) \mid \gamma(A) \geq j\}.$$

If c is a critical value of a functional $J : E \to \mathbb{R}$, we set

$$K_c = \{u \in E \mid J(u) = c,\ J'(u) = 0\}.$$

Theorem 5.6.1 *(Clark [5]) Let E be a Banach space and let $J : E \to \mathbb{R}$ be a C^1 functional which is even, satisfies PS and is such that $J(0) = 0$. For $j \geq 1$, define*

$$c_j = \inf_{A\in\Gamma_j} \sup_{v\in A} J(v).$$

Then $\{c_j\}$ is a non-decreasing sequence and, if $-\infty < c_j < 0$, c_j is a critical value of J. Further, if, for some $j \geq 1$ and some $n \geq 1$, we have

$$c_j = c_{j+1} = \ldots = c_{j+n} < 0,$$

then, $\gamma(K_{c_j}) \geq n + 1$ and thus, in this case, there are an infinite number of critical points.

Proof: Step 1. The fact that the sequence $\{c_j\}$ is non-decreasing is a straightforward consequence of the fact that $\{\Gamma_j\}$ is a decreasing family of sets. Assume that $-\infty < c_j < 0$. If c_j is not a critical value, then let ε_0 and η be as in the deformation lemma. Let $0 < \varepsilon < \min\{\varepsilon_0, |c_j|/2\}$. Choose $A \in \Gamma_j$ such that $\sup_{v\in A} J(v) < c_j + \varepsilon < 0$ so that if $B = \eta(1, A)$, then $\sup_{v\in B} J(v) \leq c_j - \varepsilon < 0$. Then B does not contain the origin. Since J is even, we can choose $\eta(t, .)$ odd so that B is symmetric with respect to the origin. Thus, $B \in \Sigma(E)$ and, as $\eta(1, .)$ is an odd homeomorphism, $\gamma(B) = \gamma(A) \geq j$. Hence $B \in \Gamma_j$ and we immediately have a contradiction to the definition of c_j. Thus c_j is a critical value.

Step 2. Assume that $c_j = c_{j+1} = \ldots = c_{j+n} < 0$. Since J satisfies PS, K_{c_j} is compact and so its genus is finite. Let $\mathcal{N}$ be given by

$$\mathcal{N} = \{x \in E \mid \rho(x, K_{c_j}) < \tau\}.$$

If $\tau > 0$ is sufficiently small, then $K = \overline{\mathcal{N}}$ is also in $\Sigma(E)$ and $\gamma(K) = \gamma(K_{c_j})$ (cf. Theorem 2.5.1). Assume that $\gamma(K) \leq n$.

Step 3. Since J satisfies PS, we can find $\varepsilon_1 > 0$ and $\delta > 0$ (in fact $\delta \leq 1$) such that if $v \in \{J \leq c_j + \varepsilon_1\}\backslash(\{J < c_j - \varepsilon_1\} \cup \mathcal{N})$, then $\|J'(v)\| \geq \delta$. Now we set

$$\begin{aligned} A &= \{J \leq c_j - \varepsilon_0\} \cup K \cup \{J \geq c_j + \varepsilon_0\} \\ B &= \{c_j - \varepsilon \leq J \leq c_j + \varepsilon\}\backslash\mathcal{N} \end{aligned}$$

where $\varepsilon_0 = \min\{\varepsilon_1, \delta^2/8, |c_j|/2\}$ and $0 < \varepsilon < \varepsilon_0$. We set

$$\varphi(x) = \frac{\rho(x, A)}{\rho(x, A) + \rho(x, B)}.$$

Now we can proceed exactly as in the proof of the deformation lemma to get η odd and such that

$$\eta(1, (\{J \leq c_j + \varepsilon\}\backslash\mathcal{N})) \subset \{J \leq c_j - \varepsilon\}.$$

Step 4. Now choose $A \in \Gamma_{j+n}$ such that

$$c_j = c_{j+n} \leq \sup_{v\in A} J(v) \leq c_j + \varepsilon.$$

Then (cf. Theorem 2.5.1)

$$\gamma(\overline{A\backslash K}) \geq \gamma(A) - \gamma(K) \geq j + n - n.$$

Set $B = \eta(1, \overline{A\backslash K})$. Since $\sup_{v\in B} J(v) \leq c_j - \varepsilon < 0$, it follows that $0 \notin B$ and, since $\eta(1, .)$ is odd, $B \in \Sigma(E)$. Further, as $\eta(1, .)$ is a homeomorphism, $\gamma(B) = \gamma(\overline{A\backslash K}) \geq j$. Thus, $B \in \Gamma_j$ and we have a contradiction to the definition of c_j. Thus $\gamma(K) \geq n + 1$ and the same is true for $\gamma(K_{c_j})$.■

Example 5.6.1 Let $n \leq 3$ and let $\Omega \subset \mathbb{R}^n$ be a bounded domain. Let $\lambda \in \mathbb{R}$. Define the functional $J : H_0^1(\Omega) \to \mathbb{R}$ by

$$J(v) = \frac{1}{2}\int_\Omega |\nabla v|^2 dx + \frac{1}{4}\int_\Omega v^4 dx - \frac{\lambda}{2}\int_\Omega v^2 dx.$$

We have already seen that this functional satisfies PS (cf. Exercise 5.3.1). We can also easily show that J is coercive and that it attains its minimum.

Let $0 < \lambda_1 < \lambda_2 \leq \lambda_3 \leq ... \leq \lambda_k \leq ... \to \infty$ denote the eigenvalues of the Laplacian in Ω with Dirichlet boundary conditions and let $\{\varphi_k\}$ denote the orthonormal basis (in $L^2(\Omega)$) of eigenfunctions, *i.e.*

$$\left.\begin{array}{rcll} -\Delta\varphi_k & = & \lambda_k\varphi_k & \text{in } \Omega \\ \varphi_k & = & 0 & \text{on } \partial\Omega \\ \int_\Omega |\varphi_k|^2 dx & = & 1. & \end{array}\right\}$$

We will show that, if $\lambda > \lambda_k$, then J has at least $2k$ critical points. Thus the problem:

$$\left.\begin{array}{rcll} -\Delta u + u^3 & = & \lambda u & \text{in } \Omega \\ u & = & 0 & \text{on } \partial\Omega \end{array}\right\}$$

will have at least $2k$ (nontrivial) solutions. (In fact, $u = 0$ is always a solution and it is easy to see that if $\lambda \leq \lambda_1$, then it is the only solution. The critical points that we shall obtain will all be nontrivial since the associated critical values will be strictly negative.)

Define c_j as in Theorem 5.6.1. Since J is bounded below, the c_j are all real valued.We will show that if $\lambda > \lambda_k$, then $c_j < 0$ for $1 \leq j \leq k$ and will thus be critical values. Further, since J is even, if u is a nontrivial critical point, so will be $-u$. Hence, if the c_j, $1 \leq j \leq k$ are all distinct, we get at least $2k$ critical points. If they are not distinct, then, by the preceding theorem, we have, in fact, infinitely many critical points.

Let $V_j = \text{span}\{\varphi_1, \varphi_2, ..., \varphi_j\}$. Let S_ε denote the sphere of radius ε in V_j. Then $S_\varepsilon \in \Sigma(E)$ and $\gamma(S_\varepsilon) = j$ (cf. Example 2.5.1). Thus, $S_\varepsilon \in \Gamma_j$. Let $v = \sum_{i=1}^j \alpha_i\varphi_i \in S_\varepsilon$. Then

$$\begin{array}{rcl} J(v) & = & \frac{1}{2}\sum_{i=1}^j(\lambda_i - \lambda)\alpha_i^2 + \frac{1}{4}\int_\Omega |\sum_{i=1}^j \alpha_i\varphi_i|^4 dx \\ & \leq & \frac{\varepsilon^2}{2}(\lambda_j - \lambda) + C\varepsilon^4. \end{array}$$

If $\lambda > \lambda_k$, then, for sufficiently small $\varepsilon > 0$, we have $\sup_{v\in S_\varepsilon} J(v) < 0$ and so $c_j < 0$ for $1 \leq j \leq k$.■

Using the notion of the genus, it is also possible to prove multiplicity results for critical points obtained via the mountain pass theorem or its variants. We give below two such results corresponding to Theorems 5.5.1 and 5.5.2. For their proofs and applications, see, for example, Kavian [11].

Theorem 5.6.2 *Let E be an infinite dimensional Banach space and let $J : E \to \mathbb{R}$ be a C^1 functional satisfying PS. Assume that J is even and, further, that:*
(i) $J(0) = 0$ and there exist $R > 0$, $a > 0$ such that, for all $\|u\| = R$, we have $J(u) \geq a$;
(ii) If $E_1 \subset E$ is a finite dimensional subspace, then the set $\{u \in E_1 \mid J(u) \geq 0\}$ is bounded.
Then J admits an unbounded sequence of critical values. ■

Theorem 5.6.3 *Let E be a Banach space. let E_0 be a finite dimensional subspace and E_2 a closed subspace such that $E = E_0 \oplus E_2$. Let $J : E \to \mathbb{R}$ be a C^1 functional satisfying PS and such that $J(0) = 0$. Assume, further, that:*
(i) there exist $R > 0$, $a > 0$ such that for all $v \in E_2$ with $\|v\| = R$, we have $J(v) \geq a$;
(ii) If $E_1 \subset E$ is a finite dimensional subspace, then $\{v \in E_1 \mid J(v) \geq 0\}$ is bounded.
Then J admits an unbounded sequence of critical values. ■

5.7 Critical Points with Constraints

In Section 4.7, we saw that by maximizing a given functional over a set of constraints, we got new solutions to the nonlinear equation in question which led to a bifurcation result. Just as in the unconstrained case, we can have several critical points which do not correspond to extremal values of the functional. Let us consider a very simple example.

Let A be an $n \times n$ real and symmetric matrix. Then it has n real eigenvalues which can therefore be numbered in increasing order as

$$\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n.$$

Let $\{u_k\}_{k=1}^n$ be an orthonormal basis of eigenvectectors. Thus, $Au_k = \lambda_k u_k$ and $(u_k, u_l) = \delta_{kl}$, $1 \leq k, l \leq n$, where $(.,.)$ denotes the inner product in $\mathbb{R}^n$. Let $\|.\|$ denote the corresponding (Euclidean) norm. If V_k is the subspace spanned by the first k eigenvectors $\{u_1, u_2, ..., u_k\}$ for $1 \leq k \leq n$, we have the following characterization of the eigenvalues which is well known.

Theorem 5.7.1 *With the preceding notations, the eigenvalues of A are characterized as follows:*

$$\begin{aligned} \lambda_k &= \max_{v \in V_k, \|v\|=1}(Av, v) \\ &= \min_{v \perp V_{k-1}, \|v\|=1}(Av, v) \\ &= \min_{\dim W = k} \max_{v \in W, \|v\|=1}(Av, v). \blacksquare \end{aligned}$$

The proof of this result is a simple exercise and is left to the reader. In particular, we have

$$\lambda_1 = \min_{\|v\|=1}(Av, v) \quad \text{and} \quad \lambda_n = \max_{\|v\|=1}(Av, v).$$

Thus, the least and the greatest eigenvalues are the extrema of the functional (Av, v) on the set of constraints which is the unit sphere in $\mathbb{R}^n$. The other eigenvalues are, in the sense of Definition 5.7.1 below, other critical values of the functional on this same set of constraints and the eigenvectors are the critical points.

Definition 5.7.1 *Let E be a Banach space and let $F : E \to \mathbb{R}$ be a C^1 functional which defines the set of constraints by*

$$S = \{v \in E \mid F(v) = 0\}. \tag{5.7.1}$$

Assume that

$$F'(v) \neq 0 \;\; \textit{for every} \;\; v \in S. \tag{5.7.2}$$

Let $J : E \to \mathbb{R}$ be a C^1 functional. Then $c \in \mathbb{R}$ is called a **critical value** *of J on S if there exists $u \in S$ and $\lambda \in \mathbb{R}$ such that $J(u) = c$ and $J'(u) = \lambda F'(u)$. The point u is called a* **critical point** *of J on S.*$\blacksquare$

As observed in Remark 1.4.3, extremal points of J over S are critical points. As seen by the example of a real symmetric matrix A, all eigenvectors of A are critical points of the functional $v \mapsto (Av, v)$ on the unit sphere of $\mathbb{R}^n$ and the corresponding eigenvalues turn out to be the critical values. In general, of course, the λ in the above definition is just a Lagrange multiplier.

Exercise 5.7.1 Let E be a Banach space and let F and J be C^1 functionals on E. Let S be defined by (5.7.1) and let (5.7.2) hold. Assume that F is weakly sequentially continuous. Assume that J is bounded below, is weakly sequentially l.s.c. and further that

$$\lim_{v \in S, \|v\| \to \infty} J(v) = +\infty.$$

Show that J attains a minimum on S.$\blacksquare$

Definition 5.7.2 *Let E be a Banach space and let F and J be C^1 functionals on E. Let S be defined by (5.7.1) and let (5.7.2) hold. Let $c \in \mathbb{R}$. We say that J satisfies the* **Palais - Smale condition** *(PS) at level c on S if for every sequence $\{(u_n, \lambda_n)\}$ in $S \times \mathbb{R}$ such that*

$$J(u_n) \to c \;\; \textit{and} \;\; J'(u_n) - \lambda_n F'(u_n) \to 0 \;\; (\textit{in } E'),$$

there exists a convergent subsequence with limit $(u, \lambda) \in S \times \mathbb{R}$.$\blacksquare$

As usual, we will say that J satisfies PS on S if it satisfies PS at all levels $c \in \mathbb{R}$.

Example 5.7.1 Let $E = H^1(\mathbb{R}^n)$. Let $F(v) = \int_{\mathbb{R}^n} |v|^2 dx - 1$. Thus

$$S = \{v \in H^1(\mathbb{R}^n) \mid \int_{\mathbb{R}^n} |v|^2 dx = 1\}.$$

Let $J(v) = \int_{\mathbb{R}^n} |\nabla v|^2 dx$. Then, J does not satisfy PS over S. Indeed, let $\varphi \in \mathcal{D}(\mathbb{R}^n) \cap S$. Define

$$u_m(x) \;=\; m^{-\frac{n}{2}}\varphi(\frac{x}{m}).$$

Then $u_m \in S$. Further $J(u_m) = m^{-2}\int_{\mathbb{R}^n} |\nabla\varphi|^2 dx \to 0$. If $J'(u_m) - \lambda_m u_m \to 0$ in $H^{-1}(\mathbb{R}^n)$, it follows that $2\int_{\mathbb{R}^n} |\nabla u_m|^2 dx - \lambda_m \to 0$ which implies that $\lambda_m \to 0$. Thus, if $(u_m, \lambda_m) \to (u, \lambda) \in S \times \mathbb{R}$, we must have $\lambda = 0$ and so, $J'(u) = \lambda u = 0$. Since $u \in S$, it follows that it is a non-zero constant, which is impossible. ■

Example 5.7.2 Let $\Omega \subset \mathbb{R}^n$ be a bounded domain. Let $E = H_0^1(\Omega)$. Let $J(v) = \int_\Omega |\nabla v|^2 dx$ and let $F(v) = \int_\Omega |v|^2 dx - 1$. Thus,

$$S \;=\; \{v \in H_0^1(\Omega) \mid \int_\Omega |v|^2 dx = 1\}.$$

Then, J satisfies PS on S. If $(u_m, \lambda_m) \in S \times \mathbb{R}$ such that $J(u_m) \to c$ and $J'(u_m) - \lambda_m u_m \to 0$ in $H^{-1}(\Omega)$, it follows, from Poincaré's inequality, that $\{u_m\}$ is bounded in $H_0^1(\Omega)$ and hence, for a subsequence, $u_m \rightharpoonup u$ weakly in $H_0^1(\Omega)$ and so, by Rellich's theorem, strongly in $L^2(\Omega)$. Thus, $u \in S$. Further, if $h_m = J'(u_m) - \lambda_m u_m$, we have, for every $w \in H_0^1(\Omega)$,

$$< h_m, w > \;=\; 2\int_\Omega \nabla u_m.\nabla w dx - 2\lambda_m \int_\Omega u_m w dx \qquad (5.7.3)$$

where $< .,. >$ denotes the duality bracket between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$. Since $h_m \to 0$ in $H^{-1}(\Omega)$, we have that λ_m is bounded and so, for a further subsequence, $\lambda_m \to \lambda$. Thus, $(u, \lambda) \in S \times \mathbb{R}$ and

$$\int_\Omega \nabla u.\nabla w dx - \lambda \int_\Omega u w dx \;=\; 0 \qquad (5.7.4)$$

for all $w \in H_0^1(\Omega)$. In particular, $\lambda = \int_\Omega |\nabla u|^2 dx$. On the other hand, setting $w = u_m$ in (5.7.3), and passing to the limit, we get

$$\lim_{m\to\infty} \int_\Omega |\nabla u_m|^2 dx - \lambda \;=\; 0.$$

Thus, $u_m \to u$ strongly in $H_0^1(\Omega)$ and J satisfies PS over S.■

Example 5.7.3 Let H be a separable Hilbert space and let L and A satisfy the conditions laid out in Section 4.7. Let $J(v) = \frac{1}{2}\|v\|^2 + \frac{1}{4}(A(v), v)$ and let $F(v) = \frac{1}{2}(Lv, v) - r$ for some $r > 0$. Thus, let

$$S \;=\; \{v \in H \mid (Lv, v) = 2r\}.$$

Then J satisfies PS over S. Let $J(u_n) \to c$ and let $J'(u_n) - \lambda_n F'(u_n) \to 0$ in H. By definition of J, it follows immediately that $\{u_n\}$ is bounded in H and so, for a subsequence, $u_n \rightharpoonup u$ weakly in H. Since L is compact, it then follows that $u \in S$. Setting $h_n = u_n + A(u_n) - \lambda_n L u_n$ which tends to zero in H, we get

$$\|u_n\|^2 + (A(u_n), u_n) - \lambda_n(Lu_n, u_n) \;=\; (h_n, u_n) \qquad (5.7.5)$$

which can be rewritten as

$$2J(u_n) + \frac{1}{2}(A(u_n), u_n) - 2r\lambda_n \;=\; (h, u_n).$$

By hypotheses and the compactness of A, we deduce that $\lambda_n \to \lambda$ and $u + A(u) - \lambda L u = 0$. Thus,

$$\|u\|^2 + (A(u), u) - \lambda(Lu, u) \;=\; 0$$

while passing to the limit in (5.7.5) yields

$$\lim_{n\to\infty} \|u_n\|^2 + (A(u), u) - \lambda(Lu, u) \;=\; 0.$$

Thus $u_n \to u$ strongly in H and we deduce that J satisfies PS on S.■

Exercise 5.7.2 Let H be a separable Hilbert space and let L and A be as in the preceding example. Let $J(v) = \frac{1}{2}(Lv, v)$ and let

$$S = \{v \in h \mid \frac{1}{2}\|v\|^2 + \frac{1}{4}(A(v), v) = r\}$$

for $r > 0$. Show that J satisfies PS on S at all levels $c > 0$ but not at $c = 0$.■

Let E be a Banach space and let $F, J : E \to \mathbb{R}$ be $\mathcal{C}^1$ functionals. Let S be defined by (5.7.1) and let (5.7.2) hold. For $v \in S$, define

$$\|J(v)\|_* = \sup\left\{ < J'(v), w > \;\middle|\; \begin{array}{c} w \in E, \|w\| = 1 \\ < F'(v), w >= 0 \end{array} \right\}.$$

Thus, $\|J'(v)\|_* = 0$ for $v \in S$ implies that $J'(v) = \lambda F'(v)$ (since $\mathrm{Ker}(F'(v)) \subset \mathrm{Ker}(J'(v))$), *i.e.* v will be a critical point of J on S. The 'norm' $\|J'(v)\|_*$ is the norm of the projection of $J'(v)$ onto the tangent plane to S at v. If E were reflexive, then $\|J'(v)\|_*$ will be attained at some point w_v which will also be unique if the norm on E were strictly convex. Thus, for a Hilbert space, w_v will be uniquely defined.

Definition 5.7.3 *Let E be a Banach space and let $F, J : E \to \mathbb{R}$ be $\mathcal{C}^1$ functionals. Let S be defined by (5.7.1) and let (5.7.2) hold. Let $u \in S$. Then $v \in S$ is said to be a* **pseudo-gradient** *of J at u tangent to S if*

$$\left.\begin{array}{rcl} \|v\| & \leq & 2\|J'(u)\|_* \\ < J'(u), v > & \geq & \|J'(u)\|_*^2 \\ < F'(u), v > & = & 0. \end{array}\right\} \quad (5.7.6)$$

Let $S_r = \{u \in S \mid J'(u) - \lambda F'(u) \neq 0 \text{ for all } \lambda \in \mathbb{R}\}$, the set of regular points of J on S. A map $V : S_r \to E$ is a **pseudo-gradient vector field** *of J tangent to S if V is locally Lipschitz on S_r and for each $u \in S_r$, $V(u)$ is a pseudo-gradient vector for J at u tangent to S.*■

Lemma 5.7.1 *Let E be a Banach space and let $F, J : E \to \mathbb{R}$ be $\mathcal{C}^1$ functionals. Let S be defined by (5.7.1) and let (5.7.2) hold. Let $F' : E \to E'$ be locally Lipschitz. Assume that J is not constant on S. Then there exists a pseudo-gradient vector field, V, for J tangent to S defined on a neighbourhood $\widetilde{S}_r$ of S_r. If, in addition, F and J are even, then $\widetilde{S}_r$ can be chosen to be symmetric with respect to the origin and V can be chosen to be odd.* ■

Lemma 5.7.2 *(Deformation Lemma) Let E, F and J be as in the preceding lemma and assume further that J satisfies PS on S. Let $c \in \mathbb{R}$ be such that it is not a critical value of J on S. Then, there exists $\varepsilon_0 > 0$ such that, for every $0 < \varepsilon < \varepsilon_0$, there exists a map $\eta : \mathbb{R} \times S \to S$ with the following properties:*
(i) For every $u \in S$, we have $\eta(0, u) = u$.
(ii) For every $t \in \mathbb{R}$ and for every $u \notin \{c - \varepsilon_0 \leq J \leq c + \varepsilon_0\}$, we have $\eta(t, u) = u$.
(iii) For every $t \in \mathbb{R}$, $\eta(t, .)$ is a homeomorphism of S into itself.
(iv) For every $u \in S$, the function $t \mapsto J(\eta(t, u))$ is decreasing.
(v) If $u \in \{J \leq c + \varepsilon\}$, then $\eta(1, u) \in \{J \leq c - \varepsilon\}$.
(vi) If J and F are even, then, for every $t \in \mathbb{R}$, the mapping $\eta(t, .)$ can be chosen to be odd. ■

For proofs of these results, see, for instance, Kavian [11]. As an immediate consequence, we have the following result.

Theorem 5.7.2 *(Min- Max Principle) Let E, F and J be as in the deformation lemma. Let $\mathcal{F}$ be a nonempty family of subsets of S. If $c \in \mathbb{R}$ is not a critical value of J on S, and if $\varepsilon > 0$ is sufficiently small, assume that for every $A \in \mathcal{F}$, we have $\eta(1, A) \in \mathcal{F}$, where η is the mapping constructed as in the deformation lemma. Set*

$$c_* = \inf_{A \in \mathcal{F}} \sup_{v \in A} J(v).$$

If $c_* \in \mathbb{R}$, then it is a critical value of J on S.■

Analogous to Theorem 5.6.1, we have the following result.

Theorem 5.7.3 *Let E be a Banach space and let $F, J : E \to \mathbb{R}$ and S be as in the deformation lemma. Assume that $0 \notin S$. Assume, in addition, that F and J are even. For each integer*

$k \geq 1$, set

$$\begin{aligned} \mathcal{F}_k &= \{A \in \Sigma(E) \mid A \subset S,\ \gamma(A) \geq k\} \\ c_k &= \inf_{A \in \mathcal{F}_k} \sup_{v \in A} J(v). \end{aligned}$$

Then,
(i) for each $k \geq 1$ such that $\mathcal{F}_k \neq \emptyset$ and $c_k \in \mathbb{R}$, we have that c_k is a critical value of J on S. Further, $c_k \leq c_{k+1}$ for all such k and if, for some integer $j \geq 1$, we have that $c_k = c_{k+j} \in \mathbb{R}$, then $\gamma(K_{c_k}) \geq j+1$, where

$$K_{c_k} = \{u \in S \mid J(u) = c_k,\ J'(u) = \lambda F'(u) \ \text{for some} \ \lambda \in \mathbb{R}\}. \tag{5.7.7}$$

(ii) If, for each $k \geq 1$, we have $\mathcal{F}_k \neq \emptyset$, and $c_k \in \mathbb{R}$, then

$$\lim_{k \to \infty} c_k = +\infty. \tag{5.7.8}$$

Proof: (i) That c_k is a critical value of J on S when $\mathcal{F}_k \neq \emptyset$ and $c_k \in \mathbb{R}$ follows immediately from the min-max principle (Theorem 5.7.2). That $c_k \leq c_{k+1}$ is obvious and the result on the genus of K_{c_k} when $c_k = c_{k+j}$ follows an argument similar to that used in the proof of Theorem 5.6.1, by modifying the proof of the deformation lemma. We omit the details and refer the reader to Kavian [11].

(ii) To prove (5.7.8), we observe, first of all, that the sequence $\{c_k\}$ cannot be stationary. For, in that case, there exists a $k \geq 1$ such that $c_k = c_{k+j}$ for all $j \geq 1$ and by the preceding arguments it will follow that $\gamma(K_{c_k}) \geq j+1$ for all j. Thus, $\gamma(K_{c_k}) = +\infty$ which is impossible since, thanks to the PS condition, K_{c_k} is compact and hence must have a finite genus. Thus, if (5.7.8) is false, the only possibility is that $c_k \to c$ where $c > c_k$ for all k.

In this case, we set

$$K = \{u \in S \mid c_1 \leq J(u) \leq c,\ J'(u) = \lambda F'(u) \ \text{for some} \ \lambda \in \mathbb{R}\}$$

which, again, thanks to PS, is compact. Let $\gamma(K) = n$. Once again ,by an argument on the lines of the proof of the deformation lemma, we can find, for some $\varepsilon > 0$, a k such that $c_k > c - \varepsilon$,

and a set $A \in \mathcal{F}_k$ for which $\sup_{v \in A} J(v) \leq c + \varepsilon$ and such that $M = \eta(1, A) \in \mathcal{F}_k$. Thus, while $\sup_{v \in M} J(v) \geq c_k$, we also have $M \subset \{J \leq c - \varepsilon\}$, which is impossible. ■

Example 5.7.4 Let $E = \mathbb{R}^n$ and $J : \mathbb{R}^n \to \mathbb{R}$ a C^1 functional which is even. Let $S = S^{n-1}$ be the unit sphere in $\mathbb{R}^n$. Then if

$$\mathcal{F}_k = \{B \subset S^{n-1} \mid B \in \Sigma(\mathbb{R}^n),\ \gamma(B) \geq k\},$$

we have that $\mathcal{F}_k \neq \emptyset$ for $1 \leq k \leq n$ while $\mathcal{F}_{n+1} = \emptyset$. Thus, since the sets involved are all compact, we can replace the sup by a max and deduce that the values

$$c_k = \inf_{B \in \mathcal{F}_k} \max_{v \in B} J(v)$$

are all real for $1 \leq k \leq n$ and are critical values of J on S. Thus, there exist at least n pairs $(u_k, \lambda_k) \in S^{n-1} \times \mathbb{R}$ such that

$$J'(u_k) = \lambda_k u_k,\ ,1 \leq k \leq n.$$

Obviously, the pairs $(-u_k, \lambda_k)$, $1 \leq k \leq n$ also have the same property. This is the theorem of Lyusternik and Schnirelmann (originally proved using the notion of the category, cf. Section 2.5).

If $J(v) = \frac{1}{2}(Av, v)$ where A is a real symmetric $n \times n$ matrix, then we get the existence of n eigenvalues of A via the c_k. If we now set

$$\mathcal{D}_k = \{W \mid W \ \text{a subspace of} \ \mathbb{R}^n,\ \dim W = k\}$$

and

$$\mu_k = \inf_{W \in \mathcal{D}_k} \max_{v \in W \cap S} J(v),$$

we can see that $\mu_k = c_k$. For, if $W \in \mathcal{D}_k$, then $\gamma(W \cap S) = k$ and so $W \cap S \in \mathcal{F}_k$ so that $c_k \leq \mu_k$. On the other hand, the span V_k of the first k eigenvectors is of dimension k and it is easily seen that

$$c_k = \max_{v \in V_k \cap S} J(v) \geq \mu_k.$$

For, $Au_k = \lambda_k u_k$, $c_k = (Au_k, u_k)/2$ and so $c_k = \lambda_k/2$. Further, it is easy to check that $\lambda_k = \max_{v \in V_k, \|v\|=1}(Av, v)$. Thus, $c_k = \mu_k = \frac{1}{2}\lambda_k$, where λ_k is the k-th eigenvalue of A in increasing order and we recover the result of Theorem 5.7.1. ■

The conditions on the sets defining the family $\mathcal{F}_k$ can vary depending on the problem that we have at hand. For instance, it could be stipulated (in infinite dimensions) that the sets are, in addition, compact or are exactly of genus k and so on. In this spirit we have the following examples.

Example 5.7.5 Let $\Omega \subset \mathbb{R}^n$ be a bounded domain. Let $A(x)$ be a symmetric matrix of order n whose coefficients, $a_{ij}(x)$ are in $L^\infty(\Omega)$. Assume that there exist constants $0 < \alpha < \beta$ such that, for all $\xi \in \mathbb{R}^n$ and almost all $x \in \Omega$, we have

$$\alpha|\xi|^2 \leq (A(x)\xi, \xi), \leq \beta|\xi|^2. \tag{5.7.9}$$

Let

$$S = \{v \in H_0^1(\Omega) \mid \int_\Omega |v|^2 dx = 1\}.$$

Then, if $J(v) = \frac{1}{2}\int_\Omega A\nabla v.\nabla v dx$, it is easy to see that J satisfies PS on S (cf. Example 5.7.2). If we set

$$\mathcal{F}_k = \{B \in \Sigma(H_0^1(\Omega)) \mid B \subset S,\ B \text{ compact and } \gamma(B) \geq k\},$$

then, since S is the unit sphere in an infinite dimensional Hilbert space, it follows that $\mathcal{F}_k \neq \emptyset$ for each $k \geq 1$. Also, by the compactness imposed on the sets B in the above definition, and the condition (5.7.9), it follows that $c_k \in \mathbb{R}$ where

$$c_k = \inf_{B \in \mathcal{F}_k} \sup_{v \in B} J(v).$$

Thus, the c_k are critical values of J on S for all $k \geq 1$ and, in fact, $c_k \to \infty$. If we set

$$\mathcal{D}_k = \{W \mid W \text{ subspace of } H_0^1(\Omega),\ \dim W = k\},$$

then, exactly as in the previous example, we can show that

$$c_k = \inf_{W \in \mathcal{D}_k} \max_{v \in W \cap S} J(v).$$

The values $\lambda_k = 2c_k$ and the corresponding critical points u_k are such that

$$\int_\Omega A(x)\nabla u_k.\nabla v dx, = \lambda_k \int_\Omega u_k v dx$$

for all $v \in H_0^1(\Omega)$ which is the variational formulation of the problem

$$\left.\begin{aligned} -\text{div}(A\nabla u) &= \lambda u \quad \text{in } \Omega \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned}\right\}$$

The characterization of the eigenvalues via the min-max principle based on k - dimensional subspaces is known in the literature as the Courant-Fischer principle. ■

In the same way, we can prove the following generalization of the Lyusternik-Schnirelmann theorem to infinite dimensions.

Theorem 5.7.4 *Let H be an infinite dimensional Hilbert space and let $J : H \to \mathbb{R}$ be a C^1 functional which is even and which is non-constant and bounded below on the unit sphere $S = \{v \in H \mid \|v\|^2 = 1\}$. Set*

$$\begin{aligned} \mathcal{F}_k &= \{A \in \Sigma(H) \mid A \subset S,\ A \text{ compact and } \gamma(A) \geq k\} \\ c_k &= \inf_{A \in \mathcal{F}_k} \sup_{v \in A} J(v). \end{aligned}$$

Then, for each $k \geq 1$, c_k is a critical value of J on S; $c_k \leq c_{k+1}$ and if, for some $j \geq 1$, we have $c_k = c_{k+j}$, then $\gamma(K_{c_k}) \geq j+1$, where K_{c_k} is defined via (5.7.7). Finally,

$$\lim_{k \to \infty} c_k = +\infty. ■$$

Example 5.7.6 Returning once more to the problem considered in Section 4.7, we can consider the functional J and the set S of Example 5.7.3. We then saw that J satisfies PS on S. Thus, as

J and F are even, we can deduce that, for each $r > 0$, there is a sequence of values $\lambda_k^r \to \infty$ as $k \to \infty$ and $u_k^r \in S$ such that

$$u_k^r - \lambda_k^r L u_k^r + A(u_k^r) = 0.$$

If we show that the λ_k^r are bounded with respect to r for each fixed k, then it follows from the relation

$$\|u_k^r\|^2 + (A(u_k^r), u_k^r) = \lambda_k^r (L u_k^r, u_k^r) = 2r\lambda_k^r$$

that $u_k^r \to 0$ as $r \to 0$. Thus if one can show that $\lambda_k^r \to \lambda_k$, the k-th characteristic value of L, as $r \to 0$, we would have proven that $(0, \lambda_k)$ is a bifurcation point for each characteristic value λ_k of L. This is the spirit of the work of Berger [2] and Berger and Fife [3] cited earlier (cf. Remark 4.7.1), who, however, use the notion of the category. ■

Bibliography

[1] Ambrosetti, A. and Rabinowitz, P. H. Dual variational methods in critical point theory and applications, *J. Funct. Anal.*, **14**, 1973, pp.349 - 381.

[2] Berger, M. S. On von Kármán's equation and the buckling of a thin elastic plate, I, *Comm. Pure Appl. Math.*, **20**, 1967, pp. 687 - 719.

[3] Berger, M. S. and Fife, P. C. On von Kármán's equation and the buckling of a thin elastic plate, II, *Comm. Pure Appl. Math.*, **21**, 1968, pp. 227 - 247.

[4] Cartan, H. **Differential Calculus**, Hermann, Paris, 1971.

[5] Clark, D. C. A variant of the Lyusternik - Schnirelmann theory, *Indiana Univ. Math. J.*, **22** 1972, pp. 65 - 74.

[6] Crandall, M. G. and Rabinowitz, P. H. Bifurcation from simple eigenvalues, *J. Funct. Anal.*, **8**, No. 2, 1971, pp.321 - 340.

[7] Deimling, K. **Nonlinear Functional Analysis**, Springer - Verlag, 1985.

[8] Dieudonné, J. **Foundations of Modern Analysis**, Academic Press, 1969.

[9] Ekeland, I. On the variational principle, *J. Math. Anal. Appl.*, **47**, 1974, pp.324 - 353.

[10] Joshi, M. C. and Bose, R. K. **Some Topics in Nonlinear Functional Analysis**, Wiley - Eastern, New Delhi, 1985.

[11] Kavian, O. **Introduction à la Théorie des Points Critiques et Applications aux Problèmes Elliptiques**, Mathématiques & Applications, **13**, Springer - Verlag, 1993.

[12] Kesavan, S. La méthode de Kikuchi appliquée aux équations de von Karman, *Numer. Math.*, **32**, 1979, pp. 209 - 232.

[13] Kesavan, S. **Topics in Functional Analysis and Applications**, Wiley-Eastern Ltd., New Delhi, 1989.

[14] Krasnoselsk'ii, M. A. **Topological Methods in the Theory of Nonlinear Integral Equations**, Pergamon Press, Oxford, 1964.

[15] Limaye, B. V. **Functional Analysis**, Second Edition, New Age International Limited, 1996.

[16] Lions, J.-L. **Quelques Méthodes de Résolution des Problèmes aux Limites Nonlinéaires**, Dunod Gauthier-Villars, 1969.

[17] Lyusternik, L. and Schnirelmann, L. **Méthodes Topologiques dans les Problèmes Variationnels**, Hermann, Paris, 1934.

[18] Milnor, J. **Morse Theory**, Princeton University Press, 1963.

[19] Nirenberg, L. **Topics in Nonlinear Functional Analysis, CLN 6**, AMS, 2001.

[20] Rabinowitz, P. H. Théorie du degré topologique et applications à des problèmes aux limites non linéaires, Notes by H. Berestycki, Publications du Laboratoire d' Analyse Numérique, Université de Paris VI, **75010**, 1975.

[21] Rabinowitz, P. H. Some critical point theorems and applications to semilinear elliptic partial differential equations, *Annali Scuola Norm. Sup. Pisa*, Classe di scienze, Serie IV, **2**, 1978, pp. 215 - 223.

[22] Sard, A. The measure of critical values of differentiable maps, *Bull. Amer. Math. Soc.*, **45**, 1942, pp.883-890.

[23] Schwartz, J. T. **Nonlinear Functional Analysis**, Gordon and Breach, New York, 1969.

[24] Stakgold, I. Branching of solutions of nonlinear equations, *SIAM Review*, **13**, 1971, pp. 289 - 332.

[25] Sunder, V. S. **Functional Analysis - Spectral Theory**, TRIM **13**, Hindustan Book Agency (India), 1997.

[26] Whitney, H. A function not constant on a connected set of critical points, *Duke Math. J.*, **1**, 1935, pp.514-517.

Index

bifurcation point, 94
Borsuk - Ulam theorem, 62
Borsuk's theorem, 60
Brouwer's theorem, 50

Carathéodory function, 4
category, 68, 167
coercive functional, 127
compact
    mapping, 69
    perturbation of identity, 70
constraints, 30, 160
convex functional, 126
Courant - Fischer principle, 169
Crandall - Rabinowitz theorem, 106
critical
    point, 29, 161
    value, 29, 161

deformation lemma, 147, 165
degree
    additivity, 45, 75
    Brouwer, 36, 42, 44
    excision, 48, 76
    homotopy invariance, 45, 75
    Leray - Schauder, 74
    product formula, 48, 78
derivative
    Fréchet, 2
    Gâteau, 2
    higher order, 18
    partial, 10
    second, 16

Ekeland variational principle, 140
Euler's equation, 30

Fredhölm operator, 97

Galerkin method, 54
genus, 64, 156, 166

hemi-continuous, 54, 131

implicit function theorem, 22
inverse function theorem, 25
isolated solution, 82
    index, 82

Jacobian, 10

Kneser - Hukuhara theorem, 91

Krasnoselsk'ii - Perov theorem, 88
Krasnoselsk'ii's theorem, 111
Ky Fan - von Neumann theorem, 134

Lagrange multipliers, 33
Lipschitz continuous, 13
Lyapunov - Schmidt method, 97
Lyusternik - Schnirelmann theorem, 167, 169

maps of finite rank, 70
mean value theorem, 12, 19
min-max principle, 150, 165
Morse's lemma, 99
mountain pass theorem, 150

Nemytskii operator, 4

Palais - Smale condition, 138, 161
pseudo-gradient, 145, 164
    vector field, 146, 164

Rabinowitz' theorem, 115
regular value, 29
relative extremum, 29

saddle point, 132
Sard's theorem, 27
Schaeffer's theorem, 80
Schauder's theorem, 79
semi-continuous, 125
    weakly, 126
    weakly sequentially, 126

Taylor's formula, 19

variational inequality, 128